



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

**Συστήματα Συστάσεων για τη γεφύρωση του χάσματος  
δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης  
στον τομέα της τεχνολογίας**

Διδακτορική Διατριβή

**Ευάγγελος Καρακόλης**

Διδάκτορας Ε.Μ.Π.

Αθήνα, Νοέμβριος 2023

---





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

**Συστήματα Συστάσεων για τη γεφύρωση του χάσματος  
δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης  
στον τομέα της τεχνολογίας**

Διδακτορική Διατριβή

**Ευάγγελος Καρακόλης**

Διδάκτορας Ε.Μ.Π.

Συμβουλευτική επιτροπή: Δημήτριος Ασκούνης  
Ιωάννης Ψαρράς  
Χρυσόστομος Δούκας

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 31<sup>η</sup> Μαΐου 2023

.....  
Δ. Ασκούνης  
Καθηγητής Ε.Μ.Π.

.....  
Ι. Ψαρράς  
Καθηγητής Ε.Μ.Π.

.....  
Χ. Δούκας  
Καθηγητής Ε.Μ.Π.

.....  
Ε. Μαρινάκης  
Επικ. Καθηγητής Ε.Μ.Π.

.....  
Θ. Βαρβαρίγου  
Καθηγήτρια Ε.Μ.Π.

.....  
Γ. Μέντζας  
Καθηγητής Ε.Μ.Π.

.....  
Κ. Τσερπές  
Επικ. Καθηγητής Χαροκοπείου  
Πανεπιστημίου

.....

Ευάγγελος Καρακόλης

Διδάκτορας Ε.Μ.Π.

Copyright © Ευάγγελος Καρακόλης, 2023

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

---



## Περίληψη

Τα τελευταία χρόνια η τεχνολογία εξελίσσεται ραγδαία και, κατά συνέπεια, το ίδιο ισχύει και για τα σχετικά επαγγέλματα, όσον αφορά τις απαιτήσεις σε συγκεκριμένες τεχνολογικές δεξιότητες. Από την άλλη πλευρά, τα τεχνολογικά εκπαιδευτικά ιδρύματα συχνά δυσκολεύονται να συμβαδίσουν με τις ανάγκες της αγοράς εργασίας. Κατ' επέκταση, καταλήγουν να προσφέρουν γνώσεις και δεξιότητες που σε πολλές περιπτώσεις θεωρούνται ξεπερασμένες. Όσον αφορά τους φοιτητές και τους δια βίου μαθητές, ο στόχος τους είναι να αποφύγουν μαθήματα που προσφέρουν ξεπερασμένες γνώσεις και δεξιότητες, καθώς μακροπρόθεσμα στοχεύουν στην εύρεση κάποιας θέσης εργασίας, στην οποία θα αξιοποιούν τις γνώσεις και τις δεξιότητες που έχουν αποκτήσει. Ακόμη, όσον αφορά τα εκπαιδευτικά ιδρύματα μια σημαντική πρόκληση είναι να συμβαδίσουν με τις ανάγκες της αγοράς εργασίας παρέχοντας μαθήματα και δεξιότητες που θα είναι χρήσιμα στην καριέρα των φοιτητών τους.

Σε αυτό το πλαίσιο, αντικείμενο της διδακτορικής διατριβής είναι η γεφύρωση του χάσματος μεταξύ των διδασκόμενων γνώσεων και δεξιοτήτων στα εκπαιδευτικά ιδρύματα και των ζητούμενων γνώσεων και δεξιοτήτων από την αγορά εργασίας στον τομέα της τεχνολογίας, με τη χρήση σύγχρονων τεχνικών ανάλυσης δεδομένων όπως είναι η εξόρυξη γνώσης, τα συστήματα συστάσεων και η επεξεργασία φυσικής γλώσσας. Συγκεκριμένα, διερευνώνται σε βάθος εργασίες και λύσεις που έχουν αναπτυχθεί προς αυτή την κατεύθυνση και αναπτύσσονται συστήματα συστάσεων που απευθύνονται σε φοιτητές και δια βίου μαθητές για την επιλογή μαθημάτων και δεξιοτήτων που είναι σε υψηλή ζήτηση από την αγορά εργασίας, και σε εκπαιδευτικούς οργανισμούς, για την αξιολόγηση και βελτίωση του προγράμματος σπουδών τους με βάση τις ανάγκες της αγοράς εργασίας. Τα συγκεκριμένα συστήματα και η προτεινόμενη μεθοδολογία εφαρμόζονται στην ελληνική αγορά εργασίας στον τομέα της ανάπτυξης λογισμικού με αξιολογικά αποτελέσματα. Η προτεινόμενη μεθοδολογία και τα εργαλεία που αναπτύχθηκαν μπορούν εύκολα να χρησιμοποιηθούν με κατάλληλες τροποποιήσεις και σε άλλα επιστημονικά πεδία.

Κατά την ανάπτυξη των προαναφερθέντων συστημάτων προέκυψαν κάποια επιπλέον προβλήματα που επίσης αποτελούν αντικείμενο της διδακτορικής διατριβής. Συγκεκριμένα, ασχοληθήκαμε με το πρόβλημα της δικαιοσύνης μεταξύ παρόχων αντικειμένων σε συστήματα συστάσεων, όσον αφορά την κάλυψη των αντικειμένων κάθε παρόχου σε χρήστες και τη διαφορετικότητα μεταξύ αυτών. Το πρόβλημα αυτό αποτελεί NP-πλήρες πρόβλημα σύμφωνα με τη βιβλιογραφία. Σε αυτό το πλαίσιο, αναπτύχθηκε ευριστικός αλγόριθμος πολυωνυμικού χρόνου που βρίσκει λύσεις πολύ κοντά στη βέλτιστη, όπως φαίνεται από τα πειραματικά αποτελέσματα. Επιπλέον, πρόβλημα της παρούσας διατριβής αποτελεί και η λειτουργία των συστημάτων συστάσεων ως προς τον αντίκτυπο που έχει στους χρήστες, και συγκεκριμένα, κατά πόσο είναι ηθική. Σε αυτό το πλαίσιο, αναλύθηκαν διάφορα ευρέως διαδεδομένα συστήματα συστάσεων τα οποία κρίνονται ηθικά προβληματικά. Για το λόγο αυτό, προτείνονται ρυθμιστικές ενέργειες για την εξυγίανση των συστημάτων συστάσεων, ως προς τον ηθικό τους αντίκτυπο.

**Λέξεις Κλειδιά:** Εκπαίδευση, Αγορά Εργασίας, Συστήματα Συστάσεων, Αναντιστοιχία Δεξιοτήτων, Επεξεργασία Φυσικής Γλώσσας, Εξόρυξη Γνώσης, Δικαιοσύνη, Ηθική.

---



## Abstract

In the last few years technology evolves rapidly, and as a result, the same applies to technology-related professions, concerning the need for specific skills. On the other hand, technological educational institutions oftentimes have difficulties to keep track and align with the labour market needs. Consequently, they sometimes end up teaching skills that are considered obsolete. Regarding students and lifelong learners, an important objective is to avoid courses that teach such skills, as they aim at landing a job in which they will utilise the skills they have acquired during their studies. Furthermore, with regards to educational institutions, an important challenge is to comply with the labour market needs, by providing courses that can be useful for the career of their students.

In this context, this dissertation's main objective is to bridge the gap between the skills taught by educational institutions and the ones required by the labour market in the technology sector, through the use of state-of-the-art data analytics techniques, including data mining, recommender systems and natural language processing. Specifically, in this dissertation, several studies and solutions that have already been proposed are examined in depth. As a next step, a skill and course recommender system has been developed for students and lifelong learners, in order to help them select courses that will improve their skills in accordance with the labour market needs. Additionally, a curriculum redesign recommender system has been developed for professors and educational institutions, for the evaluation and improvement of their organisation's curriculum, to be compliant with the needs of the labour market. These recommender systems along with their methodology have been applied in the Greek labour market and the sector of software engineering, and the results prove to be very promising. The proposed methodology and services can be applied with little to no modifications to other sectors as well.

It is worth mentioning that during the design and development phase of the aforementioned services, several additional problems have emerged. These problems and their solutions are also part of this dissertation. Specifically, this dissertation also deals with the problem of fairness among providers in recommender systems in terms of coverage and diversity in users for each provider's items. This problem is modeled as an optimisation problem in literature, and it is an NP-complete problem. To this end, we developed a heuristic algorithm with polynomial computational complexity that finds solutions very close to the optimal one, as illustrated in the experimental results. In addition, the problem of recommender system operation and particularly their ethical implications to users is thoroughly examined in this dissertation. In this context, several widely known recommender systems were analysed and proved to be ethically problematic. To this end, we propose a series of regulatory actions. Their goal is the betterment of the recommender systems' function and the harmonisation of their progress with social prosperity.

**Keywords:** Education, Labour Market, Recommender Systems, Skills Mismatch, Natural Language Processing, Data Mining, Fairness, Ethics.

---





## Πρόλογος

Η παρούσα διδακτορική διατριβή έλαβε χώρα στη Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου, στο πλαίσιο των ερευνητικών δραστηριοτήτων του Εργαστηρίου Συστημάτων Αποφάσεων και Διοίκησης, το διάστημα 10/2019 – 09/2023.

Αντικείμενο της διατριβής είναι ο σχεδιασμός και ανάπτυξη κατάλληλων εργαλείων αναλυτικής δεδομένων, και συστημάτων συστάσεων για τη γεφύρωση του χάσματος γνώσεων και δεξιοτήτων μεταξύ της αγοράς εργασίας και της εκπαίδευσης στον τομέα της τεχνολογίας. Συγκεκριμένα, αναπτύσσονται μια σειρά από υπηρεσίες για τη συγκέντρωση και επεξεργασία δεδομένων με χρήση τεχνικών εξόρυξης γνώσης από δεδομένα και επεξεργασίας φυσικής γλώσσας, για την ανάπτυξη εφαρμογών συστημάτων συστάσεων. Ακόμη αναπτύσσονται 2 συστήματα συστάσεων, εκ των οποίων το πρώτο προτείνει μαθήματα και δεξιότητες σε φοιτητές και δια βίου μαθητές, και το δεύτερο προτείνει ενημερώσεις σε ένα δοθέν πρόγραμμα σπουδών σε διδάσκοντες μαθημάτων και διοικήσεις ακαδημαϊκών οργανισμών, ώστε να συμβαδίζουν με τις ανάγκες της αγοράς εργασίας. Επιπλέον, διερευνάται το πρόβλημα της δικαιοσύνης μεταξύ των παρόχων αντικειμένων στα συστήματα συστάσεων και προτείνεται ευριστική λύση, που παρουσιάζει ικανοποιητικά αποτελέσματα στα datasets που εφαρμόστηκε. Τέλος, η διατριβή διερευνά θέματα ηθικής στα συστήματα συστάσεων. Συγκεκριμένα, αξιολογεί μια σειρά από δημοφιλή συστήματα και προτείνει ρυθμιστικές ενέργειες για τη βελτίωση τους.

Η διατριβή σχεδιάστηκε, υλοποιήθηκε και ολοκληρώθηκε υπό την επίβλεψη και την καθοδήγηση του καθηγητή της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Ε.Μ.Π. Δημητρίου Ασκούνη. Στον κύριο Ασκούνη οφείλω ιδιαίτερες ευχαριστίες για την εμπιστοσύνη και τη στήριξη καθ' όλη τη διάρκεια της εκπόνησης της διατριβής. Επιπλέον, ευχαριστώ θερμά τον καθηγητή Ι. Ψαρρά, τον καθηγητή Χ. Δούκα, τον επίκουρο καθηγητή Ε. Μαρινάκη, την καθηγήτρια Θ. Βαρβαρίγου, τον καθηγητή Γ. Μέντζα και τον επίκουρο καθηγητή Κ. Τσερπέ που είχα την τιμή να παρευρεθούν στην εξέταση υποστήριξης της διατριβής μου.

Επιπροσθέτως, ευχαριστώ θερμά τους συναδέλφους και φίλους Π. Κοκκινάκο, Π. Καψάλη, Σ. Πελέκη, Χ. Κοντζίνο και Σ. Μουζακίτη για την αगाστή συνεργασία, την ανταλλαγή απόψεων, και την προθυμία τους για βοήθεια. Ακόμα, ευχαριστώ όλους τους συναδέλφους με τους οποίους συνεργαστήκαμε όλα τα χρόνια της παρουσίας μου στο Εργαστήριο Συστημάτων Αποφάσεων και Διοίκησης.

Τέλος θα ήθελα να ευχαριστήσω τους γονείς μου Θανάση και Σούλα και τον αδερφό μου Τραϊανό, για την στήριξη τους όλα αυτά τα χρόνια, όπως επίσης και τη σύζυγό μου Κατερίνα για την αμέριστη συμπαράσταση, πίστη, υποστήριξη και υπομονή προς το πρόσωπο μου καθ' όλη τη διάρκεια εκπόνησης της παρούσας διδακτορικής διατριβής.

Ευάγγελος Καρακόλης

Αθήνα, Σεπτέμβριος 2023

---

## Περιεχόμενα

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ .....	25
1.1    Ορισμός Προβλήματος .....	25
1.1.1    Πρόβλημα 1: Το χάσμα δεξιοτήτων μεταξύ αγοράς εργασίας και τεχνολογικής εκπαίδευσης .....	25
1.1.2    Πρόβλημα 2: Δικαιοσύνη μεταξύ των παρόχων αντικειμένων στα Συστήματα Συστάσεων .....	26
1.1.3    Πρόβλημα 3: Ηθικά προβληματικές εκφάνσεις Συστημάτων Συστάσεων .....	27
1.2    Θεματική Περιοχή .....	28
1.3    Συμβολή Διατριβής .....	29
1.4    Δομή διατριβής .....	31
1.5    Αναφορές .....	33
ΚΕΦΑΛΑΙΟ 2: ΑΝΑΣΚΟΠΗΣΗ ΥΦΙΣΤΑΜΕΝΗΣ ΚΑΤΑΣΤΑΣΗΣ – ΧΑΣΜΑ ΔΕΞΙΟΤΗΤΩΝ ΜΕΤΑΞΥ ΑΓΟΡΑΣ ΕΡΓΑΣΙΑΣ ΚΑΙ ΕΚΠΑΙΔΕΥΣΗΣ .....	35
2.1    Αγορά εργασίας στην τεχνολογία .....	36
2.2    Εκπαίδευση στην τεχνολογία .....	37
2.3    Κάλυψη των απαιτήσεων της αγοράς εργασίας σε ικανότητες μέσω της εκπαίδευσης .....	38
2.4    Το έργο QualiChain .....	41
2.5    Συμπεράσματα .....	43
2.6    Αναφορές .....	44
ΚΕΦΑΛΑΙΟ 3: ΑΝΑΣΚΟΠΗΣΗ ΥΦΙΣΤΑΜΕΝΗΣ ΚΑΤΑΣΤΑΣΗΣ – ΣΥΣΤΗΜΑΤΑ ΣΥΣΤΑΣΕΩΝ .....	47
3.1    Εισαγωγή .....	47
3.2    Το Πρόβλημα .....	48
3.3    Ενδιαφερόμενα Μέρη και Χρησιμότητα Συστημάτων Συστάσεων .....	49
3.4    Τεχνικές Συστάσεων .....	50
3.4.1    Σύσταση βασισμένη στο περιεχόμενο (Content based) .....	51
3.4.2    Συνεργατικό φιλτράρισμα (Collaborative filtering) .....	51
3.4.3    Δημογραφικά συστήματα (Demographic) .....	51
3.4.4    Σύσταση βασισμένη στη γνώση (Knowledge Based) .....	51
3.4.5    Σύσταση βασισμένη στην κοινότητα (Community based) .....	52
3.4.6    Υβριδική σύσταση (Hybrid) .....	52
3.5    Αξιολόγηση Συστημάτων Συστάσεων .....	52

---

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

3.5.1	Κάλυψη (Coverage) .....	53
3.5.2	Διαφοροποίηση (Diversity) .....	57
3.5.3	Δικαιοσύνη/Δικαιότητα (Fairness) .....	62
3.5.4	Novelty, Serendipity, Unexpectedness.....	63
3.5.5	Συσχετίσεις μεταξύ των διαφορετικών μετρικών .....	66
3.6	Κοινωνική διάσταση Συστημάτων Συστάσεων .....	69
3.7	Συμπεράσματα .....	71
3.8	Αναφορές .....	72
ΚΕΦΑΛΑΙΟ 4: Η ΠΛΑΤΦΟΡΜΑ QUALICHAIN .....		77
4.1	Η Αρχιτεκτονική της πλατφόρμας QualiChain .....	77
4.2	Βασικές υπηρεσίες της πλατφόρμας QualiChain.....	80
4.2.1	Υπηρεσία ταυτοποίησης και ελέγχου πρόσβασης (Identity and Access Control Management) .....	80
4.2.2	Η υπηρεσία DeTrusty .....	83
4.2.3	Υπηρεσία επαλήθευσης OpenBadges με βάση το blockchain.....	84
4.2.4	Υπηρεσία επαλήθευσης ακαδημαϊκών τίτλων με βάση το blockchain .....	90
4.2.5	Το υποσύστημα εξαγωγής γνώσης (Knowledge Extraction).....	91
4.2.6	Το υποσύστημα ανάλυσης δεδομένων και υποστήριξης αποφάσεων .....	98
4.2.7	Front end και οπτικοποίηση δεδομένων .....	106
4.3	Σύνοψη και συμπεράσματα .....	109
4.4	Αναφορές .....	109
ΚΕΦΑΛΑΙΟ 5: Η ΥΠΗΡΕΣΙΑ ΠΡΟΤΑΣΕΩΝ ΜΑΘΗΜΑΤΩΝ ΚΑΙ ΔΕΞΙΟΤΗΤΩΝ .....		111
5.1	Εισαγωγή .....	111
5.2	Μεθοδολογία .....	112
5.2.1	Εξαγωγή αγγελιών εργασίας μέσω της υπηρεσίας JobCrawler.....	113
5.2.2	Εξαγωγή γνώσης.....	113
5.2.3	Παροχή προτάσεων με τη χρήση της τεχνολογίας ElasticSearch .....	114
5.2.4	Παροχή προτάσεων με τη χρήση συσταδοποίησης (clustering) .....	115
5.2.5	Οι τελικές συστάσεις της υπηρεσίας .....	116
5.3	Ενδεικτικά αποτελέσματα.....	117
5.3.1	Ενδιάμεσα αποτελέσματα συστάσεων με τη χρήση της τεχνολογίας ElasticSearch .....	117
5.3.2	Αποτελέσματα συστάσεων με τη χρήση τεχνικών συσταδοποίησης .....	118
5.3.3	Προτάσεις με συνδυασμό των αποτελεσμάτων .....	121

---

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

5.4	Επίδειξη υπηρεσίας και διεπαφής χρήστη .....	122
5.5	Σχολιασμός και συμπεράσματα .....	124
5.6	Αναφορές .....	125
<b>ΚΕΦΑΛΑΙΟ 6: Η ΥΠΗΡΕΣΙΑ ΑΝΑΣΧΕΔΙΑΣΜΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ ΣΠΟΥΔΩΝ.....</b>		<b>127</b>
6.1	Εισαγωγή .....	127
6.2	Μεθοδολογία .....	129
6.2.1	Η υπηρεσία JobCrawler .....	130
6.2.2	Η υπηρεσία QualiChain Knowledge Extraction .....	130
6.2.3	Η υπηρεσία Association Rules Miner .....	131
6.2.4	Η αρχιτεκτονική της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών ..	131
6.3	Ενδεικτικά Αποτελέσματα.....	133
6.4	Επίδειξη υπηρεσίας και διεπαφής χρήστη .....	137
6.5	Σχολιασμός και συμπεράσματα .....	137
6.6	Αναφορές .....	138
<b>ΚΕΦΑΛΑΙΟ 7: Η ΥΠΗΡΕΣΙΑ ΠΟΛΥΚΡΙΤΗΡΙΑΚΗΣ ΥΠΟΣΤΗΡΙΞΗΣ ΑΠΟΦΑΣΕΩΝ .....</b>		<b>141</b>
7.1	Εισαγωγή .....	141
7.2	Πολυκριτηριακή Υποστήριξη Αποφάσεων και Συστήματα Συστάσεων.....	141
7.3	Μέθοδοι πολυκριτηριακής υποστήριξης αποφάσεων .....	143
7.3.1	Η μέθοδος MAUT .....	143
7.3.2	Η μέθοδος TOPSIS .....	144
7.3.3	Η μέθοδος ELECTRE I.....	146
7.3.4	Η μέθοδος PROMETHEE II .....	147
7.4	Η υπηρεσία πολυκριτηριακής υποστήριξης αποφάσεων.....	151
7.5	Παραδείγματα χρήσης της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων .....	154
7.5.1	Περίπτωση χρήσης 1 - επιλογή δεξιοτήτων και μαθημάτων.....	154
7.5.2	Περίπτωση χρήσης 2 - Εισαγωγή νέων δεξιοτήτων στο πρόγραμμα σπουδών 159	
7.6	Σύνοψη και συμπεράσματα .....	162
7.7	Αναφορές .....	163
<b>ΚΕΦΑΛΑΙΟ 8: ΧΡΗΣΗ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ ΤΩΝ ΥΠΗΡΕΣΙΩΝ ΠΟΥ ΑΝΑΠΤΥΧΘΗΚΑΝ.....</b>		<b>165</b>
8.1	Το πιλοτικό σενάριο χρήσης .....	165
8.1.1	Γενική παρουσίαση πιλότου .....	165
8.2	Ερευνητικές ερωτήσεις και KPIs.....	167

---

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

8.3	Εκτέλεση και Αποτελέσματα του πιλότου .....	170
8.3.1	Μεθοδολογία .....	170
8.3.2	Αποτελέσματα ερωτηματολογίων .....	175
8.3.3	Ανάλυση των KPIs.....	177
8.4	Πολύτιμα μαθήματα και επεκτάσεις .....	179
8.4.1	Συγκέντρωση βιογραφικών από φοιτητές και προστασία δεδομένων .....	179
8.4.2	Ενημέρωση και βελτιστοποίηση προγράμματος σπουδών.....	180
8.4.3	Προσωποποιημένες υπηρεσίες για τους φοιτητές.....	181
8.4.4	Πολυκριτηριακή υποστήριξη αποφάσεων.....	181
8.4.5	Η πλατφόρμα QualiChain .....	182
8.5	Συμπεράσματα .....	182
8.6	Αναφορές .....	183
ΚΕΦΑΛΑΙΟ 9: ΔΙΚΑΙΟΣΥΝΗ ΜΕΤΑΞΥ ΤΩΝ ΠΑΡΟΧΩΝ ΣΤΑ ΣΥΣΤΗΜΑΤΑ ΣΥΣΤΑΣΕΩΝ.....		185
9.1	Εισαγωγή .....	185
9.2	Υπόβαθρο και σχετικές εργασίες.....	187
9.2.1	Πολυμετοχικά Συστήματα Συστάσεων (MSRS) .....	187
9.2.2	Κάλυψη.....	188
9.2.3	Διαφοροποίηση.....	189
9.2.4	Δικαιοσύνη .....	189
9.2.5	Ορισμός προβλήματος .....	191
9.3	Η ευριστική λύση.....	193
9.3.1	Η βασική λύση (πρόβλημα χωρίς περιορισμούς) .....	193
9.3.2	Η λύση για την κάλυψη .....	193
9.3.3	Η γενική λύση .....	194
9.4	Πειραματική διαδικασία και αποτελέσματα .....	197
9.4.1	Το Dataset.....	197
9.4.2	Η βασική λύση (χωρίς περιορισμούς για κάλυψη και διαφοροποίηση).....	197
9.4.3	Πάροχοι αντικειμένων ή κατηγορίες .....	197
9.4.4	Η λύση για την κάλυψη .....	198
9.4.5	Η γενική λύση .....	199
9.5	Αξιολόγηση και επεκτάσεις.....	202
9.5.1	Αξιολόγηση.....	202
9.5.2	Νέος ορισμός για τη διαφοροποίηση και νέα ευριστική προσέγγιση .....	205

---

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

9.5.3	Απαντήσεις στα ερευνητικά ερωτήματα που τέθηκαν.....	207
9.6	Συμπεράσματα και μελλοντικές επεκτάσεις.....	208
9.7	Αναφορές .....	209
ΚΕΦΑΛΑΙΟ 10: ΗΘΙΚΕΣ ΕΠΙΠΤΩΣΕΙΣ ΣΤΑ ΣΥΣΤΗΜΑΤΑ ΣΥΣΤΑΣΕΩΝ .....		213
10.1	Εισαγωγή .....	213
10.2	Θεωρητικό υπόβαθρο και σχετικές εργασίες .....	214
10.3	Ηθικές προκλήσεις στα συστήματα συστάσεων.....	215
10.4	Το προτεινόμενο πλαίσιο αξιολόγησης .....	217
10.5	Αξιολόγηση δημοφιλών συστημάτων συστάσεων .....	218
10.6	Ρυθμιστικό πλαίσιο για την εύρυθμη λειτουργία των Συστημάτων Συστάσεων	220
10.7	Συμπεράσματα και μελλοντικές επεκτάσεις.....	223
10.8	Αναφορές .....	223
ΚΕΦΑΛΑΙΟ 11: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ .....		227
11.1	Σύνοψη .....	227
11.2	Συμπεράσματα .....	229
11.2.1	Πρόβλημα 1: Το χάσμα δεξιοτήτων μεταξύ αγοράς εργασίας και τεχνολογικής εκπαίδευσης.....	229
11.2.2	Πρόβλημα 2: Δικαιοσύνη μεταξύ των παρόχων αντικειμένων στα Συστήματα Συστάσεων.....	232
11.2.3	Πρόβλημα 3: Ηθικά προβληματικές εκφάνσεις Συστημάτων Συστάσεων .....	233
11.3	Μελλοντικές επεκτάσεις .....	233
11.3.1	Πρόβλημα 1: Το χάσμα δεξιοτήτων μεταξύ αγοράς εργασίας και τεχνολογικής εκπαίδευσης.....	233
11.3.2	Πρόβλημα 2: Δικαιοσύνη μεταξύ των παρόχων αντικειμένων στα Συστήματα Συστάσεων.....	234
11.3.3	Πρόβλημα 3: Ηθικά προβληματικές εκφάνσεις Συστημάτων Συστάσεων .....	235
ΠΑΡΑΡΤΗΜΑ Ι: ΔΗΜΟΣΙΕΥΜΕΝΟ ΕΡΓΟ .....		237
I.1	Δημοσιεύσεις σε διεθνή επιστημονικά περιοδικά .....	237
I.2	Παρουσιάσεις σε διεθνή επιστημονικά συνέδρια.....	237
I.3	Υποβληθείσες εργασίες για τις οποίες εκκρεμεί η απόφαση.....	239
ΠΑΡΑΡΤΗΜΑ ΙΙ: ΥΠΗΡΕΣΙΕΣ ΤΗΣ ΠΛΑΤΦΟΡΜΑΣ QUALICHAIN ΣΤΟ GITHUB ΚΑΙ ΑΛΛΕΣ ΥΠΗΡΕΣΙΕΣ ΛΟΓΙΣΜΙΚΟΥ ΠΟΥ ΑΝΑΠΤΥΧΘΗΚΑΝ ΣΤΗΝ ΠΑΡΟΥΣΑ ΔΙΑΤΡΙΒΗ .....		241
ΠΑΡΑΡΤΗΜΑ ΙΙΙ: ΕΠΙΔΕΙΞΗ ΤΗΣ ΠΛΑΤΦΟΡΜΑΣ QUALICHAIN.....		243

---

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

III.1	Λειτουργίες για φοιτητές, δια βίου μαθητές και ανθρώπους που αναζητούν εργασία	244
III.1.1	Λειτουργίες επισκόπησης και επιλογής μαθήματος	245
III.1.2	Λειτουργίες επισκόπησης και αίτησης σε αγγελίες εργασίας	248
III.1.3	Επισκόπηση και επιδότηση με smart badges	250
III.1.4	Λειτουργία προφίλ	253
III.1.5	Συμβουλευτική καριέρας	255
III.2	Λειτουργίες για καθηγητές και διδάσκοντες μαθημάτων	256
III.2.2	Σύγκριση διπλωμάτων (Degree Comparison)	257
III.2.3	Λειτουργία προφίλ	258
III.2.4	Λειτουργία ανάλυσης και ανασχεδιασμού προγράμματος σπουδών	259
III.3	Λειτουργίες για υπεύθυνους ανθρώπινου δυναμικού	261
III.3.1	Επιλογή υποψηφίου για θέση εργασίας	261
III.3.2	Έλεγχος εγκυρότητας ακαδημαϊκού τίτλου	261
ΠΑΡΑΡΤΗΜΑ IV: ΔΕΞΙΟΤΗΤΕΣ ΠΟΥ ΕΞΑΧΘΗΚΑΝ ΚΑΙ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗΝ ΑΝΑΛΥΣΗ		265
ΠΑΡΑΡΤΗΜΑ V: ΟΙ ΔΕΞΙΟΤΗΤΕΣ ΠΟΥ ΔΕ ΔΙΔΑΣΚΟΝΤΑΙ ΣΤΑ ΠΛΑΙΣΙΑ ΤΟΥ ΔΟΘΕΝΤΟΣ ΠΡΟΓΡΑΜΜΑΤΟΣ ΣΠΟΥΔΩΝ ΑΛΛΑ ΕΙΝΑΙ ΣΕ ΥΨΗΛΗ ΖΗΤΗΣΗ ΑΠΟ ΤΗΝ ΑΓΟΡΑ ΕΡΓΑΣΙΑΣ ...		267
ΠΑΡΑΡΤΗΜΑ VI: ΚΑΝΟΝΕΣ ΣΥΣΧΕΤΙΣΗΣ ΜΕΤΑΞΥ ΤΩΝ ΔΕΞΙΟΤΗΤΩΝ		269

---



## Ευρετήριο Εικόνων

Εικόνα 1 - Διάφοροι ορισμοί του diversity στη βιβλιογραφία και αξιολόγηση τους [28].....	59
Εικόνα 2 - Αντίκτυπος του κάθε αλγορίθμου για diversity στην ποιότητα των συστάσεων σύμφωνα με τη βιβλιογραφία [28].....	60
Εικόνα 3 - Accuracy Diversity trade off (randomly select more less popular items) [8] .....	66
Εικόνα 4 - Accuracy diversity trade off (αποτελέσματα από χρήση αλγορίθμου για μικρότερη μείωση ακρίβειας) [8] .....	67
Εικόνα 5 - Η αρχιτεκτονική της πλατφόρμας QualiChain.....	79
Εικόνα 6 - Αρχιτεκτονική υπηρεσίας ταυτοποίησης και ελέγχου πρόσβασης (IAM) .....	80
Εικόνα 7 - Παράδειγμα εισόδου για την υπηρεσία DeTrusty.....	83
Εικόνα 8 - Παράδειγμα εξόδου της υπηρεσίας DeTrusty .....	84
Εικόνα 9 - High level αρχιτεκτονική της υπηρεσίας JobCrawler .....	94
Εικόνα 10 - Η αρχιτεκτονική της υπηρεσίας οπτικοποίησης των δεδομένων.....	107
Εικόνα 11 – Παραδείγματα οπτικοποίησης από την υπηρεσία οπτικοποίησης δεδομένων .	108
Εικόνα 12 – Αρχιτεκτονική της υπηρεσίας συστάσεων δεξιοτήτων και μαθημάτων.....	117
Εικόνα 13 - Τα αποτελέσματα της συσταδοποίησης και οι κοντινότερες συστάδες για δύο χρήστες.....	119
Εικόνα 14 - Ανέβασμα βιογραφικού από κάποιον φοιτητή ή δια βίου μαθητή .....	122
Εικόνα 15 - Προτεινόμενες δεξιότητες.....	123
Εικόνα 16 - Προτεινόμενα μαθήματα .....	124
Εικόνα 17 - Αρχιτεκτονική όψη των υπηρεσιών JobCrawler και εξόρυξης γνώσης που χρησιμοποιήθηκαν.....	130
Εικόνα 18 – Αρχιτεκτονική όψη και διαδικασίες της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών .....	132
Εικόνα 19 - Τα αποτελέσματα της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών για τη δεξιότητα CSS .....	135
Εικόνα 20 - Τα αποτελέσματα της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών για τη δεξιότητα JavaScript.....	135
Εικόνα 21 - Τα αποτελέσματα της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών για τη δεξιότητα Git .....	136
Εικόνα 22 - Η διεπαφή χρήστη της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών με έμφαση στη δεξιότητα CSS .....	137
Εικόνα 23 - Αλγόριθμος MAUT [13] .....	143
Εικόνα 24 - Αλγόριθμος TOPSIS [13] .....	145
Εικόνα 25 - Αλγόριθμος PROMETHEE II [13] .....	148

---

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Εικόνα 26 - Ανέβασμα βιογραφικού από έναν φοιτητή ή δια βίου μαθητή .....	155
Εικόνα 27 – Προτεινόμενες γνώσεις και δεξιότητες .....	155
Εικόνα 28 - Προτεινόμενα μαθήματα .....	156
Εικόνα 29 - Διεπαφή χρήστη για τη χρήση της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων .....	157
Εικόνα 30 - Αποτελέσματα της μεθόδου MAUT .....	158
Εικόνα 31 - Διμερείς συγκρίσεις μεταξύ δύο μαθημάτων.....	158
Εικόνα 32 - Διεπαφή χρήστη για την υπηρεσία ανασχεδιασμού προγράμματος σπουδών, δίνοντας έμφαση στη δεξιότητα CSS .....	159
Εικόνα 33 - Παράδειγμα εισόδου της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων για την εισαγωγή μίας δεξιότητάς σε ένα μάθημα .....	161
Εικόνα 34 - Αποτελέσματα της μεθόδου PROMETHEE II για την ένταξη της δεξιότητας CSS σε κάποιο μάθημα .....	161
Εικόνα 35 - Η είσοδος για τη μέθοδο TOPSIS .....	162
Εικόνα 36 - Τα αποτελέσματα της μεθόδου TOPSIS .....	162
Εικόνα 37 - Μεθοδολογία του πιλότου.....	171
Εικόνα 38 - Παραδείγματα δεξιοτήτων προς εισαγωγή σε μαθήματα με βάση τις προτάσεις της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών .....	173
Εικόνα 39 - Κάλυψη ανά πάροχο (ή κατηγορία) για $K=1$ και $K=1.5$ .....	199
Εικόνα 40 - Κάλυψη ανά πάροχο (ή κατηγορία) για $K=2$ και $K=2.5$ .....	199
Εικόνα 41 - Μέση βαθμολογία για τις διάφορες τιμές διαφοροποίησης και κάλυψης.....	200
Εικόνα 42 - Συνολική βαθμολογία για διαφορετικά κατώφλια διαφοροποίησης για $K=1.5$ .	201
Εικόνα 43 - Συνολική βαθμολογία για διαφορετικά κατώφλια διαφοροποίησης για $K=2$ ....	201
Εικόνα 44 - Συσχέτιση μεταξύ διαφοροποίησης ανά αντικείμενο και αριθμού χρηστών στους οποίους προτείνεται (με απομάκρυνση ακραίων τιμών) .....	204
Εικόνα 45 - Συσχέτιση μεταξύ διαφοροποίησης ανά αντικείμενο και αριθμού χρηστών στους οποίους προτείνεται .....	205
Εικόνα 46 - Αρχική σελίδα πλατφόρμας QualiChain.....	243
Εικόνα 47 - Σελίδα εισόδου ή εγγραφής στην πλατφόρμα .....	243
Εικόνα 48 - Αρχική σελίδα για φοιτητές και δια βίου μαθητές .....	244
Εικόνα 49 - Επιλογή από λίστα μαθημάτων .....	245
Εικόνα 50 - Επισκόπηση μαθήματος Management of the Digital Enterprise.....	246
Εικόνα 51 - Προτεινόμενα μαθήματα .....	246
Εικόνα 52 - Σελίδα πολυκριτηριακής υποστήριξης αποφάσεων.....	247

---

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Εικόνα 53 - Αποτελέσματα υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων .....	247
Εικόνα 54 - Σελίδα σύγκρισης εναλλακτικών επιλογών .....	248
Εικόνα 55 - Διαθέσιμες αγγελίες εργασίας.....	248
Εικόνα 56 - Επισκόπηση συγκεκριμένης αγγελίας εργασίας.....	249
Εικόνα 57 - Αίτηση για θέση εργασίας.....	250
Εικόνα 58 - Λίστα από διαθέσιμα smart badges.....	250
Εικόνα 59 - Σελίδα με συναδέλφους για επίδοση smart badge .....	251
Εικόνα 60 - Φόρμα για επιδότηση smart badge .....	252
Εικόνα 61 - Η σελίδα προφίλ για έναν χρήστη .....	253
Εικόνα 62 - Φόρμα δημιουργίας βιογραφικού.....	254
Εικόνα 63 - Σχεδιασμός προσωποποιημένου μονοπατιού καριέρας (career path).....	255
Εικόνα 64 - Σελίδα συμβουλευτικής καριέρας .....	255
Εικόνα 65 - Σελίδα ανάπτυξης δεξιοτήτων .....	256
Εικόνα 66 - Σελίδα επισκόπησης μαθημάτων για τους καθηγητές.....	257
Εικόνα 67 - Σελίδα σύγκρισης διπλωμάτων.....	257
Εικόνα 68 - Αποτέλεσμα υπηρεσίας σύγκρισης διπλωμάτων .....	257
Εικόνα 69 - Σελίδα για το προφίλ καθηγητή ή διδάσκοντα μαθημάτων.....	258
Εικόνα 70 - Η σελίδα ανασχεδιασμού προγράμματος σπουδών .....	259
Εικόνα 71 - Η σελίδα ανάλυσης και εύρεσης δεξιοτήτων που δε προσφέρονται στο πρόγραμμα σπουδών.....	260
Εικόνα 72 - Σελίδα επιλογής υποψηφίου .....	261
Εικόνα 73 - Σελίδα ελέγχου εγκυρότητας ακαδημαϊκού τίτλου.....	261
Εικόνα 74 - Παράδειγμα άκυρου ακαδημαϊκού τίτλου .....	262
Εικόνα 75 - Παράδειγμα έγκυρου ακαδημαϊκού τίτλου.....	262
Εικόνα 76 - Σελίδα για ανέβασμα ακαδημαϊκού τίτλου, ώστε να είναι επαληθεύσιμος.....	263

---

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



## Ευρετήριο πινάκων

Πίνακας 1 - Παράδειγμα εισόδου για τη δημιουργία ενός smart badge .....	85
Πίνακας 2 - Παράδειγμα εξόδου από την υπηρεσία δημιουργίας smart badge.....	85
Πίνακας 3 - Παράδειγμα εισόδου της υπηρεσίας για επιδότηση smart badge .....	86
Πίνακας 4 - Παράδειγμα εξόδου της υπηρεσίας επιδότησης smart badge .....	88
Πίνακας 5 - Παράδειγμα εισόδου της υπηρεσίας JobWatch.....	91
Πίνακας 6 - Παράδειγμα εξόδου της υπηρεσίας JobWatch .....	92
Πίνακας 7 - Παράδειγμα εισόδου στην υπηρεσία DOB-IE .....	94
Πίνακας 8 - Παράδειγμα εξόδου της υπηρεσίας DOB-IE.....	95
Πίνακας 9 - Παράδειγμα εισόδου της υπηρεσίας FALCON .....	97
Πίνακας 10 - Έξοδος της υπηρεσίας FALCON.....	98
Πίνακας 11 - Παράδειγμα εισόδου (περιγραφές για δύο διαφορετικά διπλώματα) στην υπηρεσία Degree Comparison. ....	99
Πίνακας 12 - Παράδειγμα εξόδου της υπηρεσίας Degree Comparison .....	101
Πίνακας 13 - Παράδειγμα εισόδου της υπηρεσίας δημιουργίας προφίλ .....	102
Πίνακας 14 - Παράδειγμα εξόδου της υπηρεσίας δημιουργίας προφίλ.....	103
Πίνακας 15 - Παράδειγμα εξόδου της υπηρεσίας πρόσληψης και εσωτερικής ανακατανομής προσωπικού .....	104
Πίνακας 16 - Παράδειγμα εξόδου της υπηρεσίας συμβουλευτικής καριέρας.....	105
Πίνακας 17 - Παράδειγμα εισόδου της υπηρεσίας οπτικοποίησης δεδομένων .....	107
Πίνακας 18 - Διάφορες εξειδικεύσεις στην τεχνολογία της πληροφορίας και οι δεξιότητες που εξάχθηκαν για καθεμιά.....	118
Πίνακας 19 - Οι προτάσεις που προκύπτουν από την προσέγγιση συσταδοποίησης για τους χρήστες User1 και User2 .....	120
Πίνακας 20 - Οι προτάσεις για τους χρήστες User1 και User2 με συνδυασμό των δύο προσεγγίσεων.....	121
Πίνακας 21 - Οι γνώσεις και δεξιότητες σε υψηλή ζήτηση που δε προσφέρονται στο δοθέν πρόγραμμα σπουδών.....	133
Πίνακας 22 - Μερικοί από τους κανόνες συσχέτισής που εξάχθηκαν .....	134
Πίνακας 23 - Παράδειγμα χρήσης της μεθόδου MAUT μέσω του API της υπηρεσίας.....	152
Πίνακας 24 - Αποτελέσματα από τη χρήση της μεθόδου MAUT.....	153
Πίνακας 25 - Σενάρια περιπτώσεων επαλήθευσης για τον πιλότο .....	168
Πίνακας 26 - KPIs για τον πιλότο.....	168
Πίνακας 27 - Ο ευριστικός αλγόριθμος που αναπτύχθηκε.....	196

---

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Πίνακας 28 - Κατανομή αντικειμένων σε παρόχους (ή κατηγορίες) ..... 198

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας







## ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

### 1.1 Ορισμός Προβλήματος

Η συγκεκριμένη διδακτορική διατριβή έχει ως πρωταρχικό στόχο τη γεφύρωση του χάσματος μεταξύ των γνώσεων και δεξιοτήτων που διδάσκονται στα εκπαιδευτικά ιδρύματα και αυτών που ζητούνται από την αγορά εργασίας στον τομέα της τεχνολογίας. Ωστόσο, κατά την ενασχόληση του συγγραφέα με το συγκεκριμένο πρόβλημα, και λόγω της τεχνολογίας που επιλέχθηκε και αναπτύχθηκε, προέκυψαν κάποια επιπρόσθετα προβλήματα σχετικά με τα Συστήματα Συστάσεων, με τα οποία επίσης καταπιάνεται η συγκεκριμένη διδακτορική διατριβή. Στις παρακάτω ενότητες παρουσιάζονται τα συγκεκριμένα προβλήματα.

#### 1.1.1 Πρόβλημα 1: Το χάσμα δεξιοτήτων μεταξύ αγοράς εργασίας και τεχνολογικής εκπαίδευσης

Είναι κοινώς αποδεκτό ότι η αγορά εργασίας, κυρίως στα τεχνολογικά επαγγέλματα, εξελίσσεται με ραγδαίους ρυθμούς. Έτσι, γνώσεις και δεξιότητες που ήταν σε υψηλή ζήτηση πριν μερικά χρόνια, θεωρούνται πλέον παρωχημένες και «παλιομοδίτικες». Λαμβάνοντας υπόψη ότι, σε μεγάλο βαθμό, οι εργαζόμενοι λαμβάνουν τις απαραίτητες γνώσεις για την αγορά εργασίας μέσω της εκπαίδευσης, το φαινόμενο αυτό έχει ως αποτέλεσμα εργαζόμενοι που ολοκλήρωσαν τις σπουδές πριν από αρκετό καιρό ή που δουλεύουν για μεγάλο χρονικό διάστημα με συγκεκριμένες τεχνολογίες (που μπορεί να ήταν σε υψηλή ζήτηση πριν κάποια χρόνια), να δυσκολεύονται να προοδεύσουν με την καριέρα τους ή να βρουν μια νέα θέση εργασίας, παρά τη σημαντική τους εργασιακή εμπειρία. Για το λόγο αυτό, ολοένα και περισσότεροι εργαζόμενοι επιλέγουν να αποκτήσουν καινούριες γνώσεις και δεξιότητες μέσω του πανεπιστημίου ή μέσω ανοιχτών μαθημάτων.

Από την άλλη πλευρά, μια σημαντική πρόκληση για τους εκπαιδευτικούς οργανισμούς είναι να συμβαδίσουν με τις ανάγκες της αγοράς εργασίας, καθότι σε περίπτωση αποτυχίας να το κάνουν, καταλήγουν να προσφέρουν παρωχημένες γνώσεις και δεξιότητες στους φοιτητές οι οποίες πλέον δε ζητούνται από την αγορά εργασίας [1], [2]. Για το λόγο αυτό είναι σημαντικότερο από ποτέ οι φοιτητές και οι δια βίου μαθητές (lifelong learners) να επιλέγουν μαθήματα τα οποία θα τους βοηθήσουν να προοδεύσουν στην αγορά εργασίας και να αποφύγουν μαθήματα που προσφέρουν γνώσεις και δεξιότητες που θεωρούνται παρωχημένες.

Επιπροσθέτως, και οι εκπαιδευτικοί οργανισμοί θα πρέπει να κινητοποιηθούν προς την κατεύθυνση της γεφύρωσης του χάσματος μεταξύ των γνώσεων και δεξιοτήτων που προσφέρουν και των γνώσεων και δεξιοτήτων που ζητούνται από την αγορά εργασίας, μιας και ο ρόλος τους είναι σε μεγάλο βαθμό, να προετοιμάσουν τους φοιτητές και τους δια βίου μαθητές για την είσοδο τους στην αγορά εργασίας. Κατά συνέπεια, αποτυχία να συμβαδίσουν με τις ανάγκες της αγοράς μπορεί να οδηγήσει στην παρακμή τους. Αυτό διότι όλο και περισσότερα ανοιχτά διαδικτυακά μαθήματα (massive open online courses - MOOCs) δημιουργούνται με σκοπό να καλύψουν την έλλειψη γνώσεων και δεξιοτήτων που ζητούνται από την αγορά. Για το λόγο αυτό οι εκπαιδευτικοί οργανισμοί στην τεχνολογία πρέπει να

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

καταβάλλουν σημαντική προσπάθεια για την αναδιάρθρωση του προγράμματος σπουδών τους με βάση τις ανάγκες της αγοράς εργασίας.

Ωστόσο, δεν είναι εύκολο για τους φοιτητές, τους εργαζομένους και τις μονάδες διοίκησης των εκπαιδευτικών οργανισμών να είναι πάντοτε ενήμεροι για τις τάσεις της αγοράς όσον αφορά τις γνώσεις και δεξιότητες που είναι σε υψηλή ζήτηση. Αυτό συμβαίνει διότι απαιτείται σημαντική προσπάθεια για την αναζήτηση και επεξεργασία των δεδομένων ώστε να βγουν τα κατάλληλα συμπεράσματα για την αγορά εργασίας, μιας και υπάρχει πληθώρα πληροφοριών για την αγορά εργασίας σε ιστοτόπους αναζήτησης εργασίας και άλλες πηγές (π.χ. το LinkedIn<sup>1</sup>). Επιπλέον, εκτός από τους τεράστιους όγκους δεδομένων, ένα ακόμα πρόβλημα αποτελεί και το γεγονός ότι τα διαθέσιμα δεδομένα παρέχονται σαν ελεύθερο κείμενο, πράγμα που καθιστά δύσκολη την επεξεργασία και την ανάλυση τους με κάποιο πρόγραμμα.

Για όλους τους παραπάνω λόγους, η χρήση τεχνικών επεξεργασίας φυσικής γλώσσας και εξόρυξης γνώσης από δεδομένα αποτελούν μονόδρομο για την απόκτηση μιας γενικής εικόνας για τις ανάγκες της αγοράς εργασίας. Επιπλέον, η χρήση συστημάτων συστάσεων και συστημάτων υποστήριξης αποφάσεων μπορούν να συμβάλλουν αποφασιστικά στη λήψη σωστών αποφάσεων, αξιοποιώντας την υπάρχουσα πληροφορία και δημιουργώντας χρήσιμες προτάσεις.

### **1.1.2 Πρόβλημα 2: Δικαιοσύνη μεταξύ των παρόχων αντικειμένων στα Συστήματα Συστάσεων**

Κατά την ενασχόληση με το πρόβλημα 1, αναπτύχθηκε μεταξύ άλλων ένα σύστημα συστάσεων μαθημάτων και δεξιοτήτων με βάση τις ανάγκες της αγοράς εργασίας όσο και την εμπειρία και τις προσωπικές προτιμήσεις του χρήστη. Τα μαθήματα που θα προτεινόταν από το σύστημα αναμενόταν να παρέχονται από διαφορετικούς οργανισμούς (παρόχους). Κατά συνέπεια, τέθηκε ως απαίτηση η επιθυμία των οργανισμών για δίκαιη κατανομή των συστάσεων μαθημάτων στους φοιτητές, με την έννοια ότι τα μαθήματα του κάθε παρόχου θα έπρεπε να προτείνονται τουλάχιστον σε έναν συγκεκριμένο αριθμό χρηστών (φοιτητών ή εργαζομένων) κατά μέσο όρο. Επιπλέον, τέθηκε και η απαίτηση για σύσταση των μαθημάτων κάθε οργανισμού σε όσο το δυνατόν πιο διαφοροποιημένο σύνολο χρηστών με δίκαιο τρόπο ως προς τους παρόχους.

Έτσι, το δεύτερο πρόβλημα με το οποίο καταπιάνεται η παρούσα διδακτορική διατριβή είναι η δικαιοσύνη μεταξύ των παρόχων, όσον αφορά την κάλυψη (coverage) σε χρήστες στους οποίους προτείνονται τα μαθήματα τους, όσο και τη διαφορετικότητα (diversity) μεταξύ αυτών των χρηστών. Οι συγκεκριμένοι περιορισμοί αποτελούν σημαντικό εμπόδιο για τη λειτουργία ενός συστήματος συστάσεων, μιας και περιορίζουν σημαντικά την ικανοποίηση που λαμβάνει ο χρήστης από τα αντικείμενα που του προτείνονται.

Επιπροσθέτως, αξίζει να σημειωθεί ότι το πρόβλημα έχει εμφανιστεί στη βιβλιογραφία και αποτελεί ένα υπολογιστικά δύσκολο πρόβλημα που εντάσσεται στην κατηγορία των NP-

---

<sup>1</sup> <https://www.linkedin.com/>

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

complete προβλημάτων [3] και κατά συνέπεια δεν μπορεί να βρεθεί η βέλτιστη λύση σε αποδεκτό χρόνο.

### **1.1.3 Πρόβλημα 3: Ηθικά προβληματικές εκφάνσεις Συστημάτων Συστάσεων**

Τέλος, λόγω της σημαντικής ενασχόλησης με τα συστήματα συστάσεων, ήρθαμε αντιμέτωποι στην υφιστάμενη βιβλιογραφία, με διάφορες ηθικά επιλήψιμες εκφάνσεις που συνδέονται με τη λειτουργία των σύγχρονων συστημάτων συστάσεων.

Συγκεκριμένα, τα τελευταία χρόνια παρατηρείται μια διαφοροποίηση των στόχων των συστημάτων συστάσεων, από την ικανοποίηση του χρήστη, στη μεγιστοποίηση του κέρδους [4]. Έτσι, ο στόχος ενός σύγχρονου συστήματος συστάσεων είναι να κρατήσει τους χρήστες στην πλατφόρμα όσο το δυνατόν περισσότερο και όχι απλά να φιλτράρει τα αντικείμενα για τα οποία ενδιαφέρονται και να τους τα προτείνει, και παράλληλα να απορρίψει αντικείμενα που και οι ίδιοι θα απέρριπταν. Ένα παράδειγμα που καταδεικνύει αυτή την αλλαγή είναι το γεγονός ότι ενώ το Netflix διοργάνωνε διαγωνισμό με βραβείο ενός εκατομμυρίου δολαρίων για τον νικητή με την προϋπόθεση ότι θα κατάφερνε μια αύξηση της τάξεως του 10% στην ακρίβεια πρόβλεψης του αλγορίθμου του Netflix, ο αλγόριθμος που κέρδισε στον διαγωνισμό δεν χρησιμοποιήθηκε ποτέ. Ο λόγος ήταν ότι μετά το διαγωνισμό το Netflix από εταιρεία ενοικίασης DVD, έγινε υπηρεσία αναπαραγωγής περιεχομένου (content streaming company) και έτσι ο στόχος δεν ήταν πλέον να μαντέψουν τις μελλοντικές προτιμήσεις του χρήστη αλλά να τον κρατήσουν προσκολλημένο στην πλατφόρμα [4].

Υπάρχουν κάποια γεγονότα που καταδεικνύουν ότι αυτή η αλλαγή στην στόχευση των συστημάτων συστάσεων μπορεί να έχει αρνητικές συνέπειες για τους χρήστες και την κοινωνία γενικότερα. Για παράδειγμα, η Frances Haugen, μια πρώην εργαζόμενη στο Facebook κατηγόρησε την εταιρεία ότι βάζει τα κέρδη της πάνω από τους ανθρώπους. Συγκεκριμένα, αποκάλυψε μεταξύ άλλων ότι το Instagram έχει αρνητικό αντίκτυπο στα έφηβα κορίτσια, ότι το σύστημα συστάσεων του Facebook είναι υπεύθυνο για την αυξανόμενη βία στις υποανάπτυκτες χώρες, ενώ ο αλγόριθμος του Facebook δίνει προτεραιότητα στο υβριστικό περιεχόμενο, έναντι του ουδέτερου [5].

Αντίστοιχα ευρήματα για το Facebook παρουσιάζονται και σε άλλες έρευνες (π.χ. στην έρευνα [6]). Επιπλέον, κάποιες έρευνες καταλήγουν ότι το Σύστημα Συστάσεων της Amazon προωθεί θεωρίες συνωμοσίας και αντιεμβολιαστικές θεωρίες, ενώ φέρεται να ενισχύει και τις πεποιθήσεις των χρηστών, αποτρέποντας τους από την έκθεση σε διαφορετικές απόψεις [7], [8].

Από τα παραπάνω γεγονότα, γίνεται εύκολα αντιληπτό ότι το πρόβλημα των ηθικών επιπτώσεων των συστημάτων συστάσεων είναι πολύ σημαντικό ενώ δεν υπήρχαν πολλές έρευνες που να προτείνουν πώς μπορεί να αντιμετωπιστεί κατά τη συγγραφή της παρούσας διδακτορικής διατριβής.

## 1.2 Θεματική Περιοχή

Η ανώτατη εκπαίδευση είναι άρρηκτα συνδεδεμένη με την αγορά εργασίας, μιας και ο ρόλος της πρώτης είναι σε μεγάλο βαθμό η προετοιμασία των φοιτητών και των δια βίου μαθητών μέσω της διδασκαλίας των απαραίτητων γνώσεων και δεξιοτήτων και προσόντων για την ομαλή εισαγωγή τους στην αγορά εργασίας και την επιτυχία σε αυτή.

Ωστόσο, τα τελευταία χρόνια με τη ραγδαία ανάπτυξη της τεχνολογίας, παρατηρείται ότι σε πολλές περιπτώσεις η εκπαίδευση με τις γνώσεις και δεξιότητες που προσφέρει αποκλίνει από αυτό που πραγματικά ζητείται από την αγορά εργασίας, ειδικά στον τομέα της τεχνολογίας. Το φαινόμενο αυτό είναι πολυδιάστατο και πρέπει να αντιμετωπιστεί αποτελεσματικά, καθώς επηρεάζει αρνητικά τόσο το εργατικό δυναμικό, όσο και την αγορά εργασίας. Αυτό διότι αφ' ενός οι δυνητικοί εργαζόμενοι δεν διαθέτουν τα απαραίτητα προσόντα για πολλές θέσεις εργασίας παρότι μπορεί να έχουν αξιόλογες σπουδές στον συγκεκριμένο τομέα και αφ' ετέρου οι εργοδότες δυσκολεύονται να καλύψουν συγκεκριμένες θέσεις εργασίας που απαιτούν υψηλή εξειδίκευση και καταλήγουν να αφιερώνουν σημαντικό χρόνο και πόρους για την προσέλκυση των κατάλληλων εργαζομένων ή την εκπαίδευση του υπάρχοντος προσωπικού.

Ο βασικότερος λόγος για την απόκλιση μεταξύ των γνώσεων και δεξιοτήτων που προσφέρονται από τα εκπαιδευτικά ιδρύματα και των γνώσεων και δεξιοτήτων που ζητούνται από την αγορά έχει να κάνει τόσο με τον ρυθμό με τον οποίο εξελίσσονται οι ανάγκες της αγοράς εργασίας, όσο και με το γεγονός ότι οι εκπαιδευτικοί οργανισμοί αλλάζουν με σχετικά αργούς ρυθμούς τα προγράμματα σπουδών τους [1]. Επιπλέον, αξίζει να σημειωθεί ότι απαιτείται σημαντική προσπάθεια για την ανάλυση των απαιτήσεων της αγοράς εργασίας, καθώς υπάρχει μια ευρεία γκάμα από πηγές δεδομένων όπως είναι οι ιστότοποι αγγελιών εργασίας, το LinkedIn κ.α.. Το πρόβλημα είναι ότι σε αυτές τις πηγές η πληροφορία δίνεται συνήθως με αδόμητο τρόπο ως ελεύθερο κείμενο ή ημιδομημένο τρόπο, ενώ τα δεδομένα παρέχονται με διαφορετικούς τρόπους από κάθε πηγή. Έτσι, η ανάλυση των αναγκών της αγοράς εργασίας καθίσταται μια χρονοβόρα διαδικασία που απαιτεί σύνθετες τεχνικές και αρκετό χρόνο για να γίνει αποτελεσματικά. Επιπροσθέτως, λόγω των ραγδαίων μεταβολών των αναγκών της αγοράς εργασίας σε κάποιους κλάδους, η ανάλυση της αγοράς εργασίας θα πρέπει να πραγματοποιείται ανά τακτά χρονικά διαστήματα.

Για όλους τους παραπάνω λόγους, ενδείκνυται η ανάπτυξη συστημάτων λογισμικού τα οποία θα χρησιμοποιούν τεχνικές εξόρυξης γνώσης από δεδομένα (data mining and web mining) από το διαδίκτυο, τεχνικές επεξεργασίας φυσικής γλώσσας (natural language processing), αναλυτικής δεδομένων (data analytics) και μηχανικής μάθησης (machine learning), για την εξαγωγή δεδομένων για την αγορά εργασίας από το διαδίκτυο, την ανάλυση τους και την εξαγωγή συμπερασμάτων μέσα από αυτά. Αντίστοιχες τεχνικές ενδείκνυται για εφαρμογή και στα προγράμματα σπουδών των πανεπιστημίων, για την εξαγωγή αντίστοιχων συμπερασμάτων για τα παρεχόμενα μαθήματα. Τέλος, για την λήψη σωστών αποφάσεων που αξιοποιούν την υπάρχουσα γνώση, μπορούν να αξιοποιηθούν τόσο τα συστήματα συστάσεων, ώστε να παρέχουν έγκυρες προτάσεις πάνω σε ενέργειες που θα ήταν χρήσιμες να γίνουν στα προγράμματα σπουδών ή άλλες αποφάσεις, όσο και τα συστήματα υποστήριξης αποφάσεων, ώστε να παρέχουν στο χρήστη τη δυνατότητα να λάβει υπόψη όλα

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

τα απαραίτητα κριτήρια για τη λήψη μιας απόφασης (υποκειμενικά και αντικειμενικά), με βάση τη σημασία που δίνει ο ίδιος σε κάθε κριτήριο.

Επιπλέον, η ανάπτυξη κατάλληλου λογισμικού που αξιοποιεί τις προαναφερθείσες τεχνικές και τεχνολογίες ενδείκνυται διότι μπορεί να αυτοματοποιήσει σε μεγάλο βαθμό τη διαδικασία της περιοδικής ανάλυσης της αγοράς εργασίας, παρέχοντας πάντα πληροφορίες οι οποίες θα βασίζονται σε έγκυρα δεδομένα.

Όσον αφορά τη λειτουργία των συστημάτων συστάσεων, τα οποία αποτελούν και την βασική τεχνολογία που αναπτύχθηκε στη συγκεκριμένη διδακτορική διατριβή, τα τελευταία χρόνια υπάρχει έντονο ενδιαφέρον, πέρα από την παροχή ικανοποιητικών συστάσεων στους χρήστες, και για τη βελτίωση άλλων παραγόντων, όπως είναι η κάλυψη (σε πόσους χρήστες προτείνεται ένα αντικείμενο), η διαφορετικότητα (πόσο διαφορετικά είναι μεταξύ τους τα αντικείμενα που προτείνονται σε έναν χρήστη ή πόσο διαφορετικοί είναι μεταξύ τους οι χρήστες στους οποίους προτείνεται ένα αντικείμενο) και η δικαιοσύνη (ίδια δικαιώματα μεταξύ των χρηστών ή των παρόχων). Κατά την υλοποίηση των συστημάτων συστάσεων, ήρθαμε αντιμέτωποι με αρκετά θέματα γύρω από τη λειτουργία των συστημάτων συστάσεων και ασχοληθήκαμε εκτενώς με αυτά.

Επομένως, η θεματική περιοχή της διατριβής περιστρέφεται γύρω από την αγορά εργασίας και την εκπαίδευση στον τομέα της τεχνολογίας, καθώς επίσης και τις τεχνικές και τεχνολογίες για την κάλυψη της απόστασης μεταξύ τους, που συμπεριλαμβάνουν την εξόρυξη γνώσης από δεδομένα, την επεξεργασία φυσικής γλώσσας, την αναλυτική δεδομένων και τη μηχανική μάθηση και κυρίως τα συστήματα συστάσεων. Τα τελευταία αναλύονται εκτενώς στα πλαίσια της παρούσας διδακτορικής διατριβής και εκτός των πρακτικών εφαρμογών που αναπτύσσονται, παρουσιάζονται επίσης γενικά, αλλά άρρηκτα συνδεδεμένα με το αντικείμενο της διατριβής, προβλήματα τα οποία απασχολούν τη βιβλιογραφία, μαζί με προτεινόμενες λύσεις στο ζήτημα της δικαιοσύνης μεταξύ των παρόχων, όσο και σε αυτό των ηθικών εκφάνσεων και του μετριασμού των αρνητικών επιπτώσεων από τη χρήση τους, όπως προέκυψαν στα πλαίσια της έρευνας της παρούσας διδακτορικής διατριβής και δημοσιεύτηκαν σε έγκριτα επιστημονικά περιοδικά και διεθνή συνέδρια.

### **1.3 Συμβολή Διατριβής**

Όπως αναλύθηκε και στις προηγούμενες ενότητες, η συμβολή της παρούσας διδακτορικής διατριβής είναι πολυδιάστατη. Συγκεκριμένα, αφ' ενός συμβάλλει στη βελτίωση της λειτουργίας των εκπαιδευτικών οργανισμών, ώστε οι γνώσεις και δεξιότητες που προσφέρονται στους τελευταίους να συμβαδίζουν με τις γνώσεις και δεξιότητες που ζητούνται από την αγορά εργασίας στον τομέα της τεχνολογίας και αφ' ετέρου στην βελτίωση της λειτουργίας των συστημάτων συστάσεων, σε ό,τι αφορά τη δικαιοσύνη μεταξύ των παρόχων αντικειμένων, όπως επίσης και στην αντιμετώπιση των αρνητικών συνεπειών που προκύπτουν από τα σύγχρονα συστήματα συστάσεων.

Πιο συγκεκριμένα, όσον αφορά τη βελτίωση της λειτουργίας των εκπαιδευτικών οργανισμών, η συγκεκριμένη διδακτορική διατριβή συμβάλλει με τους παρακάτω τρόπους:

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

- Τα εκπαιδευτικά ιδρύματα και οι φοιτητές ή οι δια βίου μαθητές μπορούν να αντλούν έγκυρη πληροφόρηση για τις ανάγκες της αγοράς εργασίας κάθε στιγμή σε λίγα μόνο δευτερόλεπτα.
- Οι φοιτητές και οι δια βίου μαθητές μπορούν να λαμβάνουν προτάσεις μαθημάτων που αφ' ενός συμβαδίζουν με τις ανάγκες της αγοράς εργασίας και αφ' ετέρου με τις προσωπικές τους προτιμήσεις.
- Οι φοιτητές και οι δια βίου μαθητές μπορούν να αποφύγουν την επιλογή μαθημάτων που προσφέρουν γνώσεις και δεξιότητες που θεωρούνται παρωχημένες από την αγορά εργασίας.
- Τα εκπαιδευτικά ιδρύματα μπορούν να γνωρίζουν ανά πάσα στιγμή ποιες γνώσεις και δεξιότητες που ζητούνται από την αγορά εργασίας δεν συμπεριλαμβάνονται στο πρόγραμμα σπουδών τους.
- Τα εκπαιδευτικά ιδρύματα και οι διδάσκοντες μαθημάτων μπορούν να γνωρίζουν ανά πάσα στιγμή εάν τα μαθήματα τους παρέχουν γνώσεις και δεξιότητες που δεν ζητούνται από την αγορά εργασίας.
- Τα εκπαιδευτικά ιδρύματα και οι διδάσκοντες των μαθημάτων μπορούν να λαμβάνουν συστάσεις για εισαγωγή συγκεκριμένων γνώσεων και δεξιοτήτων που δεν καλύπτονται από το δοθέν πρόγραμμα σπουδών σε συγκεκριμένα σχετικά μαθήματα, ενώ παράλληλα παρέχεται και ο λόγος για τον οποίο θεωρούνται σχετικά τα μαθήματα με τις γνώσεις και τις δεξιότητες αυτές.
- Όλοι οι παραπάνω χρήστες μπορούν να αξιοποιήσουν τόσο την πληροφορία που παρέχεται για τα μαθήματα και την αγορά εργασίας όσο και ένα πολυκριτηριακό σύστημα υποστήριξης αποφάσεων, ώστε να πάρουν αποφάσεις λαμβάνοντας υπόψη όλα τα κριτήρια που οι ίδιοι θεωρούν σημαντικά.

Σχετικά με το θέμα της δικαιοσύνης μεταξύ των παρόχων στα συστήματα συστάσεων η συμβολή της παρούσας διδακτορικής διατριβής έγκειται στα παρακάτω σημεία:

- Ανάπτυξη ευριστικής λύσης για το πρόβλημα της δικαιοσύνης μεταξύ των παρόχων όσον αφορά την σύσταση των αντικειμένων τους σε διαφορετικούς χρήστες (coverage) όσο και στη διαφορετικότητα μεταξύ των χρηστών αυτών.
- Τα αποτελέσματα του ευριστικού αλγορίθμου που αναπτύχθηκε είναι πολύ κοντά στο άνω όριο για τη βέλτιστη λύση και κατά συνέπεια πολύ κοντά στη βέλτιστη λύση του προβλήματος.
- Ο ευριστικός αλγόριθμος που αναπτύχθηκε απαιτεί πολυωνυμικό χρόνο για την εκτέλεση του ενώ το πρόβλημα εύρεσης της βέλτιστης λύσης αποτελεί υπολογιστικά δύσκολο πρόβλημα που εντάσσεται στην κατηγορία NP-hard.
- Εντοπίστηκαν κάποιες αδυναμίες στον αρχικό ορισμό του προβλήματος από τη βιβλιογραφία όσον αφορά τη διαφορετικότητα μεταξύ των χρηστών και προτάθηκαν και αναλύθηκαν δύο νέοι μαθηματικοί ορισμοί για την τελευταία, όπως επίσης και νέες προσεγγίσεις για την επίλυση του προβλήματος.

Τέλος, όσον αφορά τις αρνητικές επιπτώσεις που προκύπτουν από τα συστήματα συστάσεων η συμβολή της παρούσας διδακτορικής διατριβής έγκειται στα παρακάτω σημεία:

- Αναγνωρίστηκαν και κατηγοριοποιήθηκαν οι προβληματικές εκφάνσεις των συστημάτων συστάσεων.
- Προτάθηκε μέτρο για την αξιολόγηση της κρισιμότητας για ένα σύστημα συστάσεων με βάση τις κατηγορίες που αναγνωρίστηκαν.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

- Το μέτρο εφαρμόστηκε σε τέσσερα από τα πιο γνωστά συστήματα συστάσεων παγκοσμίως για την αξιολόγηση της κρισιμότητάς τους.
- Προτάθηκε ρυθμιστικό πλαίσιο για την εξυγίανση των συστημάτων συστάσεων.

## 1.4 Δομή διατριβής

Η παρούσα διδακτορική διατριβή είναι χωρισμένη σε έντεκα (11) βασικά κεφάλαια και έξι (6) παραρτήματα. Στα αρχικά κεφάλαια παρουσιάζεται το θεωρητικό υπόβαθρο και οι σχετικές εργασίες με την παρούσα διδακτορική διατριβή. Τα κεφάλαια αυτά βοηθούν στην κατανόηση των επί μέρους προβλημάτων και παρέχουν το κατάλληλο υπόβαθρο στον αναγνώστη ώστε να παρακολουθήσει ομαλά το περιεχόμενο της διατριβής. Στη συνέχεια, παρουσιάζονται ανά κεφάλαιο οι λύσεις που αναπτύχθηκαν για την απάντηση στις ερευνητικές ερωτήσεις της παρούσας διατριβής.

Συγκεκριμένα, το παρόν κεφάλαιο λειτουργεί σαν εισαγωγή, παρουσιάζοντας τα επιμέρους προβλήματα με τα οποία καταπιάνεται η παρούσα διατριβή, την ευρύτερη θεματική περιοχή, όπως επίσης και τη συμβολή της παρούσας διατριβής.

Το κεφάλαιο 2 πραγματοποιεί μια λεπτομερή ανασκόπηση της υπάρχουσας κατάστασης όσον αφορά τη σύνδεση της αγοράς εργασίας με την τεχνολογική εκπαίδευση και το χάσμα γνώσεων και δεξιοτήτων που υπάρχει μεταξύ αυτών. Επιπλέον, παρουσιάζει σχετικές εργασίες που κινούνται προς τη γεφύρωση του χάσματος αυτού.

Το κεφάλαιο 3 παρουσιάζει την υφιστάμενη κατάσταση στα συστήματα συστάσεων που αποτελεί και τη βασική τεχνολογία που αναπτύχθηκε στα πλαίσια της παρούσας διατριβής. Πιο συγκεκριμένα, παρουσιάζεται το πρόβλημα των συστάσεων, τα ενδιαφερόμενα μέρη και η χρησιμότητα των συστημάτων αυτών, οι διάφορες τεχνικές παροχής συστάσεων όπως και οι διάφορες τεχνικές αξιολόγησης, οι νέες τάσεις, καθώς επίσης και η κοινωνική διάσταση που απορρέει από τη λειτουργία τους.

Το κεφάλαιο 4 παρουσιάζει την πλατφόρμα QualiChain, καθώς οι υπηρεσίες που αναπτύχθηκαν στο πλαίσιο της παρούσας διδακτορικής διατριβής υιοθετήθηκαν από την πλατφόρμα αυτή και αξιολογήθηκαν από τους χρήστες της. Συγκεκριμένα, παρουσιάζεται η αρχιτεκτονική της πλατφόρμας, τα επιμέρους υποσυστήματα της καθώς επίσης και οι υπηρεσίες που παρέχει και οι αλληλεπιδράσεις μεταξύ τους.

Το κεφάλαιο 5 παρουσιάζει την υπηρεσία QualiChain Recommender, ένα σύστημα συστάσεων δεξιοτήτων και μαθημάτων που αναπτύχθηκε για την παροχή προσωποποιημένων προτάσεων σε φοιτητές και δια βίου μαθητές με βάση τις ανάγκες της αγοράς εργασίας. Συγκεκριμένα, παρουσιάζεται η μεθοδολογία που ακολουθήθηκε, οι επιμέρους υπηρεσίες που αναπτύχθηκαν για την εξαγωγή δεδομένων και την εξόρυξη γνώσης από αυτά, παράλληλα με το σύστημα που αναπτύχθηκε.

Το κεφάλαιο 6 παρουσιάζει την υπηρεσία ανασχεδιασμού προγράμματος σπουδών, ένα σύστημα συστάσεων που παρέχει προτάσεις για την βελτίωση ενός προγράμματος σπουδών, ώστε να συμβαδίζει με τις ανάγκες της αγοράς εργασίας. Επιπροσθέτως, παρουσιάζεται η μεθοδολογία που ακολουθήθηκε, κάποιες βοηθητικές υπηρεσίες που αναπτύχθηκαν για εξαγωγή δεδομένων και εξόρυξη γνώσης από αυτά, όπως επίσης και τα αποτελέσματα της υπηρεσίας.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Το κεφάλαιο 7 παρουσιάζει την υπηρεσία πολυκριτηριακής υποστήριξης αποφάσεων που αναπτύχθηκε για την υποβοήθηση στις αποφάσεις των χρηστών και ο συνδυασμός της με τα συστήματα συστάσεων. Πιο συγκεκριμένα, παρουσιάζονται οι μέθοδοι που αναπτύχθηκαν παράλληλα με δυο περιπτώσεις χρήσης της υπηρεσίας.

Το κεφάλαιο 8 παρουσιάζει τα αποτελέσματα που προέκυψαν από τη χρήση και την αξιολόγηση των υπηρεσιών που αναπτύχθηκαν από πραγματικούς χρήστες. Σε αυτό το πλαίσιο, παρουσιάζονται τόσο οι απαντήσεις τους στα ερωτηματολόγια που τους δόθηκαν, όσο και μερικές κατευθύνσεις για βελτίωση των υπηρεσιών που αναπτύχθηκαν.

Το κεφάλαιο 9 παρουσιάζει το πρόβλημα της δικαιοσύνης μεταξύ των παρόχων όσον αφορά την κάλυψη και τη διαφοροποίηση των αντικειμένων τους σε χρήστες. Συγκεκριμένα, παρουσιάζεται τόσο η μοντελοποίηση του προβλήματος όσο και μια ευριστική λύση που αναπτύχθηκε στα πλαίσια της παρούσας διατριβής. Επιπροσθέτως, αναγνωρίζονται μια σειρά από αδυναμίες στη μαθηματική μοντελοποίηση που προτείνεται στη βιβλιογραφία και προτείνονται δυο νέα μαθηματικά μοντέλα για το πρόβλημα.

Το κεφάλαιο 10 ασχολείται με τις ηθικές προκλήσεις που πρέπει να αντιμετωπιστούν στα συστήματα συστάσεων. Συγκεκριμένα, αναγνωρίζει με βάση τη βιβλιογραφία μια σειρά από προβληματικές εκφάνσεις στη λειτουργία των συστημάτων συστάσεων και παρουσιάζει ένα πλαίσιο αξιολόγησης της κρισιμότητας ενός συστήματος. Επιπλέον, παρουσιάζεται η εφαρμογή του προτεινόμενου πλαισίου σε μια σειρά από δημοφιλή συστήματα συστάσεων. Τέλος, προτείνεται μια σειρά από ρυθμιστικές ενέργειες που θα μπορούσαν να εξυγιάνουν τη λειτουργία των συστημάτων συστάσεων.

Τον τόμο ολοκληρώνει το κεφάλαιο 11, το οποίο συνοψίζει τις λύσεις που αναπτύχθηκαν στα πλαίσια της παρούσας διατριβής και τα συμπεράσματα που προέκυψαν. Επιπλέον, παρέχει μια σειρά από μελλοντικές επεκτάσεις για κάθε μία από τις λύσεις που αναπτύχθηκαν.

Τα Παραρτήματα της διατριβής αυτής έχουν το ρόλο των συμπληρωματικών κεφαλαίων, παρουσιάζοντας:

- Τη λίστα με τις δημοσιεύσεις του υποψήφιου διδάκτορα σε επιστημονικά περιοδικά, βιβλία και πρακτικά συνεδρίων (Παράρτημα I)
- Τη λίστα με τις υπηρεσίες που αναπτύχθηκαν στα πλαίσια της παρούσας διατριβής, όσο και του έργου QualiChain (Παράρτημα II).
- Μια επίδειξη των σημαντικότερων λειτουργιών της πλατφόρμας QualiChain μέσω της διεπαφής χρήστη που αναπτύχθηκε (Παράρτημα III).
- Την αναλυτική λίστα γνώσεων και δεξιοτήτων που εξάχθηκαν με τη χρήση των εργαλείων εξόρυξης γνώσης από δεδομένα (Παράρτημα IV).
- Την αναλυτική λίστα με τις γνώσεις και δεξιότητες οι οποίες είναι σε υψηλή ζήτηση από την αγορά αλλά δεν προσφέρονται στο δοθέν πρόγραμμα σπουδών (Παράρτημα V).
- Την αναλυτική λίστα με τους κανόνες συσχέτισης που εξάχθηκαν από την υπηρεσία εξόρυξης κανόνων συσχέτισης που χρησιμοποιήθηκε για την παροχή προτάσεων για



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

τη βελτίωση του προγράμματος σπουδών μέσω της σχετικής υπηρεσίας που αναπτύχθηκε (Παράρτημα VI).

## 1.5 Αναφορές

- [1] C. Kontzinos, O. Markaki, P. Kokkinakos, V. Karakolis, S. Skalidakis, and J. Psarras, "University process optimisation through smart curriculum design and blockchain-based student accreditation," 2019, doi: 10.33965/icwi2019\_2019131012.
- [2] C. Kontzinos, V. Karakolis, S. Skalidakis, O. Markaki, M. Androutsopoulou, and J. Psarras, "COMBINING BLOCKCHAIN, SEMANTICS AND DATA ANALYTICS FOR UNIVERSITY PROCESS OPTIMISATION," *IADIS-INTERNATIONAL J. Comput. Sci. Inf. Syst.*, vol. 14, no. 2, 2019.
- [3] I. Koutsopoulos and M. Halkidi, "Optimization of Multi-stakeholder Recommender Systems for Diversity and Coverage," in *IFIP Advances in Information and Communication Technology*, 2021, vol. 627, doi: 10.1007/978-3-030-79150-6\_55.
- [4] N. Seaver, "Captivating algorithms: Recommender systems as traps," *J. Mater. Cult.*, vol. 24, no. 4, 2019, doi: 10.1177/1359183518820366.
- [5] D. Milmo, "Facebook revelations: what is in cache of internal documents?," *The Guardian*, 2021.
- [6] J. Horwitz and D. Seetharaman, "Facebook Executives Shut Down Efforts to Make the Site Less Divisive," *Wall Str. J.*, 2020.
- [7] J. Shin and T. Valente, "Algorithms and Health Misinformation: A Case Study of Vaccine Books on Amazon," *J. Health Commun.*, vol. 25, no. 5, 2020, doi: 10.1080/10810730.2020.1776423.
- [8] E. Thomas, "Amazon's algorithms, conspiracy theories and extremist literature," *Inst. Strateg. Dialogue*, 2021.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## **ΚΕΦΑΛΑΙΟ 2: ΑΝΑΣΚΟΠΗΣΗ ΥΦΙΣΤΑΜΕΝΗΣ ΚΑΤΑΣΤΑΣΗΣ – ΧΑΣΜΑ ΔΕΞΙΟΤΗΤΩΝ ΜΕΤΑΞΥ ΑΓΟΡΑΣ ΕΡΓΑΣΙΑΣ ΚΑΙ ΕΚΠΑΙΔΕΥΣΗΣ**

Παρά την ποικιλομορφία στις ικανότητες του εργατικού δυναμικού στις αναπτυσσόμενες χώρες, τα τελευταία χρόνια όλο και περισσότερες επιχειρήσεις αναφέρουν ότι αντιμετωπίζουν δυσκολίες στην εύρεση προσωπικού με τις κατάλληλες ικανότητες για να στελεχώσουν επαρκώς τις ανοιχτές θέσεις εργασίας, ενώ θεωρούν την έλλειψη ικανών υποψηφίων ως τη μεγαλύτερη πρόκληση που έχουν να αντιμετωπίσουν [1]. Το φαινόμενο αυτό φέρεται να έχει οξυνθεί με το ξέσπασμα της πανδημίας του κορονοϊού [2]. Φυσικά, πέραν της πανδημίας, παράγοντες όπως η συνεχής τεχνολογική εξέλιξη, η παγκοσμιοποίηση και η αυξημένη κινητικότητα των εργαζομένων έχουν καθοριστική συμβολή στην ολοένα και αυξανόμενη αποσύνδεση των ικανοτήτων του εργατικού δυναμικού με τις ικανότητες που απαιτούνται από την αγορά εργασίας.

Ωστόσο, αυτή η αναντιστοιχία μεταξύ των προσφερόμενων γνώσεων και δεξιοτήτων από το διαθέσιμο εργατικό δυναμικό και των ζητούμενων γνώσεων και δεξιοτήτων από την αγορά εργασίας δεν οφείλεται αποκλειστικά στην έλλειψη επαρκούς εκπαίδευσης του εργατικού δυναμικού. Στην πραγματικότητα, σύμφωνα με τη Eurostat [3], το 2020 περισσότεροι από το 40% των Ευρωπαίων ηλικίας μεταξύ 25 και 34 ετών είχαν ολοκληρώσει τις σπουδές τους στην τριτοβάθμια εκπαίδευση. Επιπλέον, η ίδια έρευνα κατέδειξε ότι οι νεότερες γενιές λαμβάνουν ανώτερους τίτλους σπουδών σε μεγαλύτερο βαθμό σε σχέση με τις παλαιότερες. Αντίθετα, οι Ευρωπαίοι εργαζόμενοι, ως σύνολο, σύμφωνα με τη Eurostat [3], είναι υπερκαταρτισμένοι σε σχέση με τις επαγγελματικές ευκαιρίες που τους προσφέρονται ειδικά αν λάβουμε υπόψη τις θέσεις εργασίας στον αγροτικό τομέα. Επιπροσθέτως, σύμφωνα με την έρευνα για τις γνώσεις και τις δεξιότητες των Ευρωπαίων και τις θέσεις εργασίας, το 45% των ενηλίκων πιστεύει ότι οι δεξιότητες και οι γνώσεις τους θα μπορούσαν να αξιοποιηθούν ή να αναπτυχθούν καλύτερα μέσα από την εργασία τους [4].

Το φαινόμενο της αναντιστοιχίας μεταξύ των ζητούμενων από την αγορά εργασίας και των προσφερόμενων από το εργατικό δυναμικό γνώσεων και δεξιοτήτων είναι σαφώς εντονότερο στα επαγγέλματα που σχετίζονται με την τεχνολογία [5], λόγω και της ραγδαίας ανάπτυξης της τεχνολογίας κατά τις τελευταίες δεκαετίες.

Σε αυτό το πλαίσιο, θα ήταν χρήσιμο να διερευνηθούν τα βαθύτερα αίτια που οδηγούν στην αποσύνδεση όσον αφορά τις γνώσεις και δεξιότητες μεταξύ των απαιτήσεων της αγοράς εργασίας και της προσφοράς από το εργατικό δυναμικό. Στις επόμενες ενότητες θα παρουσιαστούν χωριστά η αγορά εργασίας και η εκπαίδευση στην τεχνολογία, ώστε να αναλυθούν τα αίτια του προβλήματος. Επιπλέον, θα παρουσιαστούν έρευνες που άμεσα ή έμμεσα αναγνωρίζουν και φιλοδοξούν να επιλύσουν το πρόβλημα.

## 2.1 Αγορά εργασίας στην τεχνολογία

Σύμφωνα με την Ευρωπαϊκή Ένωση σημαντικό ποσοστό των επιχειρήσεων εντός των συνόρων της καταγράφει δυσκολία στην εύρεση προσωπικού με επαρκείς γνώσεις και δεξιότητες. Παράλληλα, μεγάλος αριθμός νέων εργαζομένων στην Ευρωπαϊκή Ένωση χαρακτηρίζεται ως υπερκαταρτισμένος (overqualified) και αντιμετωπίζει μια οριζόντια αναντιστοιχία γνώσεων και δεξιοτήτων με τα καθήκοντα του, δηλαδή οι συγκεκριμένοι εργαζόμενοι δεν εργάζονται σε κάποια θέση εργασίας σχετική με το αντικείμενο των σπουδών τους. Στις περισσότερες περιπτώσεις, αυτή η αναντιστοιχία δεν αποτελεί προσωπική επιλογή των νέων, αλλά αποδίδεται κυρίως στην έλλειψη επαγγελματικών ευκαιριών, πληροφορίας ή συντονισμού. Το φαινόμενο αυτό αποτελεί εμπόδιο στην ανάπτυξη και υιοθέτηση της καινοτομίας και κατά συνέπεια στην οικονομική ανάπτυξη αλλά και στην ζωή των νέων. Για τους παραπάνω λόγους, το πρόβλημα αυτό αποτελεί μια από τις μεγαλύτερες προκλήσεις που καλείται να αντιμετωπίσει η Ευρωπαϊκή Ένωση [6].

Το πρόβλημα της αναντιστοιχίας μεταξύ των προσφερόμενων γνώσεων και δεξιοτήτων από το εργατικό δυναμικό και των ζητούμενων γνώσεων και δεξιοτήτων από την αγορά εργασίας παρατηρείται σε εντονότερο βαθμό στον τομέα της τεχνολογίας και κυρίως σε τομείς όπως οι τεχνολογίες πληροφορίας και επικοινωνιών (ICT), όπου οι απαιτούμενες γνώσεις και δεξιότητες αλλάζουν με ραγδαίους ρυθμούς και το εργατικό δυναμικό αδυνατεί να παρακολουθήσει τις εξελίξεις [7].

Το συγκεκριμένο πρόβλημα έχει μελετηθεί σε βάθος και στην εργασία [8], η οποία εξετάζει κατά πόσο το εργατικό δυναμικό στην τεχνολογία της πληροφορίας κατέχει τις γνώσεις και δεξιότητες που απαιτούνται από την αγορά εργασίας. Τα βασικά συμπεράσματα της συγκεκριμένης εργασίας δείχνουν ότι το εργατικό δυναμικό θα πρέπει να εφοδιαστεί με κατάλληλο μίγμα γνώσεων και δεξιοτήτων το οποίο θα αποτελείται από ατομικά χαρακτηριστικά όπως η ανεκτικότητα, η ενσυναίσθηση και η διπλωματία, όπως επίσης από επικοινωνιακές δεξιότητες και ικανότητα επίλυσης προβλημάτων, όπως η δυνατότητα ακρόασης, η υπομονή και η ανάλυση σύνθετων προβλημάτων σε μικρότερα και ευκολότερα υποπροβλήματα. Όλα αυτά, μαζί με τις απαιτούμενες τεχνικές γνώσεις και δεξιότητες, καθιστούν εξαιρετικά ικανούς τους εργαζομένους. Σε αυτό το πλαίσιο, ολοένα και περισσότερα μαθήματα και σχολές ασχολούνται με την αξιολόγηση αλλά και την ανάπτυξη αυτών των γνώσεων και δεξιοτήτων στους εργαζομένους.

Επιπροσθέτως, η ίδια εργασία, καταλήγει ότι λόγω της στρατηγικής φύσης των περισσότερων εργασιών στην τεχνολογία της πληροφορίας, η κατοχή σημαντικών γνώσεων και δεξιοτήτων για το συντονισμό και τη διαχείριση των έργων, αυξάνει την ικανοποίηση των εργαζομένων στην εργασία τους, πράγμα το οποίο μπορεί να έχει θετικό αντίκτυπο τόσο στην καριέρα τους όσο και στις αποφάσεις τους. Για το λόγο αυτό, προτείνονται η εκπαίδευση και η καθοδήγηση των εργαζομένων στο πλαίσιο της εργασίας τους, ώστε να αποκτήσουν τέτοιες γνώσεις και δεξιότητες [8].

Τέλος, η συγκεκριμένη εργασία σημειώνει ότι οι προαναφερθείσες γνώσεις και δεξιότητες είναι σημαντικό να προσφέρονται όσο νωρίτερα γίνεται στους εργαζομένους ή και όσο ακόμα σπουδάζουν, μιας και είναι γενικά αποδεκτό ότι οι φοιτητές και οι μαθητές είναι γενικά πιο ανοιχτοί σε νέες ιδέες. Έτσι, η συγκεκριμένη εργασία τονίζει ότι πρέπει να γίνει

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

μια συντονισμένη προσπάθεια τόσο από την αγορά εργασίας όσο και από τα εκπαιδευτικά ιδρύματα στην τεχνολογία για την κάλυψη του κενού μεταξύ των διδασκόμενων γνώσεων και δεξιοτήτων από την τεχνολογική εκπαίδευση και των ζητούμενων γνώσεων και δεξιοτήτων από την αγορά εργασίας [8].

Μία ακόμη σημαντική εργασία για τα επαγγέλματα στην τεχνολογία της πληροφορίας είναι η [9]. Στη συγκεκριμένη εργασία αναγνωρίζεται το πρόβλημα των επιχειρήσεων του συγκεκριμένου κλάδου να διατηρήσουν στις τάξεις τους το προσωπικό τους, διερευνάται ο βαθμός στον οποίο ένας οργανισμός μπορεί να διατηρήσει το προσωπικό του στις τάξεις του καθώς επίσης και ο αντίκτυπος που έχει η ηγεσία που ασκεί ο εκάστοτε manager στην επιθυμία των εργαζομένων στην τεχνολογία της πληροφορίας να διατηρήσουν μακροπρόθεσμα την εργασία τους.

Τα αποτελέσματα καταδεικνύουν ότι κάποιες συγκεκριμένες ηγετικές συμπεριφορές (δημιουργία οράματος, παροχή κατάλληλου μοντέλου εργασίας και διοίκησης και η δημιουργία κοινού στόχου εντός της ομάδας) έχουν σημαντικό αντίκτυπο στην επιθυμία του προσωπικού να μείνει στην ομάδα ενώ άλλες συμπεριφορές που πιστεύεται ότι συμβάλλουν προς αυτή την κατεύθυνση δεν φαίνεται να έχουν σημαντικό αντίκτυπο στις επιθυμίες των εργαζομένων (π.χ. επίδειξη προσδοκιών υψηλής απόδοσης, βραβεία απόδοσης), παρότι μπορεί να βελτιώνουν την καθημερινότητα τους [9].

## 2.2 Εκπαίδευση στην τεχνολογία

Όπως αναφέρθηκε και στην προηγούμενη ενότητα, η δυσκολία εύρεσης των απαραίτητων γνώσεων και δεξιοτήτων στην τεχνολογία της πληροφορίας είναι ένα παγκόσμιο φαινόμενο. Οι οργανισμοί επιδιώκουν την επίτευξη καινοτομίας και την ανάπτυξη νέων τεχνολογιών μέσω του ανθρώπινου δυναμικού τους που προσδοκούν ότι θα κατέχει σημαντικές και σύγχρονες γνώσεις και δεξιότητες. Ωστόσο, η εύρεση τέτοιου προσωπικού έχει καταστεί δύσκολη τα τελευταία χρόνια. Για το συγκεκριμένο πρόβλημα, οι οργανισμοί επιρρίπτουν μεταξύ άλλων τις ευθύνες στους φορείς εκπαίδευσης, μιας και θεωρούν ότι η ανώτατη εκπαίδευση δεν παρέχει το κατάλληλο σύνολο γνώσεων και δεξιοτήτων στους αποφοίτους – ειδικά δεδομένου ότι ο κλάδος της τεχνολογίας της πληροφορίας αλλάζει ραγδαία και κατά συνέπεια μεταβάλλονται και οι απαιτήσεις σε γνώσεις και δεξιότητες, ενώ παράλληλα δημιουργούνται και νέες επαγγελματικές ευκαιρίες. Σε αυτό το πλαίσιο, οι απόφοιτοι τεχνολογικών εκπαιδευτικών ιδρυμάτων που εισέρχονται στην αγορά εργασίας καλούνται να έχουν αναπτύξει επαρκείς γνώσεις και δεξιότητες με βάση τις ανάγκες της αγοράς εργασίας [10].

Στην εργασία [10] το θέμα των γνώσεων και δεξιοτήτων που παρέχονται μέσω της ανώτατης εκπαίδευσης στον κλάδο της τεχνολογίας της πληροφορίας και των επικοινωνιών και την αναντιστοιχία τους με τις ανάγκες της αγοράς εργασίας αναλύεται εκτενώς για τη χώρα της Νότιας Αφρικής. Η συγκεκριμένη έρευνα καταλήγει ότι οι τεχνικές γνώσεις και δεξιότητες σύντομα θεωρούνται παρωχημένες και κατά συνέπεια οι απόφοιτοι τεχνολογικών εκπαιδευτικών ιδρυμάτων πρέπει να παρακολουθούν τις εξελίξεις της τεχνολογίας και να αναπτύσσουν νέες γνώσεις και δεξιότητες, ώστε να παραμένουν ανταγωνιστικοί. Επιπροσθέτως, η συγκεκριμένη έρευνα τονίζει την ανάγκη για εκμάθηση και προσωπικών μη τεχνικών δεξιοτήτων (soft skills) στους φοιτητές, που είναι κρίσιμες για τους επαγγελματίες

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

του κλάδου. Σε αυτό το πλαίσιο, προτείνεται ένα πλαίσιο (framework) για την ενδυνάμωση τόσο των επαγγελματιών του κλάδου, όσο και των εκπαιδευτικών ιδρυμάτων και της αγοράς εργασίας.

Με το θέμα των γνώσεων και δεξιοτήτων που παρέχονται από την τεχνολογική εκπαίδευση και την επάρκεια τους σε σχέση με τις ανάγκες της αγοράς εργασίας στον τομέα της τεχνολογίας της πληροφορίας καταπιάνεται και η εργασία [11]. Πιο συγκεκριμένα, οι συγγραφείς αναγνωρίζουν τη αναντιστοιχία μεταξύ της προσφερόμενης γνώσης από την ανώτατη εκπαίδευση και των αναγκών της αγοράς εργασίας και προσπαθούν να κατανοήσουν, μέσω μιας εκτενούς βιβλιογραφικής ανασκόπησης, τις γνώσεις και δεξιότητες που πρέπει να αναπτύσσονται από τα εκπαιδευτικά ιδρύματα καθώς επίσης και στα προβλήματα που παρουσιάζουν τα προγράμματα σπουδών στο αντικείμενο της τεχνολογίας της πληροφορίας.

Πιο συγκεκριμένα, οι συγγραφείς αναγνωρίζουν ως σημαντικό πρόβλημα το γεγονός ότι παρότι οι ανάγκες της αγοράς εργασίας μεταβάλλονται με ραγδαίους ρυθμούς τα προγράμματα σπουδών παραμένουν σχετικά αμετάβλητα, με αποτέλεσμα τεχνολογίες αιχμής που είναι αναγκαίες για την αγορά εργασίας να μη διδάσκονται στα εκπαιδευτικά ιδρύματα. Επιπλέον, οι συγγραφείς αναγνωρίζουν διαφορετικές κατηγορίες γνώσεων και δεξιοτήτων που πρέπει να εισαχθούν στα προγράμματα σπουδών των εκπαιδευτικών ιδρυμάτων. Από θέματα σχετικά με την μεγάλη εικόνα (π.χ. δυνατότητες εξέλιξης, πιθανά επόμενα βήματα καριέρας (career path), την εξέλιξη του κλάδου), μέχρι διαφορετικές εξειδικεύσεις των επαγγελματιών του κλάδου της τεχνολογίας της πληροφορίας και τις διαφορετικές γνώσεις και δεξιότητες που απαιτούνται για κάθε εξειδίκευση [11].

### **2.3 Κάλυψη των απαιτήσεων της αγοράς εργασίας σε ικανότητες μέσω της εκπαίδευσης**

Όπως αναφέρθηκε και στις προηγούμενες ενότητες, η αγορά εργασίας, κυρίως στην τεχνολογία, εξελίσσεται ραγδαία, όσον αφορά τις ζητούμενες γνώσεις και δεξιότητες, με αποτέλεσμα, γνώσεις και δεξιότητες που ήταν σε μεγάλη ζήτηση πριν από λίγα χρόνια, σήμερα να θεωρούνται παρωχημένες. Συνεπώς, μερικοί εργαζόμενοι στην τεχνολογία που ολοκλήρωσαν τις σπουδές τους σε τεχνολογικά εκπαιδευτικά ιδρύματα πριν πολύ καιρό και εργάζονται μεγάλο χρονικό διάστημα σε τεχνολογίες που ήταν σε άνθιση μερικά χρόνια νωρίτερα, δυσκολεύονται να σημειώσουν πρόοδο στην καριέρα τους ή να βρουν μια νέα θέση εργασίας, παρά τη μεγάλη εμπειρία στο αντικείμενο τους. Για το λόγο αυτό, πολλοί εργαζόμενοι προσπαθούν να αποκτήσουν νέες γνώσεις και δεξιότητες μέσω του πανεπιστήμιου (π.χ. μεταπτυχιακά προγράμματα) ή ανοιχτών μαθημάτων που παρέχονται από διάφορους εκπαιδευτικούς οργανισμούς.

Από την άλλη πλευρά, μια βασική πρόκληση που έχουν να αντιμετωπίσουν τα εκπαιδευτικά ιδρύματα, είναι να καταφέρουν να συμβαδίσουν με την εξέλιξη της αγοράς εργασίας, ώστε να αποφεύγεται η διδασκαλία παρωχημένων και «παλιομοδίτικων» γνώσεων και δεξιοτήτων στους φοιτητές και τους δια βίου μαθητές [12]. Επιπλέον, είναι πολύ σημαντικό και για τους τελευταίους να επιλέγουν μαθήματα με κριτήριο τις ανάγκες της αγοράς αλλά και τις προσωπικές τους φιλοδοξίες για το μέλλον και να αποφεύγουν μαθήματα που

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

προσφέρουν παρωχημένες γνώσεις και δεξιότητες και τεχνολογίες που δε ζητούνται πλέον από την αγορά εργασίας.

Εξίσου σημαντικό είναι και για τα εκπαιδευτικά ιδρύματα και τους διδάσκοντες μαθημάτων να κινηθούν προς την μείωση της απόστασης μεταξύ των ζητούμενων γνώσεων και δεξιοτήτων στην αγορά εργασίας και των προσφερόμενων ικανοτήτων από τον εκπαιδευτικό οργανισμό, καθότι σε μεγάλο βαθμό ο ρόλος του τελευταίου είναι να προετοιμάσει τους φοιτητές και τους δια βίου μαθητές για την αγορά εργασίας και κατά συνέπεια η αδυναμία να συμβαδίσει με τις εξελίξεις της αγοράς εργασίας μπορεί να οδηγήσει στην παρακμή του. Αυτό συμβαίνει διότι τα εκπαιδευτικά ιδρύματα αντιμετωπίζουν μεγάλο ανταγωνισμό μιας και ολοένα και περισσότερα διαδικτυακά ανοιχτά μαθήματα (Massive Online Open Courses - MOOCs) δημιουργούνται με σκοπό την διδασκαλία γνώσεων και δεξιοτήτων αιχμής στους μαθητές που τα επιλέγουν. Σε αυτό το πλαίσιο, όπως αναγνωρίζεται και από την εργασία [7], τα εκπαιδευτικά ιδρύματα στην τεχνολογία πρέπει να κάνουν μεγάλη προσπάθεια για να συμβαδίσουν με τις ανάγκες της αγοράς εργασίας.

Υπάρχει πληθώρα εργασιών που κινούνται προς την κατεύθυνση της διόρθωσης της έλλειψης των ζητούμενων γνώσεων και δεξιοτήτων στο εργατικό δυναμικό στην τεχνολογία, από την τεχνολογική αγορά εργασίας, οι οποίες κινούνται σε διαφορετικούς άξονες. Συγκεκριμένα, υπάρχουν εργασίες που παρακολουθούν και αναλύουν την αγορά εργασίας όσον αφορά τις απαιτούμενες γνώσεις και δεξιότητες, άλλες που αφορούν συστήματα συστάσεων που προτείνουν γνώσεις και δεξιότητες προς ανάπτυξη ή σχετικά μαθήματα, με βάση τις ανάγκες της αγοράς εργασίας, όπως και κάποιες που καταπιάνονται με το σχεδιασμό του προγράμματος σπουδών. Τέλος, υπάρχουν εργασίες που προτείνουν τα συστήματα υποστήριξης αποφάσεων ως σημαντικό εργαλείο αποφάσεων τόσο στην εκπαίδευση όσο και στην αγορά εργασίας.

Όσον αφορά τις εργασίες για την αγορά εργασίας, διάφορα επιστημονικά άρθρα έχουν δημοσιευτεί τα τελευταία χρόνια και στο παρόν κεφάλαιο θα παρουσιαστούν τα πιο αντιπροσωπευτικά. Συγκεκριμένα, στην εργασία [13], οι συγγραφείς αναγνωρίζουν την ραγδαία εξέλιξη της αγοράς κατά τη διάρκεια της τέταρτης βιομηχανικής επανάστασης (Industry 4.0) και των σχετικών επαγγελματιών και αξιοποιούν πληροφορία από αγγελίες για θέσεις εργασίας για να αντλήσουν μια γενική εικόνα για την σχετική αγορά εργασίας με τη χρήση τεχνικών αναλυτικής κειμένου (text analytics/text mining). Η ανάλυση τους καταλήγει στο συμπέρασμα ότι υπάρχουν πολλά γενικευμένα προφίλ θέσεων εργασίας τα οποία μπορούν να χωριστούν σε δυο κατηγορίες, η πρώτη περιλαμβάνει θέσεις εργασίας που στοχεύουν στο καθαρά τεχνικό σκέλος και η δεύτερη περιλαμβάνει θέσεις σχετικές με τη διοίκηση και την επιστήμη υπολογιστών. Άλλη σχετική έρευνα [14], αξιοποιεί επίσης τεχνικές αναλυτικής κειμένου, παράλληλα με τη βάση δεδομένων O\*NET [15] για να εντοπίσει τα επαγγέλματα που είναι σε μεγαλύτερη ζήτηση στην αγορά εργασίας ανά κλάδο και περιοχή. Τέλος, στην έρευνα [16], οι συγγραφείς διερευνούν τις ικανότητες που απαιτούνται από αρχάριους εργαζόμενους ανθρωπίνων πόρων, αξιοποιώντας τεχνικές αναλυτικής κειμένου σε αγγελίες εργασίας. Το συμπέρασμα είναι ότι υπάρχουν τέσσερις κρίσιμοι παράγοντες για την επιτυχία των εργαζομένων ανθρωπίνων πόρων. Οι παράγοντες αυτοί περιλαμβάνουν συγκεκριμένες γνώσεις και δεξιότητες, ακαδημαϊκά προσόντα, εμπειρία, αλλά και συγκεκριμένες έμφυτες ικανότητες.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Αναφορικά με τα συστήματα συστάσεων, υπάρχουν πολλές εργασίες που παρουσιάζουν τέτοια συστήματα, τόσο για τη σύσταση γνώσεων και δεξιοτήτων όσο και σχετικών μαθημάτων. Επιπλέον, τα ανοιχτά διαδικτυακά μαθήματα έχουν αναπτυχθεί σημαντικά τα τελευταία χρόνια και είναι ενδεικτικό της σημασίας τέτοιων συστημάτων, ότι εταιρείες σαν το Coursera<sup>2</sup> δαπανούν σημαντικούς πόρους για την υλοποίηση ή τη βελτίωση των συστημάτων συστάσεων τους. Σε αυτό το πλαίσιο, έχουν δημοσιευτεί πολλές εργασίες στη συγκεκριμένη επιστημονική περιοχή. Συγκεκριμένα στην εργασία [17], οι ερευνητές ανέπτυξαν ένα σύστημα συστάσεων ανοιχτών διαδικτυακών μαθημάτων που βασίζεται σε καταναμημένα υπολογιστικά συστήματα. Επιπροσθέτως, στην εργασία [18] οι ερευνητές ανέπτυξαν ένα καινοτόμο σύστημα συστάσεων που προτείνει μαθήματα στους φοιτητές για να τους προετοιμάσει για άλλα πιο προχωρημένα μαθήματα που έχουν θέσει ως στόχο. Αυτό γίνεται με τη χρήση ενός αναδραστικού νευρωνικού δικτύου (recurrent neural network). Τέλος στην εργασία [19] οι συγγραφείς διερεύνησαν πώς τα χαρακτηριστικά των μαθητών και των μαθημάτων επηρεάζουν την πιθανότητα επιλογής ενός μαθήματος και ανέπτυξαν ένα σύστημα συστάσεων που λαμβάνει υπόψη του και τέτοια χαρακτηριστικά στον αλγόριθμο συστάσεων του. Τα αποτελέσματα φαίνεται να παρουσιάζουν βελτίωση σε σχέση με τις παραδοσιακές τεχνικές συστάσεων που δε λαμβάνουν υπόψη τέτοια χαρακτηριστικά.

Όσον αφορά τα συστήματα συστάσεων δεξιοτήτων, η εργασία [20] συνδυάζει σύσταση δεξιοτήτων και θέσεων εργασίας. Συγκεκριμένα, το σύστημα που αναπτύχθηκε προτείνει κατάλληλες θέσεις εργασίας για τον χρήστη, σε συνδυασμό με ικανότητες που θα βελτιώσουν τις πιθανότητες του τελευταίου να επιλεγεί για μια θέση εργασίας. Τέλος, στην εργασία [21] οι συγγραφείς ανέπτυξαν ένα σύστημα συστάσεων που προτείνει προσωποποιημένα μονοπάτια καριέρας (personalized career paths) στους χρήστες και το αποτέλεσμα παρουσιάζεται ως ένας γράφος από δεξιότητες. Τα αποτελέσματα παρουσιάζουν πολύ υψηλές τιμές στην ακρίβεια και στην ανάκληση (precision and recall) για τις παρεχόμενες προτάσεις.

Σχετικά με τον σχεδιασμό του προγράμματος σπουδών, δεν υπάρχουν πολλές εφαρμογές που να το υποστηρίζουν σε σχέση με τις ανάγκες της αγοράς εργασίας. Πιο συγκεκριμένα, η εργασία [22] εισάγει ένα πλαίσιο εργασίας (framework) για να βοηθήσει τους υπεύθυνους του προγράμματος σπουδών να αναγνωρίσουν την αναντιστοιχία σε γνώσεις και δεξιότητες μεταξύ των παρεχόμενων γνώσεων και δεξιοτήτων του εργατικού δυναμικού και των ζητούμενων γνώσεων και δεξιοτήτων από την αγορά εργασίας. Το συγκεκριμένο πλαίσιο μπορεί να χρησιμοποιηθεί για την αυτοματοποιημένη ανάλυση και αναγνώριση των απαιτούμενων γνώσεων και δεξιοτήτων από την αγορά και των παρεχόμενων από το εργατικό δυναμικό, χρησιμοποιώντας πληροφορίες τόσο από αγγελίες θέσεων εργασίας όσο και πληροφορίες σχετικές με το πρόγραμμα σπουδών, για την παροχή πιο αποτελεσματικών μαθημάτων.

Επιπροσθέτως, η εργασία [23], εξέτασε μεγάλο αριθμό αγγελιών θέσεων εργασίας στον τομέα της τεχνολογίας της πληροφορίας και τα συνέκρινε με ένα πρόγραμμα σπουδών στο ίδιο αντικείμενο και κατέληξε ότι κάποια καινούρια μαθήματα πρέπει να εισαχθούν στο πρόγραμμα σπουδών. Επίσης, εντόπισε τις σημαντικότερες εξειδικεύσεις στην αγορά

---

<sup>2</sup> <https://about.coursera.org/>



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

εργασίας, όπως επίσης και τα σημαντικότερα προσόντα. Τέλος, αναγνώρισε ότι κάποιες εξειδικεύσεις στον κλάδο της τεχνολογίας της πληροφορίας δεν μπορούσαν να υποστηριχθούν από το δοθέν πρόγραμμα σπουδών.

Ακόμη στην εργασία [24] η συγγραφέας παρουσιάζει τη μεθοδολογία που ακολουθήθηκε για τη σχεδίαση του προγράμματος σπουδών στη σχολή ανάπτυξης λογισμικού στο πανεπιστήμιο Carnegie Mellon. Τέλος, η εργασία [25], παρουσιάζει μια εκτενή βιβλιογραφική ανασκόπηση σχετικά με τις έρευνες που έχουν πραγματοποιηθεί για την σύγκλιση του προγράμματος σπουδών με τις ανάγκες της αγοράς εργασίας και συζητά εκτενώς τα σημαντικότερα προβλήματα για τον σχεδιασμό του προγράμματος σπουδών σύμφωνα με τις ανάγκες της αγοράς εργασίας.

Τέλος, όσον αφορά την υποστήριξη αποφάσεων για την ανώτατη εκπαίδευση, στην εργασία [26] αναγνωρίστηκε η ανάγκη για εργαλεία υποστήριξης αποφάσεων και πραγματοποιήθηκε ανάλυση απαιτήσεων για την ανάπτυξη τέτοιων συστημάτων βασισμένων σε πραγματικά δεδομένα. Αντίστοιχα, η εργασία [27] προτείνει μια μεθοδολογία για την ανάπτυξη εργαλείων υποστήριξης αποφάσεων για ακαδημαϊκά περιβάλλοντα τα οποία θα βελτιώσουν την ποιότητα των εκπαιδευτικών διαδικασιών και της διοίκησης.

Τέλος, αξίζει να σημειωθεί ότι τα συστήματα συστάσεων χρησιμοποιούνται σε πολλές περιπτώσεις ως συστήματα υποστήριξης αποφάσεων, καθότι προτείνουν επιλογές στο χρήστη σύμφωνα με διάφορα κριτήρια όπως επίσης και τις προτιμήσεις του [28]. Έτσι, ο συνδυασμός συστημάτων συστάσεων με συστήματα υποστήριξης αποφάσεων, όπως τα πολυκριτηριακά συστήματα υποστήριξης αποφάσεων, δείχνει να έχει σημαντικές προοπτικές.

## 2.4 Το έργο QualiChain

Στο πλαίσιο της γεφύρωσης του χάσματος μεταξύ της εκπαίδευσης και των αναγκών της αγοράς εργασίας κινείται και το Ευρωπαϊκό ερευνητικό έργο QualiChain. Συγκεκριμένα το QualiChain στοχεύει στην υλοποίηση, την αξιολόγηση και την πιλοτική χρήση μιας αποκεντρωμένης πλατφόρμας για την αποθήκευση, τον διαμοιρασμό και την επαλήθευση εκπαιδευτικών και εργασιακών προσόντων και επικεντρώνεται στην τεχνολογία του blockchain, στις αλγοριθμικές τεχνικές και στην υπολογιστική νοημοσύνη για την αναδιάρθρωση της δημόσιας εκπαίδευσης και της σύνδεσης της με την ιδιωτική εκπαίδευση και την αγορά εργασίας και τη βελτίωση των διαχειριστικών διαδικασιών στο δημόσιο τομέα.

Πιο συγκεκριμένα, το QualiChain στοχεύει στη διερεύνηση και την επιβεβαίωση ότι οι τεχνολογίες αιχμής, όπως το blockchain, οι σημασιολογικές τεχνολογίες (semantics/semantic technologies), η αναλυτική δεδομένων και η χρήση τεχνικών gamification, μπορούν να έχουν πολύ σημαντικό αντίκτυπο στη δημόσια και την ιδιωτική εκπαίδευση, την αγορά εργασίας όπως επίσης και στις διαχειριστικές διαδικασίες του δημοσίου τομέα. Αυτό επιτυγχάνεται από τη σχεδίαση, την ανάπτυξη και την πιλοτική χρήση και αξιολόγηση όσον αφορά τα οφέλη και τα ρίσκα της τεχνικής λύσης QualiChain. Πρόκειται για μια κατανεμημένη πλατφόρμα που επιτρέπει την αποθήκευση, τον διαμοιρασμό και την επαλήθευση ακαδημαϊκών και εργασιακών προσόντων και πιστοποιήσεων. Στο σημείο αυτό, αξίζει να αναφερθεί ότι παρότι η αρχική ιδέα για το έργο πυροδοτήθηκε από την ανάγκη της δημόσιας εκπαίδευσης για αποδοτική αποθήκευση και επαλήθευση πιστοποιήσεων και για αποτελεσματικότερη

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

αντιμετώπιση της απάτης που συναντάται γύρω από τις εκπαιδευτικές πιστοποιήσεις και προσόντα, η γενική ιδέα είναι πολύ πιο ευρεία, και συμπεριλαμβάνει λύσεις στα σημαντικότερα θέματα που συναντώνται στην εκπαίδευση και στην αγορά εργασίας τόσο στο δημόσιο όσο και στον ιδιωτικό τομέα. Τέτοια θέματα περιλαμβάνουν τη δια βίου μάθηση, την στελέχωση οργανισμών, την κινητικότητα του ανθρώπινου δυναμικού καθώς επίσης και την καλύτερη σύνδεση μεταξύ της αγοράς εργασίας και της εκπαίδευσης. Όλα αυτά επιτυγχάνονται καλύπτοντας τις ανάγκες διαφορετικών ενδιαφερόμενων μερών (stakeholders).

Στην πραγματικότητα οι υπηρεσίες που προσφέρει η πλατφόρμα QualiChain χωρίζονται σε δύο βασικές κατηγορίες, τις βασικές υπηρεσίες και τις υπηρεσίες προστιθέμενης αξίας. Η πρώτη κατηγορία αποτελείται από τις βασικές τεχνολογίες του QualiChain, δηλαδή του blockchain και των σημασιολογικών τεχνολογιών, οι οποίες επιτρέπουν την αποθήκευση, την αρχειοθέτηση και την επαλήθευση για τις ακαδημαϊκές πιστοποιήσεις και άλλα ακαδημαϊκά προσόντα. Επιπλέον, στην συγκεκριμένη κατηγορία συμπεριλαμβάνονται και υπηρεσίες μετάφρασης και ελέγχου ισοτιμίας πιστοποιήσεων, όπως επίσης και η διαχείριση χαρτοφυλακίου (portfolio management) προσόντων και πιστοποιήσεων.

Όσον αφορά τη δεύτερη κατηγορία υπηρεσιών (υπηρεσίες προστιθέμενης αξίας), χρησιμοποιεί ως βάση τις υπηρεσίες που παρέχονται από την πρώτη κατηγορία και προσφέρει ένα σύνολο από πιο σύνθετες υπηρεσίες, όπως μεταξύ άλλων η συμβουλευτική καριέρας (career counselling), η έξυπνη δημιουργία προφίλ (intelligent profiling), υπηρεσίες διαχείρισης προσόντων (competency management), αναλύσεις δεδομένων για την αγορά εργασίας και την εκπαίδευση, εκπαιδευτικές συμβουλές, υποστήριξη αποφάσεων για επάνδρωση επιχειρήσεων κ.ά.. Αυτές οι υπηρεσίες χρησιμοποιούν τεχνολογίες αιχμής όπως η αναλυτική δεδομένων (data analytics), η μηχανική μάθηση (machine learning), τα συστήματα υποστήριξης αποφάσεων και η παιχνιδοποίηση (gamification), για την παροχή αξιόπιστων αποτελεσμάτων και συμβουλών τόσο για την αγορά εργασίας όσο και για την εκπαίδευση.

Οι παραπάνω υπηρεσίες χρησιμοποιήθηκαν πιλοτικά από 5 οργανισμούς δημόσιους και ιδιωτικούς κατανεμημένους σε διάφορες χώρες της Ευρώπης. Στους οργανισμούς αυτούς συμπεριλαμβάνονται εκπαιδευτικά ιδρύματα, όπως επίσης και οργανισμοί προσλήψεων και οι υπηρεσίες του QualiChain εξυπηρετούν διαφορετικά σενάρια χρήσης τα οποία απεικονίζουν τις πραγματικές ανάγκες του κάθε οργανισμού.

Πιο συγκεκριμένα, τα βασικότερα σενάρια χρήσης περιλαμβάνουν α) την επαλήθευση ισοδυναμίας πτυχίων για πτυχία διαφορετικών πανεπιστημίων, β) τη δημιουργία έξυπνου προγράμματος σπουδών για ένα πανεπιστημιακό τμήμα, γ) την επάνδρωση με εργατικό δυναμικό του δημοσίου τομέα και δ) τη συμβουλευτική ανθρώπινου δυναμικού και τη διαχείριση προσόντων. Το πρώτο σενάριο (επαλήθευση ισοδυναμίας πτυχίων για πτυχία διαφορετικών πανεπιστημίων) εφαρμόζεται πιλοτικά από ένα πανεπιστήμιο στη Μεγάλη Βρετανία. Το δεύτερο σενάριο (δημιουργία έξυπνου προγράμματος σπουδών για ένα πανεπιστημιακό τμήμα) εφαρμόζεται πιλοτικά από πανεπιστήμιο στην Ελλάδα. Το τρίτο σενάριο (επάνδρωση με εργατικό δυναμικό του δημοσίου τομέα) εφαρμόζεται πιλοτικά σε μεγάλο δημόσιο οργανισμό προσλήψεων της Ελλάδας όσο και σε αντίστοιχο δημόσιο

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

οργανισμό της Πορτογαλίας. Τέλος, το τέταρτο σενάριο (συμβουλευτική ανθρώπινου δυναμικού και διαχείριση προσόντων) εφαρμόζεται σε δημόσιο οργανισμό της Πορτογαλίας.

## 2.5 Συμπεράσματα

Με βάση την ανάλυση που προηγήθηκε, επιβεβαιώνεται ότι υπάρχει σημαντική απόσταση μεταξύ των γνώσεων και δεξιοτήτων του εργατικού δυναμικού στην τεχνολογία και των απαιτούμενων γνώσεων και δεξιοτήτων στην αγορά εργασίας στην τεχνολογία. Το συγκεκριμένο φαινόμενο πηγάζει σε μεγάλο βαθμό από το γεγονός ότι τα τεχνολογικά εκπαιδευτικά ιδρύματα εξελίσσονται με πολύ αργούς ρυθμούς σε σχέση με την εξέλιξη της τεχνολογίας και των αναγκών της αγοράς εργασίας και ως εκ τούτου δυσκολεύονται να συμβαδίσουν με τις ανάγκες της τελευταίας. Αυτό συμβαίνει διότι είναι δύσκολη η συνεχής παρακολούθηση των αναγκών της αγοράς εργασίας και η άντληση σημαντικών συμπερασμάτων για τις αλλαγές που μπορεί να έχουν πραγματοποιηθεί.

Για τον μετριασμό της απόστασης μεταξύ των αναγκών της αγοράς και των γνώσεων και δεξιοτήτων του εργατικού δυναμικού, θα πρέπει αφ' ενός τα εκπαιδευτικά ιδρύματα να παρακολουθούν την αγορά εργασίας και να παρατηρούν τυχούσες αλλαγές, τόσο σε ό,τι αφορά σε νέες ανάγκες για γνώσεις και δεξιότητες, όσο και σε τάσεις μείωσης της ζήτησης για άλλες γνώσεις και δεξιότητες. Στη συνέχεια, και με βάση τις αλλαγές που αναγνωρίζονται, τα προγράμματα σπουδών και τα μαθήματα αυτών των ιδρυμάτων θα πρέπει να ενημερώνονται κατάλληλα.

Αφ' ετέρου, και οι φοιτητές και οι δια βίου μαθητές θα πρέπει να έχουν μια γενική εικόνα για τις απαιτήσεις της αγοράς εργασίας και να διαλέγουν μαθήματα που να τους προσφέρουν γνώσεις και δεξιότητες οι οποίες είναι σε μεγάλη ζήτηση από την αγορά εργασίας, ώστε να τους βοηθήσουν με την μετέπειτα επαγγελματική τους πορεία και παράλληλα να αποφεύγουν μαθήματα που προσφέρουν γνώσεις και δεξιότητες που πλέον δεν ζητούνται από την αγορά εργασίας.

Ωστόσο, είναι αρκετά δύσκολο και χρονοβόρο τόσο για τα εκπαιδευτικά ιδρύματα, όσο και για τους φοιτητές να παρακολουθήσουν τις εξελίξεις στην αγορά εργασίας, κυρίως λόγω του τεράστιου όγκου πληροφορίας που συναντάται στο διαδίκτυο, η οποία προσφέρεται από πολλές ετερογενείς πηγές με αδόμητο ή ημιδομημένο τρόπο. Ο όγκος της πληροφορίας καθιστά δύσκολη την ανθρώπινη ανάλυση, καθώς απαιτείται πολύ μεγάλη προσπάθεια για ομαδοποίηση και ερμηνεία όλης αυτής της πληροφορίας, ενώ η έλλειψη συγκεκριμένης και ενιαίας δομής στα δεδομένα καθιστά δύσκολη την εξαγωγή των δεδομένων για την ανάλυση τους με υπολογιστικές μεθόδους.

Σε αυτό το πλαίσιο, και όπως φαίνεται και από την ανάλυση που προηγήθηκε, ενδείκνυται η χρήση τεχνολογιών εξαγωγής και ανάλυσης δεδομένων τόσο για την απόκτηση των δεδομένων, την επεξεργασία και την ομογενοποίηση τους, όσο και για την εξαγωγή συμπερασμάτων και παροχής προτάσεων για όλα τα ενδιαφερόμενα μέρη (ακαδημαϊκά ιδρύματα, φοιτητές κλπ.). Τέτοιες τεχνολογίες περιλαμβάνουν την εξόρυξη δεδομένων από τον παγκόσμιο ιστό (web mining), την εξόρυξη γνώσης από δεδομένα, τη μηχανική μάθηση, την επεξεργασία φυσικής γλώσσας (Natural Language Processing - NLP), και τα συστήματα συστάσεων (Recommender Systems). Επιπλέον, σημαντικό παράγοντα αποτελεί και η δυνατότητα αυτοματοποίησης της διαδικασίας, ώστε να βασίζεται πάντα σε έγκυρα και

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

επίκαιρα δεδομένα και τυχόν αλλαγές στις απαιτήσεις της αγοράς εργασίας να αναγνωρίζονται εγκαίρως.

## 2.6 Αναφορές

- [1] B. Rios, “The skills gap – an economic burden for the EU,” 2019.
- [2] T. Hogarth, “COVID-19 and the Demand for Labour and Skills in Europe Early evidence and implications for migration policy,” 2021.
- [3] Eurostat Statistics Explained, “Educational attainment statistics,” *Eurostat*, vol. 74, no. April 2019, 2018.
- [4] CEDEFOP, “European skills and jobs survey.”, 2021
- [5] A. Avram, M. Benvenuto, C. D. Avram, and G. Gravili, “Assuring SME’s sustainable competitiveness in the digital Era: A labor policy between guaranteed minimum wage and ICT skill mismatch,” *Sustain.*, vol. 11, no. 10, 2019, doi: 10.3390/su11102918.
- [6] EU, “A sustainable future for Europe (HORIZON-CL2-2022-TRANSFORMATIONS-01) - Conditions for the successful development of skills matched to needs.” 2022.
- [7] C. Kontzinos, V. Karakolis, S. Skalidakis, O. Markaki, M. Androutopoulou, and J. Psarras, “COMBINING BLOCKCHAIN, SEMANTICS AND DATA ANALYTICS FOR UNIVERSITY PROCESS OPTIMISATION,” *IADIS-INTERNATIONAL J. Comput. Sci. Inf. Syst.*, vol. 14, no. 2, 2019.
- [8] M. Taein Eom and C. Lim, “Critical skills to be competent and relevant IT personnel: Do today’s IT personnel possess requisite skills?,” *J. Inf. Technol. Manag.*, 2012.
- [9] M. T. in Eom, “How Can Organization Retain IT Personnel? Impact of IT Manager’s Leadership on IT Personnel’s Intention to Stay,” *Inf. Syst. Manag.*, vol. 32, no. 4, 2015, doi: 10.1080/10580530.2015.1080001.
- [10] A. Harmse, “Expectations of industry regarding the ICT competencies of HEI graduates,” 2018.
- [11] T. Ngo-Ye and J. Choi, “QUALIFICATION REQUIREMENTS FOR ENTRY-LEVEL IT POSITIONS AND IS CURRICULUM: A LITERATURE REVIEW,” *Issues Inf. Syst.*, 2021, doi: 10.48009/3\_iis\_2021\_44-57.
- [12] C. Kontzinnos, O. Markaki, P. Kokkinakos, V. Karakolis, S. Skalidakis, and J. Psarras, “Using Blockchain, Semantics and Data Analytics to Optimise Qualification Certification, Recruitment and Competency Management: a Landscape Review,” 2020.
- [13] M. Pejic-bach, T. Bertonsel, M. Meško, and Ž. Krstić, “International Journal of Information Management Text mining of industry 4 . 0 job advertisements,” *Int. J. Inf. Manage.*, vol. 50, 2020.
- [14] I. Karakatsanis *et al.*, “Data mining approach to monitoring the requirements of the job market: A case study,” *Inf. Syst.*, vol. 65, 2017, doi: 10.1016/j.is.2016.10.009.
- [15] N. R. Council, “A Database for a Changing Economy: Review of the Occupational Information Network (O\*NET),” *Natl. Acad. Press*, 2010.

- [16] C. H. Chung and L. J. Chen, "Text mining for human resources competencies: Taiwan example," *Eur. J. Train. Dev.*, vol. 45, no. 6–7, 2019, doi: 10.1108/EJTD-07-2018-0060.
- [17] H. Zhang, T. Huang, Z. Lv, S. Y. Liu, and Z. Zhou, "MCRS: A course recommendation system for MOOCs," *Multimed. Tools Appl.*, vol. 77, no. 6, 2018, doi: 10.1007/s11042-017-4620-2.
- [18] W. Jiang, Z. A. Pardos, and Q. Wei, "Goal-based course recommendation," 2019, doi: 10.1145/3303772.3303814.
- [19] A. Elbadrawy and G. Karypis, "Domain-aware grade prediction and top-n course recommendation," 2016, doi: 10.1145/2959100.2959133.
- [20] V. S. Dave, B. Zhang, M. Al Hasan, K. AlJadda, and M. Korayem, "A combined representation learning approach for better job and skill recommendation.," 2018.
- [21] A. Gugnani, V. K. Reddy Kasireddy, and K. Ponnalagu, "Generating unified candidate skill graph for career path recommendation," in *IEEE International Conference on Data Mining Workshops, ICDMW*, 2019, vol. 2018-November, doi: 10.1109/ICDMW.2018.00054.
- [22] A. Almaleh, M. A. Aslam, K. Saeedi, and N. R. Aljohani, "Align my curriculum: A framework to bridge the gap between acquired university curriculum and required market skills," *Sustain.*, vol. 11, no. 9, 2019, doi: 10.3390/su11092607.
- [23] R. W. Woolridge and R. Parks, "What's In and What's Out: Defining an Industry-Aligned IS Curriculum Using Job Advertisements," *J. High. Educ. Theory Pract.*, vol. 16, no. 2, 2016.
- [24] M. Shaw, "Software Engineering for the 21st Century: A Basis for Rethinking the Curriculum," *Carnegie Mellon Univ.*, 2005.
- [25] N. R. Aljohani, A. Aslam, A. Khadidos, and S.-U. Hassan, "Bridging the skill gap between the acquired university curriculum and the requirements of the job market: A data-driven analysis of scientific literature," *J. Innov. Knowl.*, vol. 7, no. 3, 2022.
- [26] V. Gorgan, "Requirement Analysis for a Higher Education Decision Support System. Evidence from a Romanian University," *Procedia - Soc. Behav. Sci.*, vol. 197, 2015, doi: 10.1016/j.sbspro.2015.07.165.
- [27] V. P. Bresfelean, N. Ghisoiu, R. Lacurezeanu, and D. A. Sitar-Taut, "Towards the development of decision support in academic environments," 2009, doi: 10.1109/ITI.2009.5196106.
- [28] I. Nunes and D. Jannach, "A systematic review and taxonomy of explanations in decision support and recommender systems," *User Model. User-adapt. Interact.*, vol. 27, no. 3–5, 2017, doi: 10.1007/s11257-017-9195-0.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## ΚΕΦΑΛΑΙΟ 3: ΑΝΑΣΚΟΠΗΣΗ ΥΦΙΣΤΑΜΕΝΗΣ ΚΑΤΑΣΤΑΣΗΣ – ΣΥΣΤΗΜΑΤΑ ΣΥΣΤΑΣΕΩΝ

### 3.1 Εισαγωγή

Τα συστήματα συστάσεων είναι εργαλεία λογισμικού που προτείνουν στον χρήστη αντικείμενα που μπορεί να του φανούν χρήσιμα. Τα προτεινόμενα αντικείμενα σχετίζονται με μια ευρεία γκάμα αποφάσεων που ο χρήστης καλείται να λάβει, από την απόφαση ποια αντικείμενα να αγοράσει, έως την απόφαση τι μουσική να ακούσει ή τι νέα να διαβάσει στο διαδίκτυο [1].

Ως αντικείμενο σύμφωνα με τους συγγραφείς του [1] ορίζεται οτιδήποτε προτείνει ένα σύστημα συστάσεων. Συνήθως ένα σύστημα συστάσεων ασχολείται με έναν συγκεκριμένο τύπο αντικειμένων και κατά συνέπεια, είναι σχεδιασμένο έτσι ώστε να δημιουργεί βέλτιστες προτάσεις για το συγκεκριμένο τύπο αντικειμένων. Στα παραδείγματα της προηγούμενης παραγράφου τα αντικείμενα είναι τα είδη προς αγορά, τα διάφορα μουσικά κομμάτια ή οι μουσικές συλλογές, και τα διάφορα άρθρα από το διαδίκτυο.

Τα συστήματα συστάσεων, φιλτράρουν τα διαθέσιμα δεδομένα, ανάλογα με τις ανάγκες και τις προτιμήσεις του χρήστη. Για να γίνει αποτελεσματικά το φιλτράρισμα είναι απαραίτητο να υπάρχει πληροφορία για την παρελθούσα συμπεριφορά του χρήστη, τα χαρακτηριστικά του, πληροφορίες για χρήστες με παρόμοια χαρακτηριστικά, τα χαρακτηριστικά των αντικειμένων κ.ά. [1]–[3].

Έτσι, τα συστήματα συστάσεων απευθύνονται σε χρήστες που δε διαθέτουν την απαραίτητη πληροφόρηση, τις γνώσεις ή την ικανότητα να αξιολογήσουν την πληθώρα εναλλακτικών επιλογών που το διαδίκτυο προσφέρει [4]. Για παράδειγμα, ένα σύστημα συστάσεων βιβλίων, μπορεί να κατευθύνει το χρήστη να επιλέξει ένα βιβλίο κοντά στα ενδιαφέροντα του μέσα από μια μεγάλη γκάμα εναλλακτικών επιλογών [1].

Πέρα από τους χρήστες με ελλιπείς γνώσεις και ικανότητες, τα συστήματα συστάσεων απευθύνονται στον καθένα ξεχωριστά σε όλες τις εκφάνσεις της ζωής, από τη διασκέδαση και τα ψώνια μέχρι την κοινωνικοποίηση με ολοένα και αυξανόμενους ρυθμούς. Αυτό διότι κάθε χρήστης του διαδικτύου εκτίθεται σε ολοένα και αυξανόμενη πληροφορία και τα συστήματα συστάσεων δρουν ως προσωποποιημένο φίλτρο για κάθε χρήστη, αποτρέποντας την υπερφόρτωση του τελευταίου με πληροφορία (information overload) [5]. Μάλιστα, έχει γραφτεί ότι στις μέρες μας η υπερφόρτωση του χρήστη με πληροφορία είναι τόσο μεγάλη που χωρίς τα συστήματα συστάσεων να φιλτράρουν την πληροφορία που φτάνει στο χρήστη η αποτελεσματική πλοήγηση στο διαδίκτυο θα ήταν αδύνατη [6]. Ένα παράδειγμα που καταδεικνύει το συγκεκριμένο φαινόμενο είναι η απόφαση για παρακολούθηση κάποιου ταινίας ή σειράς στην πλατφόρμα Netflix. Συγκεκριμένα, εάν η πλατφόρμα δεν είχε το δικό της προηγμένο σύστημα συστάσεων, και αντ' αυτού εμφάνιζε το περιεχόμενο της με αλφαβητική ή χρονολογική σειρά, ο χρήστης θα βρισκόταν αντιμέτωπος με μια λίστα χιλιάδων αντικειμένων για να διαλέξει αυτό που τον ενδιαφέρει. Ως εκ τούτου, η επιλογή περιεχομένου προς το τέλος της παρεχόμενης λίστας θα ήταν σχεδόν απίθανη. Φυσικά, και με το υπάρχον σύστημα συστάσεων είναι πιθανό ο χρήστης

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

να μην λάβει γνώση για κάποιο αντικείμενο, ωστόσο το αντικείμενο αυτό δε θα του προταθεί γιατί απέχει πολύ από τις προτιμήσεις του.

Στην απλούστερη μορφή τους τα συστήματα συστάσεων παρέχουν προσωποποιημένες προτάσεις αντικειμένων ως ταξινομημένες λίστες αντικειμένων με βάση την προβλεπόμενη προτίμηση του εκάστοτε χρήστη. Σε αυτό το πλαίσιο γίνεται εύκολα αντιληπτό ότι η επιτυχία ενός συστήματος εξαρτάται από την ικανότητα του να προβλέπει τις προτιμήσεις του χρήστη. Για να επιτευχθούν αξιόπιστες προβλέψεις των προτιμήσεων του χρήστη τα συστήματα συστάσεων συλλέγουν πληροφορίες με βάση τις αντιδράσεις τους και τη δραστηριότητα τους στο σύστημα. Για παράδειγμα, συλλέγουν βαθμολογίες που έχει δώσει ο χρήστης στα αντικείμενα, εάν ένα αντικείμενο του κέντρισε το ενδιαφέρον, δηλαδή εάν επέλεξε να επισκεφτεί τη σελίδα του αντικειμένου, και πόσο χρόνο αφιέρωσε στην επισκόπηση του συγκεκριμένου αντικειμένου, εάν το επέλεξε κ.ά. [1].

### 3.2 Το Πρόβλημα

Όσον αφορά τον τυπικό ορισμό του προβλήματος που φιλοδοξεί να λύσει ένα σύστημα συστάσεων, το πρόβλημα των συστάσεων μπορεί να οριστεί ως πρόβλημα εκτίμησης του βαθμού προτίμησης των χρηστών για αντικείμενα που δεν έχουν ακόμα ανακαλύψει. Η εκτίμηση αυτή βασίζεται σε α) βαθμολογίες (άμεσες ή έμμεσες) που έχουν δοθεί από τον χρήστη σε άλλα αντικείμενα, β) βαθμολογίες που έχουν δοθεί από άλλους χρήστες στο συγκεκριμένο αντικείμενο και γ) άλλες πληροφορίες για τον χρήστη ή το αντικείμενο, όπως συγκεκριμένα χαρακτηριστικά του αντικειμένου, δημογραφικά χαρακτηριστικά για το χρήστη κ.ά. [7].

Έτσι το πρόβλημα των συστάσεων μπορεί να οριστεί σύμφωνα με το [8] ως εξής:

Έστω  $U$  το σύνολο των χρηστών  $U=\{u_1, u_2, \dots, u_m\}$ , και έστω  $I$  το σύνολο όλων των αντικειμένων που μπορούν να προταθούν  $I=\{i_1, i_2, \dots, i_n\}$  π.χ. μουσικά κομμάτια, εικόνες, ταινίες κλπ.

Έστω  $f$  μια συνάρτηση χρησιμότητας που μετράει τη χρησιμότητα του αντικειμένου  $i$  στο χρήστη  $u$ .

$$f: U \times I \rightarrow R$$

Όπου το  $R$  είναι ένα ταξινομημένο σύνολο. Τότε για κάθε χρήστη  $u \in U$  πρέπει να επιλεγεί ένα αντικείμενο  $i' \in I$  το οποίο μεγιστοποιεί τη συνάρτηση χρησιμότητας  $f$ , όπως φαίνεται στον παρακάτω τύπο:

$$\forall u \in U, i'_u = \arg \max f(u, i), i \in I$$

Στα συστήματα συστάσεων, η χρησιμότητα ενός αντικειμένου αναπαρίσταται συνήθως ως μια βαθμολογία, που δείχνει σε τι βαθμό ένα αντικείμενο άρεσε σε έναν συγκεκριμένο χρήστη. Για παράδειγμα ο χρήστης  $u$ , έδωσε σε ένα αντικείμενο  $i$  τον βαθμό  $R(u,i) = 3$  όπου  $R(u,i) \in \{1,2,3,4,5\}$ . Κάθε χρήστης  $u$ , έχει μια λίστα αντικειμένων  $L_u$ , τα οποία έχει βαθμολογήσει άμεσα ή έμμεσα. Φυσικά οι χρήστες δεν είναι απαραίτητο να έχουν βαθμολογήσει όλα τα αντικείμενα που μπορούν να προταθούν [7].

Αντίθετα, ο ρόλος του συστήματος συστάσεων είναι να προβλέψει τις βαθμολογίες που θα έδιναν οι χρήστες στα αντικείμενα που δεν έχουν βαθμολογήσει.



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Γενικότερα, η συνάρτηση χρησιμότητας μπορεί να είναι ακόμη και μια αυθαίρετη συνάρτηση ή μια συνάρτηση κέρδους, που ανάλογα με την εφαρμογή θα μπορούσε να εξαρτάται τόσο από τις βαθμολογίες που έχουν δώσει οι χρήστες στα αντικείμενα, όσο και από τα χαρακτηριστικά των αντικειμένων ή των χρηστών. Κάθε χρήστης και κάθε αντικείμενο μπορεί να μοντελοποιηθεί ως ένα προφίλ χαρακτηριστικών. Τα χαρακτηριστικά του χρήστη, μπορεί να περιλαμβάνουν το εισόδημα, την ηλικία, το φύλλο ή την οικογενειακή κατάσταση, ενώ τα χαρακτηριστικά του αντικειμένου μπορεί να περιλαμβάνουν πληροφορίες σχετικές με το είδος του. Για παράδειγμα, σε ένα σύστημα συστάσεων ταινιών, το προφίλ χαρακτηριστικών του αντικειμένου θα μπορούσε να περιλαμβάνει το είδος της ταινίας, τη διάρκεια, τον σκηνοθέτη ή την χρονολογία.

### 3.3 Ενδιαφερόμενα Μέρη και Χρησιμότητα Συστημάτων Συστάσεων

Πέρα όμως από τον ορισμό του προβλήματος και τα διάφορα παραδείγματα που δόθηκαν στις προηγούμενες ενότητες, θα είχε ενδιαφέρον να εξετάσουμε τους πιθανούς ρόλους που μπορεί να διαδραματίσει ένα σύστημα συστάσεων. Πρωτίστως θα πρέπει να διαχωρίσουμε το ρόλο του παρόχου αντικειμένων και υπηρεσιών με το ρόλο του χρήστη του συστήματος, ο οποίος λαμβάνει συστάσεις και προφανώς αντλεί σημαντική αξία από το σύστημα. Για παράδειγμα, ένα σύστημα συστάσεων που προτείνει ταξίδια αποτελείται από μια ενδιάμεση πλατφόρμα (π.χ. Expedia<sup>3</sup>), ή έναν οργανισμό διαχείρισης προορισμών (π.χ. visitfinland<sup>4</sup>), οι οποίοι έχουν ως στόχο την αύξηση του κέρδους, την πώληση περισσότερων πακέτων διακοπών, την κράτηση περισσότερων δωματίων ή την άνοδο του αριθμού των τουριστών σε έναν προορισμό [1].

Στην πραγματικότητα υπάρχουν πολλοί λόγοι για να χρησιμοποιήσουν τη συγκεκριμένη τεχνολογία και οι πάροχοι υπηρεσιών. Μερικοί από αυτούς, όπως αναφέρονται και στο [1] παρουσιάζονται παρακάτω:

- **Αύξηση των πωλήσεων:** Είναι πιθανότατα ο σημαντικότερος λόγος στη χρήση των εμπορικών εφαρμογών συστημάτων συστάσεων. Για παράδειγμα, η πώληση κάποιων συνοδευτικών αντικειμένων σε σύγκριση με αυτά που θα είχαν πωληθεί χωρίς τη χρήση κάποιου συστήματος συστάσεων μπορεί να επιτευχθεί με τη χρήση των τελευταίων. Αυτό συμβαίνει διότι τα προτεινόμενα αντικείμενα είναι πιθανό να ταιριάζουν στις προτιμήσεις και τις ανάγκες του χρήστη, μιας και το σύστημα γνωρίζει αρκετά για τα γούστα του. Φυσικά, και στις μη εμπορικές εφαρμογές ο στόχος είναι παρόμοιος, παρότι ο χρήστης δεν έχει κάποιο κόστος από τις επιλογές που θα κάνει. Για παράδειγμα, μια εφαρμογή ειδήσεων και περιεχομένου στοχεύει στην αύξηση των αναγνώσεων των ειδήσεων της. Γενικά, μπορούμε να πούμε ότι από την οπτική του παρόχου υπηρεσιών, ο στόχος είναι η αύξηση της κίνησης, παραδείγματος χάρη η αύξηση των χρηστών που δέχονται μια πρόταση και καταναλώνουν το αντικείμενο σε σχέση με τους απλούς χρήστες που απλά θα έψαχναν αυτό που θέλουν χωρίς τη χρήση κάποιου συστήματος συστάσεων [1].
- **Πώληση πιο ιδιαίτερων προϊόντων:** Ένας ακόμα βασικός λόγος ύπαρξης των συστημάτων συστάσεων είναι η καθοδήγηση του χρήστη στην επιλογή αντικειμένων που θα ήταν δύσκολο να βρει χωρίς μια ακριβής και στοχευμένη σύσταση. Για παράδειγμα, στο Netflix (ή σε οποιοδήποτε

---

<sup>3</sup> <https://www.expedia.com/>

<sup>4</sup> <https://www.visitfinland.com/>

σύστημα συστάσεων ταινιών) ο πάροχος της υπηρεσίας ενδιαφέρεται για την προώθηση όλων των αντικειμένων μιας συλλογής και όχι μόνο των πιο δημοφιλών. Πράγμα που φαντάζει πολύ δύσκολο χωρίς τη χρήση κάποιου συστήματος συστάσεων, μιας και είναι ασύμφορο ο πάροχος της υπηρεσίας να διαφημίσει ταινίες που δεν είναι πιθανό να αρέσουν σε κάποιον χρήστη. Συνεπώς το σύστημα προτείνει μη δημοφιλείς ταινίες στους χρήστες που θα τις εκτιμήσουν [1].

- **Αύξηση της ικανοποίησης του χρήστη:** Ένα σωστά σχεδιασμένο σύστημα συστάσεων μπορεί να βελτιώσει σημαντικά την εμπειρία του χρήστη με την ιστοσελίδα ή την εφαρμογή. Ο χρήστης θα βρει τις προτάσεις ενδιαφέρουσες, σχετικές και με σωστά σχεδιασμένη διεπαφή χρήστη και κατά συνέπεια θα ικανοποιείται από τη χρήση του συστήματος. Ο συνδυασμός αποτελεσματικών (π.χ. ακριβέστερων) συστάσεων, με μια εύχρηστη διεπαφή χρήστη μπορεί να επηρεάσει την υποκειμενική κρίση του χρήστη και κατά συνέπεια ο χρήστης να αξιολογεί καλύτερα το σύστημα. Αυτό με τη σειρά του μπορεί να αυξήσει και την πιθανότητα μία πρόταση αντικειμένου να γίνει αποδεκτή [1].
- **Αύξηση της εμπιστοσύνης του χρήστη:** Ένας χρήστης θα μείνει πιστός σε έναν ιστότοπο που όταν τον επισκέπτεται τον αναγνωρίζει ως παλιό χρήστη και του παρέχει προσωποποιημένη εμπειρία. Αυτό είναι έμφυτο χαρακτηριστικό των συστημάτων συστάσεων μιας και οι συστάσεις προκύπτουν από την αξιοποίηση της πληροφορίας που εξάγεται από τις προηγούμενες αλληλεπιδράσεις του χρήστη με τη πλατφόρμα. Επομένως, όσο περισσότερο ο χρήστης αλληλεπιδρά με τον ιστότοπο, τόσο πιο προσωποποιημένες και ικανοποιητικές θα είναι οι προτάσεις που θα λαμβάνει [1].
- **Καλύτερη αντίληψη των επιθυμιών του χρήστη:** Ένα ακόμη σημαντικό χαρακτηριστικό των συστημάτων συστάσεων που μπορεί να αξιοποιηθεί και από πολλές άλλες εφαρμογές είναι η περιγραφή των προτιμήσεων του χρήστη η οποία γίνεται από το σύστημα. Ο πάροχος της υπηρεσίας μπορεί στη συνέχεια να αποφασίσει να επαναχρησιμοποιήσει τη γνώση αυτή και για άλλους λόγους, όπως την βελτίωση της διαχείρισης αποθήκης ή της διαχείρισης παραγωγής για συγκεκριμένα αντικείμενα [1].

### 3.4 Τεχνικές Συστάσεων

Μέχρι τώρα έχει αναφερθεί ότι τα συστήματα συστάσεων βασίζονται στις βαθμολογίες των χρηστών για τα αντικείμενα και συγκεκριμένα έναν πίνακα βαθμολογιών με  $n$  σειρές και  $m$  στήλες, όπου  $n$  είναι ο συνολικός αριθμός των χρηστών και  $m$  ο συνολικός αριθμός των διαθέσιμων αντικειμένων. Σε έναν τέτοιο πίνακα το κελί  $i, j$  δείχνει τη βαθμολογία του χρήστη  $i$  για το αντικείμενο  $j$ . Είναι λογικό οι χρήστες να έχουν βαθμολογήσει περιορισμένα αντικείμενα από το σύνολο των αντικειμένων και κατά συνέπεια, ο συγκεκριμένος πίνακας είναι συνήθως αραιός. Έτσι, ο σκοπός του συστήματος συστάσεων είναι να προβλέψει τις βαθμολογίες που θα έδιναν οι χρήστες για αντικείμενα που δεν έχουν βαθμολογήσει [9].

Υπάρχουν πολλές μέθοδοι για την πρόβλεψη των βαθμολογιών που θα έδιναν οι χρήστες στα αντικείμενα που δεν έχουν βαθμολογήσει, και εν τέλει πολλές διαφορετικές τεχνικές συστάσεων. Για το λόγο αυτό, σε αυτό το κεφάλαιο θα παρουσιαστούν οι πιο διαδεδομένες, όπως παρουσιάζονται στο βιβλίο [1].

### **3.4.1 Σύσταση βασισμένη στο περιεχόμενο (Content based)**

Στην τεχνική συστάσεων βασισμένων στο περιεχόμενο, το σύστημα προτείνει στο χρήστη αντικείμενα τα οποία έχουν κοινά χαρακτηριστικά με άλλα αντικείμενα τα οποία έχει επιλέξει ο χρήστης στο παρελθόν και του άφησαν θετική εντύπωση [1]. Για τη συγκεκριμένη τεχνική το σύστημα πρέπει να έχει διαθέσιμες πληροφορίες για τα περιγραφικά χαρακτηριστικά των αντικειμένων. Επιπλέον, πρέπει να είναι διαθέσιμες και οι προηγούμενες προτιμήσεις του χρήστη για άλλα αντικείμενα. Σε περίπτωση που αυτές δεν είναι διαθέσιμες, το σύστημα μπορεί να κάνει απευθείας ερωτήσεις στον χρήστη για να παράξει έγκυρες και επιθυμητές προτάσεις [9]. Για παράδειγμα, ένα σύστημα συστάσεων ταινιών μπορεί να προτείνει σε έναν χρήστη ταινίες ενός συγκεκριμένου σκηνοθέτη σε περίπτωση που ο χρήστης έχει δείξει την προτίμηση του στον συγκεκριμένο σκηνοθέτη.

### **3.4.2 Συνεργατικό φιλτράρισμα (Collaborative filtering)**

Πρόκειται για την πιο διαδεδομένη τεχνική συστάσεων και η λογική που ακολουθεί είναι η σύσταση αντικειμένων σε έναν χρήστη εφόσον τα αντικείμενα αυτά έχουν προτιμηθεί από παρόμοιους χρήστες με παρόμοιες προτιμήσεις. Για την τεχνική αυτή δεν απαιτούνται τα χαρακτηριστικά των χρηστών και ούτε αυτά των αντικειμένων. Αντίθετα, απαιτείται ο πίνακας βαθμολογιών των χρηστών για τα αντικείμενα. Μέσα από αυτό τον πίνακα αποφασίζονται οι χρήστες με κοινές προτιμήσεις βάσει της ομοιότητας στις βαθμολογίες που έχουν δώσει σε παρόμοια αντικείμενα. Στη συνέχεια, τα αντικείμενα που προτείνονται στο χρήστη είναι αυτά που συνάδουν με τις προτιμήσεις των παρόμοιων χρηστών με αυτόν. Ένα μειονέκτημα αυτής της τεχνικής είναι ότι βασίζεται στην παραδοχή ότι οι προτιμήσεις των χρηστών δεν αλλάζουν με την πάροδο του χρόνου [1].

### **3.4.3 Δημογραφικά συστήματα (Demographic)**

Στα συστήματα αυτά οι συστάσεις παράγονται ανάλογα με τα δημογραφικά χαρακτηριστικά των χρηστών [1]. Για παράδειγμα, σε ένα τέτοιο σύστημα οι συστάσεις μπορεί να διαφοροποιούνται ανάλογα με την ηλικία, το φύλλο ή την εθνικότητα του χρήστη.

### **3.4.4 Σύσταση βασισμένη στη γνώση (Knowledge Based)**

Τα συστήματα αυτά δίνουν μεγάλη βαρύτητα στις ανάγκες του χρήστη, όσο και στα χαρακτηριστικά των αντικειμένων. Είναι πολύ χρήσιμα σε χρήστες που αγοράζουν μια φορά (one time buyers), όπου το σύστημα δε διαθέτει πληροφορίες για το χρήστη. Ένα παράδειγμα είναι η αγορά τεχνολογικού εξοπλισμού, όπου ο χρήστης συνήθως αγοράζει μια φορά κάθε λίγα χρόνια. Οι πληροφορίες για το χρήστη λαμβάνονται μέσω των επιλογών που κάνει εκείνη τη στιγμή στην πλατφόρμα, πολλές φορές μέσω ερωτήσεων που κάνει το σύστημα για να καταλάβει καλύτερα τις προτιμήσεις του για τα αντικείμενα. Αυτή η διαδικασία προϋποθέτει αναλυτική γνώση πάνω στα αντικείμενα που προτείνονται, ώστε να δίνονται έγκυρες συστάσεις και η εμπειρία του χρήστη να είναι όσο το δυνατόν πιο ομαλή, μιας και είναι απαραίτητη [1], [9].

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

### **3.4.5 Σύσταση βασισμένη στην κοινότητα (Community based)**

Η συγκεκριμένη τεχνική γνώρισε άνθιση με την άνοδο των κοινωνικών δικτύων και βασίζεται στη λογική ότι ένας χρήστης έχει προτιμήσεις παρόμοιες με αυτές των φίλων του. Σε αυτό το πλαίσιο, το σύστημα συλλέγει πληροφορίες για τις κοινωνικές σχέσεις μεταξύ των χρηστών και τα αντικείμενα που τους αρέσουν ώστε να προτείνει στο χρήστη αντικείμενα τα οποία ανήκουν στις προτιμήσεις των φίλων του [1].

### **3.4.6 Υβριδική σύσταση (Hybrid)**

Βασίζεται στον συνδυασμό τεχνικών από αυτές που παρουσιάστηκαν και βοηθάει να καλυφθούν μειονεκτήματα που προκύπτουν από μια τεχνική [1].

## **3.5 Αξιολόγηση Συστημάτων Συστάσεων**

Τα συστήματα συστάσεων στοχεύουν στην παροχή προσωποποιημένων συστάσεων αντικειμένων στους χρήστες. Ο τρόπος που το πετυχαίνουν αυτό, παραδοσιακά, είναι η πρόβλεψη της βαθμολογίας που θα έδιναν οι χρήστες σε όλα τα αντικείμενα και η σύσταση των αντικειμένων με την υψηλότερη προβλεπόμενη βαθμολογία. Κατά το πρόσφατο παρελθόν, το βασικό αντικείμενο της έρευνας ήταν πώς θα βελτιωθεί η ακρίβεια της πρόβλεψης στις προβλεπόμενες βαθμολογίες για τους χρήστες. Ωστόσο, εύλογα τίθεται το ερώτημα αν είναι επαρκής η αξιολόγηση αυτού του τύπου για ένα σύστημα συστάσεων. Μέχρι πρότινος, η αξιολόγηση γινόταν αποκλειστικά μέσω της ακρίβειας των προβλεπόμενων βαθμολογιών. Ωστόσο, αρκετές έρευνες έδειξαν ότι η ακρίβεια των συστάσεων δεν είναι η μοναδική μετρική για την αξιολόγηση της ποιότητας των συστάσεων. Αντίθετα υπάρχουν και άλλες μετρικές εξίσου σημαντικές.

Ο λόγος για τον οποίο η ακρίβεια δεν είναι πάντα επαρκής μετρική για την αξιολόγηση ενός συστήματος συστάσεων περιγράφεται στην εργασία [10]. Στη συγκεκριμένη έρευνα, καταδεικνύεται ότι οι πιο ακριβείς συστάσεις μπορεί να μην είναι και οι πιο χρήσιμες, και ότι οι μετρικές αξιολόγησης θα πρέπει να λαμβάνουν υπόψη και άλλους παράγοντες που επηρεάζουν την ποιότητα των συστάσεων, όπως η έκπληξη που προκαλεί μια σύσταση στο χρήστη. Επιπλέον, τονίστηκε ότι θα πρέπει να αξιολογούνται ολόκληρες λίστες συστάσεων και όχι τα αντικείμενα ξεχωριστά. Επιπροσθέτως, σε άλλες εργασίες τονίζεται η σημασία των διαφοροποιημένων λιστών συστάσεων [11], όπως επίσης και η σημασία της ύπαρξης καινοτόμων (novel) αντικειμένων στις λίστες συστάσεων [12]. Αυτά τα χαρακτηριστικά έχουν τεράστια σημασία σε συστήματα συστάσεων της πραγματικής ζωής, μιας και οι χρήστες το πιο πιθανό είναι να ασχοληθούν με ένα μικρό σύνολο από προτεινόμενα αντικείμενα. Συνεπώς, το σύνολο αυτό θα πρέπει να είναι όσο το δυνατόν πιο ενδιαφέρον και συμπεριληπτικό για αυτούς.

Στο παρόν υποκεφάλαιο, αναλύεται η αξιολόγηση των συστημάτων συστάσεων, δίνοντας έμφαση στην ποιότητα των προτεινόμενων προϊόντων από διαφορετικές οπτικές γωνίες όπως αποτυπώνονται μέσα από τις διάφορες μετρικές. Πιο συγκεκριμένα, οι μετρικές που αναλύονται είναι η κάλυψη (coverage), η διαφοροποίηση (diversity), η δικαιοσύνη (fairness), όπως επίσης και οι μετρικές για την ευχάριστη έκπληξη στο χρήστη (serendipity), η καινοτομία (novelty) και το απρόβλεπτο (unexpectedness), οι οποίες παρουσιάζονται μαζί μιας και έχουν πολλά κοινά μεταξύ τους. Επιπλέον, εξετάζεται η επίδραση αυτών

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

των μετρικών στην ακρίβεια των προτεινόμενων συστάσεων όπως επίσης και οι μεταξύ τους σχέσεις. Στη συνέχεια, παρουσιάζονται τα πολυμετοχικά συστήματα συστάσεων ως νέο υπόδειγμα για την αξιολόγηση των συστημάτων συστάσεων.

### 3.5.1 Κάλυψη (Coverage)

Ως κάλυψη (coverage) ορίζεται ο βαθμός στον οποίο τα προτεινόμενα αντικείμενα καλύπτουν το σύνολο των διαθέσιμων αντικειμένων (item coverage) ή ο βαθμός που οι συστάσεις καλύπτουν όλους τους διαθέσιμους χρήστες (user coverage) [13]. Ένα σύστημα συστάσεων με υψηλή κάλυψη παρέχει στο χρήστη μια πιο λεπτομερή και προσεκτική αναπαράσταση όλων των αντικειμένων που θα μπορούσαν να προταθούν. Έτσι, υψηλή κάλυψη οδηγεί σε μεγαλύτερη δυνατότητα διερεύνησης όλων των αντικειμένων και ανακάλυψης νέων επιθυμητών αντικειμένων.

Πιο συγκεκριμένα, η κάλυψη είναι ένα μέτρο της αναλογίας των αντικειμένων που συστήνονται σε χρήστες προς το συνολικό αριθμό αντικειμένων που θα μπορούσαν να προταθούν. Στη βιβλιογραφία ορίζεται είτε ως το ποσοστό των αντικειμένων για τα οποία το σύστημα δημιουργεί συστάσεις, είτε ως το ποσοστό των αντικειμένων που συστήνονται σε ένα χρήστη. Έχει μελετηθεί σε πολλές έρευνες και διακρίνεται σε κάλυψη πρόβλεψης (prediction coverage) και κάλυψη καταλόγου (catalog coverage).

Η κάλυψη πρόβλεψης εξαρτάται από την υλοποίηση του συστήματος συστάσεων και την είσοδο του. Για παράδειγμα, σε ένα σύστημα συστάσεων Collaborative Filtering, οι είσοδοι είναι βαθμολογίες αντικειμένων ενώ για content based συστήματα συστάσεων, η είσοδοι είναι κάποιοι κανόνες και οι προτιμήσεις του χρήστη. Και στις δυο περιπτώσεις το σύστημα μπορεί να δημιουργήσει συστάσεις μόνο για αντικείμενα για τα οποία έχει λάβει επαρκή πληροφορία. Οπότε, έστω  $I$  το σύνολο των διαθέσιμων αντικειμένων και  $I_p$  το σύνολο των αντικειμένων τα οποία μπορούν να προταθούν, τότε η κάλυψη πρόβλεψης δίνεται από τον τύπο:

$$Prediction\ Coverage = \frac{|I_p|}{|I|}$$

Από την άλλη πλευρά, η κάλυψη καταλόγου μπορεί να φανεί ιδιαιτέρως χρήσιμη για συστήματα που συστήνουν λίστες από αντικείμενα, μιας και η κάλυψη πρόβλεψης δε μπορεί να χρησιμοποιηθεί για αυτό το σκοπό. Η κάλυψη καταλόγου, συνήθως μετριέται σε ένα σύνολο από προτεινόμενες λίστες. Για παράδειγμα, εξετάζοντας τις προτεινόμενες λίστες κάποιων χρηστών για ένα προκαθορισμένο χρονικό διάστημα. Έστω  $I(j,L)$  το σύνολο των αντικειμένων που περιέχονται στη λίστα  $L$  και επιστρέφονται την  $j$ -οστή φορά. Έστω  $N$  ο αριθμός συστάσεων για το σύνολο των λιστών στο μελετώμενο διάστημα και  $I$  το σύνολο όλων των διαθέσιμων αντικειμένων. Η κάλυψη καταλόγου μετριέται σύμφωνα με τον τύπο:

$$Catalog\ Coverage = \frac{|\cup_{j=1..N} I(j,L)|}{|I|}$$

Μέχρι τώρα έχουμε αναφερθεί μόνο στην κάλυψη αντικειμένων (item coverage) που μετράει το βαθμό που τα προτεινόμενα αντικείμενα καλύπτουν το σύνολο των διαθέσιμων αντικειμένων του συστήματος. Ωστόσο, υπάρχει και η έννοια της κάλυψης χρηστών (user coverage), που μετράει το βαθμό στον οποίο το σύστημα καλύπτει τους χρήστες του (π.χ. ο λόγος των χρηστών για τους οποίους το σύστημα μπορεί

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

να παρέχει συστάσεις) [14]. Ωστόσο, η κάλυψη χρηστών δεν παρουσιάζει πολύ μεγάλο ενδιαφέρον και κατά συνέπεια δε συναντάται τόσο συχνά στη βιβλιογραφία.

Αξίζει να σημειωθεί, ότι σε πολλές περιπτώσεις η μέση κάλυψη για μια ομάδα από αντικείμενα παρουσιάζει μεγαλύτερο ενδιαφέρον από τη μέση κάλυψη για όλα τα αντικείμενα του συστήματος συστάσεων. Για παράδειγμα, σε ένα σύστημα συστάσεων ταινιών εκτός από τη μέση κάλυψη αντικειμένου σε όλο το σύνολο ταινιών, παρουσιάζει πολύ ενδιαφέρον η μέση κάλυψη ανά κατηγορία. Επιπλέον, οι παραγωγοί ταινιών, μπορεί να ενδιαφέρονται περισσότερο για την κάλυψη αντικειμένου των δικών τους ταινιών. Ωστόσο δεν υπάρχουν πολλές δημοσιευμένες εργασίες που να ασχολούνται με τη μέση κάλυψη μιας κατηγορίας αντικειμένων.

Μέχρι τώρα παρουσιάστηκαν διάφοροι ορισμοί για την κάλυψη σε συστήματα συστάσεων, οι οποίοι επιτρέπουν μόνο την αξιολόγηση ενός συστήματος συστάσεων με βάση τα αποτελέσματα του, και όχι τη σχεδίαση ενός συστήματος με την επιθυμητή κάλυψη. Αυτό το πρόβλημα φιλοδοξεί να λύσει η δημοσιευμένη εργασία [15], στην οποία η κάλυψη εντάσσεται στο πρόβλημα των συστάσεων. Στην προκειμένη περίπτωση, αντί να μετριέται η κάλυψη και να αξιολογείται το σύστημα ως προς αυτήν εκ των υστέρων, οι συστάσεις δημιουργούνται για να ικανοποιήσουν συγκεκριμένες απαιτήσεις για αυτή. Πιο συγκεκριμένα, στην εργασία αυτή, το πρόβλημα των συστάσεων αντιμετωπίζεται ως πρόβλημα βελτιστοποίησης και οι απαιτήσεις για συγκεκριμένη κάλυψη μπαίνουν ως περιορισμοί στο μοντέλο και έτσι το πρόβλημα των συστάσεων ανάγεται σε πρόβλημα γραμμικού προγραμματισμού. Φυσικά, απαιτήσεις για μεγάλη κάλυψη, οδηγούν σε συστάσεις με μικρότερη προβλεπόμενη βαθμολογία. Αυτό το γεγονός αξιοποιείται για να δημιουργηθεί η συνάρτηση κόστους. Συγκεκριμένα, το πρόβλημα αρχικά λύνεται χωρίς περιορισμούς με βάση τις καλύτερες προβλεπόμενες βαθμολογίες και στη συνέχεια λύνεται ως πρόβλημα γραμμικού προγραμματισμού που έχει ως στόχο η λύση να αποκλίνει κατά το δυνατόν λιγότερο από την αρχική λύση (πρόβλημα ελαχιστοποίησης) και να ικανοποιούνται οι περιορισμοί για την κάλυψη.

Για να δοθεί η μαθηματική περιγραφή του προβλήματος ως πρόβλημα γραμμικού προγραμματισμού κρίνεται χρήσιμο να παρουσιαστούν και να περιγραφούν μια προς μια οι βασικές σχέσεις και ορισμοί του προβλήματος.

Όπως έχουμε ήδη αναφέρει, η κάλυψη αντικειμένου αναφέρεται στο γεγονός ότι κάθε αντικείμενο εμφανίζεται στη λίστα συστάσεων ενός συγκεκριμένου αριθμού χρηστών. Επιπλέον, δοθέντος ενός θετικού ακέραιου  $d$  ένα αντικείμενο  $i$  ικανοποιεί τον  $d$ -coverage περιορισμό ή είναι  $d$ -covered, αν έχει προταθεί σε τουλάχιστον  $d$  χρήστες. Έτσι, αν ένα σύστημα συστάσεων  $A$  έχει  $d$ -coverage, όλα του τα αντικείμενα είναι  $d$ -covered. Συνεπώς, ο στόχος της συγκεκριμένης προσέγγισης είναι να βρεθεί μια λύση στην οποία όλα τα αντικείμενα θα είναι  $d$ -covered και επομένως νέες προτεινόμενες λίστες αντικειμένων  $\{L(u)\}$ ,  $u \in U$ , με  $u$  ένα χρήστη από το σύνολο των χρηστών  $U$ .

Επίσης, έστω ότι  $l$  είναι η θέση κατάταξης σε μια λίστα και είναι  $l \leq L$  με  $L$  τον αριθμό προτεινόμενων αντικειμένων στις λίστες  $L(u)$ . Επίσης, κάθε θέση  $l$  στη λίστα  $L(u)$  για έναν χρήστη  $u$  συνδέεται με ένα βάρος  $w(u,l)$ , που χρησιμοποιείται για να ποσοτικοποιήσει τη συμπεριφορά του χρήστη  $u$  και κατά πόσο επιλέγει ένα αντικείμενο αν βρίσκεται στη θέση  $l$ . Επιπλέον, για έναν χρήστη  $u$ , έστω  $f(u,l)$ , η βαθμολογία

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

του αντικειμένου που εμφανίζεται στη θέση  $l$  της αρχικής (baseline) λίστας συστάσεων  $L_B(u)$  του αρχικού συστήματος (χωρίς τους περιορισμούς για την κάλυψη).

Αν υποθέσουμε ότι ένα αντικείμενο  $i$  συστήνεται στο χρήστη  $u$  στη θέση  $l$  στη λίστα  $L(u)$  αντικαθιστώντας ένα αντικείμενο που είχε προταθεί σε αυτή τη θέση στο αρχικό σύστημα συστάσεων  $L_B(u)$ , τότε το κόστος ανταλλαγής της αρχικής λίστας είναι:

$$c(l, i, u) = r(i, u) - f(u, l)$$

Έστω  $x(i, u, l)$  μια δυαδική μεταβλητή, όπου  $x(i, u, l) = 1$  αν το αντικείμενο  $i$  προτείνεται στο χρήστη  $u$  στη σειρά  $l$  στη λίστα  $L(u)$  και 0 σε διαφορετική περίπτωση. Σύμφωνα με αυτό τον ορισμό η κάλυψη ενός αντικειμένου δίνεται από τον παρακάτω τύπο:

$$Cov(i, x) = \sum_{u=1}^U \sum_{l=1}^L x(i, u, l)$$

Ο στόχος είναι να κρατήσουμε το κόστος αντικατάστασης όσο το δυνατόν μικρότερο σε κάθε θέση  $l$ . Επιπλέον, η διαφορά στην απόκλιση μεταξύ ενός αντικειμένου της αρχικής λίστας με την αντικατάσταση του θέλουμε να είναι χαμηλότερη στις υψηλότερες θέσεις, μιας και τα αντικείμενα στις υψηλότερες θέσεις λαμβάνουν μεγαλύτερη προσοχή. Έτσι, για έναν χρήστη  $u$  το κόστος αντικατάστασης, όσον αφορά τη διαφορά με τις λίστες του αρχικού συστήματος είναι:

$$UDev\_Cost(u, x) = \sum_{i=1}^I \sum_{l=1}^L w(u, l) * c(i, u, l) * x(i, u, l)$$

Ενώ, το συνολικό κόστος απόκλισης για ένα αντικείμενο είναι:

$$IDev\_Cost(i, x) = \sum_{u=1}^U \sum_{l=1}^L w(u, l) * r(i, u) * (x(i, u) - I(i, u, l))$$

Όπου  $I(i, u, l) = 1$  αν το αντικείμενο  $i$  εμφανίζεται στη λίστα των προτεινόμενων αντικειμένων του χρήστη  $u$  στη θέση  $l$  στο αρχικό σύστημα.

Συνεπώς, το πρόβλημα μπορεί να περιγραφεί μαθηματικά με το ακόλουθο μοντέλο:

$$\min \sum_{u=1}^U UDev\_Cost(u, x)$$

Έτσι ώστε:

$$Cov(i, x) \geq d, \forall i \in I,$$

$$\sum_{i=1}^I x(i, u, l) = 1, \quad \forall u \in U, \forall l = 1, \dots, L$$

$$x(i, u) \in \{0, 1\}, \forall i \in I, \forall u \in U$$

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Ο πρώτος περιορισμός προσδιορίζει την ελάχιστη αποδεκτή κάλυψη αντικειμένου ενώ ο δεύτερος ότι ένα αντικείμενο μπορεί να ανατεθεί σε κάθε θέση για κάθε χρήστη μια φορά. Το πρόβλημα που παρουσιάστηκε είναι ένα πρόβλημα γραμμικού ακέραιου προγραμματισμού, το οποίο ως γνωστόν ανήκει στην κατηγορία προβλημάτων NP-complete. Παρ' όλα αυτά όπως παρουσιάζεται στη σχετική εργασία έχει υπολογιστικά εφικτή λύση ως πρόβλημα γραμμικού προγραμματισμού [15].

Αξίζει να σημειωθεί ότι στην εργασία που μελετούμε [15] παρουσιάζεται επίσης το πρόβλημα με επιπλέον περιορισμό την εξισορρόπηση του κόστους απόκλισης μεταξύ των χρηστών και το ίδιο πρόβλημα με περιορισμό εξισορρόπησης του κόστους απόκλισης μεταξύ των αντικειμένων. Ακόμη, παρουσιάζονται κάποιοι άπληστοι ευριστικοί αλγόριθμοι. Τα προαναφερθέντα προβλήματα και οι ευριστικοί αλγόριθμοι δε θα παρουσιαστούν στην παρούσα εργασία.

Όσον αφορά την κάλυψη χρήστη, μια συστηματική προσέγγιση έχει γίνει στην εργασία [16]. Πιο συγκεκριμένα, η εργασία αυτή, ασχολείται με το πρόβλημα του online learning σε online συστήματα συστάσεων και λαμβάνει υπόψη και τη δυναμική συμπεριφορά των χρηστών. Επιπλέον, στα πλαίσια της συγκεκριμένης εργασίας προτείνεται ένας αλγόριθμος βασισμένος σε γράφους για την επίλυση του προβλήματος. Η νέα μεθοδολογία ονομάζεται Graph based UCB1 bandits και δε θα αναλυθεί περαιτέρω στα πλαίσια της παρούσας διδακτορικής διατριβής.

Πέραν της θεωρητικής προσέγγισης που έχει παρουσιαστεί μέχρι τώρα, στην εργασία [17] παρουσιάζεται και μια πρακτική εφαρμογή, όπου η αύξηση της κάλυψης χρήστη χρησιμοποιείται για την αύξηση των πωλήσεων, σε ένα σύστημα συστάσεων ηλεκτρονικών πωλήσεων (e-commerce). Πιο συγκεκριμένα, ο στόχος είναι η επίλυση του προβλήματος μέγιστης κάλυψης (maximum coverage), το οποίο ανήκει στην κατηγορία NP-complete. Συνεπώς, παρουσιάζεται μια υπολογιστικά εφικτή ευριστική λύση που βασίζεται σε μια άπληστη μεθοδολογία. Το ενδιαφέρον της παρούσας δημοσίευσης είναι ότι η αξιολόγηση του αποτελέσματος έγινε σε πραγματικά δεδομένα και οδήγησε σε αύξηση των πωλήσεων κατά 3-8%, σε σχέση με τις λίστες με τα πιο δημοφιλή (best selling) προϊόντα. Επίσης, αξίζει να σημειωθεί ότι η συγκεκριμένη προσέγγιση οδήγησε και σε μεγαλύτερη διαφοροποίηση στα προτεινόμενα προϊόντα από ό,τι στα προτεινόμενα προϊόντα της αρχικής λύσης.

Μέχρι στιγμής, έχει παρουσιαστεί εκτενώς η μετρική της κάλυψης τόσο σε επίπεδο χρηστών, όσο και αντικειμένων. Ωστόσο, δεν έχει αναλυθεί η χρησιμότητά της και τι προσφέρει σε πραγματικές εφαρμογές. Η απάντηση σε αυτό το ερώτημα σχετικά με τους χρήστες είναι προφανής, μιας και μεγαλύτερος αριθμός χρηστών που παίρνουν συστάσεις από το σύστημα συνεπάγεται μεγαλύτερη πιθανότητα για επιλογή κάποιων αντικειμένων από τα προτεινόμενα από το σύστημα προϊόντα. Αντίθετα, σχετικά με τα αντικείμενα δεν είναι τόσο προφανής η απάντηση. Σύμφωνα με την εργασία [14], μεγαλύτερη κάλυψη αντικειμένου σημαίνει την έκθεση των χρηστών σε μεγαλύτερη ποικιλία αντικειμένων, το οποίο με τη σειρά του μπορεί να οδηγήσει τόσο σε μεγαλύτερη ικανοποίηση των χρηστών, όσο και να ωφελήσει τις εταιρείες πώλησης των συγκεκριμένων προϊόντων. Αυτό οφείλεται στο γεγονός ότι περισσότεροι χρήστες ανακαλύπτουν προϊόντα που ανήκουν στη μακριά ουρά (long tail) της ζήτησης και της προσφοράς [18]. Ναι μεν αυτά τα προϊόντα έχουν χαμηλή ζήτηση και είναι δυσεύρετα, αλλά είναι πολύ μεγάλος ο αριθμός τους και μπορούν να δημιουργήσουν πολύ σημαντικά έσοδα. Αυτό το φαινόμενο το έχει εκμεταλλευτεί σε πολύ μεγάλο βαθμό η Amazon, μεταξύ άλλων, για να αυξήσει τα έσοδα της, όπως φαίνεται και στην εργασία [19].



### 3.5.2 Διαφοροποίηση (Diversity)

Πέραν της κάλυψης, μια άλλη μετρική στόχος στα συστήματα συστάσεων είναι η διαφοροποίηση (diversity). Η έννοια της διαφοροποίησης προέρχεται από το επιστημονικό πεδίο της εξαγωγής πληροφορίας (Information Retrieval). Στο συγκεκριμένο επιστημονικό πεδίο, έχει αναγνωριστεί ότι η αξία του εξαγόμενου εγγράφου επηρεάζεται όχι μόνο από την ομοιότητα του με το ερώτημα (query) που τέθηκε, αλλά και την ομοιότητα με τα άλλα έγγραφα που εξάχθηκαν με αυτό. Έτσι, συνδέεται με την πιθανή ασάφεια σε ένα ερώτημα κάποιου χρήστη, μιας και με την απουσία πληροφορίας για αντιμετώπιση της ασάφειας είναι αδύνατο να γνωρίζουμε για ποια θέματα ενδιαφέρεται ο χρήστης. Για παράδειγμα, ένα ερώτημα που περιέχει τον όρο *jaguar* μπορεί να αναφέρεται στο ζώο ή σε κάποιο αυτοκίνητο της γνωστής φίρμας. Επομένως, εξασφαλίζοντας ότι τα εξαγόμενα έγγραφα καλύπτουν μια ευρεία γκάμα από πιθανά θέματα, αυξάνεται η πιθανότητα να υπάρχουν σχετικές απαντήσεις με το ερώτημα του χρήστη [14].

Όσον αφορά τα συστήματα συστάσεων, έχουν δοθεί πολλοί ορισμοί για τη διαφοροποίηση [14]. Πιο συγκεκριμένα, έχει οριστεί ως ο μέσος όρος της απόστασης κατά ζεύγη (pairwise distance) μεταξύ των προτεινόμενων αντικειμένων [20]. Επιπροσθέτως, έχει οριστεί ως η συνολική απόσταση κατά ζεύγη μεταξύ των προτεινόμενων αντικειμένων [21]. Η αποτίμηση της διαφοροποίησης ως η μέση ή η συνολική απόσταση μεταξύ των προτεινόμενων αντικειμένων έχει χρησιμοποιηθεί κατά κόρον στη βιβλιογραφία. Αυτό που διαφέρει συχνά είναι ο τρόπος υπολογισμού της απόστασης. Για παράδειγμα, όταν τα αντικείμενα αναπαρίστανται με περιγραφητές περιεχομένου (content descriptors), η απόσταση μετριέται με μια μετρική που βασίζεται σε ταξινομίες [21], ενώ όταν αναπαρίστανται με διανύσματα όρων, η απόσταση ορίζεται ως το συμπλήρωμα της ομοιότητας Jaccard (Jaccard similarity) [22], ή της συνημιτονικής ομοιότητας (cosine similarity) [23]. Από την άλλη πλευρά, όταν τα αντικείμενα αναπαρίστανται ως διανύσματα βαθμολογίας, χρησιμοποιείται η απόσταση Hamming, ή το συμπλήρωμα της ομοιότητας Pearson (Pearson similarity) [24] ή της συνημιτονικής ομοιότητας [23]. Μια άλλη προσέγγιση για τον υπολογισμό της απόστασης μεταξύ των αντικειμένων είναι η χρήση των γειτονιών ενός αντικειμένου [25]. Τέλος, η απόσταση μεταξύ αντικειμένων μπορεί να υπολογιστεί με τη μέθοδο της παραγοντοποίησης πινάκων μέσω του διανύσματος λανθάνοντων συντελεστών των χαρακτηριστικών (latent factor feature vector) [22].

Βέβαια, στη βιβλιογραφία υπάρχουν άρθρα που αμφισβητούν την αποτίμηση της διαφοροποίησης ως απόσταση μεταξύ των αντικειμένων [14]. Συγκεκριμένα, αναφέρεται ότι ο συγκεκριμένος τρόπος υπολογισμού μπορεί να αποτύχει, καθώς μια λίστα με αντικείμενα με μεγάλη ανομοιογένεια μεταξύ τους δε συμβαδίζει απαραίτητα με την άποψη των χρηστών για το αν η λίστα είναι όντως διαφοροποιημένη (diversified). Επιπλέον, τονίζεται ότι η διαφοροποίηση πρέπει να εξαρτάται από τον αριθμό των προτεινόμενων αντικειμένων.

Η σημασία της διαφορετικότητας των συστάσεων είναι πολύ μεγάλη παρότι δεν έχει μελετηθεί στο βαθμό που έχει μελετηθεί η ακρίβεια των συστάσεων [8]. Αυτό γιατί στόχος των σύγχρονων συστημάτων συστάσεων είναι να δίνουν προσωποποιημένες συστάσεις, και η διαφορετικότητα των συστάσεων ενισχύει το στοιχείο της προσωποποίησης. Η προσωποποίηση στα συστήματα συστάσεων είναι ζωτικής σημασίας, μιας και οι χρήστες χαρακτηρίζονται από διαφορετικές προτιμήσεις τόσο σε κατηγορίες

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

αντικειμένων, όσο και σε αντικείμενα της ίδιας κατηγορίας [11]. Τα συστήματα συστάσεων που έχουν ως στόχο μόνο τη μέγιστη προβλεπόμενη ωφέλεια και κυρίως τα συστήματα που βασίζονται στο περιεχόμενο (content-based recommender systems), αποτυγχάνουν να αναγνωρίσουν αυτή τη διαφορά στις προτιμήσεις των χρηστών. Επιπλέον, η διαφορετικότητα των προτεινόμενων αντικειμένων μπορεί να βοηθήσει κάποιες εταιρείες, μιας και εκθέτουν στους χρήστες και αντικείμενα που ανήκουν στη μακριά ουρά (long tail) που οδηγεί τους χρήστες σε ανακάλυψη μεγαλύτερης γκάμας αντικειμένων.

Αξίζει να σημειωθεί ότι εκτός από τη διαφοροποίηση στα προτεινόμενα αντικείμενα ενός χρήστη, πολλές φορές είναι σημαντικό να εξεταστεί και η συνολική διαφοροποίηση σε όλους τους χρήστες, μιας και υψηλή ατομική διαφοροποίηση δε σημαίνει απαραίτητα και υψηλή συνολική διαφοροποίηση. Για παράδειγμα, ένα σύστημα συστάσεων που προτείνει τα 10 πιο δημοφιλή αντικείμενα σε κάθε χρήστη, μπορεί να χαρακτηρίζεται από υψηλή ατομική διαφοροποίηση, αλλά η συνολική διαφοροποίηση θα είναι πολύ χαμηλή [8].

Από την άλλη πλευρά, υψηλότερη διαφοροποίηση (είτε ατομική είτε συνολική) επιτυγχάνεται θυσιάζοντας την ακρίβεια των συστάσεων και της αναμενόμενης ωφέλειας. Πιο συγκεκριμένα, είναι εύκολο να αυξηθεί η διαφοροποίηση απλά προτείνοντας αντικείμενα λιγότερο γνωστά. Ωστόσο, η απώλεια στην ακρίβεια των συστάσεων θα είναι πολύ σημαντική. Συνεπώς, κατά το σχεδιασμό ενός συστήματος συστάσεων, θα πρέπει το σύστημα να σχεδιάζεται ώστε να στοχεύει τόσο στην μεγιστοποίηση της ακρίβειας και της μέγιστης αναμενόμενης ωφέλειας των συστάσεων, όσο και στην μεγιστοποίηση της διαφοροποίησης [8], [26], [27].

Στην παρακάτω εικόνα, παρουσιάζονται συνοπτικά οι σημαντικότεροι ορισμοί για τη διαφοροποίηση, μαζί με μια αξιολόγηση, όπως έχουν αναλυθεί στην εργασία [28].

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Definition Equation	Evaluation	Comparison
Diversity is the average dissimilarity between all pairs of items in the result set. $D = \frac{\sum_{i=1}^n \sum_{j=1}^n (1 - \text{Similarity}(c_i, c_j))}{n/2 * (n-1)}$	Measuring diversity, similarity and relative gain of 7 RSs on a 1000 job advert dataset.	While the evaluations show the clear merit of the proposed metric it is highly dependant of the definition of item similarity (available meta-data).
Diversity is represented as the Gini coefficient - a measure of distributional inequality. $G = 1 - 2 \int_0^1 L(u)du$	The measure was evaluated using a simulated environment that simulated user buying (recommended) items. The Gini coefficient of sales with recommendations was then compared to those without.	While this measure provides an interesting contrast to the intra-list diversity it remains highly domain specific; it is questionable whether it could be applied to environments such as film/music/book recommenders.
Diversity is part of the calculation of the nDCG measure - it has a direct impact on the calculated probability value. $G[k] = \sum_{i=1}^m J(d_k, i)(1 - \alpha)^{r_i, k-1}$	Relevant document retrieval using the TREC 2006 dataset.	This measure is one of the first to combine ambiguity, diversity, redundancy and novelty into a single measure. The downside of the method is similar to that of the previous methods: it requires extensive data in order to correctly calculate the value of $f(d_k)$ .
Diversity between two items is the product of the item's relevance, similarity and places in the ranked list. $ILD(i_k   u, R) = C_i^k \sum_{i=1}^n disc(I k)p(ret i, u)dist(i_k, i_i)$	Measuring the diversity of three different RS using the MovieLens 1M dataset and several re-ranking diversifying algorithms. 20 participants answered the questionnaire about the diversity of recommended items.	This measure is fairly domain independent (authors present several different similarity calculations) and it also includes relevancy into the calculation, which is highly important. The presented results are also promising. It should be noted that this article does not offer a new definition of diversity. Instead it offers a way to collect information about how the user sees the diversity of recommended items, which has been ignored by most of the articles so far.
N/A		
Diversity is presented as a nDCG measure (see [42]), with intent replacing the 'nugget' used in the original definition. $IA - nDCG = \sum p(a u)NDCG(u a), G[k] = \sum_a r(i_k, u)(1 - \alpha)^{\sum_{i=1}^k r(i, u; a)}$	Calculation of diversity of diversified and non-diversified recommendation lists created using the MovieLens and last.fm dataset. 250 volunteers evaluated the diversity of presented recommendations using a 7 point Likert scale.	The authors put a lot of effort into expanding the definition proposed in [42] and introduced intent as a novel way of detecting the user's requirements. Their results are promising but, as with many of the articles presented in this work, they lack a user study that would confirm whether the users can actually detect / appreciate the change in diversity. This article offers a good user study that confirms that the definition used in [2] is viable for use in real-life applications.
Same as in [2], with the addition of a clearer definition of similarity between two items. $D = \frac{\sum_{i=1}^n \sum_{j=1}^n (1 - \text{Similarity}(c_i, c_j))}{n/2 * (n-1)}$	Last.fm set.	It should be noted that in this article the authors did not introduce a new diversity measure but rather used an existing one 'in reverse'. They used a change in diversity to detect a change in the user's context.
Same as in [2]. $D = \frac{\sum_{i=1}^n \sum_{j=1}^n (1 - \text{Similarity}(c_i, c_j))}{n/2 * (n-1)}$	MovieLens and Netflix dataset.	The proposed metric is novel and tries to include a lot of 'human behaviour' in its calculation by using probability functions. It looks promising but has the downside of working only with genres. Should the approach expand to cover additional metadata, the metric might perform better than the current favourite - ILD (see [21]).

Εικόνα 1 - Διάφοροι ορισμοί του diversity στη βιβλιογραφία και αξιολόγησή τους [28]

Επιπλέον, στον παρακάτω πίνακα παρουσιάζονται οι βασικοί αλγόριθμοι που έχουν χρησιμοποιηθεί για τους διάφορους ορισμούς για τη διαφοροποίηση μαζί με τα αποτελέσματα που επιτεύχθηκαν και τον αντίκτυπο τους στην ποιότητα των προτεινόμενων αντικειμένων, με τη χρήση τους, όπως παρουσιάστηκαν στην εργασία [28].

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Diversification algorithm	Impact type/measure	Impact size
Re-ranking recommendation list using: item popularity (item ratings), reverse predicted rating value, item average rating, item absolute likeability, item relative likeability	Change in distribution of rec. items in terms of popularity: best sellers/long-tail items; number of long-tail items among the items recommended across all users	With 1% precision loss, percentage of rec. long-tail items increases from 16 to 32, with 5% loss perc. increases to 58
Maximization of parameterised combined objective function, representing a trade-off between diversity and matching quality, using Greedy and Relaxation and Quantization algorithms	Increased diversity; measured by evaluation of precision and recall against the novelty of the recommendations	
Diversity determinant is genre difference among films	User satisfaction and perceived diversity; user evaluation, ratings plus additional feedback	Users noticed high-diversity items, found them interesting; especially when placed in blocks of items
Not directly used; experiment studies how Openness to Experience may affect the diversity of the recommendations given by the participant	Diversity of the recommendations measured by author, genre and themes; binary for author/genre (0/1), 3 level for themes (0/0.3/1)	Participant personality did not affect recommendations diversity
ClusDiv method, items are clustered, rec. list built by selecting items from different clusters; aims to maximize diversity without decreasing accuracy; added tunable parameter to adjust diversity levels on the rec. lists	Increased diversity with no impact on accuracy; diversity measured by calculating z-diversity, accuracy by calculating recall	Comparable diversity increase (to greedy algorithm) with little recall decrease; much lower computational complexity
Not directly used; experiment surveys the diversity, novelty, accuracy, satisfaction and degree of personalization of various rec. algorithms	Increased diversity; objective measures: RMSE for accuracy, mean popularity rank for novelty, ILD for diversity; subjective measures: user survey, comparison of rec. lists produced by different rec. algorithms	Diversity positively impacted user satisfaction and thus choice of recommendation list (rec. algorithm)
SM - probabilistic specification maximizer model	Increased diversity; measured as interlist diversity	Outperforms classic Markov-based models in terms of diversity
Pareto-efficient multi-objective ranking and rec. list build	Increased diversity; diversity measured by distance based model ([43]), accuracy measured by precision and recall	The approach has the ability to balance each of the objectives according to the desired compromise, or to maximize all 3 objectives simultaneously

Εικόνα 2 - Αντίκτυπος του κάθε αλγορίθμου για diversity στην ποιότητα των συστάσεων σύμφωνα με τη βιβλιογραφία [28]

Εκτός από τη διαφοροποίηση ως προς τα προτεινόμενα αντικείμενα για έναν ή περισσότερους χρήστες, μπορεί να οριστεί και η διαφοροποίηση ως προς τους χρήστες στους οποίους προτείνονται ένα ή περισσότερα αντικείμενα. Δηλαδή, ένα αντικείμενο να προτείνεται σε όσο το δυνατόν πιο διαφορετικούς χρήστες μεταξύ τους. Θα μπορούσε να οριστεί με ανάλογο τρόπο με τη διαφοροποίηση ως προς τα προτεινόμενα αντικείμενα δηλαδή ως η μέση ή η συνολική απόσταση μεταξύ των χρηστών στους οποίους προτείνεται ένα αντικείμενο. Ωστόσο, σε αυτό τον ορισμό υπάρχουν κάποιοι περιορισμοί. Συγκεκριμένα, στις περισσότερες περιπτώσεις δεν έχουμε στοιχεία για τους χρήστες, όπως είναι το είδος ή η περιγραφή για τα αντικείμενα, οπότε η απόσταση θα πρέπει να προσδιοριστεί αποκλειστικά από τις βαθμολογίες για τα αντικείμενα που έχουν χρησιμοποιήσει και γενικότερα από τις αντιδράσεις τους που έχουν καταγραφεί από το σύστημα. Φυσικά, η απόσταση με βάση τις βαθμολογίες σε αντικείμενα ή με βάση αντικείμενα από τα οποία έχουν εμπειρία, μπορεί εύκολα να οριστεί, ωστόσο αποτελεί σημαντικό πρόβλημα το γεγονός ότι στα συστήματα συστάσεων συνηθίζεται οι χρήστες να έχουν εμπειρία από πολύ λίγα αντικείμενα σε σχέση με το σύνολο και κάθε ένας από αυτούς να έχει εμπειρία από διαφορετικά αντικείμενα. Συνεπώς, οι περισσότεροι χρήστες τείνουν να εμφανίζονται πολύ διαφορετικοί μεταξύ τους. Μια προσέγγιση η οποία θα μπορούσε να μετριάσει κάπως το πρόβλημα αυτό σε συστήματα συστάσεων collaborative filtering, θα ήταν ο υπολογισμός της απόστασης με βάση τις εκτιμώμενες βαθμολογίες των χρηστών για αντικείμενα που δεν έχουν βαθμολογήσει. Σε αυτή την περίπτωση οδηγούμαστε από έναν πολύ αραιό πίνακα βαθμολογιών σε έναν πυκνό πίνακα. Ωστόσο, το αποτέλεσμα ίσως να μην είναι πολύ ακριβές, ακριβώς γιατί η ομοιότητα βασίζεται σε εκτιμήσεις. Επιπλέον, είναι πολύ δύσκολο να αιτιολογηθεί γιατί δύο χρήστες που δεν έχουν κοινές εμπειρίες σε αντικείμενα μπορεί να εμφανίζονται πάρα πολύ κοντά. Συνεπώς, οι ενδιαφερόμενοι φορείς, για παράδειγμα οι παραγωγοί κάποιων

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

αντικειμένων, οι οποίοι θέλουν να φτάνει το προϊόν τους σε όσο το δυνατόν πιο διαφορετικούς χρήστες, μπορούν να αμφιβάλλουν για την ακρίβεια μέτρησης της διαφορετικότητας μεταξύ των χρηστών.

Στη βιβλιογραφία έχουν παρουσιαστεί διάφορες τεχνικές για την αύξηση της διαφοροποίησης. Οι περισσότερες από αυτές βασίζονται στην αναδιαμόρφωση και επαναπροτεραιοποίηση των προτεινόμενων λιστών που δημιουργούνται από ήδη υπάρχοντα συστήματα συστάσεων, ώστε να αυξηθεί η διαφοροποίηση, διατηρώντας παράλληλα τη συνάφεια με τις προτιμήσεις του χρήστη. Από την άλλη πλευρά, κάποιες προσεγγίσεις βασίζονται στον ορισμό νέων μοντέλων για συστάσεις βασισμένες στη διαφοροποίηση [14].

Οι τεχνικές της πρώτης κατηγορίας παράγουν μια λίστα από προτεινόμενα αντικείμενα  $R$  μεγέθους  $N$  από ένα μεγαλύτερο σύνολο πιθανών συστάσεων  $C$  ( $|C| > N$ ). Τα υποψήφια προτεινόμενα αντικείμενα  $C$  παράγονται από κάποια γνωστή τεχνική συστάσεων (π.χ. collaborative filtering), και ως αποτέλεσμα παρουσιάζουν σημαντική συνάφεια. Στη συνέχεια, η αναδιαμόρφωση των λιστών ακολουθεί μια άπληστη στρατηγική. Συγκεκριμένα, σε κάθε επανάληψη το αντικείμενο από τη λίστα  $C$  που μεγιστοποιεί την αντικειμενική συνάρτηση, προστίθεται στη λίστα  $R$ . Η αντικειμενική συνάρτηση ορίζεται ως ένας συνδυασμός της συνάφειας και της σχετικής διαφοροποίησης λαμβάνοντας υπόψη τα αντικείμενα που βρίσκονται ήδη στη λίστα  $R$  [14].

Ένα μειονέκτημα των προηγούμενων τεχνικών είναι ότι αντιμετωπίζουν τους αλγορίθμους συστάσεων ως μαύρα κουτιά, μιας και επεξεργάζονται και μεταβάλλουν εκ των υστέρων λίστες που έχουν ήδη παραχθεί από γνωστούς αλγορίθμους συστάσεων. Από την άλλη πλευρά, έχουν το πλεονέκτημα ότι μπορούν πολύ εύκολα και με μικρό κόστος να εφαρμοστούν σε υπάρχοντα συστήματα και το επίπεδο διαφοροποίησης που πρέπει να επιτευχθεί μπορεί να ελεγχθεί πολύ εύκολα μέσω των επαναλήψεων του άπληστου αλγορίθμου. Ωστόσο, υπάρχει μια τάση για δημιουργία αλγορίθμων που βελτιστοποιούν τη διαφοροποίηση κατευθείαν, χωρίς να αλλάζουν μια παραχθείσα λίστα συστάσεων. Οι τεχνικές αυτές βασίζονται κυρίως σε παραγοντοποίηση πινάκων [14].

Όσον αφορά τη χρήση της διαφοροποίησης σε σύγχρονα συστήματα συστάσεων, ο Adomavicius στην εργασία [8] παρουσιάζει μια σειρά από μεθόδους που βασίζονται στην αναδιάταξη των προτεινόμενων λιστών, οι οποίες φαίνεται να προσφέρουν σημαντική βελτίωση στη διαφοροποίηση των συστάσεων, ενώ διατηρούν σε σημαντικό βαθμό την ακρίβεια. Παράλληλα, παρουσιάζονται εμπειρικά αποτελέσματα για διαφορετικές τεχνικές σε σύνολα δεδομένων του πραγματικού κόσμου. Πιο συγκεκριμένα, μια τεχνική που εφαρμόζεται είναι η αναδιάταξη των προτεινόμενων αντικειμένων βάσει της δημοτικότητας (popularity) των αντικειμένων. Με αυτή την προσέγγιση, η διαφοροποίηση αυξήθηκε κατά 3.6 φορές. Ωστόσο, η ακρίβεια των συστάσεων έπεσε από το 89% στο 69%. Μιας και η πτώση της ακρίβειας είναι πολύ μεγάλη και πιθανόν μη αποδεκτή, δοκιμάστηκε ένα κατώφλι στις βαθμολογίες (ranking threshold) των προτεινόμενων αντικειμένων. Μέσω του συγκεκριμένου κατωφλίου, μπορεί να ελεγχθεί και η πτώση στην ακρίβεια των συστάσεων. Ένα παράδειγμα, με ranking threshold=4.4 (η μέγιστη βαθμολογία στο dataset που χρησιμοποιήθηκε ήταν 5), οδήγησε σε αύξηση της διαφοροποίησης κατά 83%, ενώ η πτώση της ακρίβειας ήταν μόνο 1.32%.

Επιπροσθέτως, ο Adomavicius στην εργασία [26] καταπιάνεται με την αύξηση της συνολικής διαφοροποίησης ενός συστήματος συστάσεων σε όλους τους χρήστες, ακολουθώντας παρόμοιες

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

μεθόδους με τις προαναφερθείσες, αλλά για το σύνολο των χρηστών. Τα αποτελέσματα που παρουσιάζονται για κάθε μια από τις μεθόδους που χρησιμοποιήθηκαν, παρουσιάζουν σημαντική αύξηση της διαφοροποίησης για μικρές μειώσεις της ακρίβειας, σε διαφορετικά σύνολα δεδομένων (datasets).

Τέλος, και πάλι ο Adomavicius στην εργασία [27] παρουσιάζει μια γραφοθεωρητική προσέγγιση για την επίλυση του προβλήματος της επίτευξης σημαντικής συνολικής διαφοροποίησης για όλους τους χρήστες του συστήματος, διατηρώντας σε υψηλά επίπεδα την ακρίβεια. Η μεθοδολογία που παρουσιάζεται βασίζεται στη μέγιστη ροή (maximum flow). Όσον αφορά τα αποτελέσματα, τα οποία παρουσιάζονται τόσο για διαφορετικά σύνολα δεδομένων, όσο και για διαφορετικό αριθμό συστάσεων ανά χρήστη, φαίνονται πολύ ικανοποιητικά, σε σύγκριση με άλλες προσεγγίσεις αναδιάταξης (reranking).

### 3.5.3 Δικαιοσύνη/Δικαιότητα (Fairness)

Η δικαιοσύνη (fairness) και κυρίως η προκατάληψη συγκεντρώνουν σημαντικό ενδιαφέρον στην έρευνα γύρω από τη μηχανική μάθηση. Μια συνηθισμένη πρακτική είναι η αναγνώριση κάποιων μεταβλητών που αντιπροσωπεύουν τη συμμετοχή σε κάποια μειονότητα και η ανάπτυξη αλγορίθμων που αφαιρούν αυτή την προκατάληψη, όπως αναφέρεται και στις εργασίες [29] και [30].

Ωστόσο, για την εισαγωγή της έννοιας της δικαιοσύνης στα συστήματα συστάσεων, πρέπει να ληφθεί σοβαρά υπόψη η έννοια της προσωποποίησης. Συγκεκριμένα, ότι τα καλύτερα αντικείμενα για έναν χρήστη δεν είναι τα καλύτερα και για κάποιον άλλον. Λόγω του στόχου των προσωποποιημένων συστάσεων, η δικαιοσύνη συχνά παραβλέπεται, παρότι μπορεί να είναι πολύ σημαντική. Για παράδειγμα, σε ένα σύστημα συστάσεων θέσεων εργασίας είναι σημαντικό γυναίκες με όμοια προσόντα με άντρες χρήστες του συστήματος να λαμβάνουν και ανάλογες συστάσεις, ακόμη και αν τα δεδομένα υπαγορεύουν τη σύσταση θέσεων με υψηλότερους μισθούς για τους άντρες [29].

Επιπροσθέτως, είναι σημαντικό να σημειωθεί ότι τα συστήματα συστάσεων έχουν ως στόχο τις συναλλαγές. Συνεπώς, πολλές εφαρμογές συστημάτων συστάσεων απευθύνονται σε πολλά ενδιαφερόμενα μέρη και έτσι μπορούν να δημιουργηθούν θέματα δικαιοσύνης για περισσότερες από μια ομάδες ενδιαφερομένων. Για παράδειγμα, σε ένα σύστημα συστάσεων θέσεων εργασίας, εκτός από το να διασφαλίζεται η δικαιοσύνη από την πλευρά των χρηστών, είναι σημαντικό να διασφαλίζεται και από την πλευρά των εταιρειών που παρέχουν αυτές τις θέσεις εργασίας. Για παράδειγμα, εταιρείες που οι ιδρυτές τους ανήκουν σε μειονότητες να έχουν την ίδια παρουσία με ανάλογες εταιρείες με ιδρυτές που δεν έχουν το συγκεκριμένο χαρακτηριστικό [29].

Λόγω των διαφορετικών απαιτήσεων για δικαιοσύνη από διαφορετικά ενδιαφερόμενα μέρη, θα πρέπει να διαχωριστεί η έννοια της δικαιοσύνης για κάθε ενδιαφερόμενο. Συγκεκριμένα, στο [29] προτείνεται ο διαχωρισμός σε δικαιοσύνη πελάτη (C-fairness ή customer fairness) και δικαιοσύνη παρόχου (P-fairness ή Provider fairness). Το C-fairness έχει νόημα σε περιπτώσεις συστημάτων συστάσεων, όπου δεν υπάρχουν πολλά ενδιαφερόμενα μέρη και απαιτείται να διασφαλιστεί η δικαιοσύνη στην πλευρά του πελάτη. Αντίστοιχα, το P-fairness έχει νόημα σε περιπτώσεις που απαιτείται να μη δημιουργούνται μονοπώλια σε μια αγορά, όπως για παράδειγμα σε ένα σύστημα συστάσεων για δανειοδότηση από διαφορετικούς οργανισμούς, όπου οι οργανισμοί δανειοδότησης πρέπει να προτείνονται με δίκαιο

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

τρόπο. Σε πολλές περιπτώσεις είναι σημαντική η διασφάλιση της δικαιοσύνης τόσο από τη πλευρά του παρόχου όσο και από αυτή του πελάτη, τότε έχουμε απαιτήσεις για CP-fairness. Ένα τέτοιο παράδειγμα είναι αυτό που αναφέρθηκε στην αρχή για το σύστημα συστάσεων θέσεων εργασίας από διάφορες εταιρείες.

Μερικές πρακτικές εφαρμογές των συστημάτων συστάσεων που στοχεύουν στη δικαιοσύνη παρουσιάζονται στο [31] και στο [32]. Συγκεκριμένα στο [31], το αντικείμενο μελέτης είναι οι συστάσεις ομάδας (group recommendations), και ο στόχος είναι η μεγιστοποίηση της ικανοποίησης κάθε μέλους της ομάδας παράλληλα με την ελαχιστοποίηση της άδικης κατανομής των συστάσεων (unfairness) ανάμεσα τους. Το πρόβλημα μοντελοποιείται ως πρόβλημα μαθηματικού προγραμματισμού πολλών στόχων (multi-objective) και αποδεικνύεται ότι ανήκει στην κατηγορία προβλημάτων NP (Non deterministic Polynomial time). Ως εκ τούτου, παρουσιάζεται μια μεθοδολογία από την οπτική της αποτελεσματικότητας Pareto (Pareto Efficiency). Ο συγκεκριμένος αλγόριθμος δοκιμάστηκε σε διάφορα σύνολα δεδομένων (datasets) και τα αποτελέσματα δείχνουν ότι μπορεί να βελτιώσει σημαντικά τη δικαιοσύνη χωρίς μεγάλη απώλεια στην ακρίβεια.

Από την άλλη πλευρά, στην εργασία [32] μελετάται το πρόβλημα της δικαιοσύνης σε συστήματα συστάσεων για πακέτα διακοπών σε ομάδες τουριστών. Και πάλι το πρόβλημα μοντελοποιείται ως πρόβλημα βελτιστοποίησης που αποδεικνύεται ότι είναι NP-hard. Για την επίλυση του συγκεκριμένου προβλήματος, αυτό ανάγεται σε ένα πρόβλημα κάλυψης, για τη λύση του οποίου παρουσιάζονται κάποιοι άπληστοι αλγόριθμοι. Το αποτέλεσμα των αλγορίθμων αυτών, αποτιμάται σε διαφορετικά σύνολα δεδομένων (datasets) και παρουσιάζει σημαντική βελτίωση στη δικαιοσύνη σε σχέση με άλλες προσεγγίσεις που δε τη λαμβάνουν υπόψη, χωρίς μεγάλη απώλεια στην ποιότητα των συστάσεων.

### 3.5.4 Novelty, Serendipity, Unexpectedness

Το serendipity (θα μπορούσε να μεταφραστεί ως «ευχάριστη έκπληξη») δείχνει το βαθμό καινοτομίας που χαρακτηρίζει τα συνιστώμενα αντικείμενα και κατά πόσο οι συγκεκριμένες συστάσεις μπορούν να εκπλήξουν ευχάριστα το χρήστη [13]. Υψηλό serendipity κάνει το σύστημα να φαίνεται πιο ζωντανό και έξυπνο κάνοντας μη προφανείς και περίεργες προτάσεις που προκαλούν θετική έκπληξη στο χρήστη. Ο στόχος της συγκεκριμένης μετρικής είναι η ανακάλυψη των προτιμήσεων των χρηστών για τους οποίους το σύστημα δε γνωρίζει τις προτιμήσεις τους, καθώς επίσης και να παρέχει στο χρήστη μια αίσθηση σαν να βλέπει μια βιτρίνα και να ανακαλύπτει αντικείμενα που του αρέσουν αλλά δεν γνώριζε καν την ύπαρξη τους.

Το serendipity ορίζεται μαθηματικά ως εξής [14]:

$$\text{Serendipity}(R, u) = \frac{|R\{\text{unexpected}\} \cap R\{\text{useful}\}|}{|R|}$$

Όπου  $R$  είναι το σύνολο των προτεινόμενων αντικειμένων για τον χρήστη  $u$ ,  $R\{\text{unexpected}\}$ , ένα υποσύνολο των προτεινόμενων αντικειμένων που δε συμπεριλαμβάνονται στις προσδοκίες του χρήστη και  $R\{\text{useful}\}$ , προτεινόμενα αντικείμενα τα οποία είναι χρήσιμα για το χρήστη.

Όσον αφορά την καινοτομία (novelty), στο επιστημονικό πεδίο της εξαγωγής πληροφορίας, το novelty κατά την αναζήτηση ενός χρήστη ορίζεται ως το ποσοστό των εξαγόμενων εγγράφων, για τα οποία δεν

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

ήταν ενήμερος ο χρήστης. Ένας άλλος ορισμός είναι ότι το novelty ενός εξαγόμενου εγγράφου είναι το αντίθετο του πλεονασμού (redundancy) του εγγράφου αυτού. Και οι δύο ορισμοί συνδέονται με την ιδέα της αποτίμησης κατά πόσο ένα έγγραφο είναι καινούριο, διαφορετικό και ενδιαφέρον [14].

Σε σχέση με τα συστήματα συστάσεων, η βιβλιογραφία επικεντρώνεται κυρίως σε δύο πτυχές του novelty. Συγκεκριμένα, κατά πόσο ένα αντικείμενο είναι άγνωστο στο χρήστη και κατά πόσο το αντικείμενο αυτό είναι διαφορετικό από τα αντικείμενα που έχει δει στο παρελθόν [14].

Σύμφωνα με το [14] το novelty για μία λίστα ενός χρήστη μαθηματικά ορίζεται ως:

$$Novelty(R) = \frac{\sum_{i \in R} -\log_2 p(i)}{|R|}$$

Όπου R είναι η λίστα των προτεινόμενων αντικειμένων για ένα χρήστη, ενώ το  $p(i)$  είναι η δημοφιλία (popularity) του αντικειμένου  $i$ , η οποία ορίζεται ως ο λόγος μεταξύ των χρηστών που έχουν βαθμολογήσει το αντικείμενο  $i$  προς το σύνολο των χρηστών.

Από την άλλη πλευρά το unexpectedness (απρόβλεπτο) είναι άρρηκτα συνδεδεμένο με τις προσδοκίες του χρήστη. Τα προσδοκώμενα αντικείμενα για έναν χρήστη, είναι μια πεπερασμένη συλλογή από αντικείμενα τα οποία ο χρήστης θεωρεί προφανή και κατάλληλα για να ικανοποιήσουν τους στόχους του και προκύπτουν από την αλληλεπίδραση του με το σύστημα συστάσεων. Οι προσδοκίες του χρήστη μπορούν να οριστούν με διάφορους τρόπους, όπως για παράδειγμα μέσα από τις συστάσεις που έχει λάβει στο παρελθόν ή κάποιες τυπικές προτάσεις τις οποίες αναμένει να πάρει από το σύστημα (π.χ. top 10 items). Φυσικά, το σύνολο των προσδοκιών μπορεί να εξελίσσεται στο χρόνο. Επιπλέον, ένα αντικείμενο σχετικά προσδοκώμενο, μπορεί να μη βρίσκεται στη λίστα των προσδοκιών αλλά να μοιάζει πάρα πολύ με τα αντικείμενα της λίστας αυτής. Συνεπώς, το unexpectedness ενός αντικειμένου για ένα χρήστη, θα μπορούσε να οριστεί ως μια συνάρτηση που εξαρτάται από την απόσταση του νέου αντικειμένου από τα αντικείμενα που βρίσκονται στη λίστα προσδοκιών του χρήστη [33].

Όσον αφορά το μαθηματικό ορισμό του unexpectedness, πρώτα πρέπει να οριστεί η απόσταση ενός αντικειμένου από τη λίστα των προσδοκιών του χρήστη. Συγκεκριμένα αυτή η απόσταση ορίζεται ως εξής:

$$\delta(u, i) = d(i; E(u))$$

Όπου  $d(i; E(u))$  είναι η απόσταση μεταξύ του αντικειμένου  $i$  και της λίστας προσδοκώμενων αντικειμένων του χρήστη  $u$ . Έτσι το unexpectedness ορίζεται ως συνάρτηση της προαναφερθείσας απόστασης ως εξής:

$$\Delta(\delta(u, i); \delta^*(u))$$

Όπου  $\delta^*(u)$  είναι η προτιμώμενη απόσταση από τη λίστα προσδοκώμενων αντικειμένων  $E(u)$ .

Είναι προφανές ότι οι τρεις μετρικές παρουσιάζουν σημαντική ομοιότητα. Ωστόσο, έχουν και κάποιες διαφορές μεταξύ τους. Συγκεκριμένα, σύμφωνα με το [33] αν και μια πρόταση με μεγάλο novelty μπορεί να έχει υψηλό unexpectedness και serendipity, συνδέεται αυστηρά με την έλλειψη πρότερης εμπειρίας και γνώσης για το προτεινόμενο αντικείμενο, ενώ ένα αντικείμενο για το οποίο υπάρχει κάποια γνώση μπορεί να είναι unexpected και serendipitous αλλά δεν μπορεί να θεωρηθεί novel. Για παράδειγμα, σε ένα σύστημα συστάσεων ταινιών, ένας χρήστης που ενδιαφέρεται για ταινίες περιπέτειας και δράσης



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

μπορεί να λάβει μια πρόταση για μια ταινία δράσης του αγαπημένου του σκηνοθέτη. Αυτή η πρόταση μπορεί να θεωρηθεί novel αλλά σίγουρα δε μπορεί να θεωρηθεί unexpected ή serendipitous. Αντίθετα, μια πρόταση για μια οικογενειακή ταινία για την οποία ο χρήστης γνωρίζει αλλά δεν είναι στα γούστα του, μπορεί να θεωρηθεί unexpected και serendipitous αλλά όχι novel.

Από την άλλη πλευρά το serendipity φαίνεται να έχει μεγαλύτερη ομοιότητα με το unexpectedness. Ωστόσο, παρότι και οι δύο μετρικές υπονοούν μια θετική έκπληξη, το serendipity συνδέεται με νέα αντικείμενα και την ανακάλυψη τους κατά τύχη, χωρίς να λαμβάνει υπόψη τις προσδοκίες του χρήστη και τη συνάφεια των αντικειμένων και έτσι μια τέτοια σύσταση είναι διαφορετική και συνδέεται με μεγαλύτερο ρίσκο και ασάφεια. Η διαφορά μπορεί να φανεί καλύτερα με ένα παράδειγμα. Ας υποθέσουμε ότι στον προηγούμενο χρήστη προτείνεται μια νέα ρομαντική κωμωδία. Υπάρχουν κάποιες πιθανότητες να του αρέσει η συγκεκριμένη πρόταση κατά τύχη (serendipitous recommendation), αλλά υπάρχει και σημαντική πιθανότητα να μην του αρέσει μιας και δε λαμβάνει υπόψη τις προσδοκίες και τα γούστα του. Από την άλλη πλευρά, μια πρόταση μιας παλιάς ταινίας όπου ο αγαπημένος σκηνοθέτης του χρήστη παίζει ως ηθοποιός σε μια παλιά ταινία δράσης ενός άλλου σκηνοθέτη, μπορεί να αποτελεί μια unexpected σύσταση αλλά όχι serendipitous [33].

Εφόσον, η πληροφορία της λίστας προσδοκιών όσο και της προτιμώμενης απόστασης από αυτή είναι συνήθως άγνωστα για κάθε χρήστη και λόγω του ότι το novelty ενέχεται μερικώς στην έννοια του serendipity, στο εξής η συγκεκριμένη εργασία θα εστιάσει κυρίως στο serendipity.

Μια σειρά από έρευνες έχουν γίνει για το πώς μπορεί να αυξηθεί το serendipity στα συστήματα συστάσεων. Ωστόσο, δεν υπάρχει κάποια κοινά αποδεκτή μεθοδολογία. Για παράδειγμα, ο Iaquinta, που ήταν από τους πρώτους που ασχολήθηκαν με το συγκεκριμένο θέμα, στην εργασία [34] περιέγραψε ένα σύστημα συστάσεων βασισμένο στο περιεχόμενο (content based recommender system), όπου τα αντικείμενα περιγράφονται από περιγραφές σε μορφή κειμένου. Μια μέθοδος επιβλεπόμενης μάθησης χρησιμοποιήθηκε για να προβλέψει την πιθανότητα ένα αντικείμενο που δεν έχει δει ο χρήστης να είναι συναφές με το περιεχόμενο που αναζητεί. Αντικείμενα τα οποία δεν ήταν ξεκάθαρα αν είναι συναφή με τα ενδιαφέροντα του χρήστη θεωρούνταν serendipitous.

Επιπλέον, ο Onuma [35] σχεδίασε ένα σύστημα που χρησιμοποιεί μια γραφοθεωρητική προσέγγιση για να προτείνει αντικείμενα που προκαλούν έκπληξη. Συγκεκριμένα, η εργασία αυτή εισήγαγε την έννοια του υπολογισμού μιας βαθμολογίας γεφύρωσης μεταξύ αντικειμένων στον διμερή γράφο χρηστών, αντικειμένων. Κόμβοι αντικειμένων που ενώνουν διαφορετικές διασυνδεδεμένες περιοχές λαμβάνουν μεγάλη βαθμολογία γεφύρωσης. Τέτοια αντικείμενα με μικρή συνάφεια με το χρήστη θεωρούνται serendipitous. Μια αντίστοιχη γραφοθεωρητική προσέγγιση ακολουθήθηκε και από τον Nakatsuji [36], ο οποίος χρησιμοποίησε έναν αλγόριθμο random walk σε ένα γράφο ομοιότητας χρηστών για να εντοπίσει χρήστες που σχετίζονται αλλά δεν είναι πάρα πολύ ίδιοι με τον χρήστη που εξετάζεται και υποστηρίζει ότι τέτοιοι χρήστες είναι μια καλή πηγή για serendipitous συστάσεις.

Μια άλλη έρευνα [37] ασχολήθηκε με το serendipity σε ένα σύστημα συστάσεων μουσικής. Χρησιμοποιήθηκε η μέθοδος των γενετικών LDA μοντέλων, για τη δημιουργία συστάδων από χρήστες για την αναπαράσταση των καλλιτεχνών ως μια κατανομή σε αυτές τις συστάδες. Με την αναπαράσταση των καλλιτεχνών ως διανύσματα LDA παρέχεται η δυνατότητα υπολογισμού της ομοιότητας μεταξύ ενός

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

καλλιτέχνη και των καλλιτεχνών στη λίστα ενός χρήστη. Το σύστημα θεωρεί serendipitous τα αντικείμενα που είναι εκτός των συστάδων που ανήκει ένας χρήστης.

### 3.5.5 Συσχετίσεις μεταξύ των διαφορετικών μετρικών

Φυσικά, όλες οι μετρικές που περιγράφηκαν δεν είναι ανεξάρτητες μεταξύ τους. Αντίθετα, η μεταβολή κάποιας από αυτές έχει αντίκτυπο και στις άλλες. Ειδικότερα, η ακρίβεια των προβλέψεων μειώνεται σημαντικά με την αύξηση άλλων μετρικών. Στο παρόν κεφάλαιο, εξετάζονται οι σχέσεις μεταξύ των διαφορετικών μετρικών, με βάση την υπάρχουσα βιβλιογραφία.

Όπως περιγράφεται και στην εργασία [13] η κάλυψη είναι στενά συνδεδεμένη με την ακρίβεια των συστάσεων. Συγκεκριμένα, η κάλυψη (καταλόγου και αντικειμένων) μειώνεται με την αύξηση της ακρίβειας και το αντίστροφο. Η ιδέα είναι ότι οι ακριβείς συστάσεις είναι ανάλογες της ποσότητας των διαθέσιμων δεδομένων που έχει το σύστημα για να δημιουργήσει συστάσεις. Πιο συγκεκριμένα, σε συστήματα συστάσεων collaborative filtering όσο περισσότερες βαθμολογίες υπάρχουν διαθέσιμες για κάθε αντικείμενο και χρήστη τόσο πιο ακριβείς συστάσεις θα παρέχει το σύστημα. Γενικά, τα πιο δημοφιλή αντικείμενα τείνουν να λαμβάνουν περισσότερες βαθμολογίες και ως εκ τούτου αντικείμενα της συγκεκριμένης ομάδας συστήνονται περισσότερες φορές, πράγμα που οδηγεί προφανώς σε χαμηλότερη κάλυψη. Αντίστοιχα, αν υπάρχουν απαιτήσεις για υψηλότερη κάλυψη, το σύστημα προτείνει λιγότερο δημοφιλή αντικείμενα με αποτέλεσμα η ακρίβεια να μειώνεται σημαντικά. Η ίδια λογική ισχύει και για συστήματα συστάσεων βασισμένα στη γνώση ή στο περιεχόμενο (knowledge/content based recommender systems).

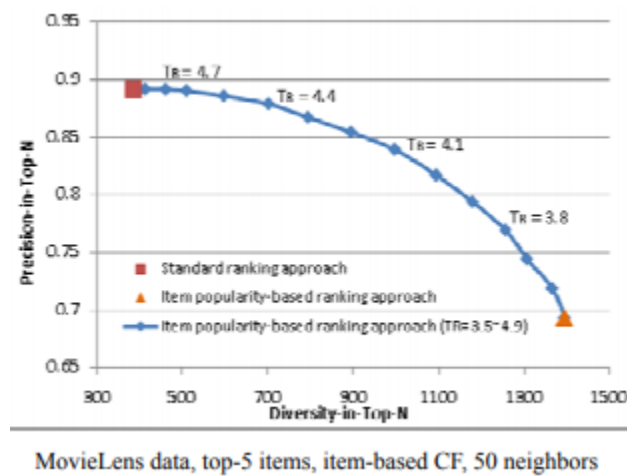
Όπως η κάλυψη, έτσι και η διαφοροποίηση συνδέεται στενά με την ακρίβεια (accuracy) των συστάσεων, και συγκεκριμένα αύξηση της διαφοροποίησης προκαλεί μείωση στην ακρίβεια και το αντίστροφο, όπως περιγράφεται και στην εργασία [8]. Αυτό είναι σχετικά προφανές, μιας και για να εξασφαλιστεί μεγαλύτερη διαφοροποίηση θα πρέπει να αντικατασταθούν προτάσεις δημοφιλών αντικειμένων με λιγότερο δημοφιλή. Στις παρακάτω εικόνες φαίνεται πειραματικά η σχέση της ακρίβειας με τη διαφοροποίηση (1) για αύξηση της διαφοροποίησης συστήνοντας τυχαία λιγότερο δημοφιλή αντικείμενα και (2) αύξηση της διαφοροποίησης μέσω κάποιων μεθοδολογιών που προσπαθούν να εξασφαλίσουν την ελάχιστη δυνατή μείωση της ακρίβειας.

Quality Metric:	Accuracy	Diversity
	Top-1 recommendation of:	
Popular Item (item with the largest number of known ratings)	82%	49 distinct items
“Long-Tail” Item (item with the smallest number of known ratings)	68%	695 distinct items

**Note.** Recommendations for 2828 users by a standard item-based collaborative filtering technique on MovieLens data.

Εικόνα 3 - Accuracy Diversity trade off (randomly select more less popular items) [8]

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 4 - Accuracy diversity trade off (αποτελέσματα από χρήση αλγορίθμου για μικρότερη μείωση ακρίβειας) [8]

Σχετικά με τη σχέση της δικαιοσύνης με την ακρίβεια ισχύει ότι και με την κάλυψη και τη διαφοροποίηση. Δηλαδή, αύξηση της δικαιοσύνης οδηγεί σε μείωση της ακρίβειας, μιας και για να επιτευχθεί μεγαλύτερη δικαιοσύνη απαιτείται να αντικατασταθούν κάποια πολύ δημοφιλή προτεινόμενα αντικείμενα με κάποια λιγότερο δημοφιλή, όπως αναφέρεται και στην εργασία [29].

Όσον αφορά τη σχέση του serendipity με την ακρίβεια, προφανώς η αύξηση του πρώτου προκαλεί σημαντική πτώση στην ακρίβεια. Ωστόσο σύμφωνα με την εργασία [13], αυτό δε σημαίνει ότι οι serendipitous συστάσεις αντικειμένων είναι ανεπιθύμητες, μιας και είναι σημαντικό ο χρήστης να έχει την ευκαιρία να ανακαλύψει νέο περιεχόμενο το οποίο φαινομενικά δεν είναι του γούστου του. Ωστόσο, η αύξηση του serendipity θα πρέπει να γίνεται στρατηγικά, ώστε να μειώνεται το ρίσκο να μπερδευτεί ο χρήστης, ή να πάψει να εμπιστεύεται το σύστημα συστάσεων. Στην ίδια εργασία παρουσιάζεται η πεποίθηση ότι το συγκεκριμένο ρίσκο μπορεί να μετριαστεί παρέχοντας στο χρήστη εξηγήσεις, γιατί του προτάθηκε ένα serendipitous αντικείμενο και κυρίως σχεδιάζοντας καλύτερα τη λίστα προτεινόμενων αντικειμένων του χρήστη, ή παρέχοντας του περισσότερες λίστες. Αυτό γιατί μια λίστα συστάσεων δίνει την εντύπωση ότι τα συνιστώμενα αντικείμενα είναι παρόμοια και το πρώτο είναι το πιο ακριβές ή επιθυμητό. Ωστόσο, αν λίγα από αυτά τα αντικείμενα είναι σημαντικά διαφορετικά από τα υπόλοιπα μπορούν να διεγείρουν την περιέργεια και το ενδιαφέρον του χρήστη και ως εκ τούτου να αυξηθεί η ποιότητα των συστάσεων, όπως την αντιλαμβάνεται ο χρήστης.

Επιπλέον αξίζει να σημειωθεί ότι η κάλυψη σχετίζεται με τη διαφοροποίηση. Συγκεκριμένα, όταν αναφερόμαστε σε κάλυψη καταλόγου είναι λογικό μια αύξηση της τελευταίας να οδηγεί και σε αύξηση της συνολικής διαφοροποίησης. Επιπροσθέτως, όταν υπάρχει η απαίτηση για κάλυψη αντικειμένων σε χρήστες, τότε μια αύξηση της απαιτούμενης κάλυψης δημιουργεί σημαντική αύξηση στη διαφοροποίηση χρηστών για το συγκεκριμένο αντικείμενο, ακόμα και αν οι νέοι χρήστες στους οποίους προτείνεται το αντικείμενο δεν είναι πολύ διαφορετικοί με τους υπόλοιπους. Αυτό συμβαίνει γιατί η διαφοροποίηση χρηστών ανά αντικείμενο έχει οριστεί ως το άθροισμα της ανομοιότητας μεταξύ των χρηστών στους οποίους έχει προταθεί το αντικείμενο, συνεπώς ένας επιπλέον χρήστης οδηγεί τη διαφοροποίηση σε αύξηση ίση με το άθροισμα της ανομοιότητας όλων των υπόλοιπων χρηστών με το νέο χρήστη στον οποίο

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

προτάθηκε το αντικείμενο. Αυτό ακριβώς συμβαίνει και με την αύξηση της κάλυψης καταλόγου στη συνολική διαφοροποίηση αντικειμένων.

Η δικαιοσύνη είναι άρρηκτα συνδεδεμένη με την κάλυψη. Πιο συγκεκριμένα, η δικαιοσύνη είναι ένας σχετικά αφηρημένος όρος και δεν υπάρχει κάποιος κοινά αποδεκτός τρόπος μέτρησης της, σε αντίθεση με την κάλυψη που ορίζεται με πιο συγκεκριμένο τρόπο. Ωστόσο, σε πολλές περιπτώσεις οι μετρικές αυτές ταυτίζονται. Για παράδειγμα στην εργασία [32] το πρόβλημα των δίκαιων συστάσεων ανάγεται σε ένα πρόβλημα κάλυψης. Επιπλέον όταν απαιτείται ίσο ποσοστό συστάσεων για παρόμοιους παρόχους αντικειμένων (P-fairness), η απαίτηση αυτή μπορεί να οριστεί και ως απαίτηση για ίδια κάλυψη αντικειμένων ανά πάροχο. Από την άλλη πλευρά σε πολλές περιπτώσεις η δικαιοσύνη δε μπορεί να αναπαρασταθεί ως πρόβλημα κάλυψης, όπως για παράδειγμα στις συστάσεις παρόμοιων αντικειμένων σε άντρες και γυναίκες με τα ίδια προσόντα. Επομένως, η σχέση της δικαιοσύνης με την κάλυψη εξαρτάται κατά κύριο λόγο από την εκάστοτε απαίτηση και τον ορισμό για τη δικαιοσύνη.

Αντίστοιχα, το ίδιο ισχύει και με τη σχέση της δικαιοσύνης με τη διαφοροποίηση. Συγκεκριμένα, θα μπορούσαν να τεθούν απαιτήσεις για δικαιοσύνη ως προς τη διαφορετικότητα των χρηστών στους οποίους προτείνεται ένα αντικείμενο, ή στη διαφορετικότητα των αντικειμένων που προτείνονται στο χρήστη. Δηλαδή, να απαιτηθεί οι χρήστες να λαμβάνουν συστάσεις επαρκώς διαφορετικές μεταξύ τους, ή τα αντικείμενα να προσφέρονται σε επαρκώς διαφορετικούς χρήστες μεταξύ τους, που συμπίπτουν με απαιτήσεις για συγκεκριμένη διαφοροποίηση αντικειμένων ανά χρήστη και διαφοροποίηση χρηστών ανά αντικείμενο, αντίστοιχα. Ωστόσο, όπως και με τη σχέση της κάλυψης με τη δικαιοσύνη, η σχέση της δικαιοσύνης με τη διαφοροποίηση εξαρτάται από την εκάστοτε απαίτηση για δικαιοσύνη.

Όσον αφορά τη σχέση μεταξύ serendipity και κάλυψης, είναι πολύ στενή. Πιο συγκεκριμένα, σύμφωνα με την εργασία [13], μια αύξηση του serendipity οδηγεί συνήθως σε αύξηση της κάλυψης καταλόγου. Ωστόσο, μια αύξηση της κάλυψης καταλόγου δε συνεπάγεται υποχρεωτικά και αύξηση του serendipity. Αυτό είναι λογικό μιας και τα serendipitous αντικείμενα είναι συνήθως αντικείμενα που δε ταιριάζουν στο προφίλ των χρηστών, και ως εκ τούτου συνήθως είναι εκτός καταλόγου. Συνεπώς, συστήνοντας τα, αυξάνεται η κάλυψη αντικειμένων των χρηστών. Αντίθετα, αυξάνοντας την κάλυψη αντικειμένων για κάποιους χρήστες, σημαίνει ότι προτείνεται μεγαλύτερη ποικιλία αντικειμένων, αλλά όχι απαραίτητα και serendipitous αντικειμένων.

Τέλος, η διαφοροποίηση συνδέεται στενά νοηματικά με το serendipity. Διαισθητικά, είναι εύκολο να σκεφτεί κανείς ότι μεγαλύτερη διαφοροποίηση οδηγεί σε μεγαλύτερο serendipity και το αντίστροφο. Σε γενικές γραμμές η σκέψη αυτή είναι σωστή, μιας και αν αυξήσουμε τη διαφοροποίηση των προτεινόμενων αντικειμένων ενός χρήστη είναι πολύ πιθανό να αυξηθεί και το serendipity. Αντίστοιχα, αν στη λίστα προτεινόμενων αντικειμένων ενός χρήστη αντικαταστήσουμε κάποια αντικείμενα με κάποια serendipitous αντικείμενα, εκτός από το serendipity μάλλον θα αυξηθεί και η διαφοροποίηση των αντικειμένων για το συγκεκριμένο χρήστη. Ωστόσο, αξίζει να σημειωθεί ότι η διαφοροποίηση εστιάζει στη διαφορετικότητα μεταξύ των προτεινόμενων αντικειμένων, ενώ το serendipity στην πρωτοτυπία και στην έκπληξη. Συνεπώς, αν το σύστημα παρείχε στο χρήστη μια λίστα με αποκλειστικά serendipitous αντικείμενα, τότε το serendipity θα ήταν μεγάλο, αλλά εάν τα αντικείμενα αυτά είχαν πολλά κοινά μεταξύ τους η διαφοροποίηση θα ήταν μικρή. Αντίστοιχα, μια λίστα με μεγάλη διαφοροποίηση δεν είναι υποχρεωτικό να συνοδεύεται και από μεγάλο serendipity μιας και τα

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

αντικείμενα μπορεί να μην προκαλούν έκπληξη στο χρήστη αλλά να είναι διαφορετικά μεταξύ τους π.χ. λόγω πολλών και διαφορετικών ενδιαφερόντων του χρήστη.

### 3.6 Κοινωνική διάσταση Συστημάτων Συστάσεων

Όπως παρουσιάστηκε και στα προηγούμενα υποκεφάλαια του συγκεκριμένου κεφαλαίου, τα συστήματα συστάσεων εμπλέκονται σχεδόν σε όλες τις πτυχές της καθημερινότητας των ανθρώπων, από τη διασκέδαση έως τα ψώνια και την κοινωνικοποίηση. Αυτό συμβαίνει διότι δρουν ως ένα προσωποποιημένο φίλτρο για κάθε χρήστη, που προστατεύει από την υπερφόρτωση με πληροφορίες που χαρακτηρίζει το σημερινό διαδίκτυο. Μάλιστα λέγεται ότι η υπερφόρτωση πληροφορίας στο διαδίκτυο είναι τόσο εκτεταμένη που χωρίς τα συστήματα συστάσεων να φιλτράρουν την πληροφορία που φτάνει στον χρήστη, θα ήταν αδύνατο ο τελευταίος να περιηγηθεί στο διαδίκτυο [6]. Έτσι, ο στόχος παραμένει η παροχή όσο το δυνατόν πιο αρεστών αντικειμένων στον χρήστη. Ωστόσο, τα τελευταία χρόνια, στην πραγματικότητα, τα συστήματα συστάσεων καθοδηγούνται από το ψηφιακό marketing και δίνεται μεγάλη έμφαση σε εμπορικούς σκοπούς αντί για την ικανοποίηση του τελικού χρήστη [38].

Έτσι, τα τελευταία χρόνια, η προσοχή έχει στραφεί από τον χρήστη στη μεγιστοποίηση του κέρδους. Για το λόγο αυτό, ο σκοπός ενός μοντέρνου συστήματος συστάσεων είναι να κρατάει τους χρήστες προσκολλημένους στην πλατφόρμα και όχι απλά να απορρίπτει αντικείμενα για τα οποία οι χρήστες δεν ενδιαφέρονται. Ένα γεγονός που καταδεικνύει αυτή την αλλαγή είναι ότι ο αλγόριθμος που κέρδισε τον διαγωνισμό του Netflix, ο οποίος προσέφερε βραβείο ενός εκατομμυρίου δολαρίων αν επιτυχανόταν 10% αύξηση της ακρίβειας πρόβλεψης δε χρησιμοποιήθηκε ποτέ από το Netflix. Ο λόγος ήταν ότι μετά το διαγωνισμό, το Netflix μετατράπηκε από εταιρεία ενοικίασης DVD σε διαδικτυακή εταιρεία αναπαραγωγής περιεχομένου (content streaming company). Κατά συνέπεια, ο στόχος δεν ήταν πλέον η πρόβλεψη των μελλοντικών προτιμήσεων του χρήστη, αλλά να τον διατηρήσει στην πλατφόρμα όσο το δυνατόν περισσότερο [39].

Φυσικά, αυτή η αλλαγή στη στόχευση των συστημάτων συστάσεων μπορεί να έχει και αρνητικές συνέπειες για τους χρήστες όσο και για την κοινωνία γενικότερα. Για παράδειγμα, η Frances Haugen, πρώην εργαζόμενη στο Facebook, κατηγορήσε την εταιρεία ότι βάζει τα κέρδη πάνω από τον άνθρωπο [40]. Συγκεκριμένα, αποκάλυψε, μεταξύ άλλων, πως το Instagram έχει αρνητικό αντίκτυπο στην ψυχική και πνευματική υγεία των έφηβων κοριτσιών, ότι το Facebook ευθύνεται για την αύξηση της βίας στις αναπτυσσόμενες χώρες, ενώ το σύστημα συστάσεων του Facebook φαίνεται να δίνει προτεραιότητα στο διχαστικό περιεχόμενο έναντι του ουδέτερου.

Ένα άλλο παράδειγμα του αρνητικού αντίκτυπου του Facebook στους χρήστες του έχει αναγνωριστεί στην εργασία [41]. Συγκεκριμένα, αναφέρεται ότι το Facebook δε φέρνει τους ανθρώπους πιο κοντά όπως διατείνεται, αλλά αντίθετα συμβάλλει στο διχασμό μεταξύ τους. Επιπλέον, αναφέρεται ότι ένας ερευνητής ανακάλυψε ότι πάνω από 33% των ομάδων του Facebook που είναι σχετικές με την πολιτική στη Γερμανία επιδείκνυε ανοδικές τάσεις σε ακραίο περιεχόμενο (extremist content), ενώ το 64% των συμμετεχόντων σε αυτές τις ομάδες, πήρε μέρος στις τελευταίες λόγω του αλγορίθμου συστάσεων του Facebook.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Αντίστοιχα παραδείγματα συναντώνται και στο σύστημα συστάσεων της Amazon. Συγκεκριμένα, μια έρευνα για τα βιβλία της Amazon σχετικά με τον εμβολιασμό έδειξε ότι ο αριθμός βιβλίων που είχαν αρνητική στάση ως προς τον εμβολιασμό ήταν διπλάσιος από τον αριθμό των βιβλίων που τον υποστήριζαν στο ηλεκτρονικό κατάστημα της Amazon. Επιπλέον, παρατηρήθηκε ότι στους χρήστες προτεινόταν βιβλία που ενισχύουν τις πεποιθήσεις τους [42]. Έτσι, εάν ένας χρήστης είχε αναζητήσει βιβλία αρνητικά προς τον εμβολιασμό, το σύστημα θα του πρότεινε περισσότερα τέτοια βιβλία. Επίσης, στο [43], οι συγγραφείς διατείνονται ότι το Σύστημα Συστάσεων της Amazon προωθεί θεωρίες συνωμοσίας. Σε αυτό το πλαίσιο, ένας χρήστης που θα αναζητήσει βιβλία για μια συγκεκριμένη θεωρία συνωμοσίας, θα λάβει προτάσεις βιβλίων που υποστηρίζουν και άλλες θεωρίες συνωμοσίας. Το ίδιο ισχύει και για βιβλία που υποστηρίζουν ρατσιστικές ή εξτρεμιστικές ιδεολογίες.

Έχοντας στο μυαλό μας όσα αναφέρθηκαν παραπάνω, έχει μεγάλο ενδιαφέρον να ερευνηθεί σε ποιο βαθμό ένα σύστημα συστάσεων μπορεί να θεωρηθεί ηθικά προβληματικό, όπως επίσης και πώς τα ηθικά προβλήματα που παρουσιάζονται θα μπορούσαν να αντιμετωπιστούν ολοκληρωτικά ή έστω μερικώς. Στις ακόλουθες παραγράφους εξετάζεται η υφιστάμενη βιβλιογραφία πάνω στο συγκεκριμένο θέμα.

Στα παραδοσιακά συστήματα συστάσεων η πλειονότητα των δημοσιευμένων εργασιών χρησιμοποιεί ως σημαντικότερη μετρική για την αξιολόγηση της απόδοσης των συστημάτων συστάσεων την ακρίβεια στην πρόβλεψη για τις προτιμήσεις των χρηστών προς τα αντικείμενα. Αυτή η προσέγγιση στοχεύει αποκλειστικά στην ικανοποίηση του χρήστη από τα αντικείμενα που του προτείνονται. Από την άλλη πλευρά, τα πιο μοντέρνα συστήματα συστάσεων δίνουν έμφαση στην διατήρηση του χρήστη και τη μεγιστοποίηση του κέρδους [39]. Ωστόσο, στις περισσότερες περιπτώσεις οι αλγόριθμοι που αποφασίζουν τις πολιτικές συστάσεων δεν αποκαλύπτονται και έτσι οι χρήστες αντιμετωπίζουν τα συστήματα συστάσεων ως «μαύρα κουτιά» [44]. Για αυτό το λόγο, έχει αναγνωριστεί ότι περισσότερες μετρικές (όπως αυτές που περιγράφηκαν στην προηγούμενη ενότητα) πρέπει να λαμβάνονται υπόψη στην αξιολόγηση ενός συστήματος συστάσεων, ενώ η προσοχή θα πρέπει να στρέφεται προς όλα τα ενδιαφερόμενα μέρη [45].

Συγκεκριμένα τα ενδιαφερόμενα μέρη σε ένα σύστημα συστάσεων είναι οι χρήστες, οι πάροχοι αντικειμένων, και ο λειτουργός ή διαχειριστής του συστήματος συστάσεων (recommender system operator). Σε αυτό το πλαίσιο, ένα νέο πρότυπο μελέτης των συστημάτων συστάσεων που ονομάζεται Πολυμετοχικά συστήματα συστάσεων (Multi-stakeholder recommender systems) έχει προταθεί στη βιβλιογραφία και η δημοτικότητα του αυξάνεται ραγδαία τα τελευταία χρόνια. Το πρότυπο αυτό μελετάει ένα σύστημα συστάσεων υπό το πρίσμα όλων των ενδιαφερόμενων μερών (μετόχων) [46].

Όσον αφορά τις εναλλακτικές μετρικές αξιολόγησης, οι σημαντικότερες είναι η κάλυψη, η διαφοροποίηση, η δικαιοσύνη και το serendipity και περιγράφηκαν αναλυτικά στην προηγούμενη ενότητα.

Ωστόσο, παρότι έχει γίνει εκτεταμένη έρευνα για την βελτίωση της απόδοσης των συστημάτων συστάσεων, πολύ λίγες δημοσιευμένες εργασίες εξετάζουν τον αντίκτυπο των τελευταίων στους χρήστες, όσο και στις ηθικές προκλήσεις που προκύπτουν λόγω και της μαζικής χρήσης τους. Πιο συγκεκριμένα, οι περισσότερες εργασίες ασχολούνται με ένα συγκεκριμένο σύστημα συστάσεων μιας συγκεκριμένης εταιρείας και τον αντίκτυπο του στους χρήστες [40], ή ένα συγκεκριμένο φαινόμενο όπως

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

είναι οι κάμαρες ηχούς (echo chambers) [47], [48] και η πόλωση [49] που συνδέεται με την λειτουργία των συστημάτων συστάσεων. Όσον αφορά τις ηθικές εκφάνσεις των συστημάτων συστάσεων, η έρευνα [50] παρουσιάζει περιπτώσεις όπου πραγματικά συστήματα συστάσεων τροποποιήθηκαν για να υπηρετήσουν διάφορες ανθρώπινες αξίες και προτείνει προσεγγίσεις που αξιοποιούν τη μεγαλύτερη συμμετοχή από όλα τα ενδιαφερόμενα μέρη για τη σύγκλιση της λειτουργίας των συστημάτων συστάσεων με τις ανθρώπινες αξίες. Επίσης, η εργασία [38] αναλύει τις ηθικές προκλήσεις που εγείρονται στα συστήματα συστάσεων για να διακρίνει τα διαφορετικά είδη κοινωνικού αντίκτυπου που προκύπτει από τη λειτουργία των συστημάτων συστάσεων. Τέλος, στην εργασία [51] οι ίδιοι συγγραφείς εξετάζουν τις ηθικές εκφάνσεις των συστημάτων συστάσεων υπό το πρίσμα των πολυμετοχικών συστημάτων συστάσεων και καταλήγουν ότι η πολυμετοχική προσέγγιση των συστημάτων συστάσεων έχει σημαντικά πλεονεκτήματα σε σύγκριση με την παραδοσιακή προσέγγιση που έχει το χρήστη στο επίκεντρο.

Κατά τη συγγραφή της παρούσας διατριβής, έλειπε από τη βιβλιογραφία τόσο ένα πλαίσιο αξιολόγησης των συστημάτων συστάσεων ως προς τις ηθικές προκλήσεις που παρουσιάζει, όσο και ένα ρυθμιστικό πλαίσιο για την αντιμετώπιση τέτοιων ηθικών ζητημάτων που μπορεί να προκύπτουν σε κάποιο σύστημα συστάσεων.

### **3.7 Συμπεράσματα**

Με βάση την ανάλυση που προηγήθηκε είναι εμφανές ότι το πρόβλημα των συστάσεων ενός συστήματος συστάσεων γίνεται αρκετά πολύπλοκο όταν ο αλγόριθμος του πρέπει να λαμβάνει υπόψη επιπλέον μετρικές εκτός από την ακρίβεια πρόβλεψης. Ένας λόγος είναι ότι υπάρχουν πολλοί και διαφορετικοί ορισμοί για τις διάφορες μετρικές και όχι κάποιος κοινά αποδεκτός ορισμός αυτών. Επιπλέον, η ποσοτικοποίηση ορισμένων μετρικών παρουσιάζει προβλήματα, μιας και μπορεί να είναι πολύ πολύπλοκη για να εισαχθεί στον αλγόριθμο κάποιου συστήματος συστάσεων, ενώ συνήθως όταν αυξάνεται κάποια άλλη μετρική εκτός της ακρίβειας, αυτό γίνεται σε βάρος της ποιότητας των συστάσεων ως προς το πόσο επιθυμητές αναμένεται να τις βρει ο χρήστης.

Επιπροσθέτως, αξίζει να σημειωθεί ότι πολλές έρευνες αξιολογούν τις επιπλέον μετρικές εκ των υστέρων (μετά από κάποια ενέργεια) και χωρίς να μπαίνουν ως απαιτήσεις στον αλγόριθμο του συστήματος, πράγμα που είναι πολύ σημαντικό για την εξασφάλιση συγκεκριμένων στόχων για τις διάφορες μετρικές. Τέλος, αξίζει να σημειωθεί ότι πολλοί ορισμοί από τη βιβλιογραφία που εντάσσουν και επιπλέον μετρικές στο πρόβλημα οδηγούν σε προβλήματα για τα οποία δεν είναι δυνατόν να βρεθεί η βέλτιστη λύση σε πολυωνυμικό χρόνο (NP-complete). Επομένως, υπάρχει μεγάλη ανάγκη είτε για απλοποίηση κάποιων από τους ορισμούς της βιβλιογραφίας είτε για ευριστικούς αλγορίθμους οι οποίοι βρίσκουν υποβέλτιστες λύσεις που όμως θα είναι κοντά στη βέλτιστη.

Ακόμη, από την ανασκόπηση της υφιστάμενης βιβλιογραφίας παρατηρείται ότι μερικά από τα πιο διαδεδομένα συστήματα συστάσεων προκαλούν αρνητικές επιπτώσεις σε ομάδες του πληθυσμού και στην κοινωνία γενικότερα. Ενδεικτικά, η λειτουργία τους συνδέεται με φαινόμενα πόλωσης, προώθηση θεωριών συνωμοσίας, και αρνητικό αντίκτυπο στην ψυχική υγεία κάποιων χρηστών. Ωστόσο, κατά τη συγγραφή της παρούσας διατριβής μόνο περιορισμένος αριθμός ερευνών υπήρχε πάνω σε αυτό το θέμα, ενώ οι υπάρχουσες έρευνες ασχολούνται κυρίως με την αναγνώριση τέτοιων φαινομένων.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Ωστόσο, απουσιάζει από τη βιβλιογραφία ένα συστηματικό πλαίσιο αξιολόγησης για τον αντίκτυπο που έχει η λειτουργία κάποιου συστήματος συστάσεων. Επίσης, απουσιάζουν εργασίες που υποδεικνύουν πώς μπορούν να αντιμετωπιστούν ή να μετριαστούν οι αρνητικές επιπτώσεις από τη λειτουργία των συστημάτων συστάσεων.

### 3.8 Αναφορές

- [1] F. Ricci, L. Rokach, and B. Shapira, "Introduction to Recommender Systems Handbook," in *Recommender Systems Handbook*, 2011.
- [2] S. S. Sohail, J. Siddiqui, and R. Ali, "Classifications of recommender systems: A review," *Journal of Engineering Science and Technology Review*, vol. 10, no. 4. 2017, doi: 10.25103/jestr.104.18.
- [3] J. Stray, I. Vendrov, J. Nixon, S. Adler, and D. Hadfield-Menell, "What are you optimizing for? Aligning Recommender Systems with Human Values." 2021.
- [4] P. Resnick and H. Varian, "Recommender Systems," *Commun. ACM*, vol. 40, no. 3, pp. 56–58, 1997.
- [5] E. Karakolis, P. F. Oikonomidis, and D. Askounis, "Identifying and Addressing Ethical Challenges in Recommender Systems," 2022.
- [6] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, *Recommender systems: An introduction*, vol. 9780521493369. 2010.
- [7] A. Lambropoulos, "Machine learning-based recommendation methods for multimedia data," 2010.
- [8] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6. 2005, doi: 10.1109/TKDE.2005.99.
- [9] Π.-Φ. Οικονομίδης and Α. Δημήτριος, "Μελέτη των Ηθικών Επιπτώσεων των Συστημάτων Συστάσεων και Πρόταση Ενός Ρυθμιστικού Πλαισίου Για Την Εξυγιάνσή Τους," 2022.
- [10] S. M. M. and J. R. and J. A. Konstan, "Accurate is not always good : How Accuracy Metrics have hurt Recommender Systems," *Search*, 2006.
- [11] S. Vargas, L. Baltrunas, A. Karatzoglou, and P. Castells, "Coverage, redundancy and size-awareness in genre diversity for recommender systems," 2014, doi: 10.1145/2645710.2645743.
- [12] J. Oh, S. Park, H. Yu, M. Song, and S. T. Park, "Novel recommendation based on Personal Popularity Tendency," 2011, doi: 10.1109/ICDM.2011.110.
- [13] M. Ge, C. Delgado-Battenfeld, and D. Jannach, "Beyond accuracy: Evaluating recommender systems by coverage and serendipity," 2010, doi: 10.1145/1864708.1864761.
- [14] M. Kaminskis and D. Bridge, "Diversity, serendipity, novelty, and coverage: A survey and empirical analysis of beyond-Accuracy objectives in recommender systems," *ACM Transactions on Interactive Intelligent Systems*, vol. 7, no. 1. 2016, doi: 10.1145/2926720.



- [15] M. Koutsopoulos, I., & Halkidi, “Efficient and Fair Item Coverage in Recommender Systems,” in *IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing*, 2018, pp. 912–918.
- [16] M. Rahman and J. C. Oh, “Graph bandit for diverse user coverage in online recommendation,” *Appl. Intell.*, vol. 48, no. 8, 2018, doi: 10.1007/s10489-017-0977-1.
- [17] M. Hammar, R. Karlsson, and B. J. Nilsson, “Using maximum coverage to optimize recommendation systems in E-commerce,” 2013, doi: 10.1145/2507157.2507169.
- [18] H. Yin, B. Cui, J. Li, J. Yao, and C. Chen, “Challenging the long tail recommendation,” *Proc. VLDB Endow.*, vol. 5, no. 9, 2012, doi: 10.14778/2311906.2311916.
- [19] E. Brynjolfsson, Y. J. Hu, and M. D. Smith, “The Longer Tail: The Changing Shape of Amazon’s Sales Distribution Curve,” *SSRN Electron. J.*, 2012, doi: 10.2139/ssrn.1679991.
- [20] B. Smyth and P. McClave, “Similarity vs. Diversity,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2001, vol. 2080, doi: 10.1007/3-540-44593-5\_25.
- [21] C.-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen, “Improving recommendation lists through topic diversification,” 2005, doi: 10.1145/1060745.1060754.
- [22] S. Vargas, P. Castells, and D. Vallet, “Intent-oriented diversity in recommender systems,” 2011, doi: 10.1145/2009916.2010124.
- [23] M. D. Ekstrand, *Collaborative Filtering Recommender Systems*. 2011.
- [24] J. P. Kelly and D. Bridge, “Enhancing the diversity of conversational collaborative recommendations: A comparison,” in *Artificial Intelligence Review*, 2006, vol. 25, no. 1–2, doi: 10.1007/s10462-007-9023-8.
- [25] C. Yu, L. V. S. Lakshmanan, and S. Amer-Yahia, “Recommendation diversification using explanations,” 2009, doi: 10.1109/ICDE.2009.225.
- [26] G. Adomavicius and Y. O. Kwon, “Improving aggregate recommendation diversity using ranking-based techniques,” *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 5, 2012, doi: 10.1109/TKDE.2011.15.
- [27] G. Adomavicius and Y. O. Kwon, “Maximizing aggregate recommendation diversity: A graph-theoretic approach,” in *CEUR Workshop Proceedings*, 2011, vol. 816.
- [28] M. Kunaver and T. Požrl, “Diversity in recommender systems – A survey,” *Knowledge-Based Syst.*, vol. 123, 2017, doi: 10.1016/j.knosys.2017.02.009.
- [29] R. Burke, “Multisided fairness for recommendation,” 2017.
- [30] R. Burke, N. Sonboli, and A. Ordoñez-Gauger, “Balanced Neighborhoods for Multi-sided Fairness in Recommendation,” 2018.
- [31] X. Lin, M. Zhang, Y. Zhang, Z. Gu, Y. Liu, and S. Ma, “Fairness-aware group recommendation with pareto-efficiency,” 2017, doi: 10.1145/3109859.3109887.
- [32] D. Serbos, S. Qi, N. Mamoulis, E. Pitoura, and P. Tsaparas, “Fairness in package-to-group recommendations,” 2017, doi: 10.1145/3038912.3052612.

- [33] P. Adamopoulos and A. Tuzhilin, "On unexpectedness in recommender systems: Or how to better expect the unexpected," *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 4, 2014, doi: 10.1145/2559952.
- [34] L. Iaquinta, M. De Gemmis, P. Lops, G. Semeraro, M. Filannino, and P. Molino, "Introducing serendipity in a content-based recommender system," 2008, doi: 10.1109/HIS.2008.25.
- [35] K. Onuma, H. Tong, and C. Faloutsos, "Tangent: A novel, 'Surprise-me', recommendation algorithm," 2009, doi: 10.1145/1557019.1557093.
- [36] M. Nakatsuji, Y. Fujiwara, A. Tanaka, T. Uchiyama, K. Fujimura, and T. Ishida, "Classical music for rock fans?: Novel recommendations for expanding user interests," 2010, doi: 10.1145/1871437.1871558.
- [37] Y. C. Zhang, D. Ó. Séaghdha, D. Quercia, and T. Jambor, "Auralist: Introducing serendipity into music recommendation," 2012, doi: 10.1145/2124295.2124300.
- [38] S. Milano, M. Taddeo, and L. Floridi, "Recommender systems and their ethical challenges," *AI Soc.*, vol. 35, no. 4, 2020, doi: 10.1007/s00146-020-00950-y.
- [39] N. Seaver, "Captivating algorithms: Recommender systems as traps," *J. Mater. Cult.*, vol. 24, no. 4, 2019, doi: 10.1177/1359183518820366.
- [40] D. Milmo, "Facebook revelations: what is in cache of internal documents?," *The Guardian*, 2021.
- [41] J. Horwitz and D. Seetharaman, "Facebook Executives Shut Down Efforts to Make the Site Less Divisive," *Wall Str. J.*, 2020.
- [42] J. Shin and T. Valente, "Algorithms and Health Misinformation: A Case Study of Vaccine Books on Amazon," *J. Health Commun.*, vol. 25, no. 5, 2020, doi: 10.1080/10810730.2020.1776423.
- [43] E. Thomas, "Amazon's algorithms, conspiracy theories and extremist literature," *Inst. Strateg. Dialogue*, 2021.
- [44] M. Millecamp, N. N. Htun, Y. Jin, and K. Verbert, "Controlling Spotify Recommendations: Effects of Personal Characteristics on Music Recommender User Interfaces," *Proc. 26th Conf. User Model. Adapt. Pers.*, 2018.
- [45] H. Abdollahpouri, R. Burke, and B. Mobasher, "Recommender systems as multistakeholder environments," 2017, doi: 10.1145/3079628.3079657.
- [46] R. Burke, H. Abdollahpouri, E. C. Malthouse, K. P. Thai, and Y. Zhang, "Recommendation in multistakeholder environments," 2019, doi: 10.1145/3298689.3346973.
- [47] W. Quattrociocchi, A. Scala, and C. R. Sunstein, "Echo Chambers on Facebook," *SSRN Electron. J.*, 2018, doi: 10.2139/ssrn.2795110.
- [48] F. Cinus, M. Minici, C. Monti, and B. F., "The Effect of People Recommenders on Echo Chambers and Polarization." 2021.
- [49] M. Badami, O. Nasraoui, W. Sun, and P. Shafto, "Detecting polarization in ratings: An automated pipeline and a preliminary quantification on several benchmark data sets," in *Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017*, 2017, vol. 2018-January, doi: 10.1109/BigData.2017.8258231.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

[50] J. Stray, "Designing Recommender Systems to Depolarize," 2021.

[51] S. Milano, M. Taddeo, and L. Floridi, "Ethical aspects of multi-stakeholder recommendation systems," *Inf. Soc.*, vol. 37, no. 1, 2021, doi: 10.1080/01972243.2020.1832636.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## ΚΕΦΑΛΑΙΟ 4: Η ΠΛΑΤΦΟΡΜΑ QUALICHAIN

Έχοντας παρουσιάσει εκτενώς:

- τόσο τα προβλήματα που είναι προς επίλυση στο πλαίσιο της παρούσας διδακτορικής διατριβής, και το σχετικό θεωρητικό υπόβαθρο που απαιτείται,
- όσο και την ανασκόπηση παρόμοιων εργασιών στον τομέα και το έργο QualiChain, στο πλαίσιο του οποίου χρησιμοποιήθηκαν και αξιολογήθηκαν οι υπηρεσίες που αναπτύχθηκαν στη συγκεκριμένη διδακτορική διατριβή

στο παρόν κεφάλαιο παρουσιάζεται η τεχνολογική λύση του έργου QualiChain, με την αρχιτεκτονική της όσο και τα επιμέρους υποσυστήματα (components) από τα οποία αποτελείται, καθώς επίσης και τις σημαντικότερες υπηρεσίες που παρέχει κάθε υποσύστημα.

### 4.1 Η Αρχιτεκτονική της πλατφόρμας QualiChain

Η πλατφόρμα QualiChain αποτελείται από αρκετές ανεξάρτητες και ετερογενείς υπηρεσίες οι οποίες επιτελούν συγκεκριμένες λειτουργικότητες. Κατά συνέπεια, η αρχιτεκτονική της είναι αρκετά πολύπλοκη όσον αφορά τόσο τα διαφορετικά υποσυστήματα, όσο και τις αλληλεπιδράσεις μεταξύ αυτών. Παρακάτω παρουσιάζεται μια γενική εικόνα της τεχνικής αρχιτεκτονικής, η οποία συμπεριλαμβάνει τα διάφορα υποσυστήματα και τις επιμέρους ενότητες (modules) αυτών, όσο και τις υπηρεσίες κάθε υποσυστήματος, παράλληλα με την επικοινωνία και τις αλληλεπιδράσεις μεταξύ των διαφορετικών υποσυστημάτων.

Συγκεκριμένα, όπως φαίνεται και στην Εικόνα 5, η πλατφόρμα QualiChain αποτελείται από 6 βασικά υποσυστήματα. Τα υποσυστήματα αυτά είναι τα ακόλουθα:

- **Το υποσύστημα για επαλήθευση και επιδότηση ακαδημαϊκών προσόντων (Academic Verification and Accreditation Component)** με τη χρήση της τεχνολογίας blockchain. Αποτελείται από υπηρεσίες για την επικοινωνία με το blockchain όσο και με το γράφο γνώσης (Knowledge graph).
- **Ο γράφος γνώσης της πλατφόρμας QualiChain (QualiChain Knowledge Graph)** αποτελείται από τις υπηρεσίες Knowledge Graph Query Engine και την οντολογία του QualiChain και αποτελεί το κεντρικό αποθετήριο δεδομένων (data repository) της πλατφόρμας QualiChain. Η οντολογία χρησιμοποιείται για την αποθήκευση και τη διαλειτουργικότητα των δεδομένων ενώ η υπηρεσία Knowledge Graph Query Engine επιτρέπει την εκτέλεση ερωτημάτων στην οντολογία. Το υποσύστημα QualiChain Knowledge Graph επικοινωνεί με όλα τα υπόλοιπα υποσυστήματα της πλατφόρμας QualiChain.
- **Το υποσύστημα εξαγωγής γνώσης (Knowledge Extraction Component)** είναι υπεύθυνο για την απόκτηση, την επισήμανση (annotation) και την εξόρυξη γνώσης από τα δεδομένα (π.χ. την εξαγωγή των απαιτούμενων γνώσεων και δεξιοτήτων σε μια αγγελία εργασίας). Αλληλοεπιδρά

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

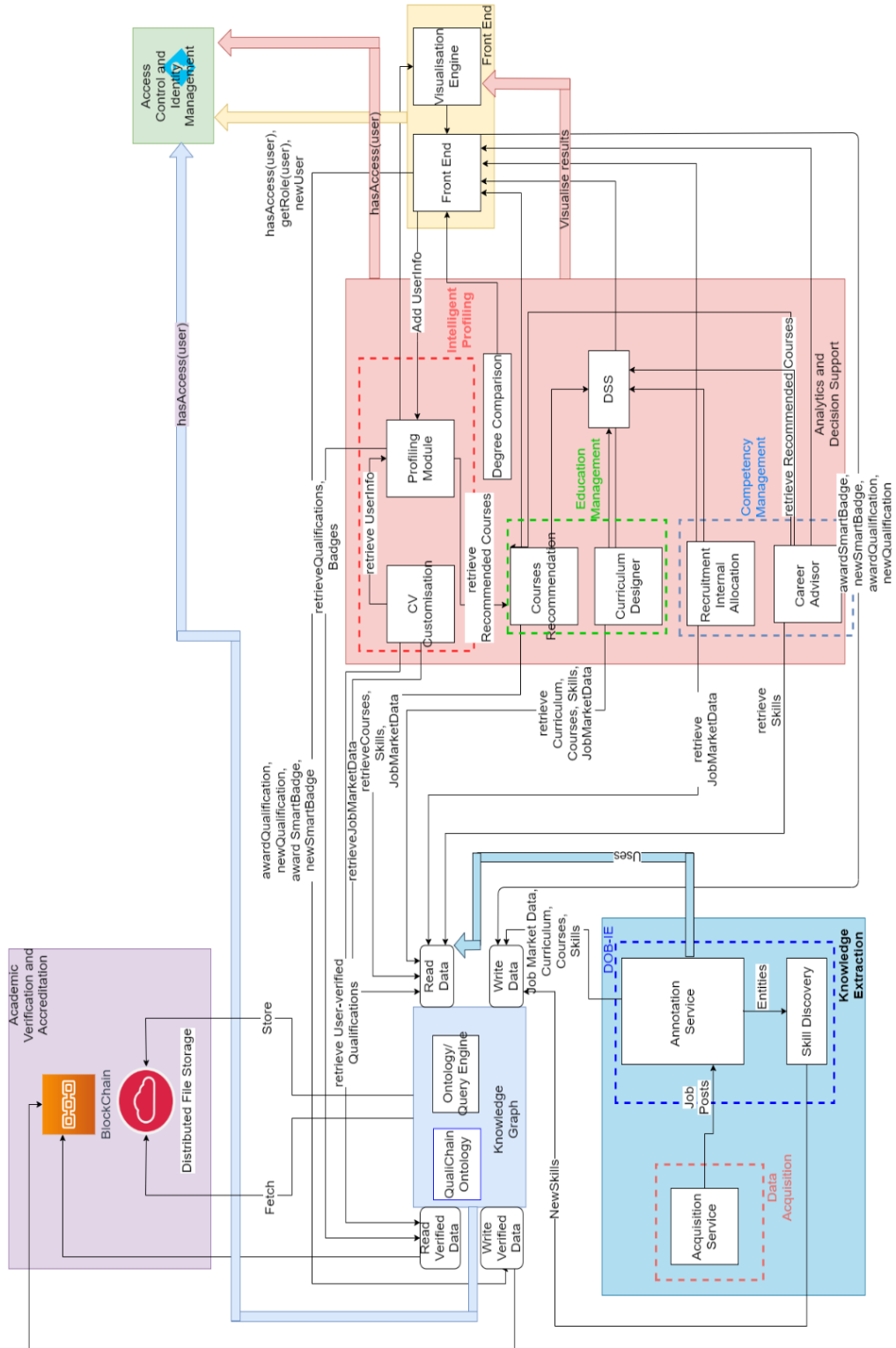
μόνο με το γράφο γνώσης της πλατφόρμας QualiChain για την αποθήκευση της εξαχθείσας γνώσης.

- **Το υποσύστημα ταυτοποίησης και ελέγχου πρόσβασης (Identity and Access Control Management component - IAM)** επικοινωνεί με όλα τα άλλα υποσυστήματα και διασφαλίζει ότι μόνο χρήστες οι οποίοι έχουν την άδεια να προσπελάσουν τα ζητούμενα δεδομένα και υπηρεσίες έχουν πρόσβαση σε αυτά.
- **Το υποσύστημα αναλυτικής δεδομένων και υποστήριξης αποφάσεων (Analytics and Decision Support component)** αποτελείται από κάποιες υπηρεσίες που βασίζονται στην αναλυτική δεδομένων και συγκεκριμένα την προσωποποιημένη προσαρμογή βιογραφικού σημειώματος (CV customisation service), τη δημιουργία προφίλ για τους χρήστες, την υπηρεσία προσωποποιημένων προτάσεων μαθημάτων σε δια βίου μαθητές, την υπηρεσία ανασχεδιασμού προγράμματος σπουδών, την υπηρεσία πρόσληψης και κάλυψης εταιρικών αναγκών σε προσωπικό (recruitment and internal allocation service), την υπηρεσία σύγκρισης πτυχίων και την υπηρεσία συμβουλευτικής καριέρας, όπως επίσης και μια υπηρεσία που παρέχει μια σειρά από μεθόδους πολυκριτηριακής ανάλυσης για την υποστήριξη αποφάσεων. Το συγκεκριμένο υποσύστημα επικοινωνεί με το γράφο γνώσης της πλατφόρμας QualiChain για την λήψη και την αποθήκευση δεδομένων.
- Τέλος, **το Front End** αποτελεί τη διεπαφή για τους χρήστες, ώστε να μπορέσουν να χρησιμοποιήσουν όλες τις λειτουργικότητες που παρέχονται από την πλατφόρμα QualiChain. Αξίζει να σημειωθεί ότι για την οπτική αναπαράσταση των δεδομένων αναπτύχθηκε μια υπηρεσία οπτικοποίησης δεδομένων (Visualisation Engine). Φυσικά, το Front End επικοινωνεί με όλες τις υπηρεσίες, ώστε να παρέχει στο χρήστη ευκολία στη χρήση των υπηρεσιών που αναπτύχθηκαν.

Όλα τα υποσυστήματα που περιγράφηκαν παραπάνω είναι ανεξάρτητα μεταξύ τους και εγκατεστημένα σε διαφορετικούς διακομιστές διαδικτύου (web servers). Η επικοινωνία μεταξύ αυτών των υποσυστημάτων πραγματοποιείται μέσω μιας αρχιτεκτονικής μικρο-υπηρεσιών (microservices architecture) και συγκεκριμένα μέσω του μοντέλου publish/subscribe, ενώ το Front End καλεί τις διαθέσιμες υπηρεσίες με τη χρήση APIs. Η επικοινωνία μεταξύ των διαφορετικών υπηρεσιών γίνεται μέσω της υπηρεσίας QualiChain Mediator.

Αξίζει να σημειωθεί ότι η Εικόνα 5 παρέχει μια πιο λεπτομερή αναπαράσταση όσων περιγράφηκαν παραπάνω, καθώς δείχνει με λεπτομέρεια τις αλληλεπιδράσεις μεταξύ των διαφόρων υποσυστημάτων, συμπεριλαμβανομένου του τύπου των μηνυμάτων που ανταλλάσσουν μεταξύ τους οι διάφορες υπηρεσίες. Επιπλέον, μια πιο λεπτομερής περιγραφή των διαφόρων υποσυστημάτων παρέχεται στην επόμενη ενότητα.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 5 - Η αρχιτεκτονική της πλατφόρμας QualiChain

## 4.2 Βασικές υπηρεσίες της πλατφόρμας QualiChain

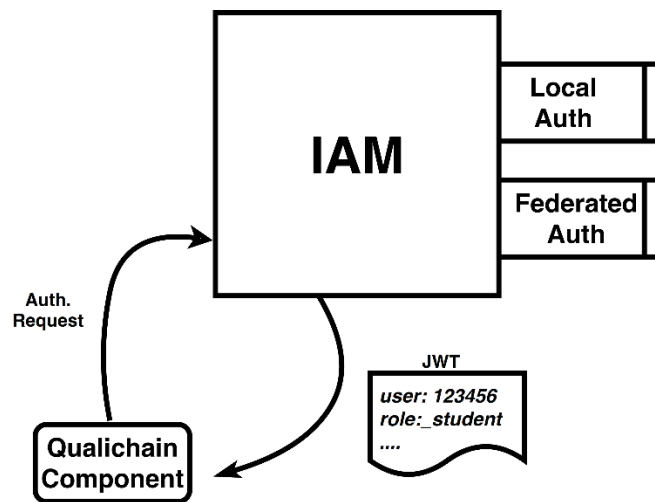
Στην παρούσα ενότητα παρουσιάζονται οι σημαντικότερες υπηρεσίες της πλατφόρμας QualiChain, οργανωμένες με βάση το υποσύστημα στο οποίο ανήκουν.

### 4.2.1 Υπηρεσία ταυτοποίησης και ελέγχου πρόσβασης (Identity and Access Control Management)

Η συγκεκριμένη υπηρεσία όπως έχει ήδη αναφερθεί είναι υπεύθυνη για την ταυτοποίηση των χρηστών και τον έλεγχο πρόσβασης σε δεδομένα και υπηρεσίες της πλατφόρμας QualiChain ανάλογα με το ρόλο του χρήστη και τα δικαιώματά του ως προς τα δεδομένα που σκοπεύει να χρησιμοποιήσει. Αποτελείται από δυο ενότητες (modules): το IAM, που είναι υπεύθυνο για την ταυτοποίηση και την ανάθεση ρόλων στους χρήστες και το Access Control Middleware που λειτουργεί ως σημείο επιβολής πολιτικών (Policy Enforcement Point - PEP).

#### 4.2.1.1 Η υπηρεσία ταυτοποίησης και ελέγχου πρόσβασης (IAM)

Η υπηρεσία IAM είναι υπεύθυνη για τη διαχείριση της ταυτότητας των χρηστών, δηλαδή με την επαλήθευση ότι οι χρήστες είναι αυτοί που ισχυρίζονται πως είναι, όσο και με την εξουσιοδότηση (authorisation), η οποία στην προκειμένη περίπτωση είναι η διαχείριση και αποθήκευση των ρόλων κάθε χρήστη. Το IAM χρησιμοποιείται από όλες τις υπηρεσίες του συστήματος για να επαληθεύσουν αν ο χρήστης έχει τα δικαιώματα να χρησιμοποιήσει κάποια υπηρεσία ή κάποια δεδομένα της πλατφόρμας QualiChain. Στην παρακάτω εικόνα παρουσιάζεται μια απλοποιημένη εικόνα της αρχιτεκτονικής της υπηρεσίας IAM.



Εικόνα 6 - Αρχιτεκτονική υπηρεσίας ταυτοποίησης και ελέγχου πρόσβασης (IAM)

Η υπηρεσία IAM υποστηρίζει την ταυτοποίηση με δύο μεθόδους, την τοπική ταυτοποίηση και την ομοσπονδιακή (federated) ταυτοποίηση. Η πρώτη μπορεί να χρησιμοποιηθεί για όλους τους ρόλους (καθηγητής, δια βίου μαθητής ή φοιτητής, υπεύθυνος πρόσληψης, ακαδημαϊκός οργανισμός, ή οργανισμός πρόσληψης) και ο χρήστης εγγράφεται μόνος του, μέσω μιας φόρμας. Η μέθοδος αυτή είναι ανεξάρτητη από άλλες υπηρεσίες. Όσον αφορά την ομοσπονδιακή ταυτοποίηση, αυτή εξαρτάται από



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

την πλατφόρμα που αναπτύχθηκε στο ερευνητικό έργο SEAL [1] για την παροχή διάφανους ταυτότητας για τους χρήστες μέσω των επαληθεύσιμων ακαδημαϊκών τους στοιχείων (eIDAS [2] και EduGain [3]). Το πρόβλημα με τη συγκεκριμένη μέθοδο είναι ότι υποστηρίζει μόνο ακαδημαϊκούς ρόλους και μάλιστα από ακαδημαϊκά ιδρύματα που έχουν υιοθετήσει τα συστήματα eIDAS και EduGain. Έτσι οι χρήστες που δεν ανήκουν στην προηγούμενη κατηγορία πρέπει να χρησιμοποιήσουν τη μέθοδο της τοπικής ταυτοποίησης.

Σχετικά με τις λεπτομέρειες υλοποίησης, λίστες ελέγχου πρόσβασης που απεικονίζουν τους ρόλους κάθε χρήστη σε ένα σύνολο δικαιωμάτων αποθηκεύονται στο σύστημα. Οι παραπάνω λίστες ορίζουν ποιες υπηρεσίες ή δεδομένα έχει δικαίωμα να χρησιμοποιήσει κάθε χρήστης. Τελικά, η ταυτοποίηση πραγματοποιείται μέσω ενός JWT token, μέσω του οποίου αναγνωρίζεται ο χρήστης.

#### 4.2.1.2 Το ενδιάμεσο λογισμικό ελέγχου πρόσβασης (Access Control Middleware)

Το ενδιάμεσο λογισμικό ελέγχου πρόσβασης (Access Control Middleware) λειτουργεί ως σημείο επιβολής πολιτικών (Policy Enforcement Point - PEP) για την πλατφόρμα QualiChain. Η συγκεκριμένη υπηρεσία είναι υπεύθυνη για τον έλεγχο κατά πόσο ένας χρήστης θα έπρεπε να έχει πρόσβαση στην υπηρεσία ή στα δεδομένα που αιτείται. Συγκεκριμένα, κατά την κλήση ενός API από έναν χρήστη, η συγκεκριμένη υπηρεσία ελέγχει εάν έχει δικαίωμα πρόσβασης στη ζητούμενη πληροφορία. Αν όντως ο χρήστης έχει το παραπάνω δικαίωμα, λαμβάνει την απάντηση που περιμένει. Σε αντίθετη περίπτωση, λαμβάνει μια απάντηση HTTP 401 (Unauthorized). Για τον έλεγχο εάν ο χρήστης έχει το δικαίωμα να χρησιμοποιήσει κάποια υπηρεσία, πρέπει να παρασχεθεί ένα access token κατά την κλήση του API. Για την απεικόνιση του παρεχόμενου Access Token σε κάποιον πραγματικό χρήστη, η υπηρεσία Access Control Middleware καλεί ένα API από την υπηρεσία IAM, το οποίο επιστρέφει πληροφορία για το χρήστη (π.χ. id, και ρόλους), σε περίπτωση που το token είναι έγκυρο.

Για την υλοποίηση της συγκεκριμένης υπηρεσίας χρησιμοποιήθηκε η προγραμματιστική τεχνοτροπία (design pattern) των decorators [4], ώστε να προστεθεί η λογική ελέγχου πρόσβασης πριν την εκτέλεση των APIs, και να είναι διαχωρισμένη από τη λογική των APIs. Έτσι, η λογική των APIs και της υπηρεσίας Access Control Middleware είναι εντελώς ανεξάρτητες. Όσον αφορά τις λεπτομέρειες υλοποίησης, τα APIs υλοποιήθηκαν με τη γλώσσα προγραμματισμού Python [5] και συγκεκριμένα με τη βιβλιοθήκη Flask-RESTful [6]. Η συγκεκριμένη βιβλιοθήκη επιτρέπει εύκολα την ενσωμάτωση decorators για έλεγχο πρόσβασης στα APIs. Σε αυτό το πλαίσιο, αναπτύχθηκαν διάφοροι decorators που εκτελούνται κατά την κλήση των APIs. Κάθε decorator, υλοποιεί διαφορετικές απαιτήσεις ταυτοποίησης και ελέγχου πρόσβασης.

Σε αυτό το πλαίσιο, αναπτύχθηκαν οι παρακάτω decorators για την κάλυψη όλων των διαφορετικών απαιτήσεων της πλατφόρμας QualiChain:

- only profile owner: Ο συγκεκριμένος decorator επιτρέπει την πρόσβαση σε ένα API αν και μόνο αν ο χρήστης που σκοπεύει να το χρησιμοποιήσει είναι και ο ιδιοκτήτης του προφίλ που ζητάει. Χρησιμοποιείται σε APIs τα οποία επιτρέπουν την ενημέρωση συγκεκριμένης προσωπικής πληροφορίας π.χ. την ενημέρωση πληροφοριών προφίλ ή την ενημέρωση πληροφοριών βιογραφικού σημειώματος.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

- only owner of notification: Ο συγκεκριμένος decorator χρησιμοποιείται για τη ρύθμιση των ειδοποιήσεων και επιτρέπει την πρόσβαση στα APIs μόνο σε χρήστες που αιτούνται να ρυθμίσουν ειδοποιήσεις που τους ανήκουν.
- only admins: Επιτρέπει την εκτέλεση APIs μόνο σε χρήστες οι οποίοι είναι διαχειριστές της πλατφόρμας QualiChain. Ο συγκεκριμένος decorator χρησιμοποιείται για ενέργειες ενημέρωσης της βάσης γνώσης της πλατφόρμας QualiChain, όπως για παράδειγμα η προσθήκη νέων γνώσεων και δεξιοτήτων στο Μοντέλο Δεδομένων του QualiChain (QualiChain Data Model).
- only lifelong learner: Ο συγκεκριμένος decorator επιτρέπει τη χρήση APIs μόνο σε δια βίου μαθητές. Οι δια βίου μαθητές αποτελούν τους χρήστες που αντιπροσωπεύουν πραγματικά πρόσωπα (φοιτητές, εργαζομένους, καθηγητές, κλπ.), ενώ δε συμπεριλαμβάνει τους διαχειριστές του συστήματος, τους ακαδημαϊκούς οργανισμούς και τους οργανισμούς ανθρωπίνου δυναμικού. Για παράδειγμα, μόνο οι δια βίου μαθητές μπορούν να εγγραφούν σε ένα μάθημα ή να κάνουν αίτηση για κάποια θέση εργασίας.
- only recruiters and profile owners: Ο συγκεκριμένος decorator επιτρέπει την εκτέλεση ενός API με την προϋπόθεση ότι ένας χρήστης είναι είτε recruiter είτε ο ιδιοκτήτης της ζητούμενης πληροφορίας. Ένα παράδειγμα χρήσης του συγκεκριμένου decorator αποτελεί η απαίτηση μόνο οι υπεύθυνοι εργατικού δυναμικού (recruiters) που έχουν δημοσιεύσει μια θέση εργασίας να έχουν πρόσβαση στο βιογραφικό ενός χρήστη που έκανε αίτηση για τη συγκεκριμένη θέση εργασίας.
- only recruiters and recruitment organizations: Ο συγκεκριμένος decorator επιτρέπει την εκτέλεση APIs υπό την προϋπόθεση ότι ένας χρήστης είναι υπεύθυνος ανθρωπίνου δυναμικού (recruiter) η οργανισμός ανθρωπίνου δυναμικού (recruiting organization). Χρησιμοποιείται μόνο για την δημιουργία αγγελιών θέσεων εργασίας.
- only recruiter creator of job: Ο συγκεκριμένος decorator επιτρέπει την εκτέλεση API αν ο χρήστης είναι υπεύθυνος ανθρωπίνου δυναμικού και παράλληλα ο ιδιοκτήτης της ζητούμενης πληροφορίας. Ένα παράδειγμα χρήσης αποτελεί η εξαγωγή όλων των αιτήσεων για μια συγκεκριμένη θέση εργασίας. Είναι προφανής η ανάγκη να έχει πρόσβαση σε τέτοια πληροφορία μόνο ο υπεύθυνος για αυτή την αγγελία.
- only professors or academic organisations: Ο συγκεκριμένος decorator επιτρέπει την εκτέλεση ενός API μόνο για διδάσκοντες μαθημάτων ή εκπαιδευτικούς οργανισμούς. Παράδειγμα χρήσης αποτελεί η δημιουργία νέου μαθήματος στην πλατφόρμα QualiChain.
- only professor or academic organisation of course: Επιτρέπει την εκτέλεση ενός API μόνο από κάποιον διδάσκοντα μαθημάτων ή εκπαιδευτικό οργανισμό, με την προϋπόθεση ότι τους ανήκει το μάθημα το οποίο ζητούν. Παράδειγμα χρήσης αποτελεί η ενημέρωση ενός μαθήματος, ή η εξαγωγή πληροφοριών για τους φοιτητές που παρακολουθούν ένα μάθημα.
- only authenticated: Επιτρέπει την εκτέλεση ενός API σε κάθε χρήστη της πλατφόρμας QualiChain, αρκεί να είναι εγγεγραμμένος (registered) στην πλατφόρμα. Χρησιμοποιείται για APIs που παρέχουν γενική πληροφορία χωρίς περιορισμούς πρόσβασης για τους χρήστες, όπως για παράδειγμα η λίστα όλων των μαθημάτων της πλατφόρμας QualiChain.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

- only profile owners and recruiters and professors: Επιτρέπει την εκτέλεση ενός API μόνον εφόσον ο χρήστης σκοπεύει να χρησιμοποιήσει πόρους που του ανήκουν, ή ο χρήστης έχει ενεργό συμμετοχή στους ζητούμενους πόρους. Χρησιμοποιείται σε APIs που δίνουν συγκεκριμένη πληροφορία για το χρήστη (π.χ. πληροφορίες προφίλ) και παρέχεται πρόσβαση σε αυτό μόνο σε χρήστες που σχετίζονται με το συγκεκριμένο προφίλ, όπως για παράδειγμα ο διδάσκων ενός μαθήματος μπορεί να βλέπει κάποιες πληροφορίες για τους φοιτητές που παρακολουθούν το μάθημα του.
- only profile owner or professor or academic organisation of course: Επιτρέπει την εκτέλεση ενός API μόνο για διδάσκοντες ή ακαδημαϊκά ιδρύματα που είναι οι ιδιοκτήτες των ζητούμενων πόρων. Χρησιμοποιείται σε APIs για τη δημιουργία και την επισκόπηση ενός χρήστη που σχετίζεται με ένα μάθημα. Συγκεκριμένα ο ιδιοκτήτης (διδάσκων ή ακαδημαϊκός οργανισμός) μπορεί να δει τους συμμετέχοντες στο μάθημα του και να ενημερώσει τη σχέση ενός συμμετέχοντα με το μάθημα (π.χ. ότι ένας φοιτητής ολοκλήρωσε το μάθημα). Με την ίδια λογική και ο φοιτητής θα πρέπει να μπορεί να δει την αντίστοιχη πληροφορία.

#### 4.2.2 Η υπηρεσία DeTrusty

Το DeTrusty είναι μια καινοτόμα τεχνολογία ερωτημάτων που επιτρέπει το συνδυασμό διαφορετικών βάσεων δεδομένων σε ένα ερώτημα. Εκτελεί ερωτήματα SPARQL σε μια αποκεντρωμένη ομοσπονδία τελικών σημείων SPARQL (decentralised federation of SPARQL endpoints). Στο πλαίσιο της πλατφόρμας QualiChain το DeTrusty χρησιμοποιείται για την συγκέντρωση δεδομένων από ετερογενείς πηγές δεδομένων. Όσον αφορά την ιδιωτικότητα όταν εκτελείται κάποιο ερώτημα σε μια ιδιωτική πηγή δεδομένων, το DeTrusty στέλνει ένα επαληθεύσιμο διαπιστευτήριο (verifiable credential) που συμπεριλαμβάνει και αποδεικτικό του δικαιώματος πρόσβασης στα ζητούμενα δεδομένα παράλληλα με το ερώτημα. Η ερωτώμενη πηγή δεδομένων είναι υπεύθυνη είτε να αποδεχθεί το επαληθεύσιμο διαπιστευτήριο και να εκτελέσει το ερώτημα είτε να το απορρίψει.

Στην παρακάτω εικόνα παρουσιάζεται ένα παράδειγμα SPARQL ερωτήματος στο DeTrusty:

```
query="
PREFIX qc: <http://qualichain.eu/vocab/>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
SELECT DISTINCT ?student ?location ?badge1 WHERE {
  ?student qc:own ?property .
  ?student qc:major ?major .
  ?student qc:badge ?badge1 .
  ?property owl:sameAs ?good .
  ?property qc:quiltYear ?year .
  ?good qc:ownerIn ?location .
  ?good qc:badge ?badge2 .
  ?location qc:state ?state .
  ?location qc:country ?country .
}
```

Εικόνα 7 - Παράδειγμα εισόδου για την υπηρεσία DeTrusty.

Όσον αφορά την επικοινωνία των υπόλοιπων υπηρεσιών με το DeTrusty, αυτή πραγματοποιείται με τη χρήση APIs. Συγκεκριμένα, η κλήση στο API γίνεται μέσω αιτήματος POST στο DeTrusty και περιλαμβάνει

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

το SPARQL ερώτημα προς εκτέλεση και το επαληθεύσιμο διαπιστευτήριο (προαιρετικά) που δείχνει ότι ο χρήστης έχει όντως δικαίωμα ανάγνωσης στις ζητούμενες πηγές δεδομένων.

```
{
  "cardinality": 3,
  "execution_time": 0.2043547869142602,
  "output_version": "2.0",
  "head": {"vars": ["student", "location", "badge1"]},
  "results": {
    "bindings": [
      {
        "student": {"type": "uri", "value": "http://qualichain.eu/vocab/stud1"},
        "location": {"type": "uri", "value": "http://qualichain.eu/vocab/53Bonn"},
        "badge1": {"type": "literal", "value": "Magic 101"},
        "__meta__": {"is_verified": True }
      },
      {
        "student": {"type": "uri", "value": "http://qualichain.eu/vocab/stud4"},
        "location": {"type": "uri", "value": "http://qualichain.eu/vocab/80Ath"},
        "badge1": {"type": "literal", "value": "Magic 104"},
        "__meta__": {"is_verified": True }
      },
      {
        "student": {"type": "uri", "value": "http://qualichain.eu/vocab/stud5"},
        "location": {"type": "uri", "value": "http://qualichain.eu/vocab/90Lisb"},
        "badge1": {"type": "literal", "value": "Magic 105"},
        "__meta__": {"is_verified": True }
      }
    ]
  }
}
```

Εικόνα 8 - Παράδειγμα εξόδου της υπηρεσίας DeTrusty

Η παραπάνω εικόνα δείχνει ένα παράδειγμα απάντησης από το API του DeTrusty:

Όπως φαίνεται και στην παραπάνω εικόνα, το DeTrusty επιστρέφει ένα JSON λεξικό (dictionary) με τα ακόλουθα κλειδιά: "cardinality" (ακέραιος αριθμός) που αποτελεί τον αριθμό των αποτελεσμάτων που εξήχθησαν, "execution\_time" (πραγματικός αριθμός) που δείχνει το χρόνο σε δευτερόλεπτα που χρειάστηκε για την εκτέλεση του ερωτήματος και τη συλλογή των αποτελεσμάτων, "output\_version" (κείμενο) που δείχνει την έκδοση της παρεχόμενης εξόδου. Το συγκεκριμένο πεδίο χρησιμοποιείται για να διαφοροποιήσει την παρούσα έξοδο με τις πιθανές παλαιότερες εκδόσεις του ίδιου ερωτήματος. Το "head" περιλαμβάνει κάποια μεταδεδομένα όπως είναι οι μεταβλητές που ζητούνται σε ένα ερώτημα. Τέλος τα "results" περιλαμβάνουν τα πεδία "bindings" ως μια λίστα από λεξικά που επιστρέφουν το τελικό αποτέλεσμα.

### 4.2.3 Υπηρεσία επαλήθευσης OpenBadges με βάση το blockchain

Η συγκεκριμένη υπηρεσία υλοποιεί τη χορήγηση και την επαλήθευση των OpenBadges με τη χρήση του blockchain της πλατφόρμας QualiChain, για τη δημιουργία και το διαμοιρασμό εκπαιδευτικών μικροδιαπιστευτηρίων (micro credentials) που είναι επαληθεύσιμα με αποκεντρωμένο και αξιόπιστο

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

τρόπο. Η υλοποίηση βασίζεται στην προσέγγιση Linkchains MerQL [7] η οποία υποστηρίζει την αποθήκευση των badges στο blockchain, όσο και στην επαλήθευση τους μέσω αυτού, και βασίζεται στα διασυνδεδεμένα δεδομένα.

Στους παρακάτω πίνακες φαίνονται παραδείγματα εισόδου και εξόδου για τη δημιουργία και την επιδότηση ενός smart badge.

Πίνακας 1 - Παράδειγμα εισόδου για τη δημιουργία ενός smart badge

```
{
  "title": "a title",
  "description": "some description",
  "issuer": {
    "issuename": "name of the issuer",
    "issuerurl": "url of the issuer",
    "issueremail": "email of the issuer",
    "issuertelephone": "telephone of the issuer",
    "issuerdescription": "description of the issuer",
    "issuerimageurl": "URL pointing to an image file for the Issuer"
  },
  "skills": "description of the skills achieved",
  "version": "version of the badge",
  "imageurl": "URL to point to the badge image"
}
```

Πίνακας 2 - Παράδειγμα εξόδου από την υπηρεσία δημιουργίας smart badge

```
{
  "@context": [
    "https://w3id.org/openbadges/v2",
    {
      "@vocab": "https://blockchain.open.ac.uk/vocab/"
    }
  ],
}
```

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
"type": "BadgeClass",
"name": "a title",
"description": "some description",
"image": "imageurl",
"version": "version",
"issuer": {
  "type": "Issuer",
  "name": "issuename",
  "description": "issuerdescription",
  "url": "issuerurl",
  "email": "issueremail",
  "telephone": "",
  "image": "issuerimageurl"
},
"criteria": {
  "type": "Criteria",
  "narrative": "The holder of this badge has achieved great things in QualiChain",
  "skills": "skills"
},
"tags": [
  "decentralisation",
  "blockchain"
]
}
```

Πίνακας 3 - Παράδειγμα εισόδου της υπηρεσίας για επιδότηση smart badge

```
{
  "badge": {
    "@context": [
      "https://w3id.org/openbadges/v2",
```

```
{
  "@vocab": "https://blockchain.open.ac.uk/vocab/"
}
],
"type": "BadgeClass",
"name": "a title",
"description": "some description",
"image": "imageurl",
"version": "version",
"issuer": {
  "type": "Issuer",
  "name": "issuename",
  "description": "issuerdescription",
  "url": "issuerurl",
  "email": "issueremail",
  "telephone": "",
  "image": "issuerimageurl"
},
"criteria": {
  "type": "Criteria",
  "narrative": "The holder of this badge has achieved great things in QualiChain",
  "skills": "skills"
},
"tags": [
  "decentralisation",
  "blockchain"
]
},
"recipient": {
  "recipientname": "name of recipient",
```

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
"recipientemail": "email address of the recipient"  
}  
}
```

Πίνακας 4 - Παράδειγμα εξόδου της υπηρεσίας επιδότησης smart badge

```
{  
  "@context": [  
    "https://w3id.org/openbadges/v2",  
    {  
      "@vocab": "https://blockchain.open.ac.uk/vocab/"  
    }  
  ],  
  "type": "Assertion",  
  "recipient": {  
    "type": "email",  
    "hashed": true,  
    "salt": "queueseal",  
    "identity": "sha256$7bec02ebbbae659711e5a093947882220dbc1f857092862e81f1c813db4f55e4",  
    "name": "name of the recipient"  
  },  
  "issuedOn": "2020-12-11T12:19:00Z",  
  "badge": {  
    "type": "BadgeClass",  
    "name": "a title",  
    "description": "some description",  
    "image": "imageurl",  
    "version": "version",  
    "issuer": {  
      "type": "Issuer",  
      "name": "issuename",  
      "description": "issuerdescription",
```



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
"url": "issuerurl",
"email": "issueremail",
"telephone": "",
"image": "issuerimageurl"
},
"criteria": {
  "type": "Criteria",
  "narrative": "The holder of this badge has achieved great things in QualiChain",
  "skills": "skills"
},
"tags": [
  "decentralisation",
  "blockchain"
]
},
"verification": {
  "type": "MerQLVerification2020"
},
"signature": {
  "type": "ETHMerQL",
  "address": "7bec02ebbbae659711e5a093947882220dbc1f857092862e81f1c813db4f55e4",
  "issuerAddress": "7bec02ebbbae659711e5a093947882220dbc1f857092862e81f1c813db4f55e4"
}
}
```

Όσον αφορά την επαλήθευση, η μέθοδος είναι πολύ απλή, καθώς δέχεται την περιγραφή ενός smart badge, σε μορφή json, όπως η παραπάνω και επιστρέφει true εάν είναι έγκυρο (ταιριάζει κρυπτογραφικά με την εγγραφή στο blockchain).

Γενικά, η συγκεκριμένη υπηρεσία επιτρέπει τη χορήγηση και τη δημιουργία smart badges από όλους τους χρήστες της πλατφόρμας, καθώς είναι σημαντικό τα badges να μπορούν να χορηγηθούν και από τους φοιτητές και όχι μόνο από το εκπαιδευτικό ίδρυμα.

#### 4.2.4 Υπηρεσία επαλήθευσης ακαδημαϊκών τίτλων με βάση το blockchain

Η συγκεκριμένη ψηφίδα λογισμικού (software component) αποτελεί μια πλατφόρμα για την επαλήθευση ακαδημαϊκών πιστοποιήσεων και διπλωμάτων με τη χρήση της τεχνολογίας blockchain. Αντίθετα με την υπηρεσία επαλήθευσης προσόντων της προηγούμενης ενότητας με τη χρήση OpenBadges, η συγκεκριμένη πλατφόρμα ασχολείται μόνο με έγγραφα πιστοποίησης τα οποία παράγονται από ανώτατα εκπαιδευτικά ιδρύματα και όχι από την πλατφόρμα. Για παράδειγμα, τέτοια αρχεία αποτελούν διπλώματα φοιτητών σε μορφή PDF.

Σε τεχνικό επίπεδο, η πλατφόρμα επαλήθευσης επίσημων ακαδημαϊκών προσόντων επιτρέπει σε μια σειρά από ανώτατα εκπαιδευτικά ιδρύματα να δημιουργήσουν ένα κρυπτογραφικό κλειδί για κάθε πιστοποίηση που εκδίδουν στο blockchain, ώστε οργανισμοί προσλήψεων και ενδιαφερόμενες εταιρείες να ελέγξουν την αυθεντικότητα και την ακεραιότητα των πιστοποιήσεων που δηλώνει ότι κατέχει ένας υποψήφιος για κάποια θέση εργασίας. Αξίζει να σημειωθεί ότι οι πιστοποιήσεις που ελέγχει η συγκεκριμένη πλατφόρμα είναι απλά αρχεία όπως τα εκδίδουν τα ανώτατα εκπαιδευτικά ιδρύματα (π.χ. PDF) και όχι αρχεία δομημένων δεδομένων, και η αυθεντικότητά τους ελέγχεται με βάση το κρυπτογραφικό κλειδί.

Η υπηρεσία αυτή αποτελείται από τις παρακάτω 3 υπηρεσίες (services) λογισμικού:

- **Η υπηρεσία για τα ανώτατα εκπαιδευτικά ιδρύματα (Higher Education Service):** Η συγκεκριμένη υπηρεσία εκτελείται από κάποιο ανώτατο εκπαιδευτικό ίδρυμα και χρησιμοποιείται για την εισαγωγή μιας πιστοποίησης στο blockchain. Συγκεκριμένα, ένα απλό αρχείο πιστοποίησης αποθηκεύεται σε έναν κοινόχρηστο φάκελο, και παράλληλα, δημιουργείται ένα κρυπτογραφικό κλειδί για αυτό και εισάγεται στο blockchain.
- **Η υπηρεσία επαλήθευσης (Recruiting Service):** Αποτελεί την υπηρεσία που εκτελείται από οργανισμούς προσλήψεων και άλλες ενδιαφερόμενες εργασίες για την επαλήθευση ότι οι ακαδημαϊκές πιστοποιήσεις των υποψηφίων για μια θέση εργασίας είναι γνήσιες. Πιο συγκεκριμένα, οι χρήστες της υπηρεσίας ανεβάζουν ένα αρχείο πιστοποίησης, και το αναγνωριστικό του αρχείου και στη συνέχεια υπολογίζεται και εξάγεται το κρυπτογραφικό κλειδί για το συγκεκριμένο αρχείο το οποίο έχει εισαχθεί σε ένα δίκτυο Ropsten Ethereum blockchain [8]. Σε περίπτωση που το δοθέν κρυπτογραφικό κλειδί (hash) ταιριάζει με το κάποιο υπάρχον κρυπτογραφικό κλειδί τότε είναι έγκυρο, που σημαίνει ότι το αρχείο είναι το αυθεντικό και δεν έχει τροποποιηθεί. Σε διαφορετική περίπτωση η πιστοποίηση θεωρείται άκυρη.
- **Η υπηρεσία κοινοπραξίας (Consortium service):** Αποτελεί μια εφαρμογή βασισμένη στα έξυπνα συμβόλαια (smart contracts) για την διαχείριση των οργανισμών που χρησιμοποιούν την εφαρμογή για να εκδίδουν και να επαληθεύουν πιστοποιήσεις. Μια κοινοπραξία είναι ένα σύνολο από ανώτατα εκπαιδευτικά ιδρύματα που χρησιμοποιούν την πλατφόρμα για να παρέχουν εγκυρότητα για τις πιστοποιήσεις που εκδίδουν. Τα smart contracts αποθηκεύουν ένα σύνολο από μέλη της κοινοπραξίας, ενώ οι αποφάσεις λαμβάνονται μέσω ψηφισμάτων. Υπάρχουν τρία είδη λειτουργιών προς ψήφιση: 1) η προσθήκη νέου οργανισμού στην

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

κοινοπραξία, 2) η απομάκρυνση οργανισμού από αυτήν και 3) η τροποποίηση του απαιτούμενου αριθμού ψήφων ώστε να ληφθεί μια απόφαση (με προκαθορισμένο αριθμό τις 2 ψήφους). Η εφαρμογή παρέχει μια διεπαφή που επιτρέπει την ψήφιση για την εισαγωγή νέων μελών στην κοινοπραξία, την απομάκρυνση κάποιου οργανισμού και την τροποποίηση του απαιτούμενου αριθμού ψήφων για τη λήψη κάποιας απόφασης.

#### **4.2.5 Το υποσύστημα εξαγωγής γνώσης (Knowledge Extraction)**

Το συγκεκριμένο υποσύστημα αποτελείται από μια σειρά από υπηρεσίες που επιτελούν διαφορετικούς ρόλους όσον αφορά την εξαγωγή γνώσης. Στη συνέχεια παρουσιάζονται οι υπηρεσίες JobWatch, JobCrawler, DOB-IE και FALCON που αποτελούν τις υπηρεσίες του συγκεκριμένου υποσυστήματος.

##### **4.2.5.1 Η υπηρεσία JobWatch**

Η υπηρεσία JobWatch είναι υπεύθυνη για την εξαγωγή δεδομένων από επιλεγμένους ιστοτόπους εύρεσης εργασίας και τον καθαρισμό των δεδομένων αυτών. Συγκεκριμένα, τα δεδομένα εξάγονται από τους ιστοτόπους JOOBLE, INDEED και ADZUNA με τη χρήση APIs. Στους παρακάτω πίνακες παρουσιάζεται ένα παράδειγμα εισόδου για τον τομέα της τεχνολογίας πληροφοριών (IT sector) από τον ιστοτόπο ADZUNA και τα αποτελέσματα της υπηρεσίας JobWatch:

**Πίνακας 5 - Παράδειγμα εισόδου της υπηρεσίας JobWatch**

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
{ "__CLASS__": "Adzuna::API::Response::JobSecurrarchResults",
  "results": [ {
    "salary_min": 55000, "location": {
      "__CLASS__": "Adzuna::API::Response::Location",
      "area": [ "UK", "London", "Central London", "The City" ],
      "display_name": "The City, Central London"
    },
    "salary_is_predicted": 0,
    "description": "... We are looking for a full-time software developer in London (or in Europe working remotely) to help us build a world-class web platform ...",
    "__CLASS__": "Adzuna::API::Response::Job", "created": "2021-11-22T18:06:22Z",
    "redirect_url": "http://www.adzuna.co.uk/jobs/land/ad/2041433515?se=Si2rr6Na7BGt29hQhSwIAw&v=7A53584F6A00B4DDE1812D83D79D625A795FD6A7",
    "contract_time": "full_time", "title": "Software Developer (Backend)", "category": {
      "__CLASS__": "Adzuna::API::Response::Category", "label": "IT Jobs", "tag": "it-jobs"
    },
    "id": "2041433515", "salary_max": 55000, "company": {
      "__CLASS__": "Adzuna::API::Response::Company", "display_name": "Programmai"
    },
    "contract_type": "permanent"
  }, ] }
```

Πίνακας 6 - Παράδειγμα εξόδου της υπηρεσίας JobWatch

```
{
  "jobPost": [
    {
      "title": "Software Developer (Backend)",
      "id": "2041433515",
      "location": "United Kingdom",
      "company": "Programmai",
      "description": "Role overview We are looking for a full-time software developer in London (or in Europe working remotely) to help us build a world-class web platform. Our Engineering team consists of an ex-Googler, a DeepMind scholar, and other really talented and friendly people. We work in a collaborative atmosphere, and value autonomy highly. About us At Programmai, we use Machine Learning to empower e-commerce companies to market online effectively using their own data. We're a London based early-phase startup that just closed our seed round. We are building an easy-to-use platform to unleash the power of Machine Learning to help our clients with digital marketing. Tech we use at present: Google Cloud Platform, managed with Terraform; Elixir for APIs, storage, and external integration; Python for the Machine Learning pipelines; React for the first version of our frontend (we are assessing ReasonML and Elm for the new version, stay tuned!) Your responsibilities: Build APIs to serve our user-facing web applications. Model the business logic in relational databases. Integrate with external APIs (e.g. Facebook, Google). Help to evolve our system, improve automation, and make important technical decisions with us. Review code and share knowledge. Ideal candidates are expected to have: Experience of writing readable, testable, and maintainable code. Solid programming skills in at least one backend language (Elixir, C++, Python, Ruby etc). Strong understanding in API design (RESTful or GraphQL). Good understanding in SQL.
```

```
Willingness to learn new languages / technologies. Bonus points if: You have experience in any functional languages (particularly Elixir). You have experience in Google Cloud Platform, Terraform, and CI/CD tools. You are interested in gaining full-stack skills. Why Programmai? You can work with highly talented, friendly and passionate people, and have some fun on the way. There are no layers of bureaucracy to work through. You can have a huge impact here. You'll have a front-row seat at an early-stage, fast-growing company that uses machine learning to disturb an existing problem area. What's in it for you? Equity in a fast-growing company. Competitive salary plus additional performance-related rewards 25 days' annual leave plus bank holidays. Well-located office with our adorable dog Billie (if you are in London). For now, everyone is fully remote, but we will return to our office in the future! Location London is the capital of the United Kingdom, situated in the South East of the island. The world's largest financial centre, London is also often referred to the culture capital of the world. The city is well-known for its heritage, including Buckingham Palace, the Tower of London, Kew Gardens, Westminster Abbey, and the historic settlement of Greenwich. London Underground is the oldest urban railway network in the world. Technology workers in London earn 60pc more than the city average, with the figure even higher for foreign born workers, who earned 28pc more than their British counterparts in tech jobs. London can be considered as a center for professional relocation. Furthermore, there are over 100 of Europe's 500 largest companies' headquarters in London. The city is famous for unpredictable weather. For instance, it could be cold, warm and rainy in one day and it's often the topic of conversation across the city. ",
  "salary": " ",
  "type": " Full-time ",
  "link": "
https://www.adzuna.co.uk/jobs/land/ad/2041433515?se=Si2rr6Na7BGt29hQhSwIAw&v=7A53584F6A00B4DDE1812D83D79D625A795FD6A7",
  "date": " 2021-11-22T18:06:22.6233124+03:00 "
}
]
}
```

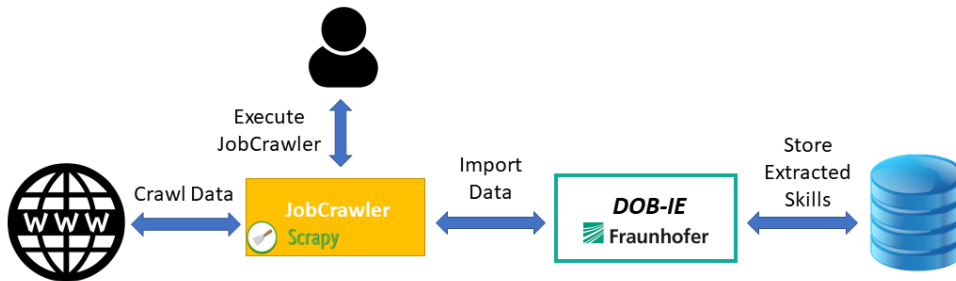
#### 4.2.5.2 Η υπηρεσία JobCrawler

Αντίστοιχη λειτουργικότητα με την υπηρεσία JobWatch παρέχει και η υπηρεσία JobCrawler, με τη διαφορά ότι χρησιμοποιεί τη μέθοδο της απόξεσης δεδομένων από το διαδίκτυο (web scraping) για την συγκέντρωση δεδομένων σχετικά με την αγορά εργασίας από αντίστοιχους ιστοτόπους εύρεσης εργασίας. Η ανάγκη για αυτή την υπηρεσία προέκυψε από το γεγονός ότι υπήρχε η ανάγκη για την απόκτηση σημαντικού όγκου δεδομένων για την ελληνική αγορά εργασίας και δεν υπήρχαν ανοιχτά APIs σε ελληνικούς ιστοτόπους εύρεσης εργασίας. Έτσι, η υπηρεσία JobCrawler αντλεί δεδομένα από τους ελληνικούς ιστοτόπους εύρεσης εργασίας Kariera, Skywalker και Indeed για την ελληνική αγορά εργασίας.

Η αρχιτεκτονική του JobCrawler, όπως φαίνεται και στην παρακάτω εικόνα, είναι όμοια με την υπηρεσία JobWatch, περιλαμβάνει τη φάση της εξαγωγής αγγελιών εργασίας, τη φάση της αποθήκευσης αυτών, καθώς επίσης και τη φάση της εξαγωγής των απαραίτητων γνώσεων και δεξιοτήτων από την κάθε αγγελία με τη χρήση της υπηρεσίας DOB-IE που παρουσιάζεται στην επόμενη ενότητα. Όπως ήδη αναφέρθηκε, σε αντίθεση με την υπηρεσία JobWatch, η εξαγωγή των αγγελιών εργασίας γίνεται με τη χρήση τεχνικών απόξεσης δεδομένων (web scraping) και όχι με τη χρήση APIs. Συγκεκριμένα για την απόξεση δεδομένων χρησιμοποιήθηκε η βιβλιοθήκη Scrapy [9] της Python. Σε αυτό το πλαίσιο, κανονικές εκφράσεις (regular expressions) και μονοπάτια XML (XPath) χρησιμοποιούνται για τον εντοπισμό σχετικής πληροφορίας σε κάθε αγγελία εργασίας. Τελικά, όλες οι αγγελίες εργασίας και οι πληροφορίες που εξάχθηκαν αποθηκεύονται σε μια σχεσιακή βάση δεδομένων.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Φυσικά, αξίζει να σημειωθεί ότι σε περίπτωση που η εσωτερική δομή ενός ιστοτόπου μεταβληθεί, θα πρέπει να μεταβληθεί αντίστοιχα και η υπηρεσία JobCrawler ώστε να εξακολουθήσει να διαβάζει σωστά τις απαιτούμενες πληροφορίες.



Εικόνα 9 - High level αρχιτεκτονική της υπηρεσίας JobCrawler

#### 4.2.5.3 Η υπηρεσία DOB-IE

Η υπηρεσία DOB-IE (Domain Ontology-Based Annotation component) είναι υπεύθυνη για την αναγνώριση εννοιών, σχέσεων και οντοτήτων και επισημαίνει τα κείμενα εισόδου με στοιχεία από την οντολογία του QualiChain. Πιο συγκεκριμένα, αναγνωρίζει διάφορα χαρακτηριστικά σε μια αγγελία εργασίας, όπως τον τίτλο, την περιοχή, τις απαιτούμενες γνώσεις και δεξιότητες, χρησιμοποιώντας τεχνικές αναλυτικής κειμένου όπως η ονομαστική αναγνώριση οντότητας (Named Entity Recognition - NER) [10] και η επισήμανση μέρους του λόγου (Part of Speech tagging – POS tagging) [11]. Έχει υλοποιηθεί στη γλώσσα προγραμματισμού Java, χρησιμοποιώντας τη βιβλιοθήκη GATE embedded [12], που παρέχει βασικές λειτουργίες αναλυτικής κειμένου όπως ο χωρισμός σε προτάσεις, η αναγνώριση μέρους του λόγου και η ονομαστική αναγνώριση οντότητας. Επιπλέον, χρησιμοποιεί την οντολογία του QualiChain για την αναγνώριση συγκεκριμένων γνώσεων και δεξιοτήτων και άλλων εννοιών στο κείμενο.

Η υπηρεσία DOB-IE δέχεται ως είσοδο κανονικοποιημένα κείμενα από αγγελίες εργασίας, όπως για παράδειγμα τα αποτελέσματα από την υπηρεσία JobWatch ή JobCrawler, ή και ελεύθερο κείμενο (π.χ. ένα βιογραφικό ή την περιγραφή κάποιου μαθήματος) και επιστρέφει τα αποτελέσματα σε μορφή RDF, που περιέχει, μεταξύ άλλων, γνώσεις και δεξιότητες που υπάρχουν στην οντολογία του QualiChain με τις περιγραφές τους, όπως επίσης και οντότητες που δεν βρέθηκαν στην οντολογία. Στους παρακάτω πίνακες παρατίθεται ένα παράδειγμα εισόδου και εξόδου της συγκεκριμένης υπηρεσίας για μια αγγελία εργασίας που έχει εξαχθεί από την υπηρεσία JobWatch για τον τομέα της υγείας στο Ηνωμένο Βασίλειο.

Πίνακας 7 - Παράδειγμα εισόδου στην υπηρεσία DOB-IE

```
{
  "jobPost": [
    {
      "title": " Official Veterinarian (Exports) ",
      "id" : " 5971025917763744836 ",
      "location": " United Kingdom ",
      "company": " HallMark Veterinary & Compliance Services ",
    }
  ]
}
```

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
"description": " HallMark provides a Veterinary Export Health Certification Service to businesses throughout the UK. Veterinarians are essential in ensuring the fulfilment of international obligations, including meeting UK legal and political responsibilities, establishing and maintaining the reputation of the UK in relation to trade controls and certification, disease control and the control of illegal trade and fraud.If you are interested in working as an Official Veterinarian certifying exports, you will be charged with ensuring the requirements of the importing country are satisfied, and that veterinary ethics and the reputation of the veterinary profession are maintained. You will verify the identity, origin and health status of the consignment and provide the reassurance that animal welfare is protected during the export process.To be considered for this role, you must be a vet registered with the RCVS, have a good level in English, a driving license and ideally hold the OCQ(V) - EX (Exports General) and OCQ(V) - PX (Product Exports) qualifications. If you would like to apply for this role or find out more information about the position, please register your details and upload your CV.Go to hallmark.recruiterbox.comVacancy posted more than 2 months ago ",
  "salary": " ",
  "type": " Full-time ",
  "link": " https://jooble.org/jdp/5971025917763744836/Official-Veterinarian-\(Exports\)-United-Kingdom?ckey=state+vet&rgn=55127&pos=74&elckey=7983466754119143160&p=4&aq=632293115856982178&age=2892&relb=100&brelb=100&bscr=3627.747&scr=3627.7470000000003 ",
  "date": " 2020-12-01T22:47:39.7470000 "
}
]
}
```

#### Πίνακας 8 - Παράδειγμα εξόδου της υπηρεσίας DOB-IE

```
@prefix so: <http://schema.org/>.
@prefix saro: <http://w3id.org/saro/>.

Job_Posting_Sample
  a          saro:JobPosting ;
  so:datePosted    "2020-12-01"^^xsd:date ;
  so:hiringOrganisation  :hallMarkVeterinaryAndComplianceServices ;
  so:jobLocation    :United_Kingdom ;
  saro:describes    :Job_Posting_Sample_JobRole .

Job_Posting_Sample_JobRole
  a          saro:JobRoleOrType ;
  rdfs:label    "Official Veterinarian"@en ;
  saro:requiresSkill  saro:animalwelfare, saro:Latin, saro:reporting,
```

```
saro:accounting, saro:excel, saro:communication
saro:statistics, saro:financialstatements,
saro:ethics .

:hallMarkVeterinaryAndComplianceServices
  a saro:Recruiter , so:Organisation .

United_Kingdom a so:Place .

saro:animalwelfare rdfs:label "animal welfare" ;
  saro:frequencyOfMention 3 .

saro:Latin rdfs:label "Latin" ;
  saro:frequencyOfMention 2 .

saro:reporting a saro:Topic ;
  rdfs:label "reporting" ;
  saro:frequencyOfMention 4 .

saro:accounting rdfs:label "accounting" ;
  saro:frequencyOfMention 4 .

saro:excel a saro:Product ;
  rdfs:label "excel" ;
  saro:frequencyOfMention 1 .

saro:communication rdfs:label "communication" ;
  saro:frequencyOfMention 1 .
```



saro:statistics a	saro:Topic ;
rdfs:label	"statistics" ;
saro:frequencyOfMention	2 .
saro:financialstatements	
rdfs:label	"financial statements" ;
saro:frequencyOfMention	1 .
saro:ethics rdfs:label	"ethics" ;
saro:frequencyOfMention	4 .

#### 4.2.5.4 Η υπηρεσία FALCON

Η υπηρεσία FALCON είναι ένα εργαλείο για τη συσχέτιση οντοτήτων και συγκεκριμένα χρησιμοποιείται για τη συσχέτιση οντοτήτων από διαφορετικές οντολογίες και γράφους γνώσης όπως οι UMLS [13], το DBPedia [14], και το Bio2RDF [15]. Η συγκεκριμένη υπηρεσία αναγνωρίζει τις οντότητες απεικονίζοντας στιγμιότυπα μιας λέξης σε ένα σύντομο κείμενο σε αντικείμενα του δοθέντος γράφου γνώσης ή της δοθείσας οντολογίας. Για την αναγνώριση και τη συσχέτιση των οντοτήτων χρησιμοποιεί μια βάση γνώσης και έναν κατάλογο κανόνων. Η βάση γνώσης χρησιμοποιεί διάφορες πηγές όπως το DBPedia, το WikiData [16] και το Wordnet [17]. Η βασικότερη λειτουργία του FALCON είναι η δυνατότητα να χωρίσει ένα σύντομο κείμενο στον ελάχιστο αριθμό αντιπροσωπευτικών οντοτήτων που περιγράφουν με ακρίβεια τις λέξεις του κειμένου. Έτσι, καλείται να λύσει το πρόβλημα βελτιστοποίησης για τη μεγιστοποίηση του αριθμού των λέξεων που συνδέονται με μια οντότητα ελαχιστοποιώντας παράλληλα τον αριθμό αναγνωρισμένων οντοτήτων. Στην πλατφόρμα QualiChain το FALCON χρησιμοποιείται για να εξαγει οντότητες από κάποιο κείμενο (αγγελία εργασίας, βιογραφικό ή περιγραφή μαθήματος) δίνοντας έμφαση σε περιπτώσεις όπου η ίδια οντότητα εκφράζεται με διαφορετικούς τρόπους. Για παράδειγμα, οι λέξεις προγραμματισμός (programming) και ανάπτυξη λογισμικού (software development) έχουν ταυτόσημη έννοια στις αγγελίες εργασίας και θα πρέπει να αναγνωρίζονται ως η ίδια οντότητα.

Στους παρακάτω πίνακες φαίνεται ένα παράδειγμα χρήσης του FALCON. Συγκεκριμένα στον πρώτο πίνακα απεικονίζεται η είσοδος ως μια λίστα από μη αναγνωρισμένες οντότητες που είναι υποψήφιος γνώσεις και δεξιότητες, όπως παρέχονται από την υπηρεσία DOB-IE και στον δεύτερο πίνακα απεικονίζεται το αποτέλεσμα που επιστρέφει το FALCON ως μια λίστα από γνώσεις και δεξιότητες που αναγνωρίστηκαν από την αρχική λίστα συνοδευόμενες από σχετικά URIs στους γράφους γνώσης DBPedia και WikiData. Αξίζει να σημειωθεί ότι τα αποτελέσματα περιλαμβάνουν μόνο οντότητες που αναγνωρίστηκαν ως γνώσεις ή δεξιότητες.

Πίνακας 9 - Παράδειγμα εισόδου της υπηρεσίας FALCON

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
{"ADDITIONAL", "Accenture", "Application", "April", "Associate", "August", "Bot", "Calm", "Chatbot", "Computer", "Different", "EXPERIENCE", "Email", "Hard", "Honest", "Hubli", "INFORMATION", "June", "Mac", "March", "Mathematics", "Networks", "Non", "November", "Operating", "Player", "Polite", "Queries", "Role", "SKILLS", "Situations", "System", "Team", "Technical", "Tolerant", "Training", "Windows", "Woodbine", "Working"}
```

Πίνακας 10 - Έξοδος της υπηρεσίας FALCON

```
{
  "Accenture": ["http://dbpedia.org/resource/Accenture", "http://www.wikidata.org/entity/Q338825"],
  "Application": ["http://dbpedia.org/resource/Application_lifecycle_management", "http://www.wikidata.org/entity/Q621590"],
  "Bot": ["http://dbpedia.org/resource/Internet_bot", "http://www.wikidata.org/entity/Q191865"],
  "Chatbot": ["http://dbpedia.org/resource/Chatbot", "http://www.wikidata.org/entity/Q870780"],
  "Computer": ["http://dbpedia.org/resource/Computer", "http://www.wikidata.org/entity/Q68"],
  "Email": ["http://dbpedia.org/resource/Email", "http://www.wikidata.org/entity/Q9158"],
  "Mac": ["http://dbpedia.org/resource/Macintosh", "http://www.wikidata.org/entity/Q75687"],
  "Mathematics": ["http://dbpedia.org/resource/Mathematics", "http://www.wikidata.org/entity/Q395"],
  "Networks": ["http://dbpedia.org/resource/Computer_network", "http://www.wikidata.org/entity/Q1301371"],
  "Player": ["http://dbpedia.org/resource/Media_player_software", "http://www.wikidata.org/entity/Q210337"],
  "Queries": ["http://dbpedia.org/resource/Query_language", "http://www.wikidata.org/entity/Q845739"],
  "System": ["http://dbpedia.org/resource/System", "http://www.wikidata.org/entity/Q58778"],
  "Technical": ["http://dbpedia.org/resource/Technical_support", "http://www.wikidata.org/entity/Q366580"],
  "Training": ["http://dbpedia.org/resource/Training", "http://www.wikidata.org/entity/Q918385"],
  "Windows": ["http://dbpedia.org/resource/Microsoft_Windows", "http://www.wikidata.org/entity/Q1406"]
}
```

## 4.2.6 Το υποσύστημα ανάλυσης δεδομένων και υποστήριξης αποφάσεων

Το υποσύστημα ανάλυσης δεδομένων και υποστήριξης αποφάσεων (Analytics and Decision Support component) είναι υπεύθυνο για τις υπηρεσίες αναλυτικής δεδομένων και υποστήριξης αποφάσεων που προσφέρει η πλατφόρμα QualiChain και μπορεί να χωριστεί σε υπηρεσίες αναλυτικής στην εκπαίδευση (Education Management Analytics services), σε υπηρεσίες διαχείρισης προσόντων (Competency Management services) και σε υπηρεσίες γενικού σκοπού. Όλες οι υπηρεσίες του συγκεκριμένου component παρουσιάζονται στις παρακάτω ενότητες.

### 4.2.6.1 Οι υπηρεσίες αναλυτικής στην εκπαίδευση

Οι υπηρεσίες της συγκεκριμένης ενότητας περιλαμβάνουν την υπηρεσία σύστασης δεξιοτήτων και μαθημάτων, την υπηρεσία ανασχεδιασμού προγράμματος σπουδών και την υπηρεσία σύγκρισης

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

διπλωμάτων (degree comparison). Οι δύο πρώτες υπηρεσίες βασίζονται στα αποτελέσματα που παρέχουν οι υπηρεσίες JobCrawler, JobWatch και DOB-IE για την εξαγωγή αγγελιών εργασίας από το διαδίκτυο και την εξαγωγή πληροφορίας από αυτές, όσο και από βιογραφικά και περιγραφές μαθημάτων. Πιο συγκεκριμένα, τα αποτελέσματα αυτά αποθηκεύονται σε βάσεις δεδομένων και χρησιμοποιούνται από τις προαναφερθείσες υπηρεσίες για περαιτέρω ανάλυση και την παροχή αξιόπιστων προτάσεων στους χρήστες. Η τελευταία υπηρεσία βασίζεται στη γλωσσική και σημασιολογική ομοιότητα μεταξύ δύο διπλωμάτων.

### **Η υπηρεσία προτάσεων μαθημάτων και δεξιοτήτων**

Η υπηρεσία πρότασης μαθημάτων και δεξιοτήτων παρέχει προσωποποιημένες προτάσεις γνώσεων και δεξιοτήτων και μαθημάτων σε φοιτητές και δια βίου μαθητές. Οι προτάσεις αυτές αφενός συμβαδίζουν με τις προσωπικές τους προτιμήσεις και αφετέρου με τις ανάγκες της αγοράς εργασίας. Η υπηρεσία αυτή χρησιμοποιεί ως είσοδο το βιογραφικό του χρήστη, τη γνώση που εξήχθη για τις ανάγκες της αγοράς εργασίας μέσω των υπηρεσιών DOB-IE, JobWatch και JobCrawler, καθώς επίσης και δεδομένα για μαθήματα και προγράμματα σπουδών.

Η συγκεκριμένη υπηρεσία θα αναλυθεί σε βάθος στο κεφάλαιο 5, τόσο ως προς τη μεθοδολογία που ακολουθεί, όσο και ως προς τα αποτελέσματα της.

### **Η υπηρεσία ανασχεδιασμού προγράμματος σπουδών (Curriculum Designer)**

Η υπηρεσία ανασχεδιασμού προγράμματος σπουδών (Curriculum Designer service) παρέχει προτάσεις για τη βελτίωση ενός προγράμματος σπουδών με βάση τις ανάγκες της αγοράς εργασίας. Συγκεκριμένα, αναγνωρίζει γνώσεις και δεξιότητες που λείπουν από το δοθέν πρόγραμμα σπουδών και προτείνει την ένταξη τους σε συγκεκριμένα σχετικά μαθήματα. Η συγκεκριμένη υπηρεσία παίρνει ως είσοδο τη γνώση που εξήχθη για την αγορά εργασίας καθώς επίσης και ένα πρόγραμμα σπουδών που περιλαμβάνει τα μαθήματα και τις περιγραφές τους, αναγνωρίζει γνώσεις και δεξιότητες που είναι σε μεγάλη ζήτηση από την αγορά εργασίας αλλά δεν προσφέρονται από το δοθέν πρόγραμμα σπουδών και χρησιμοποιεί τη μέθοδο της εξαγωγής κανόνων συσχετίσεων (association rule mining) [18] για την εύρεση σχετικών μαθημάτων για κάθε μια από τις γνώσεις και δεξιότητες προς ένταξη στο δοθέν πρόγραμμα σπουδών.

Η συγκεκριμένη υπηρεσία θα αναλυθεί σε βάθος στο κεφάλαιο 6, τόσο ως προς τη μεθοδολογία που ακολουθεί, όσο και ως προς τα αποτελέσματα της.

### **Η υπηρεσία σύγκρισης διπλωμάτων (Degree Comparison service)**

Η υπηρεσία σύγκρισης διπλωμάτων έχει ως στόχο την αναγνώριση της ισοτιμίας μεταξύ διαφορετικών διπλωμάτων που έχουν εκδοθεί από διαφορετικούς ακαδημαϊκούς οργανισμούς. Έτσι, δέχεται ως είσοδο δύο διπλώματα από διαφορετικούς χρήστες και υπολογίζει την ομοιότητα μεταξύ τους. Αυτό επιτυγχάνεται με την ανάλυση της περιγραφής των δοθέντων διπλωμάτων. Η συγκεκριμένη υπηρεσία χρησιμοποιεί ένα προκαθορισμένο json σχήμα που υποδεικνύει τη μορφή των διπλωμάτων προς σύγκριση, και τελικά υπολογίζει τη σημασιολογική ομοιότητα των περιγραφών των διπλωμάτων χρησιμοποιώντας το γλωσσικό μοντέλο BERT (BERT language model) [19]. Στους παρακάτω πίνακες παρουσιάζεται ένα παράδειγμα χρήσης της υπηρεσίας.

[Πίνακας 11 - Παράδειγμα εισόδου \(περιγραφές για δύο διαφορετικά διπλώματα\) στην υπηρεσία Degree Comparison.](#)

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

<pre>{   "properties": {     "name": "Industrial Engineering and Management",     "school": "Instituto Superior Técnico, Universidade de Lisboa",     "ects": 180,     "courses": [       {         "name": "Information Systems and Databases",         "description": "The proposed program is: database modeling techniques. Customer information systems architecture servantdevelopment of information systems in integrative environments students should have minimal programming knowledge. customer information systems architecture. The programming used will be forms and SQL, both involving only elementary programming notions.",         "ects": 6.0       },       {         "name": "Thermodynamics and Transport Phenomena",         "description": "Thermodynamic systems and borders; control volumes; interaction with envelope; concepts of heat and work. heat transmission and mechanisms; Fourier (conduction), Newton (convection) and Stefan-Boltzmann (radiation) laws; heat equation, thermal conductivity and thermal diversity; initial and border conditions; Combined modes of heat transmission. Mass conservation; energy conservation and 1st law of thermodynamics; entropy and 2nd law of thermodynamics; thermodynamics properties and models; Carnot machine, thermodynamic yield, inverted Carnot cycle. Energy conversion machines: ideal Rankine cycle and real steam cycles; Brayton cycle and real gas turbine cycles; air-cooling cycles for absorption and compression of steam; air-cooling cycles; Heat pump cycles, non-reactive mixtures of ideal gases and Psychrometry.",         "ects": 6.0       },       {         "name": "Software Engineering",         "description": "Introduction to Software Engineering. Software design. Program writing. Development process. Engineering requirements. Project Management. Verification and validation.",         "ects": 7.5       },       {         "name": "Elements of Electrothechnics",         "description": "Static electric field? Capacity of a condenser, electric power, system of more than two conductors. Stationary chains? Recitation of a piece of conductor, Joule effect, generators, electric circuits, Introduction to electrical network analysis. Stationary magnetic field? Magnetic Circuit law, magnetic properties of materials, circuits with permanent magnets. Induction law? Induction Coefficients, magnetic black. Calculation of magnetic forces? Application to electro-magnet, electrodynamic stresses, binary calculation in rotary systems. Maxwell's laws? Energy interpretation, quasi-stationary regime. Sinusoidal forced circuits? Notion of complex amplitude and symbolic calculation, notion of</pre>	<pre>{   "properties": {     "name": "Computer Science and Engineering - Alameda",     "school": "Instituto Superior Técnico, Universidade de Lisboa",     "ects": 180,     "courses": [       {         "name": "Analysis and Synthesis of Algorithms",         "description": "Introduction to Analysis and Synthesis of algorithms; mathematical fundamentals for algorithm analysis; revision of ordering algorithms: mergesort; heapsort; quicksort; non-comparison-based ordering algorithms; revision of Data Structures: lists; stacks; rows; dispersion tables; binary demand trees; balanced trees; depreciated analysis. Application examples: binomial piles; Introduction to Computational Geometry. Graphical algorithms: elementary algorithms; lower cost comprehensive trees; shorter paths; maximum flows; Introduction to Linear Programming: simplex algorithm; Algorithm synthesis techniques: Dynamic programming; greedy algorithms; algorithms for Maximum pairing; Introduction to complexity: classes P and NP; NP-complete problems; Cook theorem; Study of Some NP-complete problems; approaches algorithms for NP-workmen problems;",         "ects": 6.0       },       {         "name": "Linear Algebra",         "description": "Resolution of linear equations systems. Method of elimination of Gauss. Matrices and vectors. Linear spaces and linear transformations. Linear independence. Bases and dimension. Core and contradiction of a linear transformation. Internal products and standards. applications to linear differential equations. Orthogonal bases and Gram-Schmidt orthogonalization. Orthogonal supplements and projections. Straight equations and plans. Determinants and applications. Own values and vectors. Invariant subspaces. Diagonalization of arrays. Hermitage, anti-Hermitage and unitary transformations. Quadratic shapes.",         "ects": 6.0       },       {         "name": "Software Engineering",         "description": "Introduction to Software Engineering. Software design. Program writing. Development process. Engineering requirements. Project Management. Verification and validation.",         "ects": 7.5       },       {         "name": "Electromagnetism and Optics",         "description": "1. electrostatic field in vacuum. Coulomb's law. Overlap principle. Field and potential wedding. Electric dipole. Gauss law. Capacitor. 2.electrostatic field in matter. Dielectrics. Polarization. Electric energy3.stationary electric current. Density and current intensity. Load continuity equation. Ohm's law. Joule's law. Kirchoff's laws. Rc.4.magnetic field in vacuum circuit. Biot-Savart law. Ampere's law. Lorentz's strength.</pre>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

<p>impedance, active power, reactive and complex. Resonance. Power factor compensation. Single phase transformer? Equivalent schemes, losses. Induction law with moving bodies, principle of electromechanical energy conversion, rotary electric machines. Three-phase systems, generator and load configurations, three-phase power, Rotary Field. Direct current machine? Universal series motor. Asynchronous machine? Equivalent scheme on a permanent basis. Features. Synchronous machine? Equivalent scheme on a permanent basis. Features.",</p> <p>"ects": 6.0</p> <p>}</p> <p>]</p> <p>}</p> <p>}</p>	<p>Magnetic flux. Induction Coefficients. Bobina.5.magnetic field in matter. Magnetization. Diamagnetism, paramagnetism and ferromagnetism. Energy in magnetostatic.6.electromagnetic induction. Faraday's law. Electric motors and generators. Displacement current. Electromagnetic energy. Maxwell's rlc.7.equations circuit. Electromagnetic waves. Monochromatic flat waves. Energy and intensity of electromagnetic waves.8.electromagnetic nature of light. Dispersion, polarization, reflection, interference and diffraction. The boundary of geometric optics and the laws of reflection and refraction. Fresnel equations and Fermat principle.",</p> <p>"ects": 6.0</p> <p>}</p> <p>]</p> <p>}</p> <p>}</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Πίνακας 12 - Παράδειγμα εξόδου της υπηρεσίας Degree Comparison

<p>[</p> <p>[ "Introduction to Software Engineering. Software design. Program writing. Development process. Engineering requirements. Project Management. Verification and validation.", "Introduction to Software Engineering. Software design. Program writing. Development process. Engineering requirements. Project Management. Verification and validation.", <b>0.999999403953552</b>],</p> <p>["Static electric field? Capacity of a condenser, electric power, system of more than two conductors. Stationary chains? Recitation of a piece of conductor, Joule effect, generators, electric circuits, Introduction to electrical network analysis. Stationary magnetic field? Magnetic Circuit law, magnetic properties of materials, circuits with permanent magnets. Induction law? Induction Coefficients, magnetic black. Calculation of magnetic forces? Application to electro-magnet, electrodynamic stresses, binary calculation in rotary systems. Maxwell's laws? Energy interpretation, quasi-stationary regime. Sinusoidal forced circuits? Notion of complex amplitude and symbolic calculation, notion of impedance, active power, reactive and complex. Resonance. Power factor compensation. Single phase transformer? Equivalent schemes, losses. Induction law with moving bodies, principle of electromechanical energy conversion, rotary electric machines. Three-phase systems, generator and load configurations, three-phase power, Rotary Field. Direct current machine? Universal series motor. Asynchronous machine? Equivalent scheme on a permanent basis. Features. Synchronous machine? Equivalent scheme on a permanent basis. Features.", "1. electrostatic field in vacuum. Coulomb's law. Overlap principle. Field and potential wedding. Electric dipole. Gauss law. Capacitor. 2.electrostatic field in matter. Dielectrics. Polarization. Electric energy3.stationary electric current. Density and current intensity. Load continuity equation. Ohm's law. Joule's law. Kirchoff's laws. Rc.4.magnetic field in vacuum circuit. Biot-Savart law. Ampere's law. Lorentz's strength. Magnetic flux. Induction Coefficients. Bobina.5.magnetic field in matter. Magnetization. Diamagnetism, paramagnetism and ferromagnetism. Energy in magnetostatic.6.electromagnetic induction. Faraday's law. Electric motors and generators. Displacement current. Electromagnetic energy. Maxwell's rlc.7.equations circuit. Electromagnetic waves. Monochromatic flat waves. Energy and intensity of electromagnetic waves.8.electromagnetic nature of light. Dispersion, polarization, reflection, interference and diffraction. The boundary of geometric optics and the laws of reflection and refraction. Fresnel equations and Fermat principle.", <b>0.6857552528381348</b>]</p> <p>]</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### 4.2.6.2 Οι υπηρεσίες διαχείρισης προσόντων

##### Η υπηρεσία δημιουργίας προφίλ και προσαρμογής βιογραφικού (Profiling and CV Customisation)

Η υπηρεσία δημιουργίας προφίλ και προσαρμογής βιογραφικού παρέχει τη δυνατότητα περιγραφής βιογραφικών και αγγελιών εργασίας σύμφωνα με το μοντέλο δεδομένων της πλατφόρμας QualiChain ώστε να μπορούν να αποθηκευτούν στο γράφο γνώσης του QualiChain. Με αυτό τον τρόπο, τα βιογραφικά και οι αγγελίες εργασίας μπορούν να συγκριθούν σημασιολογικά για την επιλογή υποψηφίων για συγκεκριμένες σχετικές θέσεις εργασίας, όσο και για τη δημιουργία μονοπατιών

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

καριέρας (career paths) για τους χρήστες. Επιπλέον, η συγκεκριμένη υπηρεσία βοηθάει τους χρήστες να προσαρμόσουν το βιογραφικό τους, προτείνοντας τους γνώσεις και δεξιότητες που θα τους βοηθήσουν να είναι πιο κατάλληλοι για μια συγκεκριμένη θέση εργασίας.

Η υπηρεσία λαμβάνει ως είσοδο ένα βιογραφικό σε μορφή JSON, και παρέχει αποτελέσματα σε μορφή RDF συμβατή με την οντολογία της πλατφόρμας QualiChain. Στους παρακάτω πίνακες παρουσιάζονται παραδείγματα χρήσης της συγκεκριμένης υπηρεσίας:

Πίνακας 13 - Παράδειγμα εισόδου της υπηρεσίας δημιουργίας προφίλ

```
{
  "label": "CVTest",
  "comment": "Test comment",
  "title": "CV Title",
  "personURI": ":101",
  "userID": ":101",
  "description": "It is a test CV for RMQ",
  "reallocationInfo": "Portugal, Almada, Cova da Piedade",
  "otherInfo": "Generic info about the CV",
  "currentJob": "154",
  "targetSector": "IT",
  "skills":
  [
    { "label": "skill name", "comment": "possible skill application", "skillID": "PHP", "skillURI": ":PHP",
      "evalDate": "null(date can be set for future evaluation of skill level)", "acquiredDate": "01/09/2006",
      "skillLastUsedDate": "01/09/2020", "skillLevel": "expert", "progress": "0"
    },
    { "label": "skill name", "comment": "possible skill application", "skillID": "Javascript", "skillURI": ":
      Javascript", "evalDate": "null(date can be set for future evaluation of skill level)", "acquiredDate":
      "01/09/2006", "skillLastUsedDate": "01/09/2020", "skillLevel": "basic", "progress": "0"
    }
  ],
  "education":
  [
    { "label": "Education title", "comment": "Extra comment if necessary", "major": "the name of the
      Major this education object provides", "minor": "the name of the Minor this education provides",
      "degreeType": "Type of degree this education provides", "from": "10/08/09", "to": "10/06/12",
      "organisation": "Institution that provides the education", "description": "Description of the education
      object", "uri": "cv:localServerEducationURI", "id": "localServerEducationURI"
    }
  ],
  "workHistory":
  [
```

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
{  
  "label": "workHistoryCVInsertTest1", "comment": "Extra comment if necessary",  
  "position": "Position at the company", "from": "11/12/10", "to": "01/01/15", "employer": "Name of the  
organisation worked at", "duration": "Aproximately 4 years", "description": "Description of the work  
done at the company", "numSubordinates": "3", "isCurrent": "False, this functions as a boolean but is  
parsed as String", "jobReference": ".:JobURI", "uri": "cv:whworkHistoryCVInsertTest1",  
  "id": "whworkHistoryCVInsertTest1"}  
]  
}
```

Πίνακας 14 - Παράδειγμα εξόδου της υπηρεσίας δημιουργίας προφίλ

```
:id-auto-gen1 rdf:type cv:CV,  
  rdfs:label "CVTest",  
  rdfs:comment "Test comment",  
  cv:cvTitle "CV Title",  
  cv:aboutPerson :101,  
  cv:hasDescription "It is a test CV for RMQ",  
  saro:Sector "IT",  
  cv:hasOtherInfo "Generic info about the CV",  
  qc:currentJob :154,  
  qc:realocationInfo "Portugal, Almada, Cova da Piedade",  
  cv:hasWorkHistory cv:whworkHistoryCVInsertTest1,  
  cv:hasEducation cv:localServerEducationURI,  
  qc:hasSkillRef :id-auto-gen2, :id-auto-gen3,  
  
cv:whworkHistoryCVInsertTest1 rdf:type cv:WorkHistory,  
  rdfs:label "workHistoryCVInsertTest1",  
  rdfs:comment "Extra comment if necessary",  
  cv:jobTitle "Position at the Company",  
  cv:startDate "11/12/10",  
  cv:endDate "01/01/15",  
  qc:duration "Aproximately 4 years",  
  cv:employedIn "Name of the organisation worked at",  
  cv:jobDescription "Description of the work done at the company",  
  cv:numSubordinates "3",  
  cv:isCurrent "False",  
  qc:jobRefURI ".:JobURI"  
  
cv:localServerEducationURI rdf:type cv:Education  
  rdfs:label "Education title",  
  rdfs:comment "Extra comment if necessary",
```

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
cv:eduMajor "the name of the Major this education object provides",
cv:eduMinor "the name of the Minor this education provides",
cv:degreeType "Type of degree this education provides",
cv:eduStartDate "10/08/09",
cv:eduGradDate "10/06/12",
cv:studiedIn Institution that provides the education",
cv:eduDescription "Description of the education object"
```

### **Η υπηρεσία πρόσληψης και εσωτερικής ανακατανομής προσωπικού (Recruitment and Internal Allocation)**

Η υπηρεσία πρόσληψης και εσωτερικής ανακατανομής προσωπικού (recruitment and internal allocation service) βοηθάει τους υπεύθυνους ανθρωπίνου δυναμικού να εντοπίσουν τους καταλληλότερους υποψηφίους για μια θέση εργασίας τόσο σε θέματα πρόσληψης, όσο και εσωτερικής ανακατανομής του προσωπικού για την κάλυψη των αναγκών μιας εταιρείας. Η υπηρεσία αυτή χρησιμοποιεί ένα μηχανισμό σημασιολογικού ταιριάσματος αγγελιών εργασίας με βιογραφικά και κατατάσσει τα τελευταία για κάθε αγγελία εργασίας με βάση την ομοιότητα τους με την εκάστοτε αγγελία. Η μέθοδος σημασιολογικής ομοιότητας ελέγχει τις γνώσεις και δεξιότητες των υποψηφίων και τις απαιτήσεις της αγγελίας εργασίας, τον ορισμό τους στην οντολογία του QualiChain, και τις σχέσεις μεταξύ τους για τον υπολογισμό της σημασιολογικής ομοιότητας.

Η συγκεκριμένη υπηρεσία λαμβάνει ως είσοδο μια αγγελία εργασίας και μια λίστα από βιογραφικά και υπολογίζει τον βαθμό ομοιότητας κάθε βιογραφικού με τη συγκεκριμένη αγγελία. Ένα ενδεικτικό αποτέλεσμα της υπηρεσίας αυτής παρουσιάζεται στον παρακάτω πίνακα:

**Πίνακας 15 - Παράδειγμα εξόδου της υπηρεσίας πρόσληψης και εσωτερικής ανακατανομής προσωπικού**

```
[
{
  "id": "1",
  "cvid": "cv:id1",
  "name": "Noname",
  "role": "Solutions Architect",
  "score": 66
},
{
  "id": "5",
  "cvid": "cv:id710b71ad-bd4a-494e-8d69-9d53ed57a380",
```



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
"name": "Noname",  
"role": "Solutions Architect",  
"score": 33  
}  
]
```

### **Η υπηρεσία συμβουλευτικής καριέρας Career Advisor**

Η υπηρεσία συμβουλευτικής καριέρας αναλύει την εργασιακή εμπειρία και τις γνώσεις και δεξιότητες ενός εργαζομένου και προτείνει μια σειρά από μελλοντικές θέσεις εργασίας για τον συγκεκριμένο εργαζόμενο ως μονοπάτι καριέρας (career path). Συγκεκριμένα, δείχνει διαφορετικά πιθανά μονοπάτια στα οποία μπορεί να καταλήξει ο εργαζόμενος ανάλογα κα με τις αποφάσεις που θα λάβει για την καριέρα του. Τα αποτελέσματα βασίζονται τόσο στο βιογραφικό του χρήστη, όσο και στην πορεία καριέρας άλλων χρηστών με βάση τις γνώσεις και τις δεξιότητες που απέκτησαν κατά τη διάρκεια της καριέρας τους.

Η συγκεκριμένη υπηρεσία λαμβάνει ως είσοδο το βιογραφικό ενός χρήστη με συγκεκριμένη δομή και χρειάζεται πληροφορίες όπως την εργασιακή του εμπειρία, την εκπαίδευση που έχει λάβει καθώς επίσης και τις γνώσεις και δεξιότητες του και επιστρέφει έναν γράφο που παρουσιάζει τις πιθανότητες καριέρας για το συγκεκριμένο χρήστη. Ενδεικτικά αποτελέσματα σε json μορφή παρουσιάζονται στον παρακάτω πίνακα:

**Πίνακας 16 - Παράδειγμα εξόδου της υπηρεσίας συμβουλευτικής καριέρας**

```
[  
  {'from': 'Intern programmer', 'to': 'Back-End Developer', 'value': 10, 'missing_skills': []  
  },  
  {'from': 'Intern programmer', 'to': 'Junior programmer', 'value': 10, 'missing_skills': []  
  },  
  {'from': 'Senior advisory', 'to': 'End of career', 'value': 10, 'missing_skills': ['Final skill for end of  
  career', 'Second skill for end of career'  
  ]  
  },  
  {'from': 'Product Manager', 'to': 'Senior advisory', 'value': 10, 'missing_skills': []  
  },  
  {'from': 'Software Architect', 'to': 'Product Manager', 'value': 10, 'missing_skills': []  
  },  
  {'from': 'Full-Stack Developer', 'to': 'Product Manager', 'value': 10, 'missing_skills': []  
  },  
  {'from': 'Server Service Expert', 'to': 'Software Architect', 'value': 10, 'missing_skills': []  
  },  
  {'from': 'Front-End Developer', 'to': 'Full-Stack Developer', 'value': 10, 'missing_skills': []  
  },  
]
```

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
{'from': 'Senior Developer', 'to': 'Lead Developer', 'value': 10, 'missing_skills': []
},
{'from': 'Back-End Developer', 'to': 'Server Service Expert', 'value': 10, 'missing_skills': []
},
{'from': 'Junior programmer', 'to': 'Front-End Developer', 'value': 10, 'missing_skills': []
},
{'from': 'Junior programmer', 'to': 'Senior Developer', 'value': 10, 'missing_skills': []
}
]
```

#### 4.2.6.3 Πολυκριτηριακή υποστήριξη αποφάσεων

Η υπηρεσία πολυκριτηριακής υποστήριξης αποφάσεων (Multi-Criteria Decision Support Service - MCDSS) είναι μια υπηρεσία γενικού σκοπού που διευκολύνει τη διαδικασία λήψης αποφάσεων, λαμβάνοντας υπόψη και ποσοτικοποιώντας κατάλληλα όλα τα κριτήρια που είναι σημαντικά για έναν χρήστη. Στην πλατφόρμα QualiChain χρησιμοποιείται για την κατάταξη των υποψηφίων για μια θέση εργασίας και τελικά την επιλογή του καταλληλότερου υποψηφίου, όπως επίσης και συνδυαστικά με τις υπηρεσίες πρότασης δεξιοτήτων και μαθημάτων και την υπηρεσία ανασχεδιασμού προγράμματος σπουδών, για την επιλογή κατάλληλων μαθημάτων από φοιτητές και δια βίου μαθητές, ή την επιλογή των κατάλληλων μαθημάτων για την ένταξη μιας νέας δεξιότητας στο πρόγραμμα σπουδών μιας σχολής, λαμβάνοντας υπόψη τόσο τις προτεινόμενες επιλογές, όσο και τα υποκειμενικά κριτήρια για τους χρήστες. Συγκεκριμένα, η υπηρεσία αυτή δίνει τη δυνατότητα χρήσης των πολυκριτηριακών μεθόδων MAUT [20], TOPSIS [21], ELECTRE I [22] και PROMETHEE II [23].

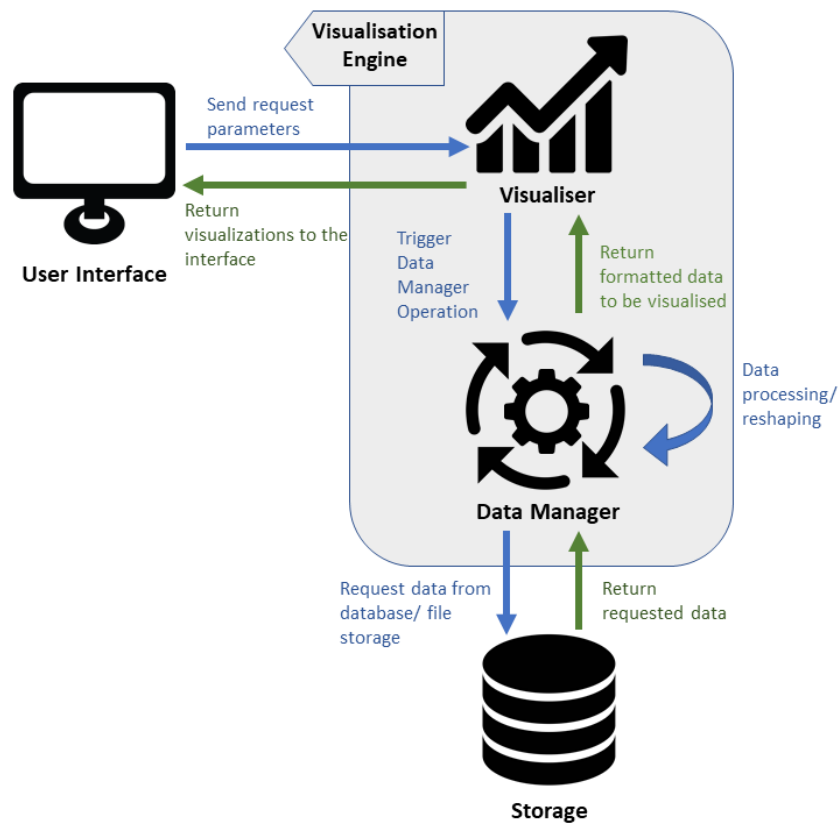
Η υπηρεσία αυτή παρουσιάζεται σε βάθος στο κεφάλαιο 7 και κατά συνέπεια δε θα αναλυθεί παραπάνω στο παρόν κεφάλαιο.

#### 4.2.7 Front end και οπτικοποίηση δεδομένων

##### 4.2.7.1 Η υπηρεσία οπτικοποίησης δεδομένων (Visualisation Engine)

Η υπηρεσία οπτικής απεικόνισης δεδομένων (Visual Analytics service) είναι υπεύθυνη για την οπτική αναπαράσταση των δεδομένων, παρέχοντας γραφήματα από τις διάφορες υπηρεσίες και διεπαφές της πλατφόρμας QualiChain χρησιμοποιώντας δεδομένα από διαφορετικές υπηρεσίες και αναλύσεις. Αποτελείται από δύο ξεχωριστά υποσυστήματα τον Data Manager και τον Visualiser. Ο Visualiser είναι υπεύθυνος για τη δημιουργία των ζητούμενων διαγραμμάτων, ενώ ο Data Manager εκτελεί ερωτήματα στη βάση δεδομένων και φέρνει τα αποτελέσματα σε κατάλληλη μορφή, ώστε να μπορούν να διαβαστούν από τον Visualiser για την δημιουργία του κατάλληλου διαγράμματος. Η τεχνολογία στην οποία βασίστηκε κυρίως η συγκεκριμένη υπηρεσία είναι η βιβλιοθήκη AmCharts 4 [24]. Η λειτουργία της υπηρεσίας παρουσιάζεται συνοπτικά στην παρακάτω εικόνα:

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 10 - Η αρχιτεκτονική της υπηρεσίας οπτικοποίησης των δεδομένων

Η συγκεκριμένη υπηρεσία παρέχει ευρεία γκάμα οπτικών απεικονίσεων για διάφορους τύπους δεδομένων, όπως για παράδειγμα διαφορετικούς τύπους από line charts, και column charts, διαγράμματα radar (radar charts), διαγράμματα πίτας, ιστογράμματα δύο διαστάσεων, διαγράμματα Sankey και άλλα.

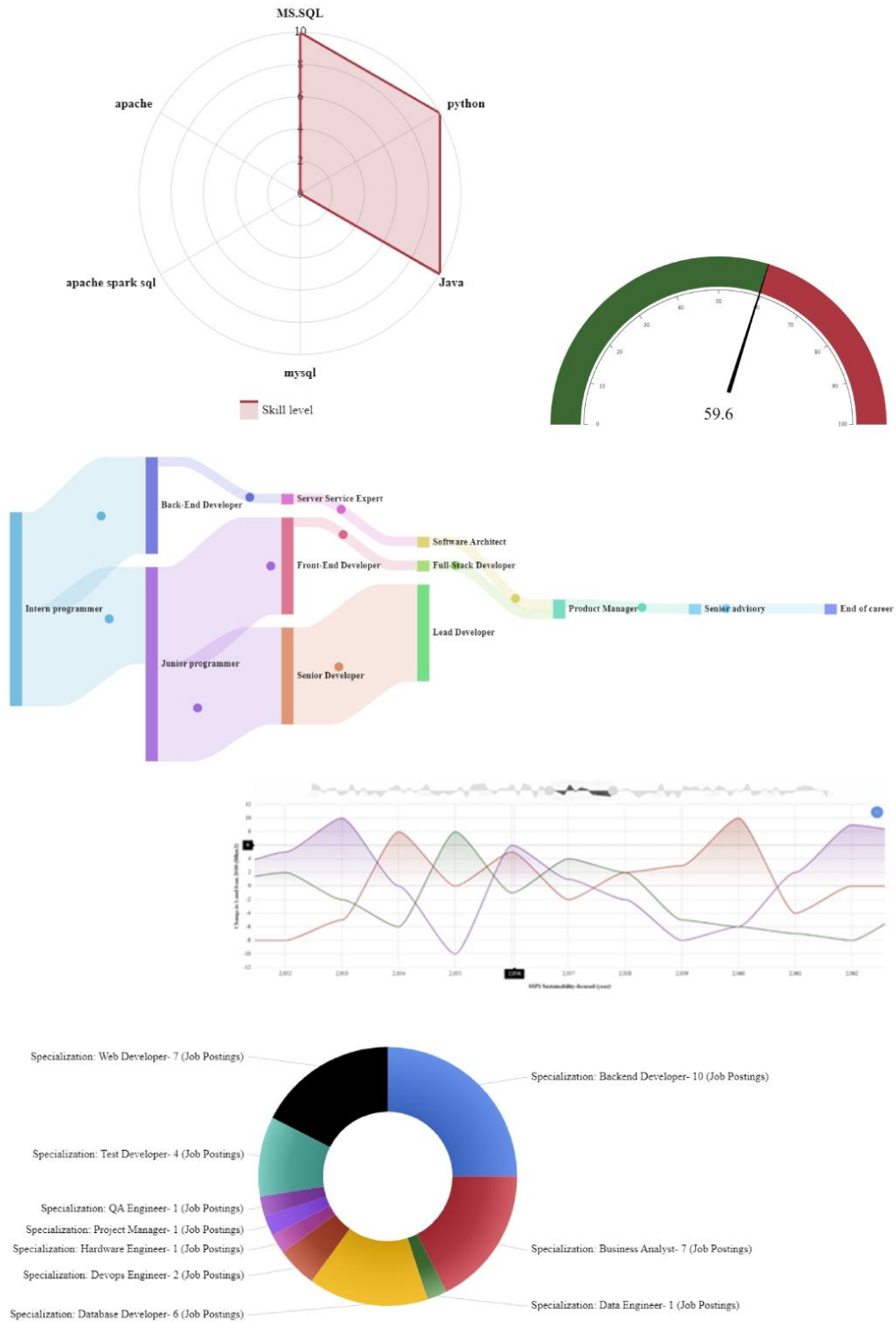
Η υπηρεσία αυτή είναι διαθέσιμη μέσω APIs και δέχεται HTTP GET αιτήματα με συγκεκριμένο μορφότυπο για να παράξει το ζητούμενο διάγραμμα. Ένα παράδειγμα κλήσης της υπηρεσίας παρουσιάζεται στον παρακάτω πίνακα:

Πίνακας 17 - Παράδειγμα εισόδου της υπηρεσίας οπτικοποίησης δεδομένων

```
curl -X GET \
'http://localhost:8000/visualiser/show_line_chart?y_var_names[]=myVariable,myVariable2&
y_var_titles[]=Variable1,Variable2&y_var_units[]=v1_unit,v2_unit&x_axis_type=time&x_axis_name=time&x_axis_title
=Time%20&x_axis_unit=&y_axis_title=YAxisVariable&color_list_request[]=blue,red&use_default_colors=false&min_
max_y_value[]=0,2000&%20x_axis_unit=-' \
```

Επιπλέον μερικά παραδείγματα εξόδου της υπηρεσίας παρουσιάζονται στην παρακάτω εικόνα:

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 11 – Παραδείγματα οπτικοποίησης από την υπηρεσία οπτικοποίησης δεδομένων

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

#### 4.2.7.2 Το Frontend της πλατφόρμας QualiChain

Το Front End της πλατφόρμας QualiChain παρέχει μια διεπαφή για τη χρήση όλων των υπηρεσιών της πλατφόρμας. Φυσικά όπως έχει αναφερθεί όλες οι υπηρεσίες είναι διαθέσιμες μέσω APIs. Τα ίδια APIs καλεί και το Front End για την επικοινωνία με τις προαναφερθείσες υπηρεσίες.

Μια επίδειξη των βασικών λειτουργιών και του Front End της πλατφόρμας QualiChain παρουσιάζεται στο Παράρτημα III.

### 4.3 Σύνοψη και συμπεράσματα

Στο παρόν κεφάλαιο παρουσιάστηκε η πλατφόρμα QualiChain, όσον αφορά τόσο την αρχιτεκτονική της, όσο και τα επιμέρους υποσυστήματα και τις υπηρεσίες καθενός εξ αυτών. Επιπλέον παρουσιάστηκε και ο τρόπος λειτουργίας του συστήματος συμπεριλαμβανομένης και της επικοινωνίας μεταξύ των υποσυστημάτων.

Η πλατφόρμα QualiChain είναι μια αρκετά πολύπλοκη πλατφόρμα που επιτελεί μια σειρά από διαφορετικές και ετερογενείς λειτουργίες εξυπηρετώντας διαφορετικούς τύπους χρηστών, με στόχο τη βελτίωση των διαδικασιών της εκπαίδευσης, όσον αφορά αφενός την αναγνώριση και επιβεβαίωση προσόντων και ακαδημαϊκών τίτλων, και αφετέρου τη βελτίωση των παρεχόμενων μαθημάτων και των προφίλ των φοιτητών και δια βίου μαθητών, με βάση τις ανάγκες της αγοράς εργασίας, αλλά και τις προσωπικές τους προτιμήσεις.

### 4.4 Αναφορές

- [1] “About SEAL | Seal.” <https://project-seal.eu/about> (accessed May 01, 2023).
- [2] “κανονισμός eIDAS | Shaping Europe’s digital future.” <https://digital-strategy.ec.europa.eu/el/policies/eidas-regulation> (accessed May 01, 2023).
- [3] “eduGAIN – enabling worldwide access.” <https://edugain.org/> (accessed May 01, 2023).
- [4] H. Mu and S. Jiang, “Design patterns in software development,” ICSESS 2011 - Proceedings: 2011 IEEE 2nd International Conference on Software Engineering and Service Science, pp. 322–325, 2011, doi: 10.1109/ICSESS.2011.5982228.
- [5] “Welcome to Python.org.” <https://www.python.org/> (accessed May 01, 2023).
- [6] “Flask-RESTful — Flask-RESTful 0.3.8 documentation.” <https://flask-restful.readthedocs.io/en/latest/> (accessed May 01, 2023).
- [7] A. Third and J. Domingue, “LinkChains: Trusted Personal Linked Data,” 2019, Accessed: May 01, 2023. [Online]. Available: <http://www.cs.toronto.edu/~consens/blocksw/papers/paper3.pdf>
- [8] “GitHub - ethereum/ropsten: Ropsten public testnet PoW chain.” <https://github.com/ethereum/ropsten> (accessed May 01, 2023).
- [9] “Scrapy | A Fast and Powerful Scraping and Web Crawling Framework.” <https://scrapy.org/> (accessed Apr. 09, 2021).
- [10] B. Mohit and I. Zitouni, “Named Entity Recognition,” pp. 221–245, 2014, doi: 10.1007/978-3-642-

45358-8\_7.

- [11] C. D. Manning, "Part-of-speech tagging from 97% to 100%: Is it time for some linguistics?," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6608 LNCS, no. PART 1, pp. 171–189, 2011, doi: 10.1007/978-3-642-19400-9\_14/COVER.
- [12] "GitHub - GateNLP/gate-core: The GATE Embedded core API and GATE Developer application." <https://github.com/GateNLP/gate-core> (accessed May 02, 2023).
- [13] O. Bodenreider, "The Unified Medical Language System (UMLS): Integrating biomedical terminology," *Nucleic Acids Res*, vol. 32, no. DATABASE ISS., Jan. 2004, doi: 10.1093/nar/gkh061.
- [14] "Home - DBpedia Association." <https://www.dbpedia.org/> (accessed May 02, 2023).
- [15] "Bio2RDF v2.7a." <https://bio2rdf.org/> (accessed May 02, 2023).
- [16] "Wikidata." [https://www.wikidata.org/wiki/Wikidata:Main\\_Page](https://www.wikidata.org/wiki/Wikidata:Main_Page) (accessed May 02, 2023).
- [17] "WordNet." <https://wordnet.princeton.edu/> (accessed May 02, 2023).
- [18] Q. Zhao and S. B. Sourav, "Association rule mining: A survey," *Nanyang Technological University*, vol. 135. Singapore, 2003.
- [19] I. Tenney, D. Das, and E. Pavlick, "BERT Rediscovered the Classical NLP Pipeline," *ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, pp. 4593–4601, May 2019, doi: 10.18653/v1/p19-1452.
- [20] J. S. Dyer, "Maut — Multiattribute Utility Theory," in *Multiple Criteria Decision Analysis: State of the Art Surveys*, New York: Springer-Verlag. doi: 10.1007/0-387-23081-5\_7.
- [21] Y. Çelikbilek and F. Tüysüz, "An in-depth review of theory of the TOPSIS method: An experimental analysis," *Journal of Management Analytics*, vol. 7, no. 2, Apr. 2020, doi: 10.1080/23270012.2020.1748528.
- [22] J. Figueira, S. Greco, and M. Ehrogott, "Electre Methods," in *Multiple Criteria Decision Analysis: State of the Art Surveys*, New York, NY: Springer New York, 2005. doi: 10.1007/b100605.
- [23] C. A. Bana e Costa, Ed., "The Promethee Methods for MCDM," in *Readings in Multiple Criteria Decision Aid*, Berlin, Heidelberg: Springer Berlin Heidelberg, 1990. doi: 10.1007/978-3-642-75935-2.
- [24] "amCharts 4 Documentation." <https://www.amcharts.com/docs/v4/> (accessed May 02, 2023).

## ΚΕΦΑΛΑΙΟ 5: Η ΥΠΗΡΕΣΙΑ ΠΡΟΤΑΣΕΩΝ ΜΑΘΗΜΑΤΩΝ ΚΑΙ ΔΕΞΙΟΤΗΤΩΝ

### 5.1 Εισαγωγή

Ο 21<sup>ος</sup> αιώνας έχει φέρει μαζί του ραγδαία αύξηση της ψηφιοποίησης σε πολλές επιστημονικές περιοχές καθώς επίσης και σε διάφορες εκφάνσεις της καθημερινότητας των ανθρώπων. Οι τεχνολογίες των επικοινωνιών και της πληροφορίας και η χρήση αναπτυσσόμενων τεχνολογιών όπως οι τεχνολογίες δεδομένων μεγάλης κλίμακας, αναλυτικής δεδομένων, τεχνητής νοημοσύνης και blockchain, μεταξύ άλλων, μπορούν να αξιοποιηθούν για την ανάπτυξη καινοτόμων τεχνικών λύσεων. Η εκπαίδευση είναι μια περιοχή όπου οι παραπάνω τεχνολογίες έχουν εφαρμοστεί και εξακολουθούν να εφαρμόζονται για τη διευκόλυνση των διαχειριστικών διαδικασιών, την καλύτερη ενημέρωση των φοιτητών και δια βίου μαθητών (lifelong learners), τη διαχείριση και ανάλυση των δεδομένων τους κ.ά. Τέτοιες λύσεις εκτός από τη βελτίωση της ποιότητας ζωής των χρηστών τους, έχουν συμβάλει και στη σημαντική αύξηση της διαθεσιμότητας των δεδομένων, όσον αφορά τις επιλογές των δια βίου μαθητών και φοιτητών, τις πληροφορίες για τα μαθήματα και τα προγράμματα σπουδών.

Το ίδιο ισχύει και για την αγορά εργασίας, η οποία παρουσιάζει επίσης σημαντική ανάπτυξη όσον αφορά τις ψηφιακές πλατφόρμες (π.χ. το LinkedIn), και τους ιστοτόπους αναζήτησης και εύρεσης εργασίας. Τόσο οι ψηφιακές πλατφόρμες όσο και οι ιστοτόποι αναζήτησης εργασίας, διαθέτουν ευρεία γκάμα δεδομένων, όσον αφορά το εργατικό δυναμικό, τους εργοδότες, βιογραφικά σημειώματα καθώς επίσης και προσόντα που απαιτούνται από διάφορα επαγγέλματα.

Η αγορά εργασίας είναι άρρηκτα συνδεδεμένη με την εκπαίδευση (και κυρίως την ανώτατη εκπαίδευση), καθώς βασική αποστολή της δεύτερης αποτελεί η ανάπτυξη των κατάλληλων γνώσεων και δεξιοτήτων στους δια βίου μαθητές, τους φοιτητές και τους εργαζομένους για την ομαλή ένταξη και εξέλιξη τους στην αγορά εργασίας.

Παρ' όλα αυτά, η σύνδεση της αγοράς εργασίας με την εκπαίδευση δεν είναι πάντοτε απόλυτα επιτυχημένη, ενώ και οι υπάρχουσες τεχνολογικές λύσεις και εργαλεία που εξυπηρετούν την αγορά εργασίας είναι, ως επί το πλείστον, αποσυνδεδεμένα από τα αντίστοιχα εργαλεία που χρησιμοποιούνται στην εκπαίδευση. Φυσικά, τα δεδομένα για την αγορά εργασίας, όπως για παράδειγμα οι απαιτήσεις για προσόντα σε συγκεκριμένα επαγγέλματα ή εξειδικεύσεις, είναι πολύ σημαντικά για την εκπαίδευση, μιας και αποτελούν τη γνώση και τις δεξιότητες που οι δια βίου μαθητές και οι φοιτητές θα πρέπει να λάβουν για να μεγιστοποιήσουν τις πιθανότητες τους να προσληφθούν σε μια θέση εργασίας που άπτεται των ενδιαφερόντων τους. Επιπλέον, οι απαιτήσεις της αγοράς εργασίας, απεικονίζουν σε σημαντικό βαθμό την πρόοδο της επιστήμης σε ένα συγκεκριμένο επιστημονικό πεδίο.

Από την άλλη πλευρά τα ανώτατα εκπαιδευτικά ιδρύματα και άλλοι εκπαιδευτικοί οργανισμοί πολύ συχνά λειτουργούν απομονωμένα και τα προγράμματα σπουδών τους σχεδιάζονται πολλές φορές με αυστηρά ακαδημαϊκά κριτήρια. Για το λόγο αυτό, καταλήγουν να προσφέρουν γνώσεις και δεξιότητες ανακόλουθες με τις ανάγκες της αγοράς εργασίας. Το φαινόμενο αυτό οδηγεί τους φοιτητές να επιλέγουν και μαθήματα τα οποία δεν θα τους βοηθήσουν στην μετέπειτα επαγγελματική τους πορεία.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Αυτό διότι οι τελευταίοι στερούνται εργαλείων που θα τους βοηθήσουν να δουν τη μεγάλη εικόνα για τις ανάγκες της αγοράς εργασίας. Επιπλέον, είναι σημαντικό, να επιλέξουν μαθήματα τα οποία προσφέρουν γνώσεις και δεξιότητες που σχετίζονται με τα ενδιαφέροντά τους, ώστε σε βάθος χρόνου να τους βοηθήσουν να προοδεύσουν στην αγορά εργασίας με βάση την επαγγελματική πορεία που οραματίζονται.

Παρότι δεν γίνεται να παραβλέψουμε τους ακαδημαϊκούς στόχους των εκπαιδευτικών ιδρυμάτων, είναι προφανές ότι οι απαιτήσεις της αγοράς εργασίας θα έπρεπε να παίζουν σημαντικό ρόλο στη διαμόρφωση του προγράμματος σπουδών ενός πανεπιστημίου ή ενός εκπαιδευτικού οργανισμού, όπως και στην καθοδήγηση των φοιτητών και των δια βίου μαθητών ως προς την απόφαση απόκτησης νέων γνώσεων και δεξιοτήτων. Σε αυτό το πλαίσιο, οι τεχνολογίες αναλυτικής δεδομένων και υποστήριξης αποφάσεων θα μπορούσαν να αξιοποιηθούν όχι μόνο για τη βελτίωση της λειτουργίας των ακαδημαϊκών ιδρυμάτων, αλλά και για τη βελτίωση της καθημερινής ζωής των δια βίου μαθητών, των φοιτητών ή των εργαζομένων, βοηθώντας τους να αξιοποιήσουν τα υπάρχοντα δεδομένα και να πάρουν σωστότερες αποφάσεις, όπως η επιλογή των κατάλληλων μαθημάτων που θα συμπληρώνουν καταλλήλως το βιογραφικό τους και παράλληλα θα είναι περιζήτητες από την αγορά εργασίας. Φυσικά, η διαδικασία αυτή θα ήταν δύσκολο να γενικευτεί για το σύνολο της αγοράς εργασίας, μιας και κάθε τομέας έχει τα δικά του χαρακτηριστικά. Σε αυτό το πλαίσιο, στην παρακάτω ενότητα εμβαθύνουμε στην αγορά εργασίας στο αντικείμενο της τεχνολογίας της πληροφορίας και τα σχετικά με αυτό επαγγέλματα. Στις επόμενες ενότητες παρουσιάζεται το σύστημα συστάσεων δεξιοτήτων και μαθημάτων που αναπτύχθηκε. Το συγκεκριμένο σύστημα συστάσεων έχει σκοπό την ανάλυση δεδομένων μεγάλης κλίμακας σε σχέση με την αγορά εργασίας, τα μαθήματα και τα προγράμματα σπουδών των πανεπιστημίων, όπως επίσης και τα δεδομένα των φοιτητών, ώστε να παρέχει προσωποποιημένες προτάσεις μαθημάτων και δεξιοτήτων που θα τους βοηθήσουν να είναι πιο ανταγωνιστικοί στην αγορά εργασίας, και θα είναι συμβατές με τις ανάγκες και τις προτιμήσεις τους.

## 5.2 Μεθοδολογία

Η μεθοδολογία που ακολουθήθηκε μπορεί να χωριστεί σε 6 διαφορετικές φάσεις. Συγκεκριμένα, αποτελείται από α) τη φάση της εξόρυξης των δεδομένων από το διαδίκτυο, β) την εξαγωγή των απαιτούμενων γνώσεων και δεξιοτήτων από τις αγγελίες εργασίας, γ) την επεξεργασία και αποθήκευση των δεδομένων σε μία βάση δεδομένων επεξεργασίας κειμένου (ElasticSearch [1]), δ) την επεξεργασία με βάση τις εξαχθείσες γνώσεις και δεξιότητες, ε) την ανάλυση και συσταδοποίηση (clustering) των αγγελιών εργασίας και, τέλος, στ) τη φάση της παροχής προτάσεων.

Κατά την πρώτη φάση, αυτή της εξόρυξης δεδομένων από το διαδίκτυο (web mining), εφαρμόστηκε η τεχνική της απόξεσης δεδομένων (web scraping), σε μια σειρά από ελληνικούς ιστοτόπους με αγγελίες εργασίας για την εξαγωγή μεγάλου όγκου αγγελιών. Αυτό επιτυγχάνεται με τη χρήση του εργαλείου JobCrawler, το οποίο περιγράφεται, με λεπτομέρεια, σε επόμενη ενότητα του παρόντος κεφαλαίου.

Στην επόμενη φάση (φάση εξαγωγής απαιτούμενων γνώσεων και δεξιοτήτων), οι πιο δημοφιλείς γνώσεις και δεξιότητες από την οπτική γωνία της αγοράς εργασίας ανιχνεύονται σε κάθε αγγελία εργασίας χρησιμοποιώντας το υποσύστημα εξαγωγής γνώσης του QualiChain (QualiChain Knowledge Extraction



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Component, που παρουσιάστηκε στο προηγούμενο κεφάλαιο) στις αγγελίες που εξήχθησαν. Το συγκεκριμένο εργαλείο περιγράφεται και επιδερμικά στη συνέχεια του παρόντος κεφαλαίου.

Στη συνέχεια, οι αγγελίες εργασίας αποθηκεύονται σε μια βάση δεδομένων επεξεργασίας κειμένου ονόματι Elasticsearch [1], η οποία επιτρέπει την εκτέλεση σύνθετων λειτουργιών αναζήτησης και ταιριάσματος κειμένου (text search and matching operations), με αποδοτικό τρόπο. Στη συνέχεια, για κάθε αγγελία εργασίας, η βασική εξειδίκευση εξάγεται ανάλογα με το κείμενο της, έτσι ώστε οι απαιτούμενες γνώσεις και δεξιότητες της κάθε εξειδίκευσης να προτείνονται σε χρήστες με αντίστοιχη εξειδίκευση σε περίπτωση που δεν τις κατέχουν.

Επόμενο βήμα αποτελεί η φάση της συσταδοποίησης (clustering), κατά το οποίο αλγόριθμοι συσταδοποίησης εφαρμόζονται στις αγγελίες εργασίας προκειμένου να οριστούν οι διάφορες εξειδικεύσεις με έναν τρόπο που βασίζεται στα δεδομένα, και είναι πλήρως αυτοματοποιημένος. Τέλος, οι προτάσεις δεξιοτήτων και μαθημάτων που λαμβάνει ο χρήστης αποτελούν συνδυασμό της σύστασης με τεχνικές συσταδοποίησης, όπου ο χρήστης λαμβάνει προτάσεις για γνώσεις και δεξιότητες από τις τρεις εξειδικεύσεις οι οποίες είναι πιο κοντά στα ενδιαφέροντα του, όσο και της σύστασης με την τεχνική του ταιριάσματος κειμένου, κατά την οποία ο χρήστης λαμβάνει συστάσεις γνώσεων και δεξιοτήτων από την εξειδίκευση στην οποία αναγνωρίζεται από το σύστημα, από το σύνολο των εξειδικεύσεων που ορίστηκαν εξ' αρχής για το σύστημα. Οι προτάσεις γνώσεων και δεξιοτήτων προσφέρονται στο χρήστη με σειρά σημαντικότητας. Όλα τα παραπάνω βήματα περιγράφονται με περισσότερη λεπτομέρεια στις επόμενες ενότητες του παρόντος κεφαλαίου.

### **5.2.1 Εξαγωγή αγγελιών εργασίας μέσω της υπηρεσίας JobCrawler**

Ο σκοπός της υπηρεσίας JobCrawler είναι η εξαγωγή μεγάλου όγκου αγγελιών εργασίας για την Ελληνική αγορά εργασίας, ώστε να χρησιμοποιηθούν από το σύστημα συστάσεων για την παροχή αξιόπιστων προτάσεων δεξιοτήτων και μαθημάτων στους φοιτητές. Η συγκεκριμένη υπηρεσία εφαρμόζει τεχνικές απόξεσης δεδομένων διαδικτύου (web scraping) σε γνωστούς ελληνικούς ιστοτόπους αναζήτησης εργασίας, όπως το Kariera [2], το ελληνικό indeed [3] και το skywalker [4].

Για την απόκτηση των δεδομένων μέσω της υπηρεσίας JobCrawler, χρησιμοποιείται η τεχνική της απόξεσης δεδομένων (web scraping), όπως έχει ήδη αναφερθεί. Πιο συγκεκριμένα, χρησιμοποιήθηκε η βιβλιοθήκη Scrapy [5] της γλώσσας προγραμματισμού Python, για την ανίχνευση των διαθέσιμων αγγελιών εργασίας στους προαναφερθέντες ιστοτόπους αναζήτησης εργασίας. Μέσω της συγκεκριμένης βιβλιοθήκης, κανονικές εκφράσεις και XPath χρησιμοποιούνται για την απόκτηση της σχετικής πληροφορίας και των κατάλληλων περιγραφών για τις αγγελίες εργασίας. Στη συνέχεια, οι εξαχθείσες αγγελίες εργασίας ομαδοποιούνται ανάλογα με το περιεχόμενό τους και αποθηκεύονται ώστε να χρησιμοποιηθούν από άλλες υπηρεσίες για περαιτέρω ανάλυση.

### **5.2.2 Εξαγωγή γνώσης**

Η υπηρεσία QualiChain Knowledge Extraction λαμβάνει ως είσοδο αγγελίες εργασίας που έχουν εξαχθεί από την υπηρεσία JobCrawler και εφαρμόζει σε αυτές τεχνικές επεξεργασίας φυσικής γλώσσας (Natural Language Processing - NLP), όπως Αναγνώριση Ονομαστικής Οντότητας (Named Entity Recognition - NER)

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

[6], αναγνώριση μερών του λόγου (Part of Speech tagging - POS) [7], όπως επίσης και τεχνικές ταιριάσματος και προεπεξεργασίας κείμενου. Στη συνέχεια, μέσω των παραπάνω τεχνικών αναγνωρίζει όλες τις απαιτούμενες γνώσεις και δεξιότητες από το κείμενο των αγγελιών εργασίας και εντοπίζει και διαγράφει τις περιττές εγγραφές (π.χ. διπλή εγγραφή της ίδιας δεξιότητας στην ίδια αγγελία εργασίας). Τέλος, κατηγοριοποιεί τις γνώσεις και δεξιότητες που αναγνωρίστηκαν σε τρεις κατηγορίες (προϊόν, εργαλείο, θεματική περιοχή – Product, Tool and Topic). Η τελευταία λειτουργία είναι πολύ χρήσιμη για το σύστημα συστάσεων, μιας και βοηθάει στην αποφυγή προτάσεων πολύ γενικών γνώσεων και δεξιοτήτων, όπως είναι η ανάπτυξη λογισμικού και δίνει έμφαση στις γνώσεις και δεξιότητες που έχουν πραγματικά μεγάλη ζήτηση, όπως συγκεκριμένες γλώσσες προγραμματισμού και άλλα χρήσιμα λογισμικά.

### 5.2.3 Παροχή προτάσεων με τη χρήση της τεχνολογίας Elasticsearch

Για τη σύσταση μαθημάτων και δεξιοτήτων βασισμένη σε τεχνικές αναλυτικής κειμένου, αναπτύχθηκε η εφαρμογή QualiChain Analyser, η οποία στέλνει αιτήματα στη βάση δεδομένων Elasticsearch. Η συγκεκριμένη εφαρμογή αναπτύχθηκε χρησιμοποιώντας το Flask framework [8] της γλώσσας προγραμματισμού Python, και τις βιβλιοθήκες Elasticsearch-DSL [9] και Pandas [10] της Python, για την παροχή όλων των λειτουργιών που είναι διαθέσιμες για την αναλυτική κειμένου (text analytics). Η εφαρμογή λαμβάνει ως είσοδο αρχεία τύπου CSV και TSV ενώ υποστηρίζει και την εισαγωγή αρχείων κατ' ευθείαν από σχεσιακές ή μη σχεσιακές βάσεις δεδομένων.

Επιπροσθέτως, για την πρόταση μαθημάτων και δεξιοτήτων με βάση τις ανάγκες της αγοράς εργασίας και τις προτιμήσεις του χρήστη, απαιτείται η ύπαρξη πληροφοριών για τα μαθήματα που διδάσκονται και για τα προγράμματα σπουδών του εκάστοτε πανεπιστημίου. Σε αυτό το πλαίσιο, το πρόγραμμα σπουδών με τα μαθήματα και τις περιγραφές τους της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών (Η.Μ.Μ.Υ.) του Εθνικού Μετσόβιου Πολυτεχνείου (Ε.Μ.Π.) εισήχθη στην εφαρμογή QualiChain Analyser η οποία είναι υπεύθυνη για την αποθήκευση της κατάλληλης πληροφορίας στη βάση δεδομένων Elasticsearch. Η συγκεκριμένη τεχνολογία παρέχει σύνθετες λειτουργίες όπως ερωτήματα πολλαπλών αντιστοιχιών (multi-match querying), ασαφή αντιστοίχιση κειμένου (fuzzy text matching) και συνδυασμένα λογικά ερωτήματα (boolean combined queries), τα οποία χρησιμοποιούνται για την εύρεση ομοιοτήτων μεταξύ αγγελιών εργασίας με λέξεις κλειδιά.

Ακόμη, οι αγγελίες εργασίας που εξήχθησαν μέσω της υπηρεσίας JobCrawler αποθηκεύονται στην προαναφερθείσα βάση δεδομένων και στη συνέχεια συνδέονται με τις γνώσεις και δεξιότητες οι οποίες έχουν αντιστοιχιστεί με αυτά μέσω των ερωτημάτων αντιστοιχίας κειμένου που επιτρέπει η τεχνολογία Elasticsearch, για την δημιουργία ομάδων από αγγελίες εργασίας που συνδέονται με συγκεκριμένες εξειδικεύσεις και ρόλους. Οι ομάδες αγγελιών εργασίας που προκύπτουν εισάγονται στην υπηρεσία εξαγωγής γνώσης, η οποία χρησιμοποιώντας τεχνικές επεξεργασίας φυσικής γλώσσας και αναλυτικής κειμένου επιστρέφει μια λίστα με τις πιο σημαντικές γνώσεις και δεξιότητες για κάθε μια ομάδα εξειδίκευσης.

Παρόμοια διαδικασία ακολουθείται και με το βιογραφικό σημείωμα του χρήστη, το οποίο περιλαμβάνει τις αποκτηθείσες γνώσεις και δεξιότητες του χρήστη, την εργασιακή του εμπειρία, μαθήματα που έχει παρακολουθήσει κ.ά.. Όλες αυτές οι πληροφορίες αποθηκεύονται στην πλατφόρμα QualiChain. Ο στόχος

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

της συγκεκριμένης ανάλυσης όπως προαναφέρθηκε είναι η βελτίωση της ικανότητας του χρήστη στην αγορά εργασίας, μέσω της πρότασης γνώσεων και δεξιοτήτων που δεν διαθέτει και είναι σύμφωνες τόσο με τις ανάγκες της αγοράς εργασίας, όσο και με την εξειδίκευση του και τις ανάγκες του. Η πρόταση δεξιοτήτων και μαθημάτων γίνεται με την αξιοποίηση των λιστών με τις γνώσεις και δεξιότητες για κάθε εξειδίκευση.

#### **5.2.4 Παροχή προτάσεων με τη χρήση συσταδοποίησης (clustering)**

Για τη βελτίωση των παρεχόμενων προτάσεων μέσω της προσέγγισης που περιγράφηκε στην προηγούμενη ενότητα, όπου οι προτάσεις δεξιοτήτων και μαθημάτων ήταν κυρίως αποτέλεσμα ανάλυσης και αντιστοιχίας κειμένου στις εξειδικεύσεις που αναγνωρίστηκαν εκ των προτέρων, αναπτύχθηκε και ένας εναλλακτικός τρόπος παροχής προτάσεων δεξιοτήτων και μαθημάτων, που εκμεταλλεύεται τη μηχανική μάθηση και συγκεκριμένα την τεχνική της συσταδοποίησης (clustering) [11]. Στη συγκεκριμένη προσέγγιση οι εξειδικεύσεις αναγνωρίζονται δυναμικά με βάση τα διαθέσιμα δεδομένα, και κατά συνέπεια τα αποτελέσματα είναι σε μεγάλο βαθμό διαφορετικά από αυτά που προκύπτουν από την προσέγγιση που περιγράφεται στην προηγούμενη ενότητα.

Συγκεκριμένα, δοκιμάστηκαν διάφορες τεχνικές συσταδοποίησης (clustering) στις αγγελίες εργασίας, όπως ο αλγόριθμος K-means [12], ο αλγόριθμος DBScan [13] και η μέθοδος των LDA (Latent Dirichlet Allocation) [14], για την αναγνώριση των διαφορετικών εξειδικεύσεων και των σημαντικότερων γνώσεων και δεξιοτήτων που αυτές απαιτούν. Τα αποτελέσματα εξετάστηκαν διεξοδικά και στα αρχικά πειράματα, τα αποτελέσματα δεν ήταν τα επιθυμητά, καθώς οι εξειδικεύσεις που αναγνωρίστηκαν δεν ήταν αυτές που θα περίμενε κάποιος που γνωρίζει το συγκεκριμένο επιστημονικό αντικείμενο. Ο βασικός λόγος για αυτό ήταν η μεγάλη διαφορά σε πλήθος μεταξύ των διαφορετικών εξειδικεύσεων, που είχε ως αποτέλεσμα μόνο λίγες αντιπροσωπευτικές εξειδικεύσεις και μια τεράστια συστάδα, που περιελάμβανε όλες τις αγγελίες που δεν μπορούσαν να ενταχθούν στις προηγούμενες συστάδες (clusters). Η τελευταία συστάδα στην πραγματικότητα περιελάμβανε πολλές και λιγότερο δημοφιλείς εξειδικεύσεις. Έτσι, έγινε γρήγορα αντιληπτό ότι η καταλληλότερη μέθοδος ήταν η ιεραρχική συσταδοποίηση (hierarchical/agglomerative clustering) [15].

Πιο συγκεκριμένα, αντί να χρησιμοποιηθεί κάποια έτοιμη βιβλιοθήκη για τον αλγόριθμο της ιεραρχικής συσταδοποίησης, χρησιμοποιήθηκε αναδρομικά ο αλγόριθμος K-means clustering. Για την ακρίβεια, για κάθε συστάδα που ήταν πολύ ετερογενής και πολυπληθής, εφαρμόζοταν ο αλγόριθμος K-means αναδρομικά στη συγκεκριμένη συστάδα. Με αυτό τον τρόπο, υπήρχε μεγαλύτερος έλεγχος του αποτελέσματος για τις συστάδες, και για το βάθος του αλγορίθμου. Φυσικά, σε περίπτωση που η αναδρομή συνεχιστεί έως ότου δεν υπάρχουν ετερογενείς συστάδες, το αποτέλεσμα θα είναι πολλές μικρές συστάδες με πολύ παρόμοια χαρακτηριστικά στις αγγελίες της ίδιας συστάδας. Έτσι, όταν ο αριθμός των συστάδων είναι αρκετά μεγάλος και οι περισσότερες συστάδες είναι αρκετά αντιπροσωπευτικές κάποιας εξειδίκευσης, η αναδρομή θα πρέπει να σταματήσει. Πιο συγκεκριμένα, το κριτήριο για την διακοπή της αναδρομής περιλαμβάνει δύο συνιστώσες. Πρωτίστως, ορίζεται εξ αρχής ένας μέγιστος αριθμός συστάδων (στην περίπτωση μας αποφασίστηκε μετά από πειραματισμό ότι 11 συστάδες θα είναι ο μέγιστος αριθμός συστάδων). Δεύτερον, ορίζεται ο μέγιστος βαθμός ετερογένειας μιας συστάδας. Συγκεκριμένα, ορίστηκε ότι δεν πρέπει να υπάρχουν κοινές γνώσεις και δεξιότητες για

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

τουλάχιστον το 20% του πληθυσμού της συστάδας, οπότε σε αυτή την περίπτωση η συστάδα θα πρέπει να διαχωριστεί περαιτέρω.

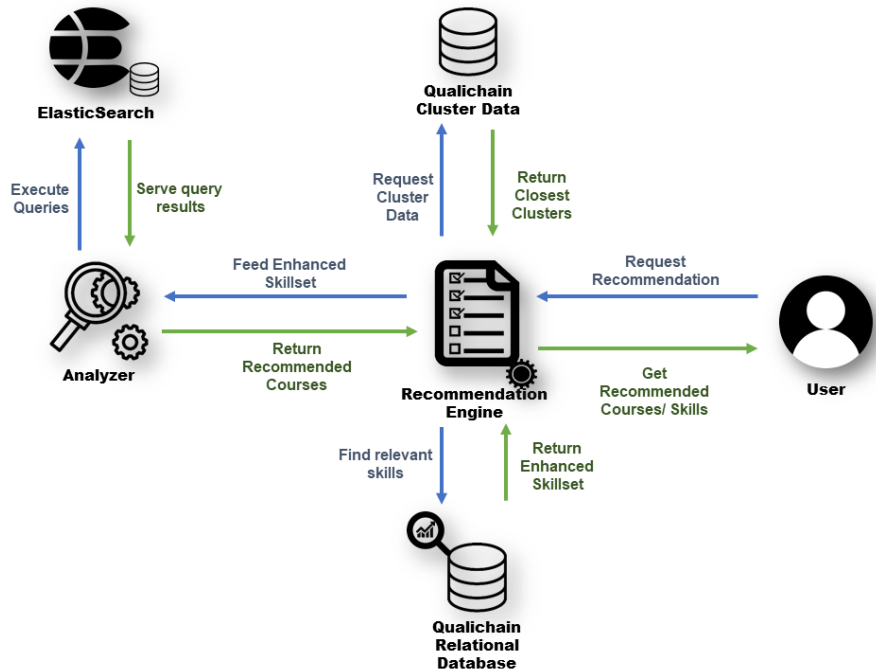
Όσον αφορά την παράμετρο  $K$  του αλγορίθμου  $K$ -means, αποφασίζεται σε κάθε βήμα με τη μέθοδο του βαθμού silhouette (silhouette score method) [16]. Η συγκεκριμένη μέθοδος υποδεικνύει τον βέλτιστο αριθμό συστάδων με βάση την ποιότητα τους και την εμπιστοσύνη διαχωρισμού. Το αποτέλεσμα της συσταδοποίησης είναι ένα σύνολο από συστάδες μαζί με την πιθανότητα κάθε δεξιότητα να εμφανίζεται σε κάθε συστάδα. Βεβαίως, αξίζει να σημειωθεί ότι η διαδικασία που περιγράφηκε για τη συσταδοποίηση απαιτεί σημαντικό χρόνο για την εκτέλεση της, ενώ παράλληλα η αγορά εργασίας, να μεν μεταβάλλεται γρήγορα αλλά όχι και σε καθημερινή βάση, έτσι η διαδικασία που περιγράφεται δε χρειάζεται και δε γίνεται να εκτελείται σε κάθε χρήση της υπηρεσίας. Αντίθετα, εκτελείται περιοδικά και τα αποτελέσματα αποθηκεύονται στον διακομιστή (server) και φορτώνονται από εκεί σε κάθε αίτημα κάποιου χρήστη για πρόταση μαθημάτων και δεξιοτήτων.

Όσον αφορά τις προτάσεις δεξιοτήτων και μαθημάτων, σε κάθε αίτημα, το προφίλ κάθε χρήστη μεταφράζεται σε ένα διάνυσμα δεξιοτήτων. Επιπλέον, τα αποτελέσματα της συσταδοποίησης, αποθηκεύονται και με τη μορφή διανύσματος για κάθε συστάδα. Η συγκεκριμένη αναπαράσταση αναπαριστά το κεντροειδές (centroid) της συστάδας. Για την παροχή των τελικών προτάσεων, υπολογίζεται η συνημιτονική ομοιότητα (cosine similarity) [17] του χρήστη με κάθε συστάδα και τελικά προτείνονται στο χρήστη γνώσεις και δεξιότητες τις οποίες δεν κατέχει αλλά είναι σε μεγάλη ζήτηση από τις τρεις κοντινότερες σε αυτόν συστάδες (εξειδικεύσεις).

### 5.2.5 Οι τελικές συστάσεις της υπηρεσίας

Και οι δυο προσεγγίσεις για τις προτάσεις δεξιοτήτων και μαθημάτων που παρουσιάστηκαν στις προηγούμενες ενότητες παρέχουν ακριβείς προτάσεις στις περισσότερες περιπτώσεις. Ωστόσο, οι προτάσεις που παρέχονται με κάθε μέθοδο έχουν σημαντικές διαφορές. Καθότι καμία από τις δυο προσεγγίσεις δε φαίνεται να υπερτερεί ως προς το αποτέλεσμα, αποφασίστηκε το σύστημα συστάσεων να παρέχει προτάσεις που προέρχονται και από τις δύο μεθόδους συνδυαστικά. Συγκεκριμένα, καλείται τόσο η υπηρεσία σύστασης βασισμένη στην αντιστοίχιση κειμένου όσο και αυτή που βασίζεται στη συσταδοποίηση μετά το αίτημα ενός χρήστη και τα προτεινόμενα σύνολα γνώσεων και δεξιοτήτων συνδυάζονται και προτείνονται με βάση τη σημαντικότητά τους, παράλληλα με τα μαθήματα στα οποία διδάσκονται οι συγκεκριμένες γνώσεις και δεξιότητες. Η ροή εργασιών (workflow) του συστήματος συστάσεων παρουσιάζεται στην ακόλουθη εικόνα.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 12 – Αρχιτεκτονική της υπηρεσίας συστάσεων δεξιοτήτων και μαθημάτων

### 5.3 Ενδεικτικά αποτελέσματα

Τα αποτελέσματα από την υπηρεσία JobCrawler περιλαμβάνουν 1777 αγγελίες εργασίας από την ελληνική αγορά εργασίας για τα επαγγέλματα που σχετίζονται με την τεχνολογία της πληροφορίας και την ανάπτυξη λογισμικού. Στη συνέχεια, η υπηρεσία εξαγωγής γνώσης η οποία εκτελέστηκε στις προηγούμενες αγγελίες εργασίας αναγνώρισε 345 διαφορετικές γνώσεις ή δεξιότητες και τις ταξινόμησε σε μία από τις κατηγορίες “Topic”, “Product”, και “Tool”. Καθώς πολλές από τις συγκεκριμένες γνώσεις και δεξιότητες δεν εμφανίζονται σε επαρκή αριθμό αγγελιών εργασίας, μόνο 94 γνώσεις και δεξιότητες κρατήθηκαν για τις συστάσεις που βασίζονται στη συσταδοποίηση. Από την άλλη πλευρά 200 γνώσεις και δεξιότητες εντοπίστηκαν στο πρόγραμμα σπουδών της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου, οι οποίες σχετίζονταν με την επιστήμη των υπολογιστών και την επιστήμη των δεδομένων. Στις επόμενες ενότητες παρουσιάζονται ενδεικτικά αποτελέσματα από τις δυο τεχνικές προτάσεων που παρουσιάστηκαν, όσο και από τον συνδυασμό τους.

#### 5.3.1 Ενδιάμεσα αποτελέσματα συστάσεων με τη χρήση της τεχνολογίας ElasticSearch

Όπως έχει αναφερθεί και σε προηγούμενη ενότητα, η υπηρεσία QualiChain Analyser συνδέει τις εξαχθείσες γνώσεις και δεξιότητες με εξειδικεύσεις της επιστήμης υπολογιστών με τη χρήση αντιστοίχισης κειμένου που προσφέρει η τεχνολογία ElasticSearch. Ο παρακάτω πίνακας απεικονίζει τις εξειδικεύσεις που ορίστηκαν και ένα υποσύνολο από τις πιο σημαντικές γνώσεις και δεξιότητες που είναι συνυφασμένες με τις εξειδικεύσεις αυτές.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Πίνακας 18 - Διάφορες εξειδικεύσεις στην τεχνολογία της πληροφορίας και οι δεξιότητες που εξάχθηκαν για καθεμιά

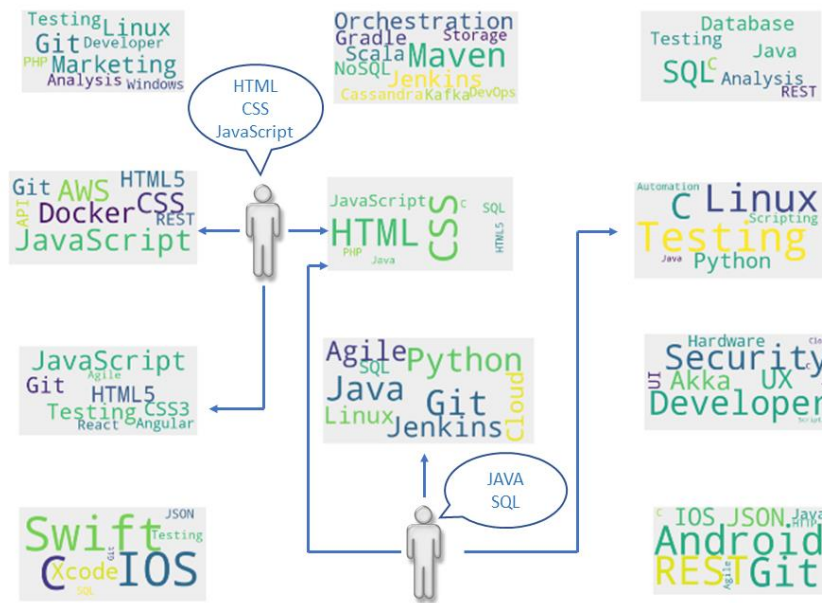
IT roles	User1 Recommended Courses
Backend Developer	Php, Databases, Go, SQL, Linux, Python, AWS, Docker
Frontend Developer	JavaScript, Ajax, CSS3, jQuery, React.js, Node.js, React Native, UX
Database Developer	Oracle, SQL, Automation. Cloud, RDBMS, NoSQL, Parallel Processing
Hardware Engineer	C++, TCP, C, Networking, Compiler, Linux Kernel, ARM, FPGA
Data Analyst	Analysis, Python, Modelling, Visualization, SQL, R, Statistics, Data Science
Machine Learning Engineer	Python, Machine Learning, Analytics, R, Java, Statistics, Databases, SQL
Data Engineer	Python, Big Data, SQL, Hadoop, Java, Spark, Scala, Kafka, Hive
Network Architect	Linux, Network, Cisco, Security, Wireless, VPN, CentOS, IPsec, TCP
DevOps Engineer	Automation, Go, Docker, Python, Security, Ansible, Jenkins, Continuous Integration
Security Engineer	Networking, Firewall, Linux, Windows, Cisco, VPN, Cloud, SIEM
Web Developer	JavaScript, CSS, HTML, jQuery, PHP, SQL, UI, MVC, React.js
Mobile Developer	Android, Rest, Java, Git, JSON, UI, Swift, Design Patterns

Όσον αφορά τους τελικούς χρήστες, τους προτείνονται γνώσεις και δεξιότητες από την κοντινότερη εξειδίκευση με το προφίλ τους από τον παραπάνω πίνακα. Με την εκτέλεση συνδυαστικών ερωτημάτων στο Elasticsearch για την αντιστοίχιση των γνώσεων και δεξιοτήτων του χρήστη με μαθήματα και γνώσεις και δεξιότητες που δεν κατέχει, ο χρήστης λαμβάνει μια λίστα από σχετικές δεξιότητες και μαθήματα. Για παράδειγμα, ένας χρήστης που κατέχει τις γνώσεις/δεξιότητες Java, Python και SQL θα λάβει προτάσεις να επιλέξει τα μαθήματα των καταναμημένων συστημάτων, της μηχανικής λογισμικού και των προχωρημένων θεμάτων βάσεων δεδομένων, τα οποία σχετίζονται με γνώσεις και δεξιότητες που περιλαμβάνουν τις βάσεις δεδομένων, τις μη σχεσιακές βάσεις δεδομένων (NoSQL databases), το cloud computing, τις διαδικτυακές υπηρεσίες της Amazon (Amazon Web Services - AWS), τη γλώσσα προγραμματισμού Go, το εργαλείο Selenium, τη στατιστική και την αναλυτική δεδομένων.

### 5.3.2 Αποτελέσματα συστάσεων με τη χρήση τεχνικών συσταδοποίησης

Όπως έχει ήδη αναφερθεί, οι συστάσεις με τη μέθοδο της συσταδοποίησης, απαιτούν την συσταδοποίηση στις αγγελίες εργασίας, για την εύρεση των διαφορετικών εξειδικεύσεων της αγοράς εργασίας για ένα συγκεκριμένο επιστημονικό πεδίο. Στη συνέχεια, οι χρήστες παραθέτουν τις γνώσεις και τις δεξιότητες τους μέσω μιας φόρμας και βάσει της συνημητονικής ομοιότητας με τα κεντροειδή της κάθε εξειδίκευσης, λαμβάνουν προτάσεις από τις τρεις κοντινότερες σε αυτούς εξειδικεύσεις. Στην παρακάτω εικόνα, παρουσιάζονται οι διαφορετικές εξειδικεύσεις/συστάδες, με τις σημαντικότερες γνώσεις και δεξιότητες που απαιτούν, καθώς και τη σημαντικότητα της κάθε μιας που αναπαρίσταται από το μέγεθος και τη θέση της στη συστάδα. Επιπλέον, παρουσιάζονται δύο χρήστες και οι τρεις κοντινότερες σε αυτούς εξειδικεύσεις/συστάδες, από τις οποίες θα λάβουν προτάσεις δεξιοτήτων.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 13 - Τα αποτελέσματα της συσταδοποίησης και οι κοντινότερες συστάδες για δύο χρήστες

Όπως φαίνεται και στην προηγούμενη εικόνα, η ιεραρχική συσταδοποίηση καταλήγει σε 11 συστάδες. Όπως έχει αναφερθεί, το μέγεθος μιας δεξιότητας σε μια συστάδα εκφράζει τη σημαντικότητα της δεξιότητας μέσα στη συστάδα. Έτσι, η πρώτη συστάδα περιλαμβάνει πολύ γενικές δεξιότητες και είναι αρκετά ετερογενής. Στην πραγματικότητα, περιλαμβάνει όλες τις αγγελίες εργασίας οι οποίες δεν μπορούν να συμπεριληφθούν σε καμία από τις υπόλοιπες συστάδες. Στη δεύτερη συστάδα, οι σημαντικότερες δεξιότητες είναι το εργαλείο Maven, η ενορχήστρωση υπηρεσιών, το εργαλείο Gradle, η γλώσσα προγραμματισμού Scala, το εργαλείο αυτοματισμού Jenkins κ.ά.. Αυτές οι γνώσεις/δεξιότητες ταιριάζουν σε μεγάλο βαθμό με την εξειδίκευση του μηχανικού λογισμικού στα DevOps. Οι σημαντικότερες δεξιότητες της τρίτης συστάδας περιλαμβάνουν τη γλώσσα προγραμματισμού SQL, τις βάσεις δεδομένων καθώς και εργαλεία για τη διαχείριση δεδομένων. Συνεπώς, η συγκεκριμένη εξειδίκευση θα μπορούσε να περιγράψει έναν μηχανικό λογισμικού SQL, έναν προγραμματιστή βάσεων δεδομένων, ή έναν αναλυτή δεδομένων. Η σημαντικότερη δεξιότητα της τέταρτης συστάδας είναι η γλώσσα προγραμματισμού JavaScript, ακολουθούμενη από τις γλώσσες HTML5, CSS, την τεχνολογία Docker και τις διαδικτυακές υπηρεσίες της Amazon (AWS). Αυτές οι δεξιότητες αναπαριστούν ένα προγραμματιστή διαδικτύου με εξειδίκευση στο front end. Στην πέμπτη συστάδα κυριαρχούν οι δεξιότητες HTML και CSS, κι έτσι η συγκεκριμένη συστάδα αναπαριστά έναν μηχανικό λογισμικού front end ή έναν προγραμματιστή διεπαφής χρήστη (UI developer). Η έκτη συστάδα περιλαμβάνει δεξιότητες όπως το λειτουργικό σύστημα Linux, η γλώσσα προγραμματισμού C, το testing, η γλώσσα προγραμματισμού Python, ο αυτοματισμός και ο προγραμματισμός σεναρίων κι έτσι αναπαριστά έναν μηχανικό λογισμικού στα Operations. Η έβδομη συστάδα περιλαμβάνει τη γλώσσα προγραμματισμού JavaScript ως κυρίαρχη δεξιότητα και ακολουθείται από δεξιότητες όπως τα frameworks React, και Angular, αλλά και τις γλώσσες προγραμματισμού HTML5 και CSS3 κι έτσι αναπαριστά έναν μηχανικό λογισμικού front end με εξειδίκευση στη JavaScript. Η όγδοη συστάδα περιλαμβάνει δεξιότητες που

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

αποδίδονται στον μηχανικό λογισμικού στο back end, όπως είναι οι γλώσσες προγραμματισμού Java, και Python και οι τεχνολογίες Linux, Git και Jenkins. Η ένατη συστάδα αναφέρεται σε μηχανικούς λογισμικού στην ασφάλεια και έχει ως κυρίαρχες δεξιότητες την ασφάλεια, το framework Akka και το UX, μεταξύ άλλων. Τέλος, οι δυο τελευταίες συστάδες αναφέρονται σε προγραμματιστές κινητών τηλεφώνων, εκ των οποίων η πρώτη συστάδα περιλαμβάνει την εξειδίκευση στις τεχνολογίες Swift και XCode ενώ η δεύτερη στο λειτουργικό σύστημα Android.

Όσον αφορά τους χρήστες User1 και User2 της προηγούμενης εικόνας, ο πρώτος διαθέτει τις δεξιότητες Java και SQL κι έτσι οι κοντινότερες σε αυτόν εξειδικεύσεις που αναγνωρίζει το σύστημα είναι αυτές του Μηχανικού λογισμικού στο back end, του Μηχανικού λογισμικού SQL και του Μηχανικού λογισμικού Front end. Συνεπώς, ο User1 θα λάβει προτάσεις από τις προηγούμενες εξειδικεύσεις. Από την άλλη πλευρά ο χρήστης User2 διαθέτει τις δεξιότητες HTML, CSS και JavaScript και οι κοντινότερες σε αυτόν εξειδικεύσεις είναι αυτές του προγραμματιστή διαδικτύου με ειδίκευση στο front end, του μηχανικού λογισμικού front end με εξειδίκευση στη JavaScript και του μηχανικού λογισμικού front end με ειδίκευση στις γλώσσες προγραμματισμού HTML και CSS, από τις οποίες του προτείνονται και δεξιότητες από το σύστημα. Οι προτεινόμενες δεξιότητες και τα προτεινόμενα μαθήματα για τους χρήστες User1 και User2 παρουσιάζονται στον παρακάτω πίνακα.

**Πίνακας 19 - Οι προτάσεις που προκύπτουν από την προσέγγιση συσταδοποίησης για τους χρήστες User1 και User2**

User1 Recommended Skills	User1 Recommended Courses	User2 Recommended Skills	User2 Recommended Courses
Database	Databases	Git	Databases
Analysis	Distributed Systems	HTML5	Advanced Topics in Database Systems
Testing	Internet Programming	Testing	Internet Programming
C		CSS3	Analysis and Design of Information Systems
REST		Angular	
Git		React	
Python		Agile	
Jenkins		SQL	
Agile		PHP	
Linux		AWS	
Cloud		Docker	
HTML		REST	
CSS		API	
JavaScript			
PHP			
HTML5			



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Στον παραπάνω πίνακα είναι φανερό ότι ένας μεγάλος αριθμός γνώσεων και δεξιοτήτων προτείνεται και στους δύο χρήστες, καθώς μοιράζονται μια κοινή εξειδίκευση. Επιπλέον, σε κάθε έναν προτείνεται μια μικρή λίστα από σχετικά μαθήματα που προσφέρουν τις προτεινόμενες γνώσεις και δεξιότητες. Είναι φανερό ότι και στις δύο περιπτώσεις τόσο οι προτεινόμενες δεξιότητες όσο και τα μαθήματα είναι πολύ κοντά στις γνώσεις και δεξιότητες του χρήστη.

### 5.3.3 Προτάσεις με συνδυασμό των αποτελεσμάτων

Όπως αναφέρθηκε και στις προηγούμενες ενότητες, τα αποτελέσματα και από τις δύο μεθόδους σύστασης δείχνουν πολύ ακριβή και κανένα από τα δυο δε φαίνεται να υπερτερεί του άλλου γενικά. Επιπλέον, σε πολλές περιπτώσεις τα προτεινόμενα μαθήματα είναι λιγότερα από όσα θα περίμενε κανείς. Σε αυτό το πλαίσιο, αποφασίστηκε ο συνδυασμός των αποτελεσμάτων και από τις δύο μεθόδους για την παροχή μεγαλύτερης ποικιλίας προτάσεων δεξιοτήτων και μαθημάτων στους χρήστες του συστήματος. Στον παρακάτω πίνακα παρουσιάζονται τα αποτελέσματα του συνδυασμού των μεθόδων για τους χρήστες User1 και User2.

Πίνακας 20 - Οι προτάσεις για τους χρήστες User1 και User2 με συνδυασμό των δύο προσεγγίσεων

User1 Recommended Skills	User1 Recommended Courses	User2 Recommended Skills	User2 Recommended Courses
Database	Programming Languages I	PHP	Digital Systems Laboratory
Scripting	Internet Programming	Angular	Software Engineering
Jira	Advanced Topics in Database Systems	REST	Databases
JavaScript	Multimedia Technology	React	Internet Programming
Automation	Computer Systems Performance	Agile	Programming Languages I
Jenkins	Databases	API	Compilers
HTML	Project Management	AWS	Management of Information Systems
AWS	Software Engineering	CSS3	Management Game
Python	Distributed Systems	Git	Telecommunication Electronics
Docker	Telecommunication Electronics	Docker	
CSS		SQL	
Cloud		Tools	
Ansible		HTML5	
Azure		Design	
Git		Testing	
NoSQL		Version Control	
PHP		Backend	
Programming		JQuery	
Agile		C	

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Selenium		UI	
Android		UX	
REST		Marketing	
JSON		Web Applications	

## 5.4 Επίδειξη υπηρεσίας και διεπαφής χρήστη

Όπως έχει ήδη αναφερθεί σε προηγούμενες ενότητες, ο χρήστης προκειμένου να λάβει προτάσεις μαθημάτων και δεξιοτήτων, θα πρέπει να εισάγει αναλυτικές πληροφορίες σχετικές με το βιογραφικό του, συμπεριλαμβανομένων των γνώσεων και δεξιοτήτων που κατέχει και της επαγγελματικής του εμπειρίας. Στο παράδειγμα της παρακάτω εικόνας, ο χρήστης θέλει να λάβει προτάσεις για τον τομέα της τεχνολογίας της πληροφορίας και κατέχει τις δεξιότητες Java και MySQL, σε μέτριο βαθμό την πρώτη, και σε εξαιρετικό βαθμό τη δεύτερη, ενώ διαθέτει δύο χρόνια επαγγελματικής εμπειρίας στη γλώσσα προγραμματισμού Java και τρία στη γλώσσα προγραμματισμού βάσεων δεδομένων MySQL.

The image shows a user profile form with two skill entries. Each entry has a 'Delete' button, a 'Label' field, a 'Proficiency level' dropdown, and a 'Comment' field. The first skill is 'Java' with a 'Medium' proficiency level and '2 years of experience'. The second skill is 'MySQL' with an 'Advanced' proficiency level and '3 years of experience'. At the bottom of the form are 'Save' and 'Close' buttons.

Εικόνα 14 - Ανέβασμα βιογραφικού από κάποιον φοιτητή ή δια βίου μαθητή

Μετά την αποθήκευση της πληροφορίας που εισάγει ο χρήστης από την παραπάνω φόρμα, η υπηρεσία εντοπίζει τις καταλληλότερες γνώσεις και δεξιότητες για να προτείνει στον χρήστη, ανάλογα με το προφίλ του. Οι προτεινόμενες γνώσεις και δεξιότητες για το χρήστη φαίνονται στην παρακάτω εικόνα.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

### Recommended skills

Filter

Skill	Action
git	<a href="#">View</a>
docker	<a href="#">View</a>
jira	<a href="#">View</a>
sql	<a href="#">View</a>
database	<a href="#">View</a>
azure	<a href="#">View</a>
cloud	<a href="#">View</a>
jenkins	<a href="#">View</a>

Εικόνα 15 - Προτεινόμενες δεξιότητες

Από την παραπάνω εικόνα φαίνεται ότι οι γνώσεις και δεξιότητες που προτείνονται στο χρήστη περιλαμβάνουν τις τεχνολογίες Git, Docker, Jira, SQL, Databases, Azure, Cloud και Jenkins. Υπάρχουν περισσότερες γνώσεις και δεξιότητες που προτείνονται στο χρήστη, ωστόσο παρουσιάζονται μόνο οι σημαντικότερες, όσον αφορά τις προτιμήσεις της αγοράς εργασίας με φθίνουσα σειρά ανάλογα με τη ζήτηση τους από την αγορά εργασίας.

Στη συνέχεια, ο χρήστης μπορεί είτε να αποφασίσει ποιες γνώσεις και δεξιότητες θα προσπαθήσει να αποκτήσει είτε να λάβει προτάσεις μαθημάτων από τη σχολή του. Φυσικά, εφόσον η επιλογή γνώσεων και δεξιοτήτων είναι άρρηκτα συνδεδεμένη με την επιλογή μαθημάτων, ο χρήστης επιλέγει να λάβει και προτάσεις μαθημάτων όπως παρουσιάζεται στην παρακάτω εικόνα.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## Recommended courses

Filter

Name	Description	Score	Action
Databases	Database Management Systems and their architecture. Data Structures for Databases. Modeling - The E-R model. Reference to classic Database Models (Hierarchical, Network). The Relational Model. Database Languages - The SQL language. File Systems and Physical Database Design. Logical Design and Normalization. Management and Operation Issues (integrity, optimization, reorganization, security, functionality, etc.). Other Issues (Object Oriented Systems, Multi-Systems, Personal Computer Systems, etc.).	1	<a href="#">View</a>
Operating Systems	This course covers the general principles of Operating Systems. Evolution of Operating Systems, Input/Output, concurrent processes, critical section, process synchronization and communication. CPU scheduling, Memory management (static and dynamic allocation, virtual memory, paging, segmentation). File management, disk scheduling, deadlocks.	1	<a href="#">View</a>
Distributed Systems	Synchronization: Natural, logical and vector clocks, natural and logical clock synchronization, Lamport's algorithm. Distributed Coordination: Mutual Exclusion Algorithms, Election Algorithms, Distributed Agreement, Interactive Consistency, Byzantine Protocols, Distributed Consensus Algorithm Paxos. Tasks and concurrency control: ACID properties, embedded transactions, strict two-phase locking, concurrent problems, timer layout, optimistic concurrency control. Distributed Transactions and Timing Control. Distributed Individual Transactions, Distributed Timing Control, Deadlock Detection, Error Recovery. Data Backup: Error Tolerance, Primary / Secondary Backup, Passive / Active Backup, Network Partition Management, High Availability, Consistency Levels, CAP Theorem. Peer-to-peer networks, distributed hash tables, distributed storage systems, Map Reduce and Bulk Synchronous Parallel distributed processing models.	1	<a href="#">View</a>
Operating Systems Laboratory	This course offers laboratory projects to assimilate critical concepts from the previous courses "Computer Architecture" and "Operating Systems". The course includes 3-4 exercises in distributed UNIX programming, development of device drivers, kernel-level programming, and extensions of the Linux OS.	1	<a href="#">View</a>
Human-Computer Interaction	The course deals with the design, development and evaluation methodologies of computer systems that interact with users to a significant degree. The purpose is to review theoretical models, technologies, methods and tools for the design and development of interactive computing systems. The course includes (1) Human modeling of computer system and conceptual interaction framework. (2) Cognitive models, perception and representation, attention and memory representation and organization of knowledge(3) Interactive Systems Design Methodologies: Human-centric Design, Usability Requirements, Task Analysis, GOMS Models, Dialogue Methods, Interface Design, Usability and Accessibility of Internet Services, Interactive Systems Evaluation Techniques. (4) Overview of Interactive Technologies and Applications: Distributed Intelligence Systems, Ubiquitous Computer, Augmented Reality Systems, Synthetic Animation, Accessibility Interface Systems, Collaboration Applications.	0.87	<a href="#">View</a>

Εικόνα 16 - Προτεινόμενα μαθήματα

Όπως φαίνεται στην παραπάνω εικόνα, τα προτεινόμενα μαθήματα περιλαμβάνουν τα μαθήματα των Βάσεων Δεδομένων, των Λειτουργικών Συστημάτων, των Κατανεμημένων Συστημάτων, το Εργαστήριο Λειτουργικών Συστημάτων, και την Αλληλεπίδραση Ανθρώπου Μηχανής. Επίσης, ανάμεσα στις προτάσεις περιλαμβάνονται και τα μαθήματα της Τεχνολογίας Πολυμέσων και της Επίδοσης Υπολογιστικών Συστημάτων, αλλά δε φαίνονται στην παραπάνω εικόνα. Αξίζει να σημειωθεί πως μαζί με το μάθημα παρατίθεται η περιγραφή του, αλλά και ένας βαθμός συνάφειας με τις προτιμήσεις του χρήστη, με βάση τον οποίο ταξινομείται το αποτέλεσμα.

## 5.5 Σχολιασμός και συμπεράσματα

Το παρόν κεφάλαιο παρουσιάζει το σύστημα συστάσεων μαθημάτων και δεξιοτήτων που αναπτύχθηκε, το οποίο παρέχει προσωποποιημένες προτάσεις δεξιοτήτων και μαθημάτων σε φοιτητές, δια βίου μαθητές και υποψήφιους εργαζομένους, χρησιμοποιώντας δεδομένα που εξάγονται τόσο για την αγορά εργασίας όσο και για την εκπαίδευση. Στο κεφάλαιο αυτό παρουσιάζεται η λειτουργία του συστήματος για την αγορά εργασίας στον επιστημονικό κλάδο της επιστήμης υπολογιστών και της τεχνολογίας της πληροφορίας, και τα μαθήματα που προτείνονται προέρχονται από το πρόγραμμα σπουδών της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου, ωστόσο μπορεί εύκολα να επεκταθεί με επιπλέον σχολές, πανεπιστήμια και άλλους εκπαιδευτικούς οργανισμούς, ή και διαφορετικούς επιστημονικούς κλάδους με μικρές τροποποιήσεις.

Τα δεδομένα αποκτώνται από μια σειρά από ιστοτόπους αγγελιών εργασίας στην Ελλάδα μέσω μιας υπηρεσίας που βασίζεται στην απόξεση δεδομένων διαδικτύου (web scraping). Στη συνέχεια, οι αγγελίες εργασίας που εξάχθηκαν, αναλύονται από την υπηρεσία εξαγωγής γνώσης του QualiChain, η οποία αναγνωρίζει τις βασικές γνώσεις και δεξιότητες που απαιτούνται από κάθε αγγελία εργασίας. Επίσης, με

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

τον ίδιο τρόπο αναγνωρίζονται και οι βασικές γνώσεις και δεξιότητες που προσφέρονται σε κάθε μάθημα από το πρόγραμμα σπουδών που μελετάται. Ακολούθως, εφαρμόζονται τεχνικές συσταδοποίησης (clustering) στις αγγελίες εργασίας για τον εντοπισμό των εξειδικεύσεων του επιστημονικού πεδίου που μελετάται. Οι τελικές προτάσεις παρέχονται τόσο με βάση τα αποτελέσματα της συσταδοποίησης όσο και με βάση τεχνικές και τεχνολογίες αντιστοιχίας κειμένου.

Τα αποτελέσματα φαίνονται λογικά και ικανοποιητικά για τους χρήστες. Επιπλέον, οι εξειδικεύσεις που αναγνωρίστηκαν με τη συσταδοποίηση είναι αντιπροσωπευτικές και αντιστοιχίζονται με την αγορά εργασίας και οι παρεχόμενες προτάσεις θα μπορούσαν ρεαλιστικά να βοηθήσουν έναν χρήστη να βελτιώσει τις γνώσεις και τις δεξιότητες του κατά τέτοιο τρόπο που θα είναι πιο ελκυστικός στην αγορά εργασίας. Περισσότερες πληροφορίες σχετικά με την αξιολόγηση και τη χρήση της συγκεκριμένης υπηρεσίας παρουσιάζονται στο κεφάλαιο 8.

## 5.6 Αναφορές

- [1] “Elasticsearch: The Official Distributed Search & Analytics Engine | Elastic.” <https://www.elastic.co/elasticsearch/> (accessed Apr. 09, 2021).
- [2] “Kariera Hiring Solutions.” <https://www.kariera.gr> (accessed Apr. 09, 2021).
- [3] “Indeed.” <https://gr.indeed.com/> (accessed Apr. 09, 2021).
- [4] “Skywalker.” <https://www.skywalker.gr> (accessed Apr. 09, 2021).
- [5] “Scrapy | A Fast and Powerful Scraping and Web Crawling Framework.” <https://scrapy.org/> (accessed Apr. 09, 2021).
- [6] B. Mohit and I. Zitouni, “Named Entity Recognition,” pp. 221–245, 2014, doi: 10.1007/978-3-642-45358-8\_7.
- [7] C. D. Manning, “Part-of-speech tagging from 97% to 100%: Is it time for some linguistics?,” Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 6608 LNCS, no. PART 1, pp. 171–189, 2011, doi: 10.1007/978-3-642-19400-9\_14/COVER.
- [8] “Welcome to Flask — Flask Documentation (2.3.x).” <https://flask.palletsprojects.com/en/2.3.x/> (accessed May 02, 2023).
- [9] “Elasticsearch DSL — Elasticsearch DSL 7.2.0 documentation.” <https://elasticsearch-dsl.readthedocs.io/en/latest/> (accessed May 02, 2023).
- [10] “pandas - Python Data Analysis Library.” <https://pandas.pydata.org/> (accessed May 02, 2023).
- [11] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering,” ACM Computing Surveys (CSUR), vol. 31, no. 3, pp. 264–323, Sep. 1999, doi: 10.1145/331499.331504.
- [12] A. Likas, N. Vlassis, and J. J. Verbeek, “The global k-means clustering algorithm,” Pattern Recognit, vol. 36, no. 2, Feb. 2003, doi: 10.1016/S0031-3203(02)00060-2.
- [13] K. Khan, S. U. Rehman, K. Aziz, S. Fong, S. Sarasvady, and A. Vishwa, “DBSCAN: Past, present and future,” 5th International Conference on the Applications of Digital Information and Web

- Technologies, ICADIWT 2014, pp. 232–238, 2014, doi: 10.1109/ICADIWT.2014.6814687.
- [14] H. Jelodar et al., “Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey,” *Multimed Tools Appl*, vol. 78, no. 11, pp. 15169–15211, Jun. 2019, doi: 10.1007/S11042-018-6894-4/TABLES/11.
- [15] Daniel Müllner, “Modern hierarchical, agglomerative clustering algorithms.” 2011.
- [16] D. T. Dinh, T. Fujinami, and V. N. Huynh, “Estimating the Optimal Number of Clusters in Categorical Data Clustering by Silhouette Coefficient,” *Communications in Computer and Information Science*, vol. 1103 CCIS, pp. 1–17, 2019, doi: 10.1007/978-981-15-1209-4\_1/TABLES/4.
- [17] S. van Dongen and A. J. Enright, “Metric distances derived from cosine similarity and Pearson and Spearman correlations.” 2012.

## ΚΕΦΑΛΑΙΟ 6: Η ΥΠΗΡΕΣΙΑ ΑΝΑΣΧΕΔΙΑΣΜΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ ΣΠΟΥΔΩΝ

### 6.1 Εισαγωγή

Τα τελευταία χρόνια η αγορά εργασίας, κυρίως στην τεχνολογία, εξελίσσεται πιο ραγδαία από ποτέ και σε αυτό το πλαίσιο τόσο οι εργαζόμενοι και οι φοιτητές ή οι δια βίου μαθητές, όσο και οι εκπαιδευτικοί οργανισμοί αγωνίζονται να συμβαδίσουν με τις τελευταίες εξελίξεις. Από τη μια πλευρά, οι πρώτοι δεν είναι πάντα ενήμεροι για τις γνώσεις και δεξιότητες που είναι σε υψηλή ζήτηση ή θα είναι σε υψηλή ζήτηση στο μέλλον ώστε να επενδύσουν σε αυτές, ενώ ακόμα και όταν είναι ενήμεροι για αυτές, συχνά δεν γνωρίζουν πού και πώς θα μπορούσαν να διδαχθούν τις συγκεκριμένες γνώσεις και δεξιότητες. Από την άλλη πλευρά, η επιτυχία και η μακροημέρευση για τα τεχνολογικά εκπαιδευτικά ιδρύματα εξαρτάται σε μεγάλο βαθμό από την διδασκαλία γνώσεων και δεξιοτήτων που είναι σε υψηλή ζήτηση από την αγορά εργασίας, μιας και συχνά η ποιότητα και η αποτελεσματικότητα τους αξιολογούνται με βάση το χρόνο που χρειάζεται κάποιος απόφοιτος να βρει κάποια εργασία σχετική με το αντικείμενο του. Για το λόγο αυτό, τα εκπαιδευτικά ιδρύματα θα πρέπει όσο γίνεται να συμβαδίζουν με τις ανάγκες της αγοράς εργασίας, για να βελτιώσουν την αποτελεσματικότητά τους και να διατηρήσουν ανταγωνιστικό πλεονέκτημα έναντι των ανταγωνιστών τους. Ωστόσο, συχνά τα εκπαιδευτικά ιδρύματα αποτυγχάνουν να συμβαδίσουν με την αγορά εργασίας όσον αφορά τις διδασκόμενες γνώσεις και δεξιότητες λόγω του ότι αφ' ενός τα προγράμματα σπουδών και τα μαθήματα σε πολλές περιπτώσεις διαμορφώνονται με βάση αυστηρά εκπαιδευτικά κριτήρια σε ένα απομονωμένο περιβάλλον χωρίς να λαμβάνουν υπόψη τις ανάγκες της αγοράς εργασίας και αφ' ετέρου σε πολλά εκπαιδευτικά ιδρύματα, τα προγράμματα σπουδών και τα μαθήματα αλλάζουν πολύ σπάνια σε σχέση με την αγορά εργασίας, κι έτσι δεν μπορούν να ενσωματώσουν στο πρόγραμμα σπουδών τους τις πρόσφατες αλλαγές στην αγορά εργασίας [1].

Η υπηρεσία ανασχεδιασμού προγράμματος σπουδών αποτελεί ένα εργαλείο αναλυτικής δεδομένων που φιλοδοξεί να φέρει την τεχνολογική εκπαίδευση πιο κοντά με την αγορά εργασίας λαμβάνοντας υπόψη τις πιο πρόσφατες επιστημονικές και τεχνολογικές εξελίξεις της αγοράς εργασίας. Συγκεκριμένα, η υπηρεσία αυτή είναι πάντα ενημερωμένη για τις απαιτήσεις της αγοράς εργασίας, μιας και βασίζεται σε μια σειρά εργαλείων που έχουν αναπτυχθεί για εξόρυξη δεδομένων διαδικτύου και αναγνώρισης γνώσεων και δεξιοτήτων μέσα από το εξαγόμενο ελεύθερο κείμενο. Σε αυτό το πλαίσιο, η υπηρεσία ανασχεδιασμού προγράμματος σπουδών μπορεί να αναγνωρίσει ποιες γνώσεις και δεξιότητες λείπουν από ένα δοθέν πρόγραμμα σπουδών ή μια λίστα μαθημάτων και προτείνει την εισαγωγή των γνώσεων και δεξιοτήτων αυτών σε σχετικά μαθήματα, όπου θα μπορούσαν να ενταχθούν οι συγκεκριμένες γνώσεις και δεξιότητες, αξιοποιώντας διάφορες τεχνικές εξόρυξης γνώσης από δεδομένα.

Τα τελευταία χρόνια υπάρχει σημαντικό επιστημονικό και επιχειρηματικό ενδιαφέρον γύρω από τη σχεδίαση προγράμματος σπουδών και τις προκλήσεις που πρέπει να διευθετηθούν ώστε ένα πρόγραμμα σπουδών ενός εκπαιδευτικού οργανισμού να θεωρηθεί επιτυχημένο. Συγκεκριμένα, οι Zablich και Azad [2] ασχολήθηκαν με την αλληλοκάλυψη της ύλης μεταξύ διαφορετικών μαθημάτων, καθώς επίσης και πώς μπορούν να αντιμετωπιστούν τέτοιες αλληλοκαλύψεις. Ο Olney και η ομάδα του [3] παρουσίασαν μια μεθοδολογία για την σχεδίαση προγραμμάτων σπουδών, που μπορούν να χρησιμοποιηθούν από

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

εκπαιδευτικούς για τη δημιουργία εκπαιδευτικών σχεδίων χρησιμοποιώντας οπτικοποιήσεις (visualisations) και άλλα σύνολα δεδομένων. Επιπλέον, ο Shalin Hai-Jew [4] διερευνά τρόπους για την σχεδίαση προγραμμάτων σπουδών με οπισθοδρόμηση (backwards curriculum design). Η συγκεκριμένη μεθοδολογία μπορεί να χρησιμοποιηθεί για την απόκτηση πολύπλοκων γνώσεων και δεξιοτήτων από τους δια βίου μαθητές. Ακόμη, ο Chang και η ομάδα του [5] καταδεικνύουν πώς οι τεχνολογίες αναλυτικής δεδομένων μπορούν να συμβάλλουν στην αποτελεσματικότερη σχεδίαση προγράμματος σπουδών. Επιπροσθέτως, οι συγγραφείς στην έρευνα [6] αναγνωρίζουν την ανάγκη για ορθολογικότερη σχεδίαση των προγραμμάτων σπουδών των πανεπιστημίων αξιοποιώντας τεχνικές προχωρημένης υποστήριξης αποφάσεων και αναλυτικής δεδομένων.

Αξίζει να σημειωθεί ότι υπάρχει πληθώρα επιστημονικών δημοσιεύσεων που τονίζουν τη δυναμική της επιστήμης της εξόρυξης γνώσης από δεδομένα (data mining) και της αναλυτικής δεδομένων (data analytics) στην εκπαίδευση. Πιο συγκεκριμένα, ο Romero και η ομάδα του [7] παρουσίασαν την υφιστάμενη κατάσταση που επικρατεί στον τομέα της εκπαιδευτικής εξόρυξης γνώσης από δεδομένα (Educational Data Mining - EDM), παράλληλα με τις υφιστάμενες εφαρμογές, τα εργαλεία και τις μελλοντικές προσδοκίες. Από την άλλη πλευρά, ο Daniel [8] διερευνά τη δυναμική των μεγάλων δεδομένων και της αναλυτικής δεδομένων για την διευθέτηση των προκλήσεων που αντιμετωπίζουν τα ιδρύματα ανώτατης εκπαίδευσης.

Τέλος, σημαντική δουλειά έχει γίνει και σχετικά με την ανάλυση της αγοράς εργασίας και των αναγκών της. Για παράδειγμα, ο Brãndas και η ομάδα του [9] διερευνούν τις γνώσεις και δεξιότητες που απαιτούνται στην αγορά εργασίας αναλύοντας αγγελίες εργασίας μέσω της εξόρυξης δεδομένων από το διαδίκτυο (web mining) και αναλύοντας τα εξαγόμενα δεδομένα με τεχνικές εξόρυξης γνώσης από δεδομένα (data mining). Επιπλέον, ο Wowczko [10] εφάρμοσε μια παρόμοια μεθοδολογία, δίνοντας έμφαση στις ανάγκες της αγοράς εργασίας και τους πιο επιδραστικούς παράγοντες για τις ανάγκες αυτές, όπως για παράδειγμα η γεωγραφία, η χρονική συγκυρία και ο επαγγελματικός κλάδος.

Η καινοτομία του εργαλείου που περιγράφεται στο παρόν κεφάλαιο έγκειται στο γεγονός ότι συνδυάζονται όλα τα επιστημονικά πεδία που παρουσιάστηκαν παραπάνω, και συγκεκριμένα η σχεδίαση προγράμματος σπουδών, η εξόρυξη δεδομένων από το διαδίκτυο, η αναλυτική κειμένου, η εξόρυξη γνώσης από δεδομένα και η ανάλυση της αγοράς εργασίας. Έτσι παρουσιάζεται μια καινοτόμα μεθοδολογία σχετική με το πώς μπορεί να σχεδιαστεί ένα πρόγραμμα σπουδών καθώς επίσης και ένα συμβουλευτικό εργαλείο που μπορεί να προσφέρει σημαντική προστιθέμενη αξία σε εκπαιδευτικά ιδρύματα κατά τη σχεδίαση ενός προγράμματος σπουδών.

Στο πλαίσιο της παρούσας διδακτορικής διατριβής, ο τομέας που μελετήθηκε είναι αυτός της ανάπτυξης λογισμικού στην Ελλάδα, ενώ το πρόγραμμα σπουδών για το οποίο αναπτύχθηκε η πρωτότυπη εφαρμογή είναι το πρόγραμμα σπουδών της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών (Η.Μ.Μ.Υ.) του Εθνικού Μετσόβιου Πολυτεχνείου (Ε.Μ.Π.). Η επιλογή του τομέα της ανάπτυξης λογισμικού έγινε διότι αποτελεί έναν από τους πιο ραγδαία αναπτυσσόμενους επιστημονικούς κλάδους, ενώ η σχολή Η.Μ.Μ.Υ. του Ε.Μ.Π. επιλέχθηκε ως μια από τις μεγαλύτερες και πιο αναγνωρισμένες τεχνολογικές σχολές στην Ελλάδα, ενώ παράλληλα προσφέρει ένα ευρύ και πολύπλοκο πρόγραμμα σπουδών που περιλαμβάνει περισσότερα από εκατό μαθήματα.



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Στη συνέχεια παρουσιάζεται η μεθοδολογία καθώς επίσης και τα επιμέρους εργαλεία που αναπτύχθηκαν για εξαγωγή δεδομένων από το διαδίκτυο, την ανάλυση δεδομένων και τις δυνατότητες υποστήριξης αποφάσεων που παρέχει η υπηρεσία. Στη συνέχεια, παρουσιάζονται ενδεικτικά αποτελέσματα, μερικά παραδείγματα χρήσης της εφαρμογής και τέλος τα βασικά συμπεράσματα και οι μελλοντικές επεκτάσεις της εφαρμογής.

## 6.2 Μεθοδολογία

Στην παρούσα ενότητα παρουσιάζεται η μεθοδολογία που ακολουθήθηκε για την ανάπτυξη της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών, από την εξόρυξη διαδικτυακών δεδομένων έως την ανάλυση και τις τελικές προτάσεις του συστήματος. Συγκεκριμένα, η μεθοδολογία περιλαμβάνει τα ακόλουθα βήματα:

1. **Εξαγωγή δεδομένων από διάφορους ιστοτόπους εύρεσης εργασίας:** Τεχνικές εξαγωγής διαδικτυακών δεδομένων χρησιμοποιήθηκαν για να εξάγουν δεδομένα από διάφορους ιστοτόπους εύρεσης εργασίας στην Ελλάδα. Η εξαγωγή των δεδομένων έγινε μέσω της υπηρεσίας JobCrawler που αναπτύχθηκε για την υποστήριξη της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών και περιγράφεται σε επόμενη ενότητα.
2. **Αναγνώριση των σημαντικότερων γνώσεων και δεξιοτήτων που απαιτούνται για κάθε αγγελία εργασίας:** Η υπηρεσία QualiChain Knowledge Extraction χρησιμοποιήθηκε για την αναγνώριση των γνώσεων και δεξιοτήτων από το ελεύθερο κείμενο των αγγελιών εργασίας. Η συγκεκριμένη υπηρεσία παρουσιάζεται σε επόμενη ενότητα.
3. **Αναγνώριση γνώσεων και δεξιοτήτων που είναι μεν σε μεγάλη ζήτηση από την αγορά εργασίας αλλά δεν καλύπτονται από το δοθέν πρόγραμμα σπουδών.** Αυτό γίνεται αφ' ενός αναλύοντας, συναθροίζοντας και μετρώντας τη ζήτηση των γνώσεων και δεξιοτήτων από την αγορά εργασίας και αφ' ετέρου αναλύοντας το δοθέν πρόγραμμα σπουδών. Αξίζει να σημειωθεί ότι για την ανάλυση αυτή σχεδιάστηκε και αναπτύχθηκε βάση δεδομένων με λεπτομερή δεδομένα για το πρόγραμμα σπουδών που μελετούμε. Η συγκεκριμένη βάση δεδομένων περιλαμβάνει αναλυτικές περιγραφές μαθημάτων, εξειδικεύσεις της σχολής, καθώς επίσης και γνώσεις και δεξιότητες που προσφέρονται στα πλαίσια του κάθε μαθήματος.
4. **Εύρεση μαθημάτων σχετικών με τις γνώσεις και δεξιότητες οι οποίες είναι σε υψηλή ζήτηση αλλά δεν καλύπτονται από το δοθέν πρόγραμμα σπουδών:** Για την εύρεση τέτοιων μαθημάτων εφαρμόζονται τεχνικές εξόρυξης κανόνων συσχέτισης (association rules mining) [11] μεταξύ των γνώσεων και δεξιοτήτων που προσφέρονται στα μαθήματα και αυτών που δε προσφέρονται στο πρόγραμμα σπουδών αλλά είναι σε υψηλή ζήτηση από την αγορά εργασίας. Η διαδικασία εξόρυξης κανόνων συσχέτισης παρουσιάζεται εκτενέστερα σε επόμενη ενότητα.
5. **Πρόταση για εισαγωγή των γνώσεων και δεξιοτήτων που λείπουν από το πρόγραμμα σπουδών αλλά είναι σε υψηλή ζήτηση από την αγορά εργασίας σε συγκεκριμένα μαθήματα:** Αυτό γίνεται με βάση την πιθανότητα συσχέτισης των προαναφερθέντων γνώσεων και δεξιοτήτων με τις γνώσεις και δεξιότητες που προσφέρονται σε κάθε μάθημα, αξιοποιώντας τους εξαχθέντες κανόνες συσχέτισης.

Στις επόμενες ενότητες παρουσιάζονται οι υπηρεσίες JobCrawler, QualiChain Knowledge Extraction, και Association Rules Miner, καθώς επίσης και οι αλληλεπιδράσεις μεταξύ τους.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

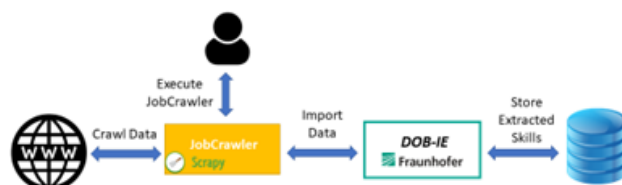
### 6.2.1 Η υπηρεσία JobCrawler

Η υπηρεσία JobCrawler εξάγει μεγάλο όγκο δεδομένων για την ελληνική αγορά εργασίας από διάφορους ιστοτόπους αγγελιών εργασίας συμπεριλαμβανομένων των kariera.gr [12], του ελληνικού indeed [13] και του skywalker [14]. Η εξαγωγή των δεδομένων από τους συγκεκριμένους ιστοτόπους γίνεται μέσω της τεχνικής της απόξεσης δεδομένων διαδικτύου (web scrapping) και πιο συγκεκριμένα χρησιμοποιώντας το framework scrapy [15] και την αντίστοιχη βιβλιοθήκη της γλώσσας προγραμματισμού Python. Οι εξαχθείσες αγγελίες εργασίας μπορούν να χρησιμοποιηθούν στη συνέχεια από την υπηρεσία ανασχεδιασμού προγράμματος σπουδών, για την ανανέωση ενός προγράμματος σπουδών. Στα πλαίσια της συγκεκριμένης διδακτορικής διατριβής παρουσιάζονται αποτελέσματα μόνο για τον τομέα της ανάπτυξης λογισμικού, ωστόσο, η υπηρεσία JobCrawler μπορεί να χρησιμοποιηθεί για την εξαγωγή δεδομένων για οποιοδήποτε επάγγελμα, αρκεί να είναι διαθέσιμο στους προαναφερθέντες ιστοτόπους.

### 6.2.2 Η υπηρεσία QualiChain Knowledge Extraction

Η υπηρεσία QualiChain Knowledge Extraction είναι υπεύθυνη για την αναγνώριση και εξαγωγή γνώσεων και δεξιοτήτων μέσα από το ελεύθερο κείμενο των αγγελιών εργασίας. Λαμβάνει είσοδο από διαφορετικές πηγές δεδομένων σε διαφορετικές μορφοποιήσεις και στη συνέχεια κανονικοποιεί την είσοδο και παράγει μια κανονικοποιημένη έκδοση της σε μορφοποίηση JSON-LD με τη χρήση της οντολογίας δεξιοτήτων και πρόσληψης ανθρωπίνων πόρων (Skills and Recruitment Ontology – SARO) [16]. Στη συνέχεια, λαμβάνει χώρα η διαδικασία του χαρακτηρισμού των δεδομένων (data annotation process). Η φάση αυτή περιλαμβάνει μια αρχική επεξεργασία φυσικής γλώσσας, από διαχωρισμό κειμένου (text splitting), αναγνώριση μερών του λόγου (Part of Speech tagging – POS tagging) [17], και ονομαστική αναγνώριση οντότητας (Named Entity Recognition - NER) [18]. Το αποτέλεσμα είναι μια λίστα από γνώσεις και δεξιότητες για κάθε αγγελία εργασίας. Ακολούθως, αναγνωρίζονται και αφαιρούνται διπλότυπες εγγραφές. Τα αποτελέσματα αποθηκεύονται τόσο στην Οντολογία του QualiChain όσο και σε μια σχεσιακή βάση δεδομένων.

Στην παρακάτω εικόνα παρουσιάζονται οι ροές εργασίας της εξαγωγής δεδομένων, και του χαρακτηρισμού των γνώσεων και δεξιοτήτων. Συγκεκριμένα, ο χρήστης κάνει ένα αίτημα στην υπηρεσία JobCrawler, η οποία εξάγει αγγελίες εργασίας από διάφορους ιστοτόπους και το αποτέλεσμα εισάγεται ως είσοδος στην υπηρεσία Knowledge Extraction η οποία με τη σειρά της αναγνωρίζει τις γνώσεις και δεξιότητες μέσα από τα κείμενα των αγγελιών εργασίας, αφαιρεί διπλότυπες αγγελίες και αποθηκεύει τα αποτελέσματα.



Εικόνα 17 - Αρχιτεκτονική όψη των υπηρεσιών JobCrawler και εξόρυξης γνώσης που χρησιμοποιήθηκαν

### 6.2.3 Η υπηρεσία Association Rules Miner

Η υπηρεσία Association Rules Miner είναι υπεύθυνη για την αναγνώριση συσχετίσεων μεταξύ των διαφορετικών γνώσεων και δεξιοτήτων που εμφανίζονται στις διάφορες αγγελίες εργασίας. Εφόσον η αναγνώριση συσχετίσεων μεταξύ διαφορετικών γνώσεων και δεξιοτήτων δεν είναι μια εύκολη διαδικασία, αξίζει να παρουσιαστεί η μεθοδολογία που ακολουθήθηκε σε μεγαλύτερο βάθος. Η εξόρυξη κανόνων συσχετίσεων (association rules mining) [11], είναι μια μέθοδος μηχανικής μάθησης που βασίζεται στους κανόνες και στοχεύει στην ανακάλυψη ενδιαφέροντων συσχετίσεων μεταξύ μεταβλητών σε βάσεις δεδομένων μεγάλης κλίμακας. Έχει ως στόχο την αναγνώριση ισχυρών κανόνων συσχέτισης σε βάσεις δεδομένων μέσω της χρήσης διαφόρων μετρικών του μεγέθους του ενδιαφέροντος (interestingness). Στο πλαίσιο της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών οι μεταβλητές είναι οι γνώσεις και οι δεξιότητες, η βάση δεδομένων είναι το σύνολο των αγγελιών εργασίας που έχουν συγκεντρωθεί μέσω της υπηρεσίας JobCrawler και οι μετρικές ενδιαφέροντος είναι οι ακόλουθες:

- **Υποστήριξη (Support)** είναι η πιθανότητα δύο οι περισσότερες γνώσεις και δεξιότητες να εμφανίζονται μαζί σε μία αγγελία εργασίας.
- **Εμπιστοσύνη (Confidence)** είναι η πιθανότητα μια αγγελία εργασίας που περιέχει μια ή περισσότερες συγκεκριμένες γνώσεις και δεξιότητες να περιέχει επίσης και μια άλλη συγκεκριμένη γνώση ή δεξιότητα.
- **Ανύψωση (Lift):** αποτελεί μια μετρική που δείχνει εάν δύο γνώσεις ή δεξιότητες εμφανίζονται μαζί σε αγγελίες πιο συχνά από ό,τι θα περίμενε κανείς εάν εμφανίζονταν τυχαία και ορίζεται ως ο λόγος μεταξύ της υποστήριξης των συγκεκριμένων γνώσεων ή δεξιοτήτων προς την υποστήριξη που θα είχαν οι συγκεκριμένες γνώσεις ή δεξιότητες εάν ήταν ανεξάρτητες μεταξύ τους. Φυσικά αν η τιμή της ανύψωσης είναι μεγαλύτερη της μονάδας οι γνώσεις/δεξιότητες εμφανίζονται μαζί πιο συχνά από το αναμενόμενο.

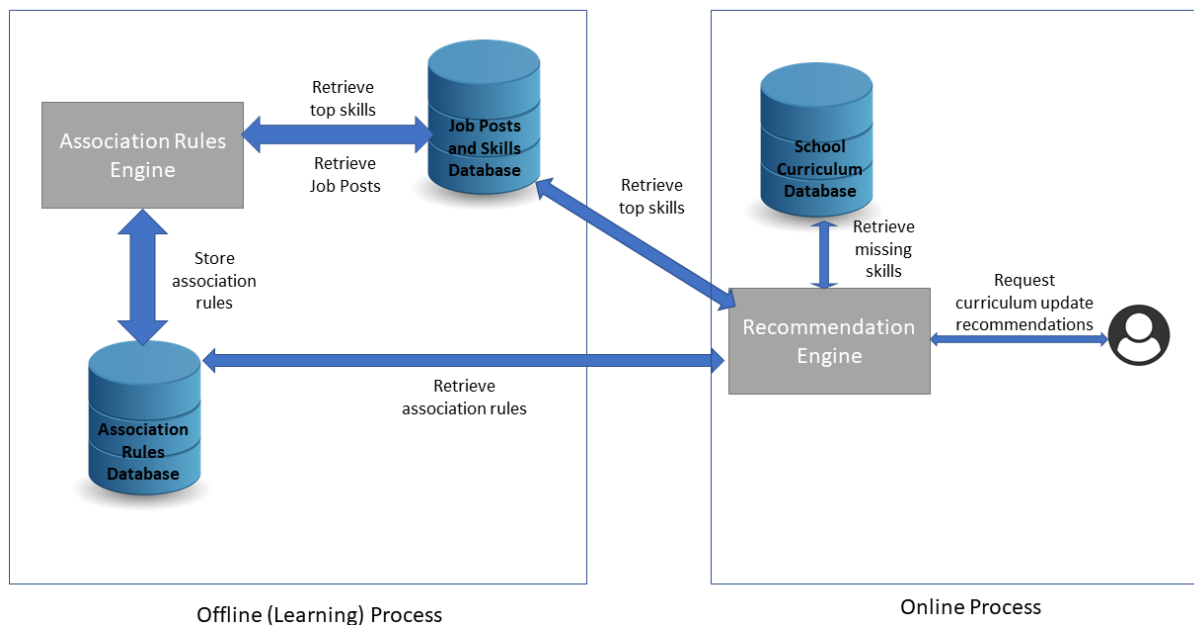
Είναι εύκολο να αντιληφθεί κανείς ότι δεν είναι εφικτός ο υπολογισμός αυτών των μετρικών για κάθε πιθανό συνδυασμό όλων των διαφορετικών γνώσεων και δεξιοτήτων. Για το λόγο αυτό, εφαρμόστηκε ο αλγόριθμος Apriori [19]. Πιο συγκεκριμένα, ο αλγόριθμος αυτός αναγνωρίζει τα συχνότερα εμφανιζόμενα αντικείμενα (γνώσεις και δεξιότητες) στο σύνολο των αγγελιών εργασίας και τα επεκτείνει σε όλο και μεγαλύτερα σύνολα αντικειμένων (γνώσεων και δεξιοτήτων), με την προϋπόθεση ότι εμφανίζονται επαρκώς συχνά στο σύνολο των αγγελιών εργασίας. Στην περίπτωση της συγκεκριμένης υπηρεσίας, ορίζουμε την επαρκή συχνότητα ως ένα κατώφλι στην υποστήριξη ενός συνόλου αντικειμένων. Το συγκεκριμένο κατώφλι, κατόπιν πειραματισμού ορίστηκε στο 4%. Έτσι, σύνολα αντικειμένων με υποστήριξη μικρότερη του 4% απορρίπτονται.

### 6.2.4 Η αρχιτεκτονική της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών

Όσον αφορά την εκτέλεση της εφαρμογής ανασχεδιασμού προγράμματος σπουδών, εφόσον η όλη διαδικασία είναι πολύπλοκη και απαιτεί σημαντικό χρόνο για να εκτελεστεί, η εκτέλεση της είναι προφανές ότι δεν μπορεί να γίνεται με κάθε αίτημα στην υπηρεσία. Αντίθετα, η διαδικασία χωρίζεται σε δυο φάσεις την φάση εκτός σύνδεσης ή φάση εκμάθησης (offline or learning phase) και τη φάση σε σύνδεση (online phase). Κατά την φάση εκμάθησης οι κανόνες συσχέτισης υπολογίζονται και αποθηκεύονται σε μια βάση δεδομένων και σε έναν αποθηκευτικό χώρο αρχείων. Για λόγους απλότητας

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

και επαναχρησιμοποίησης, και έχοντας υπόψη ότι η φάση συγκέντρωσης δεδομένων και εξαγωγής δεξιοτήτων από αυτά (data acquisition and skill extraction phase) είναι ανεξάρτητη από τη φάση εξόρυξης κανόνων συσχετίσεων, τα αποτελέσματα από τη φάση συγκέντρωσης δεδομένων και εξαγωγής δεξιοτήτων από αυτά χρησιμοποιούνται κατευθείαν, αντί να εκτελούνται οι υπηρεσίες εξαγωγής δεδομένων από το διαδίκτυο και η υπηρεσία εξαγωγής γνώσης από τα δεδομένα, μιας και για να εκτελεστούν οι τελευταίες απαιτείται σημαντικός χρόνος και μνήμη. Από την άλλη πλευρά, κατά την online διαδικασία στη βάση δεδομένων με το δοθέν πρόγραμμα σπουδών ή τη δοθείσα λίστα μαθημάτων, όπως επίσης και στη βάση δεδομένων για την αγορά εργασίας εκτελούνται ερωτήματα για την εξαγωγή τόσο των πιο επιθυμητών γνώσεων και δεξιοτήτων από την αγορά εργασίας όσο και των πιο σημαντικών γνώσεων και δεξιοτήτων που λείπουν από το δοθέν πρόγραμμα σπουδών ή τη δοθείσα λίστα μαθημάτων. Για τις γνώσεις και δεξιότητες που λείπουν από το δοθέν πρόγραμμα σπουδών, εκτελούνται ερωτήματα στη βάση δεδομένων με τους κανόνες συσχετίσεων, για την ανάκτηση γνώσεων και δεξιοτήτων που προσφέρονται στο δοθέν πρόγραμμα σπουδών και σχετίζονται με την γνώση/δεξιότητα προς ένταξη σε αυτό. Σε περίπτωση εύρεσης τέτοιων συσχετίσεων, οι γνώσεις/δεξιότητες προς ένταξη συνδέονται με μαθήματα που προσφέρουν σχετικές γνώσεις και δεξιότητες και ως εκ τούτου συνίσταται η εισαγωγή των γνώσεων και δεξιοτήτων προς ένταξη στα σχετικά μαθήματα, μαζί με τη βαθμολογία ταιριάσματος. Σε περίπτωση που δεν βρεθούν σχετικές γνώσεις και δεξιότητες με τις γνώσεις και δεξιότητες προς ένταξη στο δοθέν πρόγραμμα σπουδών ή στη δοθείσα λίστα μαθημάτων, προτείνεται η δημιουργία νέου μαθήματος που προσφέρει τις γνώσεις και δεξιότητες προς ένταξη παράλληλα με άλλες σχετικές γνώσεις και δεξιότητες. Η διαδικασία που περιεγράφηκε αποδίδεται σχηματικά στην παρακάτω εικόνα.



Εικόνα 18 – Αρχιτεκτονική όψη και διαδικασίες της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών

### 6.3 Ενδεικτικά Αποτελέσματα

Η χρήση της υπηρεσίας JobCrawler έδωσε ως αποτέλεσμα 1777 αγγελίες εργασίας από την ελληνική αγορά εργασίας στον τομέα του λογισμικού. Η υπηρεσία QualiChain Knowledge Extraction εκτελέστηκε στις προαναφερθείσες αγγελίες εργασίας και τελικά εξήγαγε 345 διαφορετικές γνώσεις και δεξιότητες, οι οποίες κατατάσσονται σε μια από τις τρεις διαθέσιμες κατηγορίες. Οι κατηγορίες είναι οι εξής:

**Θέμα (topic):** η κατηγορία αυτή περιλαμβάνει γενικευμένες γνώσεις και δεξιότητες που περιγράφουν γενική ικανότητα σε κάποιο αντικείμενο (π.χ. προγραμματισμός).

**Προϊόν (product):** η κατηγορία αυτή περιλαμβάνει γνώσεις και δεξιότητες σε συγκεκριμένα προϊόντα, όπως για παράδειγμα το Microsoft Office.

**Εργαλείο (tool):** η κατηγορία αυτή περιλαμβάνει γνώσεις και δεξιότητες στη χρήση δημοφιλών εργαλείων που χρησιμοποιούνται για το πεδίο που μελετούμε (π.χ. κάποια γλώσσα προγραμματισμού).

Εφόσον, η κατηγορία "θέμα" περιέχει πολύ γενικευμένες γνώσεις και δεξιότητες όπως ο προγραμματισμός, η ανάπτυξη λογισμικού και ο έλεγχος ορθότητας (testing), αποφασίστηκε να μη ληφθεί υπόψη για την υπηρεσία ανασχεδιασμού προγράμματος σπουδών. Έτσι, ένας σημαντικός αριθμός από τις αναγνωρισμένες γνώσεις και δεξιότητες απορρίφθηκαν. Επιπλέον, απορρίφθηκαν γνώσεις και δεξιότητες που ήταν πάρα πολύ ειδικές, αφού βρέθηκαν σε πολύ μικρό αριθμό αγγελιών εργασίας.

Τελικά, από τις 345 γνώσεις και δεξιότητες που αναγνωρίστηκαν από την υπηρεσία QualiChain Knowledge Extraction, μόνο 57 χρησιμοποιήθηκαν για την ανάλυση και τη λειτουργία της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών. Από την άλλη πλευρά, γύρω στις 200 γνώσεις και δεξιότητες αναγνωρίστηκαν στο πρόγραμμα σπουδών της σχολής Η.Μ.Μ.Υ. του Ε.Μ.Π. που είναι σχετικές με την επιστήμη των υπολογιστών και την επιστήμη των δεδομένων. Φυσικά η επιστήμη των υπολογιστών και η επιστήμη των δεδομένων είναι πιο ευρείες από τον τομέα της ανάπτυξης λογισμικού. Ωστόσο, οι επιστήμες αυτές σε συνδυασμό αποτελούν υπερσύνολο του τομέα της ανάπτυξης λογισμικού και στόχος είναι ο εντοπισμός γνώσεων και δεξιοτήτων ανάπτυξης λογισμικού που λείπουν από το πρόγραμμα σπουδών, που σημαίνει ότι δεν υπάρχει λόγος να διαχωρίσουμε περαιτέρω το πρόγραμμα σπουδών. Το τελικό αποτέλεσμα έδειξε ότι 23 από τις 57 γνώσεις και δεξιότητες με την υψηλότερη ζήτηση δεν συμπεριλαμβάνονται στο πρόγραμμα σπουδών. Αυτές οι γνώσεις και δεξιότητες παρουσιάζονται παρακάτω (Πίνακας 21):

Πίνακας 21 - Οι γνώσεις και δεξιότητες σε υψηλή ζήτηση που δε προσφέρονται στο δοθέν πρόγραμμα σπουδών

Javascript	git	css
Docker	json	go
R	android	html5
Scripting	angular	linux
css3	tcp	aws

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Api	rest	jenkins
React	ios	php
Akka	windows	

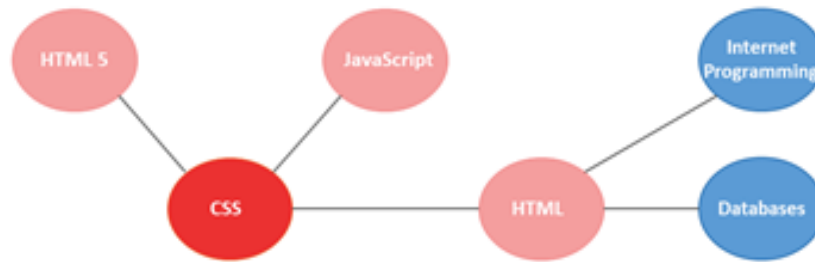
Φυσικά, η εύρεση των γνώσεων και δεξιοτήτων που είναι σε υψηλή ζήτηση από την αγορά και δε προσφέρονται από το δοθέν πρόγραμμα σπουδών είναι πολύ σημαντικό βήμα. Ωστόσο, η δυσκολότερη πρόκληση είναι η πρόταση συγκεκριμένων μαθημάτων για την ένταξη των συγκεκριμένων γνώσεων και δεξιοτήτων. Όπως έχει ήδη αναφερθεί, για αυτό το σκοπό χρησιμοποιήθηκε η υπηρεσία Association Rules Miner και η τεχνική εξόρυξης κανόνων συσχέτισης. Το αποτέλεσμα της συγκεκριμένης υπηρεσίας είναι πολύ μακροσκελές καθώς υπάρχουν πολλά μαθήματα και γνώσεις/δεξιότητες προς εισαγωγή, για το λόγο αυτό δε θα παρουσιαστούν ολοκληρωμένα στα πλαίσια του παρόντος κεφαλαίου, αντίθετα στον επόμενο πίνακα παρουσιάζονται μόνο μερικοί από τους κανόνες συσχέτισης που εξάχθηκαν. Η πλήρης λίστα παρουσιάζεται στο Παράρτημα VI.

Πίνακας 22 - Μερικοί από τους κανόνες συσχέτισης που εξάχθηκαν

Antecedents	Consequents	Support	Confidence	Lift
CSS3	HTML5	0.046	0.965	10.78
HTML5, CSS	JavaScript	0.045	0.964	5.826
HTML, JavaScript	CSS	0.042	0.872	8.658
Jenkins	Git	0.051	0.857	3.635
CSS	JavaScript	0.083	0.821	4.963
Android	IOS	0.041	0.802	10.182
Docker	Git	0.062	0.797	3.381
JSON	REST	0.051	0.728	6.045
NoSQL	Python	0.041	0.598	4.945
IOS	Android	0.041	0.521	10.182
Java	SQL	0.055	0.447	2.557
Java	JavaScript	0.04	0.329	1.988

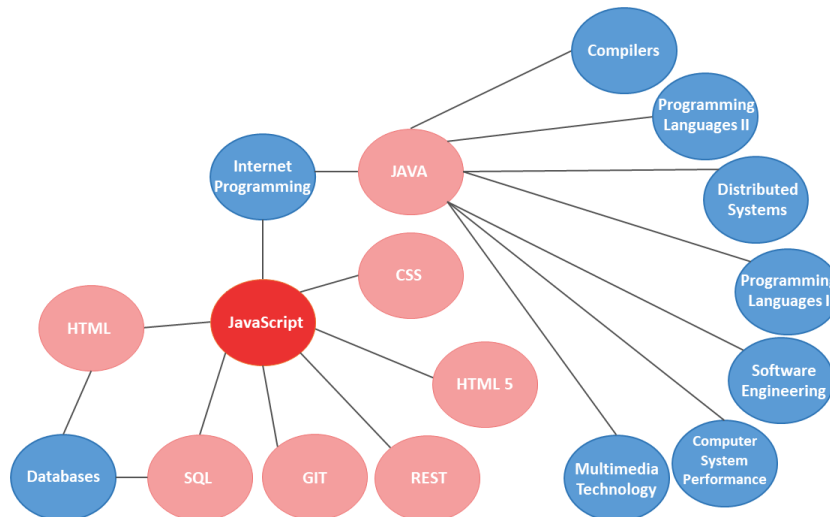
Στις παραπάνω εικόνες παρουσιάζονται συγκεκριμένα παραδείγματα των αποτελεσμάτων της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών, με τη χρήση γραφικής αναπαράστασης γράφων. Οι παρακάτω εικόνες παρουσιάζουν τις πολιτικές που προτείνονται για τις δεξιότητες css, javascript και git. Αξίζει να σημειωθεί ότι η γνώση/δεξιότητα προς ένταξη στο πρόγραμμα σπουδών αναπαρίσταται με κόκκινο χρώμα, οι σχετικές γνώσεις και δεξιότητες παρουσιάζονται με ροζ, ενώ το μπλε χρώμα αναπαριστά τα μαθήματα που συσχετίζονται με τη δεξιότητα προς εισαγωγή. Οι ακμές του γράφου δείχνουν πώς συσχετίζεται η γνώση/δεξιότητα προς εισαγωγή με τα μαθήματα στα οποία προτείνεται. Συγκεκριμένα, οι ακμές μεταξύ γνώσεων/δεξιοτήτων δείχνουν ότι υπάρχει κανόνας συσχέτισης μεταξύ των δυο γνώσεων/δεξιοτήτων, ενώ η ακμή που συνδέει ένα μάθημα με μια γνώση/δεξιότητα, δείχνει ότι η γνώση/δεξιότητα προσφέρεται στο συγκεκριμένο μάθημα.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 19 - Τα αποτελέσματα της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών για τη δεξιότητα CSS

Όπως φαίνεται στην εικόνα 3, η γλώσσα προγραμματισμού CSS που είναι σε μεγάλη ζήτηση από την αγορά εργασίας αλλά δε διδάσκεται από το δοθέν πρόγραμμα σπουδών, σχετίζεται με τις γλώσσες προγραμματισμού HTML, HTML5 και JavaScript. Από αυτές τις γνώσεις/δεξιότητες μόνο η HTML προσφέρεται στο συγκεκριμένο πρόγραμμα σπουδών στα μαθήματα των βάσεων δεδομένων και του διαδικτυακού προγραμματισμού. Έτσι, το σύστημα προτείνει την ένταξη της γλώσσας προγραμματισμού CSS σε κάποιο από τα προαναφερθέντα μαθήματα. Βέβαια τα αποτελέσματα δεν είναι τόσο απλά για κάθε γνώση/δεξιότητα. Για παράδειγμα, σε κάποιες περιπτώσεις υπάρχουν πολλές σχετικές γνώσεις/δεξιότητες με τη γνώση/δεξιότητα που λείπει από το πρόγραμμα σπουδών και κατά συνέπεια και αρκετά μαθήματα στα οποία θα μπορούσε να ενταχθεί. Τέτοια παραδείγματα αποτελούν οι γνώσεις/δεξιότητες JavaScript και Git, όπως φαίνεται και στις παρακάτω εικόνες.



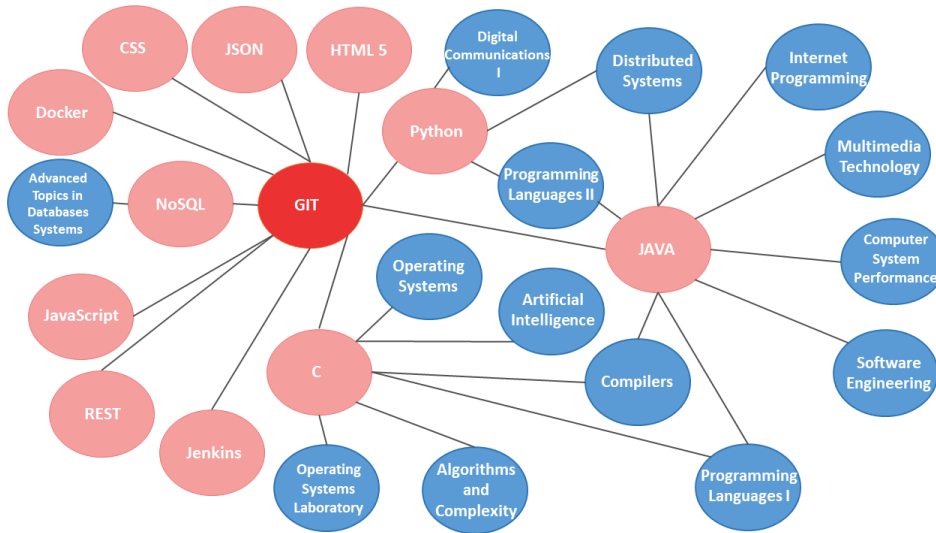
Εικόνα 20 - Τα αποτελέσματα της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών για τη δεξιότητα JavaScript

Όπως φαίνεται στην εικόνα 4, η γλώσσα προγραμματισμού JavaScript έχει πολλές σχετικές γνώσεις και δεξιότητες. Πιο συγκεκριμένα συνδέεται με τις γνώσεις/δεξιότητες CSS, HTML, HTML5, SQL, GIT, REST και τη γλώσσα προγραμματισμού JAVA. Παρότι κάποιες από τις προαναφερθείσες γνώσεις/δεξιότητες δε προσφέρονται στο πλαίσιο κάποιου μαθήματος, η γλώσσα προγραμματισμού SQL διδάσκεται στο μάθημα των βάσεων δεδομένων, η HTML στις βάσεις δεδομένων και στον διαδικτυακό προγραμματισμό,

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

ενώ η γλώσσα JAVA διδάσκεται στο διαδικτυακό προγραμματισμό, στους μεταγλωττιστές, στις γλώσσες προγραμματισμού I και II, στα κατανεμημένα συστήματα, στην τεχνολογία λογισμικού, στην επίδοση υπολογιστικών συστημάτων και στην τεχνολογία πολυμέσων, και άρα η γλώσσα JavaScript είναι σχετική και μπορεί να διδαχθεί σε οποιοδήποτε από τα παραπάνω μαθήματα.

Όσον αφορά το εργαλείο GIT ο γράφος είναι ακόμα πιο πολύπλοκος όπως φαίνεται και στην εικόνα 5.



Εικόνα 21 - Τα αποτελέσματα της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών για τη δεξιότητα Git

Συγκεκριμένα, η παραπάνω εικόνα δείχνει ότι το εργαλείο Git σχετίζεται με όλες τις γλώσσες προγραμματισμού και κατά συνέπεια με όλα τα μαθήματα που ασχολούνται με τον προγραμματισμό. Έτσι, οι σχετικές γνώσεις/δεξιότητες με το εργαλείο GIT είναι οι γλώσσες προγραμματισμού Python, JAVA, JavaScript, C, SQL, καθώς επίσης και οι γνώσεις/δεξιότητες REST, NoSQL, JSON, CSS, Jenkins, HTML5 και Docker και τα μαθήματα όπου θα μπορούσε να διδαχθεί το εργαλείο GIT περιλαμβάνουν τις γλώσσες προγραμματισμού I και II, το διαδικτυακό προγραμματισμό, τα κατανεμημένα συστήματα, την τεχνολογία πολυμέσων, την τεχνολογία λογισμικού, την επίδοση υπολογιστικών συστημάτων, τους μεταγλωττιστές, τα λειτουργικά συστήματα και το εργαστήριο λειτουργικών συστημάτων, το μάθημα των αλγορίθμων και πολυπλοκότητας, την τεχνητή νοημοσύνη, τα προχωρημένα θέματα βάσεων δεδομένων και τις ψηφιακές τηλεπικοινωνίες.

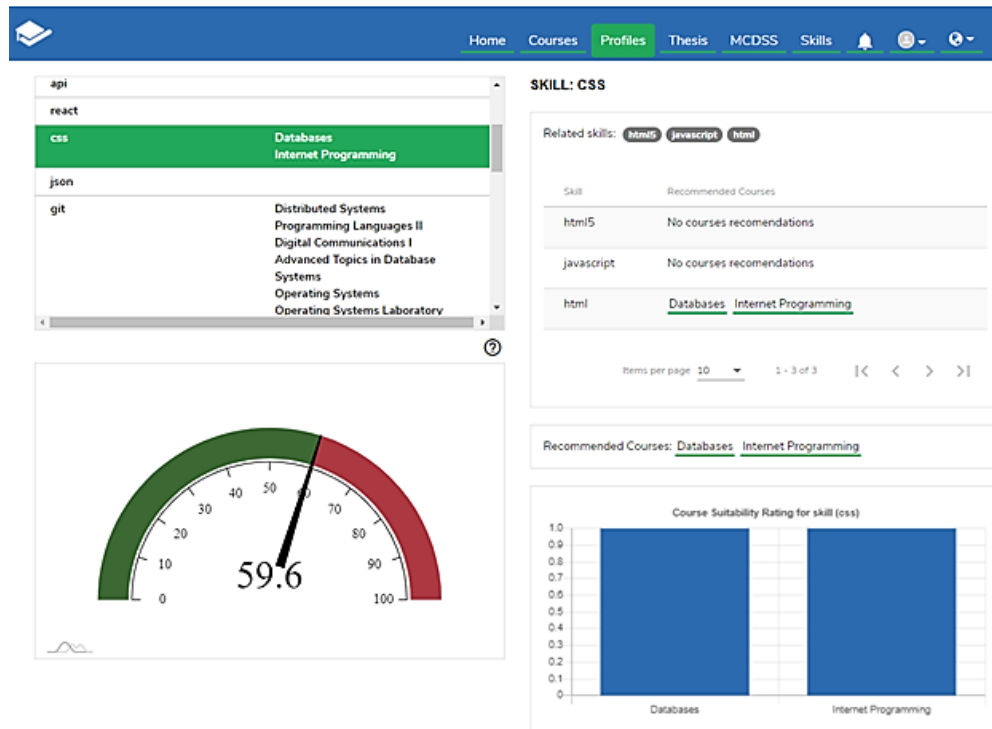
Επιπροσθέτως, μερικές ακόμα προτάσεις του συστήματος είναι η δημιουργία μαθήματος που θα επικεντρώνεται στην ανάπτυξη εφαρμογών έξυπνων κινητών τηλεφώνων. Πιο συγκεκριμένα, οι γνώσεις/δεξιότητες iOS και Android είναι στενά συνδεδεμένες και σε μεγάλη ζήτηση από την αγορά εργασίας ενώ δε προσφέρονται στα πλαίσια του δοθέντος προγράμματος σπουδών και ούτε υπάρχει κάποιο σχετικό μάθημα για την εισαγωγή τους στην ύλη του. Συνεπώς, η δημιουργία τέτοιου μαθήματος θα πρόσθετε σημαντική αξία στο πρόγραμμα σπουδών. Παρόμοιο παράδειγμα αποτελούν οι γνώσεις/δεξιότητες REST APIs, και JSON που συνδέονται με τη γλώσσα προγραμματισμού JavaScript, ενώ η ζήτηση για προγραμματιστές APIs είναι πολύ μεγάλη. Έτσι, η προσθήκη μαθήματος για ανάπτυξη APIs θα πρόσθετε επίσης σημαντική αξία στο δοθέν πρόγραμμα σπουδών.



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## 6.4 Επίδειξη υπηρεσίας και διεπαφής χρήστη

Στην παρούσα ενότητα παρουσιάζεται η διεπαφή που αναπτύχθηκε για την εφαρμογή ανασχεδιασμού προγράμματος σπουδών, για την ελληνική αγορά εργασίας στην ανάπτυξη λογισμικού και το πρόγραμμα σπουδών της σχολής Η.Μ.Μ.Υ. του Ε.Μ.Π. Στην εικόνα 22 φαίνεται η διεπαφή της εφαρμογής. Συγκεκριμένα, βλέπουμε ότι περίπου 60% των γνώσεων/δεξιοτήτων που είναι σε υψηλή ζήτηση από την αγορά εργασίας προσφέρονται από το δοθέν πρόγραμμα σπουδών, ενώ πάνω αριστερά στη σελίδα φαίνονται οι γνώσεις/δεξιότητες σε υψηλή ζήτηση που δε προσφέρονται στα πλαίσια κάποιου μαθήματος από το δοθέν πρόγραμμα σπουδών. Μαζί με τις γνώσεις/δεξιότητες αυτές παρέχονται και κάποιες προτάσεις για μαθήματα στα οποία θα μπορούσαν να ενταχθούν.



Εικόνα 22 - Η διεπαφή χρήστη της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών με έμφαση στη δεξιότητα CSS

Πιο συγκεκριμένα οι δεξιότητες APIs, React και JSON δε φαίνεται να σχετίζονται με κανένα μάθημα της σχολής Η.Μ.Μ.Υ. η γλώσσα CSS όπως παρουσιάστηκε και προηγουμένως είναι σχετική με τα μαθήματα των βάσεων δεδομένων και του διαδικτυακού προγραμματισμού (κάτω δεξιά στη σελίδα), με ίδιο βαθμό συσχέτισης, όσο και με τις δεξιότητες HTML5, JavaScript και HTML (πάνω δεξιά στη σελίδα). Από την άλλη πλευρά το GIT προτείνεται για ένταξη σε πληθώρα μαθημάτων, όπως έχει αναφερθεί και σε προηγούμενη ενότητα.

## 6.5 Σχολιασμός και συμπεράσματα

Το παρόν κεφάλαιο παρουσιάζει την υπηρεσία συστάσεων ανασχεδιασμού προγράμματος σπουδών, η οποία παρέχει προτάσεις για την εισαγωγή γνώσεων και δεξιοτήτων σε ένα δοθέν πρόγραμμα σπουδών ή μια λίστα μαθημάτων. Οι συγκεκριμένες γνώσεις και δεξιότητες προτείνονται διότι αφ' ενός είναι σε

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

μεγάλη ζήτηση από την αγορά εργασίας και αφ' ετέρου δε προσφέρονται στο δοθέν πρόγραμμα σπουδών ή στη δοθείσα λίστα μαθημάτων.

Στο κεφάλαιο αυτό παρουσιάζεται η λειτουργία του συστήματος για την αγορά εργασίας στον επιστημονικό κλάδο της ανάπτυξης λογισμικού, ενώ το πρόγραμμα σπουδών στο οποίο προτείνονται οι αλλαγές είναι το πρόγραμμα σπουδών της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου. Ωστόσο, η εφαρμογή μπορεί εύκολα να επεκταθεί για επιπλέον εκπαιδευτικούς οργανισμούς και ιδρύματα, ή και για διαφορετικούς επιστημονικούς κλάδους με μικρές τροποποιήσεις. Τα δεδομένα αποκτώνται από μια σειρά από ιστοτόπους αγγελιών εργασίας στην Ελλάδα μέσω μιας υπηρεσίας που βασίζεται στην απόξεση δεδομένων ιστοτόπου (web scraping). Στη συνέχεια, οι αγγελίες εργασίας που εξάχθηκαν, αναλύονται από την υπηρεσία εξαγωγής γνώσης του QualiChain, η οποία αναγνωρίζει τις βασικές γνώσεις και δεξιότητες που απαιτούνται από κάθε αγγελία εργασίας. Επίσης, με τον ίδιο τρόπο αναγνωρίζονται και οι βασικές γνώσεις και δεξιότητες που προσφέρονται σε κάθε μάθημα από το πρόγραμμα σπουδών που μελετάται. Στη συνέχεια εφαρμόζονται τεχνικές εξόρυξης συσχετίσεων (association rules mining) στις αγγελίες εργασίας για τον εντοπισμό συσχετίσεων μεταξύ διαφορετικών γνώσεων και δεξιοτήτων.

Τα αποτελέσματα φαίνονται λογικά και ικανοποιητικά για τους χρήστες. Επιπλέον, οι κανόνες συσχέτισης που εξάχθηκαν φαίνονται λογικές και παρόμοιες με τις συσχετίσεις που θα φανταζόταν ένας πεπειραμένος επαγγελματίας στο αντικείμενο της ανάπτυξης λογισμικού. Η υπηρεσία ανασχεδιασμού προγράμματος σπουδών μπορεί να χρησιμοποιηθεί συμβουλευτικά για την ενημέρωση ενός προγράμματος σπουδών κάποιου εκπαιδευτικού οργανισμού, ενώ ένας τέτοιος οργανισμός μπορεί να αντλήσει σημαντική προστιθέμενη αξία από τη χρήση της, καθώς το σύστημα είναι πάντα ενημερωμένο για τις τελευταίες εξελίξεις στην αγορά εργασίας. Περισσότερες πληροφορίες σχετικά με την αξιολόγηση και τη χρήση της συγκεκριμένης υπηρεσίας παρουσιάζονται στο κεφάλαιο 8.

## 6.6 Αναφορές

- [1] C. Kontzinos, V. Karakolis, S. Skalidakis, O. Markaki, P. Kokkinakos, and J. Psarras, "Combining blockchain, semantics and data analytics for university process optimization," *IADIS International Journal on Computer Science & Information Systems*, vol. 14, no. 2, pp. 59–77, 2019.
- [2] F. Zabliith and B. Azad, "Unpacking Course Overlap: Using Digital Tools to Incorporate Students' Voice in Curriculum Design," <https://doi.org/10.5465/AMBPP.2020.11846abstract>, vol. 2020, no. 1, p. 11846, Jul. 2020, doi: 10.5465/AMBPP.2020.11846ABSTRACT.
- [3] "Learning Analytics in the Classroom: Translating Learning Analytics Research ... - Βιβλία Google." [https://books.google.gr/books?hl=el&lr=&id=XiBtDwAAQBAJ&oi=fnd&pg=PT106&dq=%5B3%5D%09T.+Olney,+B.+Rienties,+and+L.+ToeteneL,+%22Gathering,+visualising+and+interpreting+learning+design+analytics+to+inform+classroom+practice+and+curriculum+design,%E2%80%9D+in+Le+arning+Analytics+in+the+Classroom:+Translating+Learning+Analytics+Research+for+Teach&ots=I0I009DoFo&sig=8YL-VWD\\_yOWvgtC2cNzWrFZQIDU&redir\\_esc=y#v=onepage&q&f=false](https://books.google.gr/books?hl=el&lr=&id=XiBtDwAAQBAJ&oi=fnd&pg=PT106&dq=%5B3%5D%09T.+Olney,+B.+Rienties,+and+L.+ToeteneL,+%22Gathering,+visualising+and+interpreting+learning+design+analytics+to+inform+classroom+practice+and+curriculum+design,%E2%80%9D+in+Le+arning+Analytics+in+the+Classroom:+Translating+Learning+Analytics+Research+for+Teach&ots=I0I009DoFo&sig=8YL-VWD_yOWvgtC2cNzWrFZQIDU&redir_esc=y#v=onepage&q&f=false) (accessed May 02, 2023).
- [4] S. Hai-Jew, "Designing Integrated Learning Paths for Individual Lifelong Learners and/or Small Groups: Backwards Curriculum Design From Target Complex-Skill Capabilities (for Nonformal

- Informal Learning),” <https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-5225-9833-6.ch005>, pp. 68–104, Jan. 1AD, doi: 10.4018/978-1-5225-9833-6.CH005.
- [5] J. K. Chang, N. L. Tsao, C. H. Kuo, and H. H. Hsu, “Curriculum design - A data-driven approach,” *Proceedings - 2016 5th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2016*, pp. 492–496, Aug. 2016, doi: 10.1109/IIAI-AAI.2016.163.
- [6] C. Kontzinos, V. Karakolis, S. Skalidakis, O. Markaki, M. Androutopoulou, and J. Psarras, “University process optimisation through smart curriculum design and blockchain-based student accreditation,” in *Proceedings of 18th International Conference on WWW/Internet, Torino, 2019*.
- [7] C. Romero and S. Ventura, “Data mining in education,” *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 3, no. 1, pp. 12–27, Jan. 2013, doi: 10.1002/WIDM.1075.
- [8] B. Daniel, “Big Data and analytics in higher education: Opportunities and challenges,” *British Journal of Educational Technology*, vol. 46, no. 5, pp. 904–920, Sep. 2015, doi: 10.1111/BJET.12230.
- [9] C. Brândaș and C. Pânzaru, “Analysing of the Labour Demand and Supply Using Web Mining and Data Mining,” in *Big Data and the Complexity of Labour Market Policies: New Approaches in Regional and Local Labour Market Monitoring for Reducing Skills Mismatches*, in München: Hampp Verlag, 2015, pp. 71–87.
- [10] I. A. Wowczko, “Skills and Vacancy Analysis with Data Mining Techniques,” *Informatics 2015, Vol. 2, Pages 31-49*, vol. 2, no. 4, pp. 31–49, Nov. 2015, doi: 10.3390/INFORMATICS2040031.
- [11] Q. Zhao and S. B. Sourav, “Association rule mining: A survey,” *Nanyang Technological University*, vol. 135. Singapore, 2003.
- [12] “Kariera Hiring Solutions.” <https://www.kariera.gr> (accessed Apr. 09, 2021).
- [13] “Indeed.” <https://gr.indeed.com/> (accessed Apr. 09, 2021).
- [14] “Skywalker.” <https://www.skywalker.gr> (accessed Apr. 09, 2021).
- [15] “Scrapy | A Fast and Powerful Scraping and Web Crawling Framework.” <https://scrapy.org/> (accessed Apr. 09, 2021).
- [16] E. M. Sibarani, S. Scerri, C. Morales, S. Auer, and D. Collarana, “Ontology-guided job market demand analysis: A cross-sectional study for the data science field,” *ACM International Conference Proceeding Series*, vol. 2017-September, pp. 25–32, Sep. 2017, doi: 10.1145/3132218.3132228.
- [17] C. D. Manning, “Part-of-speech tagging from 97% to 100%: Is it time for some linguistics?,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6608 LNCS, no. PART 1, pp. 171–189, 2011, doi: 10.1007/978-3-642-19400-9\_14/COVER.
- [18] B. Mohit and I. Zitouni, “Named Entity Recognition,” pp. 221–245, 2014, doi: 10.1007/978-3-642-45358-8\_7.
- [19] C. Borgelt and R. Kruse, “Induction of Association Rules: Apriori Implementation,” in *Compstat, Heidelberg: Physica-Verlag HD, 2002*. doi: 10.1007/978-3-642-57489-4\_59.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## **ΚΕΦΑΛΑΙΟ 7: Η ΥΠΗΡΕΣΙΑ ΠΟΛΥΚΡΙΤΗΡΙΑΚΗΣ ΥΠΟΣΤΗΡΙΞΗΣ ΑΠΟΦΑΣΕΩΝ**

### **7.1 Εισαγωγή**

Οι υπηρεσίες για σύσταση μαθημάτων και δεξιοτήτων και για ανασχεδιασμό του προγράμματος σπουδών που αναπτύχθηκαν στο πλαίσιο της παρούσας διδακτορικής διατριβής και παρουσιάστηκαν στα προηγούμενα κεφάλαια, βοηθούν ουσιαστικά τους χρήστες να λάβουν αποφάσεις βασισμένες στη γνώση για την πραγματική εικόνα των αναγκών της αγοράς εργασίας. Ωστόσο, σε πολλές περιπτώσεις, οι τελευταίοι χρειάζεται να λάβουν υπόψη τους επιπλέον κριτήρια, πέρα από τη ζήτηση από την αγορά εργασίας για συγκεκριμένες γνώσεις και δεξιότητες προκειμένου να λάβουν τις αποφάσεις τους, τα οποία πολλές φορές μπορεί να είναι και υποκειμενικά.

Για παράδειγμα, στο πρόβλημα της επιλογής μαθημάτων από έναν φοιτητή το σύστημα συστάσεων θα του προτείνει μια λίστα μαθημάτων, η οποία αφ' ενός θα περιέχει γνώσεις και δεξιότητες συμβατές με τις ανάγκες της αγοράς εργασίας και αφ' ετέρου με τη μέχρι τώρα εμπειρία του, όπως αυτή παρουσιάζεται στο βιογραφικό του. Ωστόσο, ο φοιτητής μπορεί να θεωρεί ότι υπάρχουν ανερχόμενες γνώσεις και δεξιότητες που θα ήθελε να μάθει μιας και πιστεύει ότι στο μέλλον θα είναι και αυτές σε υψηλή ζήτηση. Επιπροσθέτως, παρότι η υπηρεσία ανασχεδιασμού προγράμματος σπουδών μπορεί να προτείνει την ένταξη κάποιας γνώσης/δεξιότητας σε ένα συγκεκριμένο μάθημα, οι διδάσκοντες του μαθήματος μπορεί να είναι διστακτικοί να προχωρήσουν με την αλλαγή για λόγους που η υπηρεσία δεν λαμβάνει υπόψη. Τέτοιο παράδειγμα μπορεί να είναι ο φόρτος εργασίας του μαθήματος.

Φυσικά, τα συστήματα συστάσεων που αναπτύχθηκαν έχουν συμβουλευτικό ρόλο και οι χρήστες μπορούν κάλλιστα να διαλέξουν κάποια εναλλακτική η οποία βρίσκεται χαμηλότερα (ή δεν βρίσκεται καθόλου) στη λίστα προτάσεων του συστήματος συστάσεων. Ωστόσο, αυτή η προσέγγιση πολλές φορές αφήνει τους χρήστες αβέβαιους για την επιλογή τους. Για το λόγο αυτό, κρίθηκε απαραίτητο να παρασχεθούν στους χρήστες υπηρεσίες για επιπλέον υποστήριξη στις αποφάσεις τους, ώστε να μπορούν να ποσοτικοποιήσουν όλα τα κριτήρια που οι ίδιοι θεωρούν σημαντικά προκειμένου να προχωρήσουν στην τελική τους απόφαση με σιγουριά.

Για την υποστήριξη των χρηστών με τις αποφάσεις τους επιλέχθηκε η χρήση μεθόδων ανάλυσης αποφάσεων πολλαπλών κριτηρίων (Multi-Criteria Decision Support Systems - MCDSS). Η ανάλυση αποφάσεων πολλαπλών κριτηρίων αποτελεί κλάδο της επιστήμης της επιχειρησιακής έρευνας (Operations Research), η οποία αξιολογεί και ποσοτικοποιεί πολλαπλά κριτήρια για τη λήψη μιας απόφασης τα οποία είναι συχνά αντικρουόμενα [1].

### **7.2 Πολυκριτηριακή Υποστήριξη Αποφάσεων και Συστήματα Συστάσεων**

Τα συστήματα συστάσεων μπορούν να θεωρηθούν και ως συστήματα υποστήριξης αποφάσεων καθώς διευκολύνουν το χρήστη να επιλέξει κάποιο αντικείμενο (ή κάποια αντικείμενα) από μια πολυπληθή λίστα αντικειμένων με βάση κάποια συνάρτηση χρησιμότητας που λαμβάνει υπόψη συγκεκριμένα χαρακτηριστικά [2]. Σημαντική απαίτηση για την επιτυχία ενός συστήματος συστάσεων είναι η

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

εμπιστοσύνη των χρηστών στις προτάσεις του συστήματος και ως εκ τούτου η διαφάνεια γύρω από τη λειτουργία του.

Σε αυτό το πλαίσιο και λαμβάνοντας υπόψη ότι σε αρκετές περιπτώσεις είναι αναγκαίο να λαμβάνονται υπόψη πολλαπλά και αντικρουόμενα κριτήρια, πολλές εργασίες ασχολούνται με την ένταξη μεθόδων πολυκριτηριακής υποστήριξης αποφάσεων στον αλγόριθμο κάποιου συστήματος συστάσεων.

Για παράδειγμα, η εργασία [3] παρουσιάζει ένα σύστημα συστάσεων που βασίζεται σε πολλαπλά κριτήρια και χρησιμοποιεί τη μέθοδο πολυκριτηριακής υποστήριξης αποφάσεων PROMETHEE σε ένα σύστημα συστάσεων για την επιλογή μεταχειρισμένου αυτοκινήτου, λαμβάνοντας υπόψη όλα τα κριτήρια που η πλειονότητα θεωρεί σημαντικά για την επιλογή αυτοκινήτου. Τα αποτελέσματα αξιολογούνται με βάση την ακρίβεια (precision) και το βαθμό ανάκλησης (recall) και κρίνονται αρκετά ικανοποιητικά.

Η εργασία [4] παρουσιάζει ένα σύστημα συστάσεων για την επιλογή ειδικών στην επίλυση προβλημάτων στη βιομηχανική συντήρηση (industrial maintenance). Στην προκειμένη περίπτωση, η μέθοδος PROMETHEE II χρησιμοποιείται για να αναλύσει περαιτέρω τα αποτελέσματα που παρέχονται από το προτεινόμενο σύστημα συστάσεων. Το αποτέλεσμα αξιολογείται και πάλι με βάση την ακρίβεια (precision) και το επίπεδο ανάκλησης (recall) και κρίνεται ικανοποιητικό.

Παρόμοια προσέγγιση ακολουθούν και οι εργασίες [5] και [6]. Συγκεκριμένα, η πρώτη χρησιμοποιεί τη μέθοδο πολυκριτηριακής ανάλυσης ELECTRE για την παραγωγή προτάσεων σε φοιτητές σχετικά με τα πεδία που πρέπει να ακολουθήσουν στις σπουδές τους. Η δεύτερη χρησιμοποιεί τη μέθοδο ELECTRE για να εισάγει σημασιολογικά κριτήρια σε ένα σύστημα συστάσεων τουριστικών προορισμών.

Τέλος, η εργασία [7] παρουσιάζει ένα σύστημα συστάσεων το οποίο χρησιμοποιεί το σύνολο Pareto και εφαρμόζει τη μέθοδο PROMETHEE στο σύνολο Pareto, για την ακριβέστερη αξιολόγηση των αποτελεσμάτων, και την επίτευξη μεγαλύτερης διαφοροποίησης μεταξύ των αντικειμένων της λίστας προτάσεων. Τα πειραματικά αποτελέσματα της συγκεκριμένης έρευνας παρουσιάζουν βελτίωση τόσο σε ακρίβεια όσο και σε διαφοροποίηση, συγκριτικά με άλλους γνωστούς αλγορίθμους συστημάτων συστάσεων.

Αξίζει να σημειωθεί ότι για την επιτυχία ενός συστήματος συστάσεων ως σύστημα υποστήριξης αποφάσεων είναι αναγκαίο να είναι γνωστά εκ των προτέρων τα κριτήρια για τη λήψη μιας απόφασης, όσο και η σημαντικότητα του κάθε κριτηρίου. Ωστόσο, τόσο στην περίπτωση της υπηρεσίας συστάσεων μαθημάτων και δεξιοτήτων, όσο και στην υπηρεσία ανασχεδιασμού προγράμματος σπουδών που αναπτύχθηκαν, δεν ήταν γνωστά εκ των προτέρων τα κριτήρια των χρηστών. Επιπλέον, είναι λογικό ότι διαφορετικοί χρήστες έχουν και διαφορετικά κριτήρια. Κατά συνέπεια, αποφασίστηκε η δημιουργία μιας υπηρεσίας υποστήριξης αποφάσεων η οποία μπορεί να χρησιμοποιηθεί μαζί με τα αποτελέσματα από τα προαναφερθέντα συστήματα συστάσεων και στην οποία οι χρήστες θα μπορούν να θέσουν τα δικά τους υποκειμενικά κριτήρια.

### 7.3 Μέθοδοι πολυκριτηριακής υποστήριξης αποφάσεων

Υπάρχει πληθώρα μεθόδων πολυκριτηριακής υποστήριξης αποφάσεων. Η πλειοψηφία αυτών βασίζονται είτε στη θεωρία πολυκριτηριακής χρησιμότητας (multi-attribute utility theory) [8] είτε στη θεωρία σχέσεων υπεροχής (outranking relations theory) [9]. Επιπλέον, υπάρχουν μέθοδοι πολυκριτηριακής ανάλυσης αποφάσεων που βασίζονται στην προσέγγιση διαχωρισμού προτιμήσεων (preference disaggregation approach) [10]. Στην ενότητα αυτή παρουσιάζονται μερικές από τις σημαντικότερες μεθόδους πολυκριτηριακής ανάλυσης αποφάσεων.

#### 7.3.1 Η μέθοδος MAUT

Η μέθοδος MAUT (Multi-Attribute Utility Theory) [11], [12] αποτελεί ίσως την απλούστερη μέθοδο πολυκριτηριακής ανάλυσης αποφάσεων και αποτελεί μια δομημένη μεθοδολογία που αναπτύχθηκε αρχικά για το συμβιβασμό μεταξύ πολλαπλών αντικειμενικών συναρτήσεων. Στοχεύει στην ποσοτικοποίηση της έκφρασης προτιμήσεων των αποφασιζόντων σε μία ενιαία αντικειμενική συνάρτηση αδιάστατης κλίμακας, αξιοποιώντας όλες τις διαθέσιμες συναρτήσεις χρησιμότητας.

Η μέθοδος αυτή βασίζεται στην παραδοχή ότι σε κάθε πρόβλημα απόφασης υπάρχει μια πραγματική συνάρτηση χρησιμότητας  $U$  την οποία οι αποφασίζοντες επιθυμούν να μεγιστοποιήσουν.

Ο αλγόριθμος της μεθόδου MAUT παρουσιάζεται παρακάτω:

---

```
input:  $n$  (alternatives),  $q$  (criteria),  $w_i$  (weights),  $f_i(a_j)$  (evaluation table);
for all alternatives  $a_i, i \in \{1, \dots, n\}$  do
  for all criteria  $f_k, k \in \{1, \dots, q\}$  do
    | compute normalised decision matrix  $x_k(a_i)$ 
  end
end
for all alternatives  $a_i, i \in \{1, \dots, n\}$  do
  for all criteria  $f_k, k \in \{1, \dots, q\}$  do
    | compute integrated DM's attitude  $u_k(a_i)$ 
  end
end
for all alternatives  $a_i, i \in \{1, \dots, n\}$  do
  | compute utility  $U(i) = \text{sum}(w_k u_k(a_i))$ 
end
FinalRanking = SortDesc( $U$ )
Result: FinalRanking
```

---

Εικόνα 23 - Αλγόριθμος MAUT [13]

Συγκεκριμένα κατά την εκτέλεση της μεθόδου MAUT ακολουθούνται τα παρακάτω βήματα [13], [14]:

- 1) **Κανονικοποιημένος πίνακας απόφασης:** Έστω  $f_k(a_{\min})$ ,  $f_k(a_{\max})$  που αντιπροσωπεύουν την ελάχιστη και τη μέγιστη αξία για το κριτήριο  $k$ . Ο πίνακας αποτίμησης κανονικοποιείται ως εξής: Για κριτήρια προς μεγιστοποίηση:

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

$$x_k(a_i) = \frac{f_k(a_i) - f_k(a_{min})}{f_k(a_{max}) - f_k(a_{min})}$$

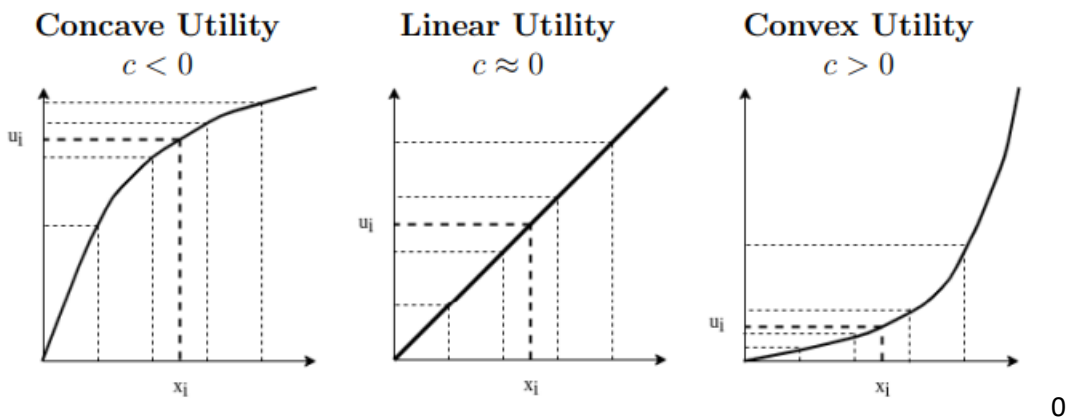
Και για κριτήρια προς ελαχιστοποίηση:

$$x_k(a_i) = \frac{f_k(a_{max}) - f_k(a_i)}{f_k(a_{max}) - f_k(a_{min})}$$

- 2) **Ενσωμάτωση της προτίμησης του αποφασίζοντα:** Η προτίμηση του αποφασίζοντα ενσωματώνεται στον πίνακα απόφασης με τον ακόλουθο τρόπο:

$$u_k(a_i) = \frac{1 - e^{cx_i}}{1 - e^c}$$

Όπου  $c$  είναι ένας αριθμός που συμβολίζει την προτίμηση του αποφασίζοντα όπως φαίνεται στην παρακάτω εικόνα [13]:



- 3) **Συνάρτηση χρησιμότητας:** η συνάρτηση χρησιμότητας (utility function) υπολογίζεται ως εξής:

$$U_i = \sum_{k=1}^q w_k u_k(a_i)$$

Η τελική κατάταξη της MAUT γίνεται με την ταξινόμηση των εναλλακτικών αποφάσεων με φθίνουσα σειρά με βάση την τιμή της συνάρτησης χρησιμότητας για κάθε εναλλακτική απόφαση.

### 7.3.2 Η μέθοδος TOPSIS

Η μέθοδος TOPSIS [15] (Technique for Order of Preference by Similarity to Ideal Solution) είναι μια μέθοδος πολυκριτηριακής ανάλυσης αποφάσεων που βασίζεται στη γεωμετρική απόσταση ανάμεσα στην θετική ιδανική λύση (Positive Ideal Solution - PIS) και στην αρνητική ιδανική λύση (Negative Ideal Solution - NIS). Μία καλή επιλογή ανάμεσα στις εναλλακτικές έχει μικρή απόσταση από την PIS και μεγάλη απόσταση από την NIS.

Ο αλγόριθμος της μεθόδου TOPSIS παρουσιάζεται παρακάτω:



---

```

input:  $n$  (alternatives),  $q$  (criteria),  $w_i$  (weights),  $f_i(a_j)$  (evaluation table);
for all alternatives  $a_i, i \in \{1, \dots, n\}$  do
  | for all criteria  $f_k, k \in \{1, \dots, q\}$  do
  | | compute normalised score  $r_k(a_i)$ 
  | end
end
for all alternatives  $a_i, i \in \{1, \dots, n\}$  do
  | for all criteria  $f_k, k \in \{1, \dots, q\}$  do
  | | compute weighted normalised score  $t_k(a_i)$ 
  | end
end
compute positive ideal solution  $A^+$ 
compute negative ideal solution  $A^-$ 
for all alternatives  $a_i, i \in \{1, \dots, n\}$  do
  | compute separation distance from positive ideal solution  $S^+(i)$ 
  | compute separation distance from negative ideal solution  $S^-(i)$ 
  | compute relative closeness to the positive ideal solution  $C^-(i)$ 
end
FinalRanking = SortDesc( $C$ )
Result: FinalRanking

```

---

Εικόνα 24 - Αλγόριθμος TOPSIS [13]

Συγκεκριμένα κατά την εκτέλεση της μεθόδου TOPSIS ακολουθούνται τα παρακάτω βήματα [13], [14]:

- 1) **Κανονικοποιημένος πίνακας αποφάσεων:** Δοθέντος ενός πίνακα αποφάσεων  $n \times q$  δημιουργείται ένας νέος κανονικοποιημένος πίνακας αποφάσεων μέσω της ακόλουθης μεθόδου κανονικοποίησης:

$$r_k(a_i) = \frac{x_k(a_i)}{\sqrt{\sum_{j=1}^n x_k(a_j)^2}}$$

- 2) **Κανονικοποιημένος πίνακας αποφάσεων με βάρη:** Στο βήμα αυτό ποσοτικοποιείται η σχετική σημαντικότητα για κάθε κριτήριο. Ο κανονικοποιημένος πίνακας αποφάσεων προκύπτει πολλαπλασιάζοντας κάθε στοιχείο κάθε στήλης του κανονικοποιημένου πίνακα αποφάσεων με τα αντίστοιχα αντισταθμίσιμα (offsets), ως εξής:

$$t_k(a_i) = r_k(a_i) * w_k, \text{ όπου } w_k = W_k / \sum_{j=1}^q W_j, \text{ με } j = 1, 2, \dots, q$$

- 3) **Θετική και αρνητική ιδανική λύση:** Η θετική ιδανική λύση  $A^+$  και η αρνητική ιδανική λύση  $A^-$  ορίζονται μέσω του κανονικοποιημένου πίνακα αποφάσεων ως εξής:

$$A^+ = \{(\min(t_k(a_i) \mid i = 1, 2, \dots, n) \mid k \in J_-), (\max(t_k(a_i) \mid i = 1, 2, \dots, n) \mid k \in J_+)\}$$

$$A^- = \{(\max(t_k(a_i) \mid i = 1, 2, \dots, n) \mid k \in J_-), (\min(t_k(a_i) \mid i = 1, 2, \dots, n) \mid k \in J_+)\}$$

Όπου για κριτήρια προς μεγιστοποίηση:

$$J_+ = \{k = 1, 2, \dots, q \mid k\}$$

Και για κριτήριο προς ελαχιστοποίηση:

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

$$J_- = \{k = 1, 2, \dots, q | k\}$$

- 4) **Απόσταση διαχωρισμού από την ιδανική και τη μη ιδανική λύση:** Έστω  $t_k^+$  η θετική ιδανική τιμή και  $t_k^-$  η αρνητική ιδανική τιμή για το κριτήριο  $k$ . Η απόσταση διαχωρισμού (L2 - Distance) για κάθε εναλλακτική από την ιδανική και τη μη ιδανική λύση υπολογίζεται ως εξής:

$$S_i^+ = \sqrt{\sum_{k=1}^q (t_k(a_i) - t_k^+)^2}, \quad i = 1, 2, \dots, n$$

$$S_i^- = \sqrt{\sum_{k=1}^q (t_k(a_i) - t_k^-)^2}, \quad i = 1, 2, \dots, n$$

- 5) **Σχετική εγγύτητα με την ιδανική λύση:** Τέλος, για κάθε εναλλακτική υπολογίζεται η σχετική εγγύτητα με την ιδανική λύση. Έτσι, η τελική κατάταξη της μεθόδου TOPSIS προκύπτει από την ταξινόμηση των εναλλακτικών με φθίνουσα σειρά με βάση το βαθμό σχετικής εγγύτητας από την ιδανική λύση.

$$C_i = S_i^- / (S_i^+ + S_i^-), \quad 0 \leq C_i \leq 1, i = 1, 2, \dots, n$$

### 7.3.3 Η μέθοδος ELECTRE I

Η μέθοδος ELECTRE I [16], [17] χρησιμοποιείται περισσότερο για παιδαγωγικούς λόγους, καθώς δεν έχει μεγάλο πρακτικό ενδιαφέρον, μιας και η φύση των εφαρμογών του πραγματικού κόσμου συνήθως περιλαμβάνει ευρεία γκάμα ποσοτικών και ποιοτικών επιπτώσεων που συνδέονται με μια απόφαση, οδηγώντας σε μια σειρά αντίθετων και ετερογενών κριτηρίων είτε αριθμητικών είτε τακτικών (ordinal). Επιπλέον, η γνώση για τα προβλήματα του πραγματικού κόσμου περιλαμβάνει ένα βαθμό ανακρίβειας ή αβεβαιότητας. Η συγκεκριμένη μέθοδος αναπτύχθηκε για λόγους πληρότητας και στην παρούσα διδακτορική διατριβή.

Η μέθοδος αυτή είναι πολύ απλή και πρέπει να εφαρμόζεται όταν όλα τα κριτήρια είναι κωδικοποιημένα σε αριθμητικές κλίμακες με ίδιο εύρος. Στην περίπτωση αυτή μπορούμε να ισχυριστούμε ότι «η απόφαση  $a$  υπερτερεί της απόφασης  $b$ » (ή η  $a$  είναι τουλάχιστον όσο καλή είναι και η  $b$ ) με το συμβολισμό  $aSb$  με την προϋπόθεση ότι ισχύουν δύο συνθήκες.

Πρωτίστως, η δύναμη του συνασπισμού συμφωνίας (concordant coalition) να είναι αρκετή. Δηλαδή, το άθροισμα των βαρών που συνδέεται με τα κριτήρια που συναποτελούν τον συνασπισμό. Μπορεί να οριστεί μέσω του δείκτη συμφωνίας ως εξής:

$$c(aSb) = \sum_{\{j: g_j(a) \geq g_j(b)\}} w_j$$

Όπου υποθέτουμε για λόγους απλότητας ότι:

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

$$\sum_{j \in J} w_j = 1$$

Όπου  $J$  είναι το σύνολο δεικτών στα κριτήρια, και  $\{j: g_j(a) \geq g_j(b)\}$  είναι το σύνολο δεικτών για όλα τα κριτήρια που ανήκουν στο συνασπισμό συμφωνίας με την σχέση υπεροχής  $aSb$ .

Με άλλα λόγια η τιμή του δείκτη συμφωνίας πρέπει να είναι μεγαλύτερη ή ίση με ένα δοθέν επίπεδο συμφωνίας  $s$ , του οποίου η τιμή γενικά έχει το εύρος τιμών  $[0.5, 1 - \min_{j \in J} w_j]$ , παραδείγματος χάρη  $c(aSb) \geq s$ .

Από την άλλη πλευρά δε θα πρέπει να υπάρχει διαφωνία (discordance) κατά του ισχυρισμού «η απόφαση  $a$  είναι τουλάχιστον όσο καλή όσο και η  $b$ ». Το επίπεδο διαφωνίας μετριέται με τη χρήση ενός ελάχιστου επιπέδου διαφωνίας ως εξής:

$$d(aSb) = \max_{\{j: g_j(a) < g_j(b)\}} \{g_j(b) - g_j(a)\}$$

Το επίπεδο διαφωνίας μετράει τη δύναμη του συνασπισμού διαφωνίας, με την έννοια ότι αν η τιμή του ξεπεράσει ένα δοθέν επίπεδο  $v$ , ο ισχυρισμός δεν ισχύει. Ο συνασπισμός διαφωνίας δεν έχει καμία ισχύ όταν  $d(aSb) \leq v$ .

Τόσο ο δείκτης συμφωνίας όσο και ο δείκτης διαφωνίας, πρέπει να υπολογιστούν για όλα τα ζεύγη εναλλακτικών αποφάσεων  $(a,b)$  στο σύνολο  $A$ , για  $a \neq b$ .

Είναι εύκολο να συμπεράνουμε ότι μια τέτοια υπολογιστική διαδικασία οδηγεί σε δυαδικές σχέσεις στο σύνολο  $A$ . Συνεπώς για κάθε ζεύγος εναλλακτικών αποφάσεων  $(a, b)$  μόνο μια από τις ακόλουθες καταστάσεις μπορεί να ισχύει.

- $aSb$  και όχι  $bSa$ , οπότε  $aPb$  δηλαδή η  $a$  προτιμάται αυστηρά σε σχέση με τη  $b$ .
- $bSa$  και όχι  $aSb$ , οπότε  $bPa$  δηλαδή η  $b$  προτιμάται αυστηρά σε σχέση με την  $a$ .
- $aSb$  και  $bSa$ , οπότε  $aIb$  δηλαδή υπάρχει αδιαφορία (indifference) αν θα επιλεγεί η  $a$  ή η  $b$ .
- Όχι  $aSb$  και όχι  $bSa$ , οπότε  $aRb$  δηλαδή η  $a$  δε μπορεί να συγκριθεί με τη  $b$ .

Η συγκεκριμένη μεθοδολογία προτίμησης- αδιαφορίας με την πιθανότητα να καταλήξει και σε αδυναμία σύγκρισης μεταξύ εναλλακτικών, δε βοηθάει καθόλου στο πώς να επιλεγεί η επόμενη ενέργεια ή το σύνολο ενεργειών στο οποίο θα πρέπει να επικεντρωθεί ο αποφασίζων. Το μόνο που μπορούμε να συμπεράνουμε από τη μέθοδο ELECTRE I είναι μόνο η σχέση υπεροχής μεταξύ κάποιων εναλλακτικών.

### 7.3.4 Η μέθοδος PROMETHEE II

Η μέθοδος PROMETHEE II (Preference Ranking Organization METHod for Enrichment of Evaluations) [18] είναι μια μέθοδος πλήρους κατάταξης που βασίζεται στην πολυκριτηριακή καθαρή ροή (multicriteria net flow). Περιλαμβάνει ένα κατώφλι προτίμησης και ένα κατώφλι αδιαφορίας τα οποία θα εξηγηθούν αναλυτικά στις επόμενες παραγράφους.

Ο αλγόριθμος της μεθόδου PROMETHEE II παρουσιάζεται παρακάτω:

---

```

input:  $n$  (alternatives),  $q$  (criteria),  $f_i(a_j)$  (evaluation table),  $P_i$  (preference
function);
for all criteria  $f_k, k \in \{1, \dots, q\}$  do
|  $d_k(a_i, a_j) = f_k(a_i) - f_k(a_j)$ 
end
for all pairs of alternatives  $a_i, a_j, i, j \in \{1, \dots, n\}$  do
|  $\pi_k(a_i, a_j) = P_k[d_k(a_i, a_j)]$ 
end
for all criteria  $f_k, k \in \{1, \dots, q\}$  do
|  $\pi(a_i, a_j) = \text{sum}(\pi_k(a_i, a_j))$ 
end
for all pairs of alternatives  $a_i, a_j, i, j \in \{1, \dots, n\}$  do
|  $\phi^+(a_i) = \text{sum}(\pi(a_i, a_j))$ 
|  $\phi^-(a_i) = \text{sum}(\pi(a_j, a_i))$ 
end
for all alternatives  $a_i, i \in \{1, \dots, n\}$  do
|  $\phi(a_i) = \phi^+(a_i) - \phi^-(a_i)$ 
end
FinalRanking = SortDesc( $\phi$ )
Result: FinalRanking

```

---

Εικόνα 25 - Αλγόριθμος PROMETHEE II [13]

Συγκεκριμένα κατά την εκτέλεση της μεθόδου PROMETHEE II ακολουθούνται τα παρακάτω βήματα [13], [14]:

- 1) **Σύγκριση κατά ζεύγη:** Ως πρώτο βήμα πραγματοποιούνται συγκρίσεις για όλα τα ζεύγη εναλλακτικών για κάθε κριτήριο. Ως  $d_k(a_i, a_j)$ , ορίζεται η διαφορά μεταξύ των εναλλακτικών  $a_i, a_j$  για το κριτήριο  $f_k$ :

$$d_k(a_i, a_j) = f_k(a_i) - f_k(a_j)$$

- 2) **Βαθμός προτίμησης:** Οι διαφορές που υπολογίστηκαν στο βήμα 1 μεταφράζονται σε βαθμούς προτίμησης για κάποιο κριτήριο ως εξής:

$$\pi(a_i, a_j) = P_k[d_k(a_i, a_j)]$$

Όπου  $P_k : R \rightarrow [0,1]$  είναι μια θετική και όχι φθίνουσα συνάρτηση προτίμησης έτσι ώστε  $P_j(0) = 0$ . Έξι διαφορετικοί τύποι συναρτήσεων προτίμησης προτείνονται από τη μέθοδο PROMETHEE II, οι οποίοι παρουσιάζονται μετά τα βήματα εκτέλεσης της μεθόδου.

- 3) **Πολυκριτηριακός βαθμός προτίμησης:** Η ανά ζεύγη σύγκριση των εναλλακτικών πραγματοποιείται υπολογίζοντας τον πολυκριτηριακό βαθμό προτίμησης (multicriteria preference degree) για κάθε ζεύγος, ως εξής:

$$\pi(a_i, a_j) = \sum_{i=1}^q \pi(a_i, a_j) * w_k$$

Όπου το  $w_k$  αναπαριστά το βάρος για το κριτήριο  $f_k$ , και ισχύει ότι  $w_k \geq 0$  και  $\sum_{k=1}^q w_k = 1$

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

- 4) **Πολυκριτηριακές ροές προτίμησης:** Ορίζονται δύο διαφορετικά μεγέθη για τις ροές κυριαρχίας (outranking flows). Συγκεκριμένα:

Η θετική ροή κυριαρχίας εκφράζει κατά πόσο μια εναλλακτική κυριαρχεί έναντι όλων των άλλων και δείχνει το χαρακτήρα κυριαρχίας της. Μεγαλύτερος βαθμός θετικής ροής κυριαρχίας σημαίνει και μια καλύτερη εναλλακτική λύση. Η θετική ροή κυριαρχίας ορίζεται ως εξής:

$$\varphi^+(a) = \frac{1}{n-1} \sum_{x \in A} \pi(a, x)$$

Η αρνητική ροή κυριαρχίας εκφράζει κατά πόσο μια εναλλακτική κυριαρχείται από όλες τις άλλες. Χαμηλότερος βαθμός αρνητικής ροής κυριαρχίας σημαίνει μια καλύτερη εναλλακτική. Η αρνητική ροή κυριαρχίας ορίζεται ως εξής:

$$\varphi^-(a) = \frac{1}{n-1} \sum_{x \in A} \pi(x, a)$$

Μια ιδανική εναλλακτική θα είχε θετική ροή κυριαρχίας ίση με 1 και αρνητική ροή κυριαρχίας ίση με το 0. Η θετική και η αρνητική ροή κυριαρχίας συνδυάζονται για να δώσουν την καθαρή ροή προτίμησης (net preference flow) ως εξής:

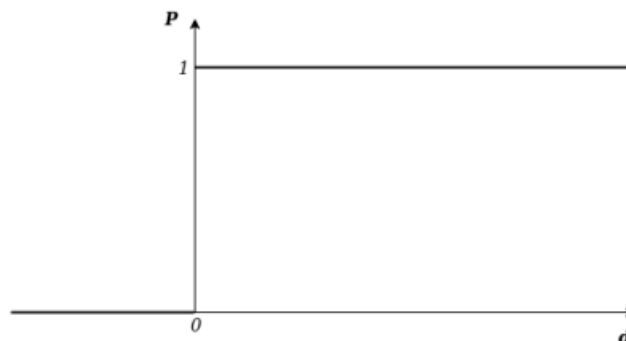
$$\varphi(a) = \varphi^+(a) - \varphi^-(a)$$

Η τελική κατάταξη της μεθόδου PROMETHEE II δίνεται με την ταξινόμηση των εναλλακτικών με βάση τις τιμές των καθαρών ροών προτίμησης σε φθίνουσα σειρά.

**Οι συναρτήσεις προτίμησης της μεθόδου PROMETHEE II:** Έστω  $d_j$  η διαφορά μεταξύ δύο εναλλακτικών  $a_i, a_j$ , και έστω  $q_j, p_j$  τα κατώφλια αδιαφορίας και προτίμησης αντίστοιχα. Όταν η διαφορά  $d_j$  είναι μικρότερη από το κατώφλι αδιαφορίας  $q_j$  θεωρείται αμελητέα και ο βαθμός προτίμησης είναι 0. Αντίθετα, αν η διαφορά  $d_j$  είναι μεγαλύτερη από το κατώφλι προτίμησης  $p_j$ , τότε θεωρείται σημαντική και ο βαθμός προτίμησης γίνεται 1. Διαφορετικά, εάν η διαφορά  $d_j$  είναι μεταξύ των δύο κατωφλίων  $q_j, p_j$ , ο βαθμός προτίμησης χρησιμοποιείται με γραμμική παρεμβολή (linear interpolation). Παρακάτω παρουσιάζονται τα έξι κριτήρια της μεθόδου PROMETHEE II:

#### Σύνηθες κριτήριο

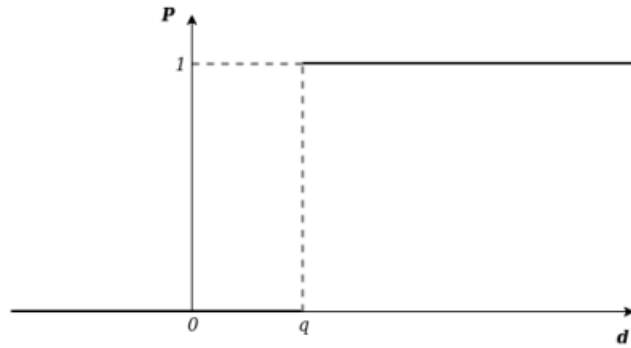
$$P_j(d_j) = \begin{cases} 0, & \text{αν } d_j \leq 0 \\ 1, & \text{αν } d_j > 1 \end{cases}$$



#### Κριτήριο σχήματος U

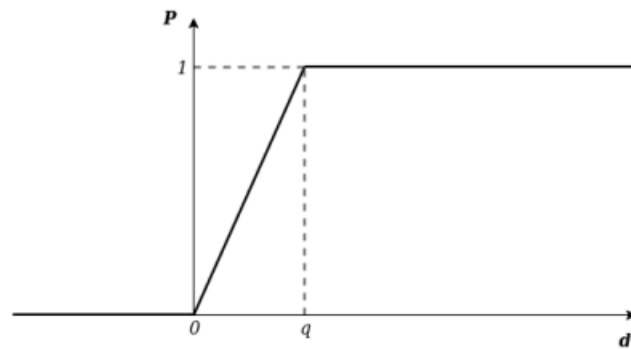
Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

$$P_j(d_j) = \begin{cases} \mathbf{0}, & \text{αν } |d_j| \leq q_j \\ \mathbf{1}, & \text{αν } |d_j| > q_j \end{cases}$$



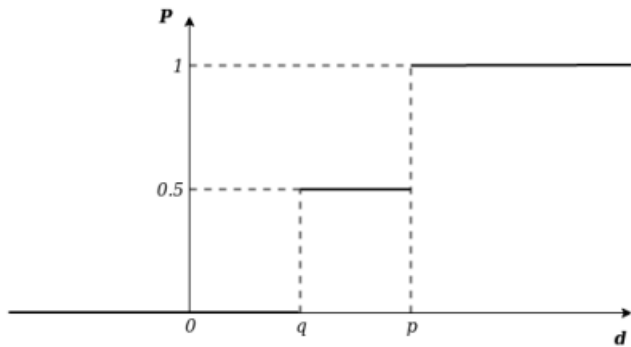
**Κριτήριο σχήματος V**

$$P_j(d_j) = \begin{cases} \frac{|d_j|}{p_j}, & \text{αν } |d_j| \leq p_j \\ \mathbf{1}, & \text{αν } |d_j| > p_j \end{cases}$$



**Κριτήριο επιπέδου**

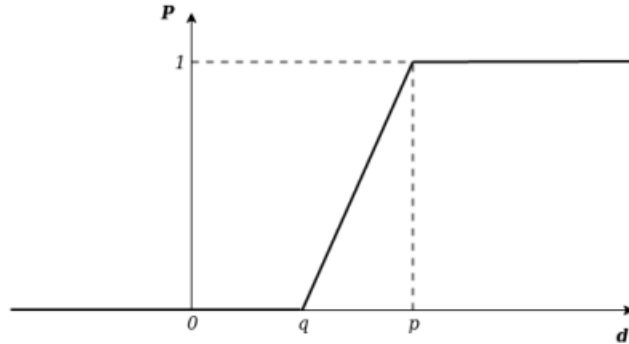
$$P_j(d_j) = \begin{cases} \mathbf{0}, & \text{αν } |d_j| \leq q_j \\ \frac{1}{2}, & \text{αν } q_j < |d_j| \leq p_j \\ \mathbf{1}, & \text{αν } |d_j| > p_j \end{cases}$$



**Γραμμικό κριτήριο**

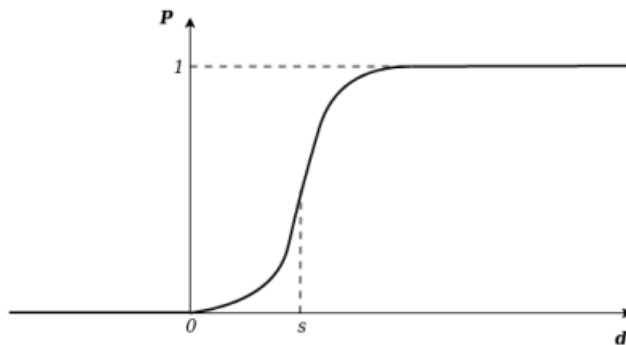
Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

$$P_j(d_j) = \begin{cases} 0, & \text{αν } |d_j| \leq q_j \\ \frac{|d_j| - q_j}{p_j - q_j}, & \text{αν } q_j < |d_j| \leq p_j \\ 1, & \text{αν } |d_j| > p_j \end{cases}$$



**Γκαουσιανό κριτήριο**

$$P_j(d_j) = 1 - e^{-\frac{d_j^2}{2s_j^2}}$$



Μέχρι στιγμής παρουσιάστηκαν μερικές από τις πιο αντιπροσωπευτικές μεθόδους πολυκριτηριακής ανάλυσης αποφάσεων. Φυσικά, υπάρχουν αρκετές επιπλέον μέθοδοι που δεν θα παρουσιαστούν σε βάθος στο συγκεκριμένο κεφάλαιο όπως για παράδειγμα η AHP [19], η ELECTRE III , η ELECTRE Tri [9], μεταξύ άλλων.

## 7.4 Η υπηρεσία πολυκριτηριακής υποστήριξης αποφάσεων

Η υπηρεσία πολυκριτηριακής υποστήριξης αποφάσεων είναι ένα εργαλείο γενικού σκοπού που υποβοηθά τη διαδικασία λήψης αποφάσεων λαμβάνοντας υπόψη όλα τα κριτήρια τα οποία είναι σημαντικά για το χρήστη. Οι περισσότερες μέθοδοι πολυκριτηριακής ανάλυσης αποφάσεων κατατάσσονται σε μια εκ των μεθόδων πολυκριτηριακής ωφέλειας (multicriteria utility) και μεθόδων κυριαρχίας (outranking). Οι μέθοδοι αυτοί προέρχονται από τη θεωρία πολυκριτηριακής ωφέλειας (Multi-Criteria Utility Theory) και τη θεωρία σχέσεων υπεροχής (Outranking Relations Theory) αντίστοιχα.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Οι μέθοδοι πολυκριτηριακής υποστήριξης αποφάσεων που αναπτύχθηκαν για την συγκεκριμένη υπηρεσία παρουσιάστηκαν στην προηγούμενη ενότητα.

Συγκεκριμένα αναπτύχθηκαν οι μέθοδοι MAUT, TOPSIS, ELECTRE I και PROMETHEE II. Οι μέθοδοι MAUT, TOPSIS και PROMETHEE II επιστρέφουν μια λίστα κατάταξης με ταξινομημένες τις εναλλακτικές αποφάσεις από την καλύτερη προς τη χειρότερη, ενώ η ELECTRE I που όπως έχει ήδη αναφερθεί δεν μπορεί να χρησιμοποιηθεί για κατάταξη, παρέχει έναν πίνακα κυριαρχίας, που δείχνει διμερώς ποιες μέθοδοι κυριαρχούν έναντι ποιων άλλων.

Η υπηρεσία είναι διαθέσιμη τόσο μέσω μιας διεπαφής χρήστη (user interface), όσο και μέσω APIs. Ένα παράδειγμα χρήσης της υπηρεσίας μέσω API παρουσιάζεται παρακάτω.

Συγκεκριμένα, στον παρακάτω πίνακα παρουσιάζεται ένα cURL request στο API της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων για τη μέθοδο MAUT:

#### Πίνακας 23 - Παράδειγμα χρήσης της μεθόδου MAUT μέσω του API της υπηρεσίας

```
curl --location --request POST 'http://127.0.0.1:7070/mcdss/maut' \  
--header 'Content-Type: application/json' \  
--data-raw '  
{  
  "Decision_Matrix":  
    {  
      "Number_of_alternatives": 3,  
      "Number_of_criteria": 5,  
      "Criteria": ["Criterion 1", "Criterion 2", "Criterion 3", "Criterion  
4", "Criterion 5"],  
      "Alternatives": [  
        {  
          "Name": "Alternative 1",  
          "Values": [85, 75, 75, 65, 75]  
        },  
        {  
          "Name": "Alternative 2",  
          "Values": [80, 65, 75, 62.5, 75]  
        },  
        {  
          "Name": "Alternative 3",
```



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

```
        "Values": [50, 17.5, 70, 6.12, 60]
      }
    ]
  },
  "Criteria_Details":
  {
    "Number_of_criteria": 5,
    "Weights": [70, 75, 42.5, 77.5, 75],
    "Optimization_Type": [0, 0, 0, 0, 0]
  }
}
```

Στον επόμενο πίνακα παρουσιάζονται τα αποτελέσματα της μεθόδου MAUT για το προηγούμενο request στο API.

**Πίνακας 24 - Αποτελέσματα από τη χρήση της μεθόδου MAUT**

```
[
  {
    "Alternative": "Alternative 3",
    "Ranking": 1,
    "Score": 0.93
  },
  {
    "Alternative": "Alternative 1",
    "Ranking": 2,
    "Score": 0.19669999999999999
  },
  {
    "Alternative": "Alternative 2",
    "Ranking": 3,
    "Score": -0.37660000000000005
  },
  {
    "Alternative": "Alternative 4",
    "Ranking": 4,
    "Score": -0.75
  }
]
```

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Στην επόμενη ενότητα παρουσιάζεται η διεπαφή χρήστη και τα αποτελέσματα της υπηρεσίας πάνω στα αποτελέσματα του συστήματος συστάσεων μαθημάτων και δεξιοτήτων, για να βοηθήσουν στην απόφαση επιλογής μαθημάτων με βάση τις ανάγκες της αγοράς εργασίας και το προφίλ του χρήστη. Επιπροσθέτως, παρουσιάζεται το σενάριο της εισαγωγής μιας νέας γνώσης/δεξιότητας σε ένα υπάρχον πρόγραμμα σπουδών αξιοποιώντας της προτάσεις της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών, παράλληλα με τα αποτελέσματα της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων, για την υποβοήθηση τόσο της επιλογής της νέας γνώσης/δεξιότητας όσο και για την επιλογή του καταλληλότερου μαθήματος για να εισαχθεί.

## **7.5 Παραδείγματα χρήσης της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων**

Στην ενότητα αυτή παρουσιάζονται δυο παραδείγματα χρήσης της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων, ένα για κάθε μια από τις υπηρεσίες συστάσεων που αναπτύχθηκαν (σύσταση μαθημάτων και δεξιοτήτων και ανασχεδιασμός προγράμματος σπουδών).

### **7.5.1 Περίπτωση χρήσης 1 - επιλογή δεξιοτήτων και μαθημάτων**

Όπως έχει ήδη αναφερθεί στο κεφάλαιο 5, για να παράγει συστάσεις δεξιοτήτων και μαθημάτων, το σύστημα συστάσεων μαθημάτων και δεξιοτήτων που αναπτύχθηκε απαιτεί ως είσοδο από το χρήστη το βιογραφικό του σε κωδικοποιημένη μορφή, το οποίο περιλαμβάνει τον επαγγελματικό τομέα στον οποίο στοχεύει, τις γνώσεις και δεξιότητες που κατέχει και την εργασιακή του εμπειρία. Στο παράδειγμα μας, ως υποθέσουμε ότι ένας χρήστης (δια βίου μαθητής) με το όνομα Mark στοχεύει στον τομέα της τεχνολογίας της πληροφορίας και έχει εμπειρία στις γλώσσες προγραμματισμού Java και MySQL. Έτσι, συμπληρώνει αναλόγως και τα απαιτούμενα πεδία στη φόρμα εισαγωγής του βιογραφικού. Συγκεκριμένα, όπως φαίνεται στην παρακάτω εικόνα έχει μέτρια εμπειρία σε Java (δύο χρόνια εμπειρίας) και μεγάλη εμπειρία σε MySQL (τρία χρόνια εμπειρίας).

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

The screenshot shows a form for adding skills. It contains two skill entries:

- Skill 1:** Label: Java, Proficiency level: Medium, Comment: 2 years of experience.
- Skill 2:** Label: MySQL, Proficiency level: Advanced, Comment: 3 years of experience.

Each skill entry has a red 'Delete' button. At the bottom of the form are green 'Save' and red 'Close' buttons.

Εικόνα 26 - Ανέβασμα βιογραφικού από έναν φοιτητή ή δια βίου μαθητή

Με το πάτημα του κουμπιού “Save” το βιογραφικό αποθηκεύεται στην πλατφόρμα QualiChain, η οποία στη συνέχεια υπολογίζει τις πιο κατάλληλες γνώσεις και δεξιότητες για τον Mark με βάση το βιογραφικό του, όσο και με βάση τις ανάγκες της αγοράς εργασίας. Έτσι οι προτεινόμενες γνώσεις και δεξιότητες για τον Mark παρουσιάζονται στην παρακάτω εικόνα.

## Recommended skills

Filter

Skill	Action
git	<a href="#">View</a>
docker	<a href="#">View</a>
jira	<a href="#">View</a>
sql	<a href="#">View</a>
database	<a href="#">View</a>
azure	<a href="#">View</a>
cloud	<a href="#">View</a>
jenkins	<a href="#">View</a>

Εικόνα 27 – Προτεινόμενες γνώσεις και δεξιότητες

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Όπως φαίνεται στην παραπάνω εικόνα, οι προτεινόμενες γνώσεις και δεξιότητες για τον Mark (τις οποίες δεν κατέχει ακόμα σύμφωνα με το βιογραφικό του) περιλαμβάνουν εργαλεία όπως το Git, το Docker, το Jira, η γλώσσα προγραμματισμού SQL, οι βάσεις δεδομένων, η πλατφόρμα Azure και το cloud computing γενικότερα, και η πλατφόρμα Jenkins. Υπάρχουν περισσότερες προτάσεις δεξιοτήτων για τον Mark, ωστόσο παρουσιάστηκαν μόνο οι σημαντικότερες όσον αφορά τη ζήτηση από την αγορά εργασίας και παρουσιάζονται σε φθίνουσα σειρά με βάση το ίδιο κριτήριο.

Στη συνέχεια, ο Mark μπορεί είτε να διαλέξει ποιες γνώσεις και δεξιότητες θα ήθελε να καλλιεργήσει με τη βοήθεια της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων είτε να λάβει προτάσεις μαθημάτων τα οποία προσφέρουν τις προτεινόμενες γνώσεις και δεξιότητες. Εφόσον, τα μαθήματα που προτείνει το σύστημα συστάσεων μαθημάτων είναι άρρηκτα συνδεδεμένα με τις προτεινόμενες γνώσεις και δεξιότητες, ο Mark επιλέγει να λάβει επίσης προτάσεις μαθημάτων όπως φαίνεται στην παρακάτω εικόνα.

## Recommended courses

Filter

Name	Description	Score	Action
Databases	Database Management Systems and their architecture. Data Structures for Databases. Modeling - The E-R model. Reference to classic Database Models (Hierarchical, Network). The Relational Model. Database Languages - The SQL language. File Systems and Physical Database Design. Logical Design and Normalization. Management and Operation Issues (Integrity, optimization, reorganization, security, functionality, etc.). Other Issues (Object Oriented Systems, Multi-Systems, Personal Computer Systems, etc.).	1	<a href="#">View</a>
Operating Systems	This course covers the general principles of Operating Systems. Evolution of Operating Systems, Input/Output, concurrent processes, critical section, process synchronization and communication. CPU scheduling, Memory management (static and dynamic allocation, virtual memory, paging, segmentation), File management, disk scheduling, deadlocks.	1	<a href="#">View</a>
Distributed Systems	Synchronization: Natural, logical and vector clocks, natural and logical clock synchronization, Lamport's algorithm. Distributed Coordination: Mutual Exclusion Algorithms, Election Algorithms, Distributed Agreement, Interactive Consistency, Byzantine Protocols, Distributed Consensus Algorithm Paxos. Tasks and concurrency control: ACID properties, embedded transactions, strict two-phase locking, concurrent problems, timer layout, optimistic concurrency control. Distributed Transactions and Timing Control: Distributed Individual Transactions, Distributed Timing Control, Deadlock Detection, Error Recovery. Data Backup: Error Tolerance, Primary / Secondary Backup, Passive / Active Backup, Network Partition Management, High Availability, Consistency Levels, CAP Theorem. Peer-to-peer networks, distributed hash tables, distributed storage systems, Map Reduce and Bulk Synchronous Parallel distributed processing models.	1	<a href="#">View</a>
Operating Systems Laboratory	This course offers laboratory projects to assimilate critical concepts from the previous courses "Computer Architecture" and "Operating Systems". The course includes 3-4 exercises in distributed UNIX programming, development of device drivers, kernel-level programming, and extensions of the Linux OS.	1	<a href="#">View</a>
Human-Computer Interaction	The course deals with the design, development and evaluation methodologies of computer systems that interact with users to a significant degree. The purpose is to review theoretical models, technologies, methods and tools for the design and development of interactive computing systems. The course includes (1) Human modeling of computer system and conceptual interaction framework, (2) Cognitive models, perception and representation, attention and memory representation and organization of knowledge(3) Interactive Systems Design Methodologies: Human-centric Design, Usability Requirements, Task Analysis, GOMS Models, Dialogue Methods, Interface Design, Usability and Accessibility of Internet Services, Interactive Systems Evaluation Techniques, (4) Overview of interactive Technologies and Applications: Distributed Intelligence Systems, Ubiquitous Computer, Augmented Reality Systems, Synthetic Animation, Accessibility Interface Systems, Collaboration Applications.	0.87	<a href="#">View</a>

Εικόνα 28 - Προτεινόμενα μαθήματα

Όπως φαίνεται στην παραπάνω εικόνα, τα προτεινόμενα μαθήματα από το σύστημα συστάσεων είναι οι βάσεις δεδομένων (Databases), τα λειτουργικά συστήματα (Operating Systems), τα κατανεμημένα συστήματα (Distributed Systems), το εργαστήριο λειτουργικών συστημάτων (Operating Systems Laboratory) και η αλληλεπίδραση ανθρώπου μηχανής. Επιπλέον, η τεχνολογία πολυμέσων (Multimedia Technology), και η επίδοση υπολογιστικών συστημάτων περιλαμβάνονται στα αποτελέσματα, αλλά δε φαίνονται στην προηγούμενη εικόνα. Τα προτεινόμενα μαθήματα παρουσιάζονται με σειρά σχετικότητας με το χρήστη παράλληλα με το δείκτη σχετικότητας (Relevance score).

Ο Mark ενδιαφέρεται περισσότερο να παρακολουθήσει τα μαθήματα των Βάσεων Δεδομένων, των λειτουργικών συστημάτων, των κατανεμημένων συστημάτων και της αλληλεπίδρασης ανθρώπου

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

μηχανής. Ωστόσο, έχει το περιθώριο να επιλέξει μόνο ένα από αυτά τα μαθήματα λόγω φόρτου εργασίας. Επιπλέον, σκοπεύει να επιλέξει μαθήματα, όχι μόνο με βάση το βαθμό σχετικότητας με τα ενδιαφέροντα του αλλά και με βάση τη δυσκολία που πιστεύει ότι θα έχει το κάθε μάθημα, τα προσωπικά του ενδιαφέροντα αλλά και το κατά πόσο οι γνώσεις και δεξιότητες που προσφέρει το κάθε μάθημα είναι γνώσεις/δεξιότητες «αιχμής». Για το λόγο αυτό, χρησιμοποιεί την υπηρεσία πολυκριτηριακής ανάλυσης αποφάσεων, για να αποφασίσει ποιο μάθημα να διαλέξει, όπως φαίνεται και στην παρακάτω εικόνα.

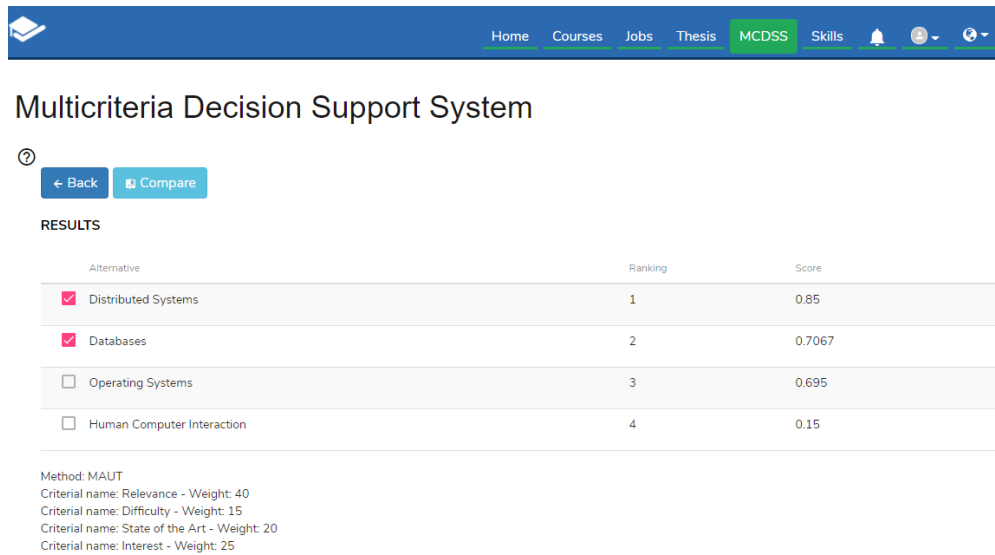
+ Add criteria + Add alternative > Send data

MAUT  
TOPSIS  
ELECTRE I  
PROMETHEE II

	Criteria name *	Criteria name *	Criteria name *	Criteria name *
	Relevance	Difficulty	State of the Art	Interest
	Weight *	Weight *	Weight *	Weight *
	40	15	20	25
	Type *	Type *	Type *	Type *
	0	1	0	0
Alternative name *	Value *	Value *	Value *	Value *
Databases	100	50	50	60
Operating Systems	100	40	50	55
Distributed Systems	100	80	80	70
Human Computer Interaction	87	30	40	40

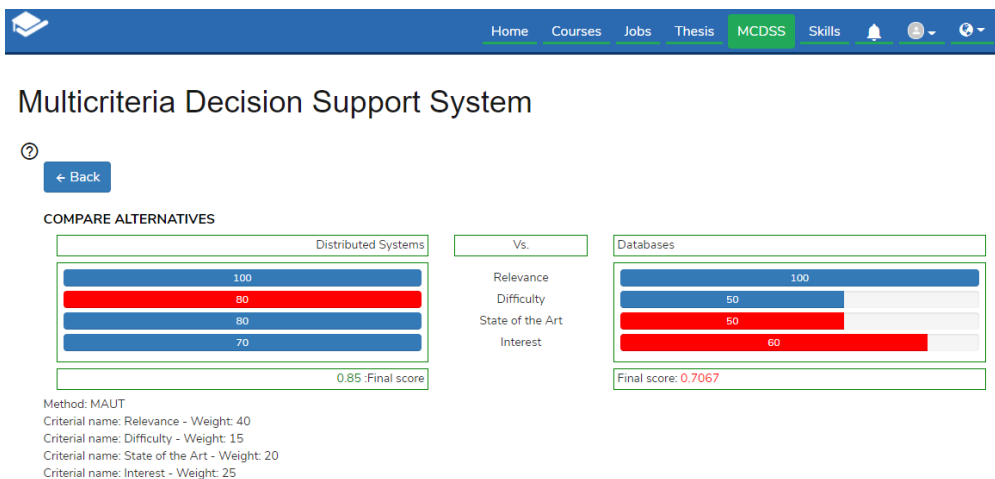
Εικόνα 29 - Διεπαφή χρήστη για τη χρήση της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων

Όπως φαίνεται στην παραπάνω εικόνα, ο Mark επιλέγει τη μέθοδο MAUT και τοποθετεί ως βάρη σημαντικότητας τις εξής τιμές: 40% για τη σχετικότητα (Relevance) κάθε μαθήματος (ο βαθμός σχετικότητας κάθε μαθήματος παρέχεται κατευθείαν από το σύστημα συστάσεων), 15% για τη δυσκολία (Difficulty), 20% για το κατά πόσο το κάθε μάθημα παρέχει γνώσεις και δεξιότητες «αιχμής» (State of the art) και 25% για το πόσο ενδιαφέρον θεωρεί το μάθημα ο Mark (Interest). Όλες οι μετρικές εκτός από το βαθμό σχετικότητας συμπληρώνονται από το χρήστη με βάση την προσωπική του αντίληψη για κάθε κριτήριο και κάθε εναλλακτική. Το αποτέλεσμα της μεθόδου MAUT παρουσιάζεται στην παρακάτω εικόνα.



Εικόνα 30 - Αποτελέσματα της μεθόδου MAUT

Όπως φαίνεται στην παραπάνω εικόνα, τα καταλληλότερα μαθήματα για τον Mark με βάση τα κριτήρια που έθεσε και τη σημασία του καθενός από αυτά, είναι τα κατανεμημένα συστήματα και οι Βάσεις δεδομένων, ενώ τα λειτουργικά συστήματα ακολουθούν με μικρή διαφορά και η αλληλεπίδραση ανθρώπου μηχανής είναι η τελευταία επιλογή με μεγάλη διαφορά σε σχέση με τις υπόλοιπες. Φυσικά, ο Mark ενδεχομένως να χρειάζεται επιπλέον ανάλυση για να λάβει την τελική του απόφαση. Οπότε επιλέγει τις δυο κορυφαίες επιλογές (βάσεις δεδομένων και κατανεμημένα συστήματα) και έχει τη δυνατότητα να πραγματοποιήσει διμερείς συγκρίσεις για κάθε κριτήριο. Στην παρακάτω εικόνα φαίνεται ένα παράδειγμα σύγκρισης των προαναφερθέντων μαθημάτων.



Εικόνα 31 - Διμερείς συγκρίσεις μεταξύ δύο μαθημάτων

Όπως φαίνεται στην παραπάνω εικόνα, ο Mark βρίσκει τις βάσεις δεδομένων ευκολότερο μάθημα από τα κατανεμημένα συστήματα, αλλά από την άλλη πλευρά θεωρεί τις πρώτες λιγότερο ενδιαφέρουσες από τα δεύτερα. Επιπλέον, θεωρεί ότι οι βάσεις δεδομένων παρέχουν γνώσεις και δεξιότητες «αιχμής»

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

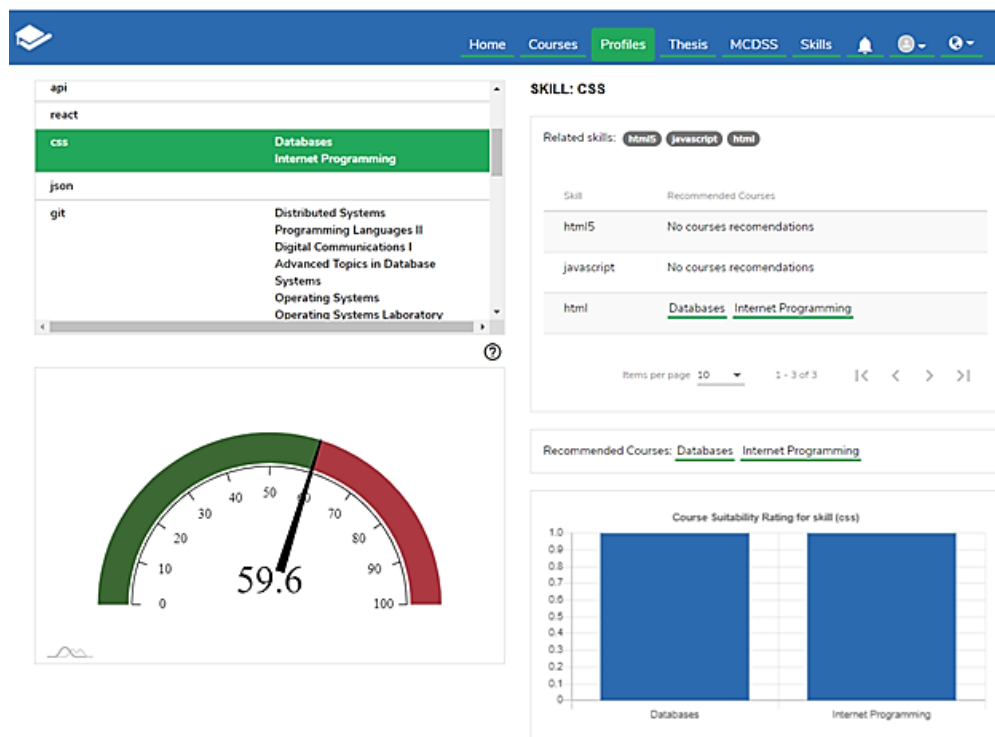
σε μικρότερο βαθμό από ό,τι τα κατανεμημένα συστήματα. Για όλους αυτούς τους λόγους θεωρεί ότι τα κατανεμημένα συστήματα αποτελούν την καλύτερη επιλογή για αυτόν και επιλέγει το μάθημα.

Φυσικά, παρόμοια είναι και η διαδικασία επιλογής γνώσεων και δεξιοτήτων προς εκμάθηση για έναν δια βίου μαθητή. Για το λόγο αυτό δε θα παρουσιαστούν στα πλαίσια του παρόντος κεφαλαίου.

## 7.5.2 Περίπτωση χρήσης 2 - Εισαγωγή νέων δεξιοτήτων στο πρόγραμμα σπουδών

Στη συγκεκριμένη ενότητα παρουσιάζεται η διαδικασία ανασχεδιασμού ενός προγράμματος σπουδών με βάση τις ανάγκες της αγοράς εργασίας. Συγκεκριμένα, το παράδειγμα αναφέρεται στην ελληνική αγορά εργασίας και στον τομέα της τεχνολογίας της πληροφορίας και των επικοινωνιών (ICT sector), ενώ το πρόγραμμα σπουδών προς ενημέρωση είναι το πρόγραμμα γνωστής σχολής παρόμοιου γνωστικού αντικειμένου από μεγάλο πανεπιστήμιο της Ελλάδας.

Στο παράδειγμα μας, έστω ότι η διεύθυνση της σχολής σκοπεύει να εξετάσει ποιες γνώσεις και δεξιότητες είναι σε υψηλή ζήτηση από την αγορά εργασίας αλλά δεν προσφέρονται στα πλαίσια του δοθέντος προγράμματος σπουδών. Στην παρακάτω εικόνα, και με τη χρήση της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών, φαίνεται ότι περίπου το 60% των γνώσεων και δεξιοτήτων που είναι σε υψηλή ζήτηση από την αγορά εργασίας προσφέρονται στα πλαίσια του δοθέντος προγράμματος σπουδών, ενώ πάνω αριστερά στην εικόνα διακρίνονται οι πιο σημαντικές γνώσεις και δεξιότητες που λείπουν από το δοθέν πρόγραμμα σπουδών, παράλληλα με προτάσεις για μαθήματα στα πλαίσια των οποίων θα μπορούσε να ενταχθεί η εκάστοτε γνώση ή δεξιότητα.



Εικόνα 32 - Διεπαφή χρήστη για την υπηρεσία ανασχεδιασμού προγράμματος σπουδών, δίνοντας έμφαση στη δεξιότητα CSS

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Συγκεκριμένα, η γλώσσα προγραμματισμού CSS, και το προγραμματιστικό εργαλείο Git είναι ανάμεσα στις γνώσεις και δεξιότητες που λείπουν από το δοθέν πρόγραμμα σπουδών και συνοδεύονται από κάποια προτεινόμενα μαθήματα που θεωρούνται σχετικά. Αντίθετα τα APIs, το framework React της γλώσσας προγραμματισμού JavaScript, όπως και η γλώσσα json αποτελούν δεξιότητες σε μεγάλη ζήτηση από την αγορά εργασίας, ωστόσο δε θεωρούνται σχετικές με κάποιο μάθημα του προγράμματος σπουδών ώστε να ενταχθούν σε αυτό.

Όσον αφορά τη γλώσσα προγραμματισμού CSS είναι σχετική με τα μαθήματα των βάσεων δεδομένων (databases) και του διαδικτυακού προγραμματισμού (Internet Programming), και όπως φαίνεται στο δεξί μέρος της σελίδας, η δεξιότητα αυτή είναι εξίσου σχετική με τα δύο προαναφερθέντα μαθήματα, καθώς και τα δύο διδάσκουν τη γλώσσα προγραμματισμού HTML.

Σχετικά με το προγραμματιστικό εργαλείο Git, προτείνεται η ένταξη του σε πολλά μαθήματα, όπως μεταξύ άλλων τα κατανεμημένα συστήματα (Distributed Systems), οι γλώσσες προγραμματισμού II (Programming Languages II), τα προχωρημένα θέματα βάσεων δεδομένων (Advanced Topics in Database Systems), και τα λειτουργικά συστήματα (Operating Systems).

Στην παρακάτω εικόνα παρουσιάζεται η είσοδος της υπηρεσίας για την απόφαση εισαγωγής της γλώσσας CSS είτε στο μάθημα των βάσεων δεδομένων είτε σε αυτό του διαδικτυακού προγραμματισμού. Η μέθοδος πολυκριτηριακής ανάλυσης που επιλέχθηκε για το συγκεκριμένο πρόβλημα είναι η PROMETHEE II, ενώ τα κριτήρια για την απόφαση είναι η σχετικότητα (Relevance) με το κάθε μάθημα που είναι και το πιο σημαντικό (βαθμός σημαντικότητας 45%), ο φόρτος (Workload) κάθε μαθήματος (βαθμός σημαντικότητας 30%) και ο βαθμός ευκολίας (Convenience) για να προστεθεί μια νέα δεξιότητα στο κάθε μάθημα (βαθμός σημαντικότητας 25%). Όπως φαίνεται, η δεξιότητα CSS είναι πολύ πιο σχετική με το μάθημα του διαδικτυακού προγραμματισμού από ό,τι με αυτό των βάσεων δεδομένων, ωστόσο ο φόρτος εργασίας των βάσεων δεδομένων είναι μικρότερος από αυτόν του διαδικτυακού προγραμματισμού. Επίσης, θεωρείται πιο εύκολο από τη διοίκηση να προστεθεί η γλώσσα CSS στο μάθημα των βάσεων δεδομένων παρά σε αυτό του διαδικτυακού προγραμματισμού. Αξίζει να αναφέρουμε ότι η τιμή της προτίμησης (Preference) δείχνει την ένταση της προτίμησης για μία επιλογή έναντι κάποιας άλλης, ενώ η αδιαφορία (Indifference) αποτελεί ένα κατώφλι (threshold), το οποίο δείχνει το μέγεθος της αποδεκτής διαφοράς ώστε μια εναλλακτική να θεωρείται όντως ανώτερη από κάποια άλλη.



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Select a method  
PROMETHEE II

Criterion name *	Weight *	Type *	Preference *	Indifference *	Criteria *
Relevance	45	0	20	10	usual
Workload	30	1	25	0	usual
Convenience	25	0	10	20	usual
Alternative name *	Value *	Value *	Value *		
Internet Programming	100	80	50		
Databases	50	70	60		

Εικόνα 33 - Παράδειγμα εισόδου της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων για την εισαγωγή μίας δεξιότητάς σε ένα μάθημα

Όπως φαίνεται στην παρακάτω εικόνα το μάθημα των βάσεων δεδομένων θεωρείται ελαφρώς καλύτερη επιλογή για να εισαχθεί η γλώσσα CSS στην διδακτική του ύλη, παρότι είναι λιγότερο σχετικό με τη συγκεκριμένη δεξιότητα. Ωστόσο, τα αποτελέσματα θα μπορούσαν να είναι διαφορετικά σε περίπτωση που επιλέγονται διαφορετικές τιμές για τον βαθμό προτίμησης ή τον βαθμό αδιαφορίας, ή αν επιλεγόταν κάποια άλλη μέθοδος πολυκριτηριακής ανάλυσης αποφάσεων.

## Multicriteria Decision Support System



← Back

Compare

### RESULTS

Alternative	Ranking	Score
<input checked="" type="checkbox"/> Databases	1	0.10000000000000003
<input checked="" type="checkbox"/> Internet Programming	2	-0.10000000000000003

Method: PROMETHEE II  
 Criterial name: Relevance - Weight: 45  
 Criterial name: Workload - Weight: 30  
 Criterial name: Convenience - Weight: 25

Εικόνα 34 - Αποτελέσματα της μεθόδου PROMETHEE II για την ένταξη της δεξιότητας CSS σε κάποιο μάθημα

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Για παράδειγμα στην παρακάτω εικόνα παρουσιάζεται η είσοδος της μεθόδου TOPSIS για τις ίδιες εναλλακτικές και τα ίδια κριτήρια.

	Criterial name *	Weight *	Type *
	Relevance	45	0
	Workload	30	1
	Convenience	25	0
Alternative name *	Value *	Value *	Value *
Internet Programming	100	80	50
Databases	50	70	60

Εικόνα 35 - Η είσοδος για τη μέθοδο TOPSIS

Όπως φαίνεται στα αποτελέσματα το μάθημα του διαδικτυακού προγραμματισμού φαίνεται να είναι καλύτερη επιλογή σύμφωνα με τη συγκεκριμένη μέθοδο.

#### RESULTS

Alternative	Ranking	Score
<input type="checkbox"/> Internet Programming	1	0.8251
<input type="checkbox"/> Databases	2	0.1749

Method: TOPSIS  
 Criterial name: Relevance - Weight: 45  
 Criterial name: Workload - Weight: 30  
 Criterial name: Convenience - Weight: 25

Εικόνα 36 - Τα αποτελέσματα της μεθόδου TOPSIS

## 7.6 Σύνοψη και συμπεράσματα

Η υπηρεσία πολυκριτηριακής υποστήριξης αποφάσεων παρέχει στους χρήστες της πλατφόρμας ένα εργαλείο που τους βοηθάει να είναι πιο σίγουροι για τις αποφάσεις που λαμβάνουν, ταξινομώντας τις εναλλακτικές τους με βάση όλα τα κριτήρια τα οποία οι ίδιοι θεωρούν σημαντικά. Συγκεκριμένα, στο πλαίσιο του έργου QualiChain χρησιμοποιήθηκε για την υποβοήθηση της απόφασης επιλογής υποψηφίου για μια θέση εργασίας, όσο και για την επιλογή μαθημάτων και δεξιοτήτων από φοιτητές

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

και δια βίου μαθητές, όπως επίσης και για την υποστήριξη αποφάσεων για την ενημέρωση ενός προγράμματος σπουδών με την εισαγωγή νέων γνώσεων και δεξιοτήτων.

Ένας περιορισμός που παρατηρήθηκε είναι ότι πολλοί χρήστες της πλατφόρμας δεν ήταν εξοικειωμένοι με την πολυκριτηριακή υποστήριξη αποφάσεων και τις μεθόδους της και κατά συνέπεια χρειάστηκε να προστεθούν επεξηγήσεις στην διεπαφή χρήστη της εφαρμογής, όπως επίσης να πραγματοποιηθούν πολλές παρουσιάσεις για το πώς μπορεί να χρησιμοποιηθεί αποτελεσματικά. Επιπλέον, αξίζει να σημειωθεί ότι είχε σχεδιαστεί η ανάπτυξη περισσότερων μεθόδων πολυκριτηριακής ανάλυσης αποφάσεων στην υπηρεσία, αλλά λόγω της έλλειψης εξοικείωσης από τους χρήστες, αποφασίστηκε να δοθεί περισσότερη έμφαση στην βελτίωση της διεπαφής και στην παροχή επεξηγήσεων και παραδειγμάτων, αντί για την ανάπτυξη νέων πιο πολύπλοκων μεθόδων. Περισσότερες πληροφορίες σχετικά με την αξιολόγηση και τη χρήση της συγκεκριμένης υπηρεσίας παρουσιάζονται στο κεφάλαιο 8.

## 7.7 Αναφορές

- [1] L. A. Franco and G. Montibeller, “Problem Structuring for Multicriteria Decision Analysis Interventions,” *Wiley Encyclopedia of Operations Research and Management Science*, Jan. 2011, doi: 10.1002/9780470400531.EORMS0683.
- [2] I. Nunes and D. Jannach, “A systematic review and taxonomy of explanations in decision support and recommender systems,” *User Modeling and User-Adapted Interaction* 2017 27:3, vol. 27, no. 3, pp. 393–444, Oct. 2017, doi: 10.1007/S11257-017-9195-0.
- [3] H. Mahgfuri, R. D. Perdanakusuma, F. Basfianto, A. A. Sukmandhani, and I. P. Saputro, “A Recommendation System for Buying a used Car using the Promethee Method,” 2022 International Conference on Sustainable Islamic Business and Finance, SIBF 2022, pp. 209–213, 2022, doi: 10.1109/SIBF56821.2022.9939732.
- [4] N. S. Houari and N. Taghezout, “An Effective Tool for the Experts’ Recommendation Based on PROMETHEE II and Negotiation: Application to the Industrial Maintenance,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 6, pp. 67–77, 2021, doi: 10.9781/IJIMAI.2021.01.002.
- [5] L. Marlinda, T. Baidawi, Y. Durachman, and Akmaludin, “A multi-study program recommender system using ELECTRE multicriteria method,” 2017 5th International Conference on Cyber and IT Service Management, CITSM 2017, Oct. 2017, doi: 10.1109/CITSM.2017.8089271.
- [6] S. Buirra and A. Valls, “A Method to Integrate Semantic Criteria into a Recommender System Based on ELECTRE Outranking Relations,” *Frontiers in Artificial Intelligence and Applications*, vol. 308, pp. 146–155, 2018, doi: 10.3233/978-1-61499-918-8-146.
- [7] Z. Chai, Y. Li, and S. Zhu, “P-MOIA-RS: a multi-objective optimization and decision-making algorithm for recommendation systems,” *J Ambient Intell Humaniz Comput*, vol. 12, no. 1, pp. 443–454, Jan. 2021, doi: 10.1007/S12652-020-01997-X/TABLES/8.
- [8] D. Von Winterfeldt and G. W. Fischer, “Multi-Attribute Utility Theory: Models and Assessment Procedures,” *Utility, Probability, and Human Decision Making*, pp. 47–85, 1975, doi: 10.1007/978-94-010-1834-0\_3.

- [9] B. Roy, "The outranking approach and the foundations of electre methods," *Theory Decis*, vol. 31, no. 1, pp. 49–73, 1991, doi: 10.1007/BF00134132/METRICS.
- [10] E. Jacquet-Lagrèze and Y. Siskos, "Preference disaggregation: 20 years of MCDA experience," *Eur J Oper Res*, vol. 130, no. 2, pp. 233–245, Apr. 2001, doi: 10.1016/S0377-2217(00)00035-7.
- [11] J. S. Dyer, P. C. Fishburn, R. E. Steuer, J. Wallenius, and S. Zionts, "Multiple Criteria Decision Making, Multiattribute Utility Theory: The Next Ten Years," <https://doi.org/10.1287/mnsc.38.5.645>, vol. 38, no. 5, pp. 645–654, May 1992, doi: 10.1287/MNSC.38.5.645.
- [12] J. S. Dyer, "MAUT-multiattribute utility theory," *International Series in Operations Research and Management Science*, vol. 78, pp. 265–295, 2005, doi: 10.1007/0-387-23081-5\_7/COVER.
- [13] E. Sarmas, P. Xidonas, and H. Doukas, "The Proposed Methodology," *Springer Optimization and Its Applications*, vol. 163, pp. 45–77, 2020, doi: 10.1007/978-3-030-53743-2\_5/FIGURES/4.
- [14] E. Sarmas, P. Xidonas, and H. Doukas, "Multicriteria Portfolio Construction with Python," vol. 163, 2020, doi: 10.1007/978-3-030-53743-2.
- [15] Y. Çelikbilek and F. Tüysüz, "An in-depth review of theory of the TOPSIS method: An experimental analysis," *Journal of Management Analytics*, vol. 7, no. 2, Apr. 2020, doi: 10.1080/23270012.2020.1748528.
- [16] D. BOUYSSOU and P. VINCKE, "Ranking Alternatives on the Basis of Preference Relations: A Progress Report with Special Emphasis on Outranking Relations," *Journal of Multi-Criteria Decision Analysis*, vol. 6, no. 2, Mar. 1997, doi: 10.1002/(SICI)1099-1360(199703)6:2<77::AID-MCDA144>3.0.CO;2-I.
- [17] J. Figueira, S. Greco, and M. Ehrogott, "Electre Methods," in *Multiple Criteria Decision Analysis: State of the Art Surveys*, New York, NY: Springer New York, 2005. doi: 10.1007/b100605.
- [18] J. P. Brans and B. Mareschal, "The Promethee Methods for MCDM; The Promcalc, Gaia And Bankadviser Software," *Readings in Multiple Criteria Decision Aid*, pp. 216–252, 1990, doi: 10.1007/978-3-642-75935-2\_10.
- [19] A. Darko, A. P. C. Chan, E. E. Ameyaw, E. K. Owusu, E. Pärn, and D. J. Edwards, "Review of application of analytic hierarchy process (AHP) in construction," <https://doi.org/10.1080/15623599.2018.1452098>, vol. 19, no. 5, pp. 436–452, Sep. 2018, doi: 10.1080/15623599.2018.1452098.

## **ΚΕΦΑΛΑΙΟ 8: ΧΡΗΣΗ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ ΤΩΝ ΥΠΗΡΕΣΙΩΝ ΠΟΥ ΑΝΑΠΤΥΧΘΗΚΑΝ**

### **8.1 Το πιλοτικό σενάριο χρήσης**

Στην ενότητα αυτή παρουσιάζεται το πιλοτικό σενάριο χρήσης του έργου QualiChain με τίτλο «Βελτιστοποίηση των διαδικασιών του πανεπιστημίου με τη χρήση της υπηρεσίας Έξυπνου Ανασχεδιασμού Προγράμματος Σπουδών και τη διαπίστευση φοιτητών» (University Process Optimisation through Smart Curriculum Design and Student Accreditation). Συγκεκριμένα εξετάζεται η χρήση της πλατφόρμας QualiChain στη λειτουργία της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών (Η.Μ.Μ.Υ.) του Εθνικού Μετσόβιου Πολυτεχνείου (Ε.Μ.Π.) για τη βελτιστοποίηση κάποιων διαδικασιών όσο και για την ενίσχυση του ακαδημαϊκού και επαγγελματικού προφίλ των φοιτητών. Η εκτέλεση του πιλότου περιλαμβάνει τη χρήση του QualiChain από ομάδες φοιτητών και καθηγητών της σχολής και παρείχε πολύ ενδιαφέροντα αποτελέσματα, τα οποία θα παρουσιαστούν σε βάθος στις επόμενες υποενότητες.

#### **8.1.1 Γενική παρουσίαση πιλότου**

Ο συγκεκριμένος πιλότος επικεντρώνεται στη βελτιστοποίηση της διαδικασίας διδασκαλίας, του προγράμματος σπουδών και της λειτουργίας του πανεπιστημίου καθώς επίσης και στην αναγνώριση και επαλήθευση των φοιτητών του πανεπιστημίου. Όπως έχει ήδη αναφερθεί, ο πιλότος εφαρμόστηκε στη σχολή Η.Μ.Μ.Υ. του Ε.Μ.Π.. Το Ε.Μ.Π. είναι ένα από τα παλαιότερα ακαδημαϊκά ιδρύματα στην τεχνική εκπαίδευση στην Ελλάδα, και δεσμεύεται να είναι πάντα στην αιχμή της τεχνολογίας σε ό,τι αφορά τις επιστήμες των μηχανικών και να παρέχει πάντα εκπαίδευση υψηλής ποιότητας στους φοιτητές του. Η σχολή Η.Μ.Μ.Υ., από την άλλη, αριθμεί περισσότερους από 2.000 προπτυχιακούς φοιτητές, 800 διδακτορικούς φοιτητές και 75 καθηγητές. Σε αυτό το πλαίσιο η συγκεκριμένη περίπτωση χρήσης (use case), επέτρεψε την δοκιμή της πλατφόρμας QualiChain σε ένα ακαδημαϊκό ίδρυμα που περιλαμβάνει μαθήματα τεχνολογίας σε διάφορα πεδία και κατά συνέπεια παρέχει στους φοιτητές του τη δυνατότητα να ακολουθήσουν διάφορα εναλλακτικά μονοπάτια στην καριέρα τους.

Πιο αναλυτικά, ο συγκεκριμένος πιλότος ασχολείται με τις ακόλουθες δραστηριότητες:

- Παροχή εργαλείων στους φοιτητές (προπτυχιακούς και διδακτορικούς) που θα τους βοηθήσουν να βελτιώσουν το ακαδημαϊκό και το επαγγελματικό τους προφίλ. Οι λειτουργίες αναλυτικής δεδομένων (data analytics) και υποστήριξης αποφάσεων, παρέχουν προτάσεις για τη βελτίωση του προφίλ ενός φοιτητή.
- Παροχή αναλυτικής δεδομένων και προτάσεων σε καθηγητές και διοίκηση για τον ανασχεδιασμό του προγράμματος σπουδών της σχολής Η.Μ.Μ.Υ.
- Αξιοποίηση ενός έμπιστου, αμετάβλητου (immutable) και ασφαλούς Blockchain ledger για την επιβράβευση των φοιτητών με smart badges για την αναγνώριση της απόδοσης και της εμπειρίας τους.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Συγκεκριμένα, ο πιλότος αυτός αξιοποιεί την υπολογιστική νοημοσύνη που παρέχεται από το υποσύστημα της αναλυτικής δεδομένων και την υπηρεσία υποστήριξης αποφάσεων για την πρόταση συγκεκριμένων μαθημάτων στους φοιτητές, όσο και την αναδιάρθρωση του προγράμματος σπουδών της σχολής Η.Μ.Μ.Υ., ώστε αφ' ενός οι φοιτητές να βελτιώσουν το προφίλ τους, και αφ' ετέρου το πρόγραμμα σπουδών να εκμοντερνιστεί με βάση τις ανάγκες της αγοράς εργασίας, προσφέροντας τις πιο δημοφιλείς γνώσεις και δεξιότητες. Επιπλέον, τόσο οι προπτυχιακοί όσο και οι διδακτορικοί φοιτητές της σχολής είχαν τη δυνατότητα να λάβουν προσωποποιημένες προτάσεις για την αποτελεσματική επιλογή πανεπιστημιακών μαθημάτων, ανοιχτών μαθημάτων και άλλων εξωσχολικών δραστηριοτήτων. Οι προτάσεις αυτές γίνονταν με βάση τα μαθήματα που είχε ολοκληρώσει ο χρήστης, τις προτιμήσεις του, και τις επιθυμητές επιλογές καριέρας. Επιπροσθέτως, ο συγκεκριμένος πιλότος εκμεταλλεύτηκε την υπηρεσία για αναγνώριση του φοιτητή μέσω της παροχής smart badges που αποδεικνύουν την εμπειρία και ικανότητα του χρήστη σε συγκεκριμένες γνώσεις και δεξιότητες με τη χρήση μικρο-διαπιστεύσεων (micro-accreditations). Το QualiChain μέσω του συγκεκριμένου πιλότου επέτρεψε την επαλήθευση όχι μόνο «σκληρών δεξιοτήτων» (hard skills) (όπως π.χ. η ικανότητα προγραμματισμού στη γλώσσα προγραμματισμού Python) αλλά και «μαλακών δεξιοτήτων» (soft skills) (π.χ. επικοινωνιακές δεξιότητες, ομαδικό πνεύμα), που είναι πολύ δύσκολο να κριθούν από τους εργοδότες ή από εταιρείες εύρεσης ανθρώπινου δυναμικού. Μέχρι στιγμής, τέτοιες δεξιότητες αξιολογούνται μέσω συνεντεύξεων και διαγωνισμάτων που αξιολογούν την συναισθηματική νοημοσύνη και τις «μαλακές δεξιότητες» των υποψηφίων. Με την παροχή ενδείξεων κατοχής τέτοιων δεξιοτήτων, με επαληθεύσιμο τρόπο, το QualiChain έχει τη δυνατότητα να αλλάξει τον τρόπο αξιολόγησης και επιλογής υποψηφίων εργαζομένων. Οι βασικοί χρήστες της συγκεκριμένης περίπτωσης χρήσης είναι οι προπτυχιακοί φοιτητές, οι διδακτορικοί φοιτητές, οι καθηγητές και η σχολή Η.Μ.Μ.Υ. του Ε.Μ.Π. ως σύνολο.

- **Προπτυχιακοί φοιτητές:** Μια από τις βασικές ανάγκες των φοιτητών είναι η επιλογή μαθημάτων που είναι όσο το δυνατόν πιο ενδιαφέροντα για αυτούς και παράλληλα μπορούν να βοηθήσουν στη βελτίωση των δεξιοτήτων και των γνώσεων τους έτσι ώστε να είναι καλύτερα προετοιμασμένοι για την ένταξη τους στην αγορά εργασίας με την αποφοίτησή τους. Μέχρι σήμερα δεν υπήρχε τεχνική υποδομή, οργανωμένες βάσεις δεδομένων και συστήματα συστάσεων που βοηθούν στην επιλογή μαθημάτων. Συνεπώς, οι φοιτητές πρέπει να κάνουν τη συγκεκριμένη δουλειά μόνοι τους, αφιερώνοντας σημαντικό χρόνο για την επιλογή των καταλληλότερων μαθημάτων για αυτούς. Επιπροσθέτως, δεν υπάρχει τρόπος επαλήθευσης συγκεκριμένων γνώσεων και δεξιοτήτων από τους φοιτητές, όπως για παράδειγμα προχωρημένη γνώση μια γλώσσας προγραμματισμού, πράγμα που αποτελεί μεγάλο πρόβλημα όσον αφορά την απόδειξη ότι όντως διαθέτουν τις γνώσεις και δεξιότητες που λένε ότι διαθέτουν σε πιθανούς εργοδότες. Η πλατφόρμα QualiChain παρέχει στους φοιτητές την ευκαιρία να έχουν το προσωπικό τους προφίλ που επιτρέπει την γρήγορη επαλήθευση των γνώσεων και δεξιοτήτων και των προσόντων που έχουν κερδίσει μέσα από τα διάφορα μαθήματα και εργασίες που έχουν ολοκληρώσει. Παράλληλα, το προφίλ βοηθάει τους χρήστες να λάβουν σωστές και αξιόπιστες αποφάσεις, όσον αφορά τα μαθήματα και τις εξωτερικές δραστηριότητες που θα επιλέξουν.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

- **Διδακτορικοί φοιτητές:** Οι περισσότεροι διδακτορικοί φοιτητές στο πλαίσιο των σπουδών τους εκτός από την έρευνα εκτελούν και άλλα καθήκοντα (π.χ. διόρθωση διαγωνισμάτων, επικουρική διδασκαλία). Ωστόσο, πολλές φορές δεν αναγνωρίζεται η εκπλήρωση αυτών των επιπλέον καθηκόντων - ή απαιτείται άτυπη διαδικασία όπως π.χ. η σύνταξη κατάλληλων συστατικών επιστολών. Το QualiChain επιτρέπει τη χρήση smart badges για την έμμεση αναγνώριση αυτών των υπηρεσιών. Επομένως, έχουν τη δυνατότητα να χτίσουν το προφίλ τους και να το βελτιώσουν με προσόντα που μέχρι στιγμής δεν ήταν επαληθεύσιμα.
- **Καθηγητές πανεπιστημίου:** Μέχρι σήμερα, η ανανέωση ενός μαθήματος, είναι μια πολύ χρονοβόρα διαδικασία, μιας και οι καθηγητές δε διαθέτουν έναν αυτοματοποιημένο και δομημένο τρόπο να γνωρίζουν τις ανάγκες και τις εξελίξεις της αγοράς εργασίας και να τις συγκρίνουν με τις γνώσεις και δεξιότητες που προσφέρονται στα μαθήματα τους. Οι λειτουργικότητες αναλυτικής δεδομένων και υποστήριξης αποφάσεων του QualiChain επιτρέπουν στους καθηγητές γρήγορα και αποτελεσματικά να έχουν επίγνωση για τις γνώσεις και δεξιότητες και τις γνώσεις που θα μπορούσαν να προστεθούν στα μαθήματα τους. Επιπλέον, οι καθηγητές μπορούν να δίνουν smart badges σε προπτυχιακούς και διδακτορικούς φοιτητές για την αναγνώριση των γνώσεων και δεξιοτήτων και της συνεισφοράς τους μέσω της πλατφόρμας QualiChain.
- **Η σχολή Η.Μ.Μ.Υ:** Παρότι η σχολή Η.Μ.Μ.Υ. δεν έχει συγκεκριμένο ρόλο στο πλαίσιο του πιλότου αυτού, αναμένεται να επωφεληθεί από τις λειτουργικότητες που παρέχονται στους φοιτητές και στους καθηγητές. Η ανανέωση του προγράμματος σπουδών μπορεί να οδηγήσει σε καλύτερη φήμη στην αγορά εργασίας, ενώ η επαλήθευση ακαδημαϊκών προσόντων, και γνώσεων και δεξιοτήτων μπορεί να βοηθήσει τους φοιτητές να βρουν κάποια θέση εργασίας και να μειώσει το χρόνο των διαχειριστικών διαδικασιών.

## 8.2 Ερευνητικές ερωτήσεις και KPIs

Στην ενότητα αυτή παρουσιάζονται οι σημαντικότερες ερευνητικές ερωτήσεις που φιλοδοξεί να απαντήσει ο συγκεκριμένος πιλότος, οι περιπτώσεις επαλήθευσης (test cases) και τα KPIs που σχεδιάστηκαν για την αξιολόγηση της απόδοσης και της χρησιμότητας του συγκεκριμένου πιλότου.

Οι περισσότεροι φοιτητές σε όλο τον κόσμο δε διαθέτουν ένα δομημένο και έξυπνο τρόπο για να επιλέξουν μαθήματα και να λάβουν αποφάσεις αξιοποιώντας σημαντικές πληροφορίες για την κατάσταση της αγοράς εργασίας που θα τους βοηθήσουν να ακολουθήσουν την καριέρα που επιθυμούν. Για το λόγο αυτό, ο συγκεκριμένος πιλότος προσπαθεί να αποσαφηνίσει κατά πόσο η αναλυτική δεδομένων και η υποστήριξη αποφάσεων που παρέχουν τα εργαλεία της πλατφόρμας QualiChain μπορούν να απλοποιήσουν και να ενισχύσουν τις επιλογές των φοιτητών και να τους βοηθήσουν να λάβουν πληροφορημένες αποφάσεις για τη μελλοντική καριέρα τους. Επιπλέον, στην Ελλάδα υπάρχει σημαντικό κενό μεταξύ της ακαδημαϊκής εκπαίδευσης και της αγοράς εργασίας. Πιο συγκεκριμένα, πολλοί φοιτητές αισθάνονται ότι οι σπουδές τους δεν τους παρέχουν τις απαραίτητες γνώσεις και

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

δεξιότητες για να είναι ανταγωνιστικοί στην αγορά εργασίας. Επομένως, ο συγκεκριμένος πιλότος προσπαθεί να ερευνήσει εάν η πλατφόρμα QualiChain μπορεί σε κάποιο βαθμό να καλύψει το κενό που υπάρχει μεταξύ των ακαδημαϊκών ιδρυμάτων και της αγοράς εργασίας, με την παροχή προτάσεων για τη βελτίωση του προγράμματος σπουδών της σχολής Η.Μ.Μ.Υ. του Ε.Μ.Π.. Ακόμη, ο πιλότος αυτός προσπαθεί να διερευνήσει τη σημασία και την προστιθέμενη αξία από τη χρήση των smart badges σε ένα πανεπιστήμιο. Τέλος, μια ερευνητική ερώτηση σχετική με τον πιλότο αυτόν που προκύπτει, αλλά διερευνάται στα πλαίσια άλλων πιλότων του QualiChain, είναι κατά πόσο το blockchain έχει τη δυναμική να παρέχει μια νέα λύση για αποκεντρωμένη επαλήθευση προσόντων και πτυχίων. Στο πλαίσιο της συγκεκριμένης διδακτορικής διατριβής δε θα ασχοληθούμε με την επαλήθευση προσόντων και την τεχνολογία blockchain. Για το λόγο αυτό, δε θα παρουσιαστούν οι περιπτώσεις επαλήθευσης και τα KPIs που σχετίζονται με αυτά.

Ο ακόλουθος πίνακας παρέχει μια γενική εικόνα των περιπτώσεων επαλήθευσης που σχεδιάστηκαν για το συγκεκριμένο πιλότο.

Πίνακας 25 - Σενάρια περιπτώσεων επαλήθευσης για τον πιλότο

Περιπτώσεις χρήσης	Στόχος	Εμπλεκόμενοι
A	Οι φοιτητές δημιουργούν το προσωπικό τους προφίλ και χρησιμοποιούν τεχνικές ανάλυσης δεδομένων και υποστήριξης αποφάσεων για να βελτιώσουν τις γνώσεις και δεξιότητες τους και τις γνώσεις τους, επιλέγοντας κατάλληλα μαθήματα και εξωσχολικές δραστηριότητες	<ul style="list-style-type: none"> <li>• Προπτυχιακοί φοιτητές</li> <li>• Δια βίου μαθητές</li> </ul>
B	Καθηγητές της σχολής Η.Μ.Μ.Υ. χρησιμοποιούν τα εργαλεία ανάλυσης δεδομένων και υποστήριξης αποφάσεων για να ανανεώσουν τα μαθήματά τους	<ul style="list-style-type: none"> <li>• Καθηγητές</li> <li>• Η σχολή Η.Μ.Μ.Υ. του Ε.Μ.Π.</li> </ul>

Ο παρακάτω πίνακας παρουσιάζει τα KPIs που ορίστηκαν για τις δυο πρώτες περιπτώσεις χρήσης για τη μέτρηση του αποτελέσματος του συγκεκριμένου πιλότου.

Πίνακας 26 - KPIs για τον πιλότο

KPI	Πώς συλλέγεται η πληροφορία;	Από ποιον;	Πόσο συχνά;	Αναμενόμενη αλλαγή στο χρόνο
<b>Περίπτωση χρήσης A</b>				
Χρόνος πρότασης ενός μαθήματος	Με τη χρήση μετρητών χρόνου (time trackers)	Από την πλατφόρμα QualiChain	Κάθε εξάμηνο	Μείωση
Βελτίωση των γνώσεων και δεξιοτήτων ενός χρήστη (αξιολογώντας το προφίλ του	Από τη βάση δεδομένων	Από την πλατφόρμα QualiChain	Κάθε εξάμηνο	Αύξηση



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

πριν και μετά την εφαρμογή του πιλοτικού σεναρίου) όσον αφορά τον αριθμό γνώσεων και δεξιοτήτων όσο και του βαθμού στον οποίο κατέχει κάθε γνώση/δεξιότητα				
Ποσοστό επιτυχημένων προτάσεων μαθημάτων (ποσοστό φοιτητών που ακολούθησαν την πρόταση του συστήματος)	Ερωτηματολόγιο	Από τους φοιτητές της σχολής Η.Μ.Μ.Υ.	Κάθε εξάμηνο	Αύξηση
Ικανοποίηση των φοιτητών από τη διαδικασία επιλογής μαθημάτων	Ερωτηματολόγιο	Από τους φοιτητές της σχολής Η.Μ.Μ.Υ.	Κάθε εξάμηνο	Αύξηση
<b>Περίπτωση χρήσης Β</b>				
Χρόνος για τη διερεύνηση μιας πιθανής ενημέρωσης μαθήματος με βάση τις ανάγκες της αγοράς εργασίας	Με τη χρήση μετρητών χρόνου (time trackers)	Από την πλατφόρμα QualiChain	Ετησίως	Μείωση
Ικανοποίηση των καθηγητών από την διαδικασία ανανέωσης μαθημάτων	Ερωτηματολόγιο	Από τους καθηγητές της σχολής Η.Μ.Μ.Υ.	Κάθε εξάμηνο	Αύξηση

Τα KPIs του συγκεκριμένου πιλότου σχεδιάστηκαν από τις πρώτες φάσης υλοποίησης του έργου QualiChain και ενημερώνονταν καθ' όλη τη διάρκεια υλοποίησης του έργου ώστε να λαμβάνουν υπόψη τις ανάγκες του πιλότου καθώς επίσης και τα αποτελέσματα. Σημαντική αλλαγή στην καθημερινότητα των φοιτητών και των εργαζομένων αποτέλεσε η πανδημία του Covid-19, η οποία προκάλεσε σημαντικές αλλαγές στη λειτουργία των πανεπιστημίων γενικά. Το γεγονός ότι οι καθηγητές, καθώς επίσης και οι προπτυχιακοί και διδακτορικοί φοιτητές δε χρειαζόταν να έχουν φυσική παρουσία στα μαθήματα του πανεπιστημίου προκάλεσε σημαντικές δυσκολίες όσον αφορά τη διάχυση της πλατφόρμας σε μεγάλο αριθμό ενδιαφερόμενων της σχολής Η.Μ.Μ.Υ..

Μια από τις αρχικές φιλοδοξίες του συγκεκριμένου πιλότου ήταν η μαζική προσέγγιση φοιτητών μέσω της ενημέρωσης και επίδειξης της πλατφόρμας κατά την παράδοση μαθημάτων στους φοιτητές. Παρότι τέτοιες δραστηριότητες πραγματοποιήθηκαν κανονικά κατά την εκτέλεση του πιλότου, η πανδημία Covid-19 κατέστησε δύσκολη την προσέλευση αρκετών ενδιαφερόμενων. Στην πράξη, οι ενδιαφερόμενοι προσεγγίστηκαν και ενημερώθηκαν μέσω email, ανακοινώσεων στα διαδικτυακά κανάλια της σχολής και παρουσιάσεων του έργου σε διαδικτυακά μαθήματα του εργαστηρίου συστημάτων αποφάσεων, το οποίο ήταν υπεύθυνο για την υλοποίηση του συγκεκριμένου πιλότου.

Κατά την εκτέλεση του πιλότου και με βάση τους περιορισμούς που έθετε η πανδημία, πραγματοποιήθηκαν εργαστήρια (workshops) σε ομάδες των 10 περίπου προπτυχιακών φοιτητών, ενώ

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Οι καθηγητές ενημερώθηκαν ατομικά για την πλατφόρμα QualiChain και τις υπηρεσίες που αναπτύχθηκαν. Το γεγονός αυτό δημιούργησε αντικειμενικές δυσκολίες στην οργάνωση και επίδειξη πιο προχωρημένων λειτουργιών του QualiChain, όπως για παράδειγμα η ανανέωση ενός μαθήματος. Παρότι αρκετοί καθηγητές της σχολής έδειξαν ενδιαφέρον για την πλατφόρμα και τη χρησιμοποίησαν για να τους βοηθήσει να ανανεώσουν κάποια από τα μαθήματα τους, ήταν αδύνατο να παρακολουθήσουμε τις πραγματικές αντιδράσεις των φοιτητών στις αλλαγές αυτές, καθώς λόγω πανδημίας υπήρχε μικρή συμμετοχή στα μαθήματα και ο διαθέσιμος χρόνος για την αξιολόγηση του πιλότου ήταν πολύ περιορισμένος. Επίσης, ενώ κάποιοι καθηγητές χρησιμοποίησαν την πλατφόρμα για την ανανέωση των μαθημάτων τους και πείστηκαν για τη χρησιμότητα των προτάσεων του συστήματος, η ανανέωση ενός πανεπιστημιακού μαθήματος δεν είναι απλή διαδικασία, ειδικά αν σκεφτεί κανείς τον περιορισμένο χρόνο εκτέλεσης ενός πιλοτικού προγράμματος στο πλαίσιο ενός ερευνητικού έργου. Μια τέτοια δραστηριότητα θα έπρεπε να περιλαμβάνει την πλειονότητα των καθηγητών της σχολής, ειδικά σε ό,τι αφορά περιπτώσεις στις οποίες μια γνώση/δεξιότητα ταιριάζει και προτείνεται σε διαφορετικά μαθήματα, ώστε η προσθήκη κάποιων γνώσεων και δεξιοτήτων να γίνει πιο αποτελεσματικά. Επιπροσθέτως, οι καθηγητές έπρεπε να σκεφτούν το φόρτο εργασίας κάθε μαθήματος, τις τροποποιήσεις που θα έπρεπε να γίνουν στο πρόγραμμα και τη διδασκαλία του μαθήματος, όπως επίσης και στις εργασίες, ώστε να περιλαμβάνουν και τις νέες γνώσεις και δεξιότητες. Οι καθηγητές που συμμετείχαν στο συγκεκριμένο πιλότο δήλωσαν ότι κάτι τέτοιο δεν ήταν ρεαλιστικό στο χρονικό πλαίσιο του πιλότου και στο περιβάλλον απομακρυσμένης μάθησης που προκάλεσε η πανδημία.

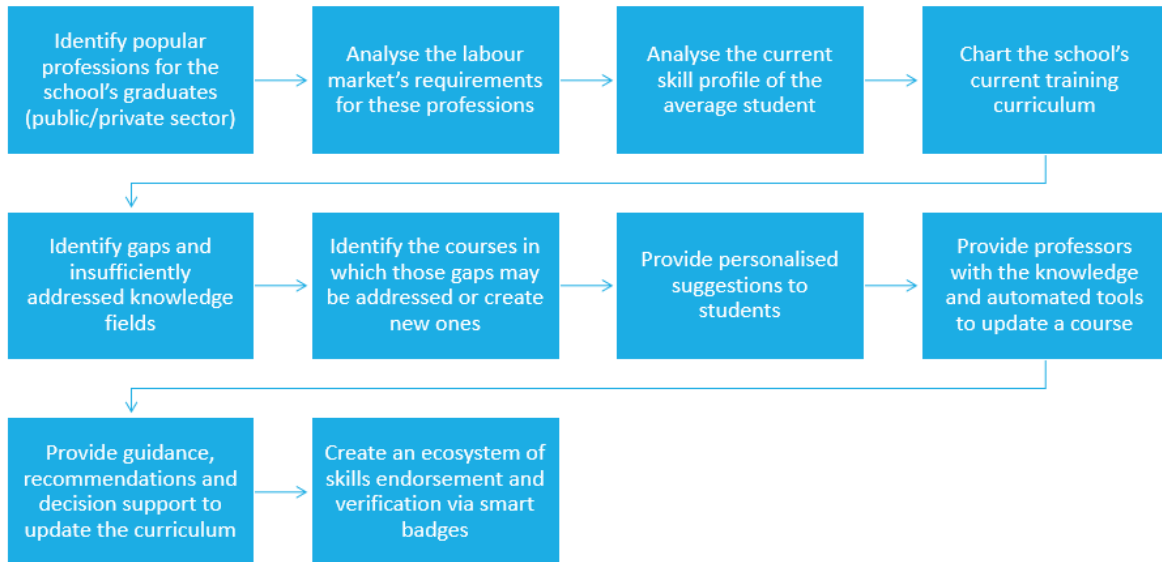
Ενώ πολλά από τα KPIs που ορίστηκαν, μετρήθηκαν κατευθείαν από την πλατφόρμα, άλλα KPIs αφορούσαν τη γνώμη των χρηστών και κατά συνέπεια, απαιτούσαν τη δημιουργία ερωτηματολογίων. Σε αυτό το πλαίσιο, δυο ερωτηματολόγια δημιουργήθηκαν για την εκτέλεση του συγκεκριμένου πιλότου. Το ένα αφορούσε τους φοιτητές (προπτυχιακούς και διδακτορικούς) και το άλλο τους καθηγητές.

## **8.3 Εκτέλεση και Αποτελέσματα του πιλότου**

### **8.3.1 Μεθοδολογία**

Η παρούσα ενότητα περιγράφει τη μεθοδολογία που ακολουθήσαμε στην εκτέλεση του συγκεκριμένου πιλότου. Συγκεκριμένα, παρέχει μια σύντομη περιγραφή για κάθε βήμα που ακολουθήθηκε πριν την εκτέλεση του πιλότου. Επίσης παρουσιάζει την εμπλοκή των διαφορετικών χρηστών με την πλατφόρμα. Η ακόλουθη εικόνα δείχνει τη διαδικασία που ακολουθήθηκε από την αρχή του έργου για την προετοιμασία και ολοκλήρωση του συγκεκριμένου πιλότου.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 37 - Μεθοδολογία του πιλότου

Το πρώτο βήμα ήταν η ανάλυση των απαιτήσεων της αγοράς εργασίας για κάποια δημοφιλή επαγγέλματα για τους αποφοίτους της σχολής, σε σχέση με το προφίλ γνώσεων και δεξιοτήτων, γνώσεων και εμπειρίας του μέσου (προπτυχιακού) φοιτητή της σχολής. Σε αυτό το πλαίσιο, οι υπηρεσίες αναλυτικής δεδομένων του έργου QualiChain, παρείχαν διάφορους δείκτες σχετικά με τα εναλλακτικά επαγγέλματα που μπορούν να ακολουθήσουν οι απόφοιτοι της σχολής, παράλληλα με πληροφορίες για τη δημοφιλία τους. Με βάση τη διαδικασία που περιεγράφηκε, στα πλαίσια του πιλότου αναλύθηκε σε βάθος το πρόγραμμα σπουδών της σχολής H.M.M.Y. ώστε να ανιχνευθούν γνώσεις και δεξιότητες που δεν προσφέρονται από το πρόγραμμα σπουδών αλλά είναι σε υψηλή ζήτηση από την αγορά εργασίας, ώστε να εισαχθούν στο πρόγραμμα σπουδών στην ύλη κάποιου μαθήματος. Όλα τα παραπάνω δεδομένα χρησιμοποιήθηκαν ως είσοδος στις υπηρεσίες ανάλυσης δεδομένων και υποστήριξης αποφάσεων της πλατφόρμας QualiChain, ώστε να παρέχουν προσωποποιημένες συστάσεις στους φοιτητές, ως προς τις γνώσεις και δεξιότητες και τα μαθήματα που θα ήταν καλό να επιλέξουν, όπως επίσης και στους καθηγητές που θα ήθελαν να ενημερώσουν τα μαθήματα τους και γενικά να λάβουν καθοδήγηση και υποστήριξη στις αποφάσεις τους για την ανανέωση του προγράμματος σπουδών της σχολής. Με βάση τη συγκεκριμένη μεθοδολογία, τα διακριτά βήματα που ακολουθήθηκαν κατά την εκτέλεση του πιλότου είναι τα εξής:

1. Αναγνώριση δημοφιλών επαγγελμάτων για τους αποφοίτους της σχολής. Η ομάδα έργου όρισε πέντε κατηγορίες επαγγελμάτων (Επιστήμη Υπολογιστών και δεδομένων, Επιστήμη ηλεκτρολόγου μηχανικού, Οργάνωση και διοίκηση (Management), Μηχανική τηλεπικοινωνιών, και Διδασκαλία), που περιλαμβάνουν όλες τις εναλλακτικές πορείες καριέρας (career trajectories) που μπορεί να ακολουθήσει ένας απόφοιτος, με βάση τις γνώσεις που προσφέρονται από το πρόγραμμα σπουδών της σχολής. Κάθε μία από τις παρακάτω κατηγορίες αντιπροσωπεύει μια σειρά από συγκεκριμένα επαγγέλματα τα οποία χρησιμοποιήθηκαν για την

αναζήτηση και εξαγωγή πληροφοριών από διάφορους ιστοτόπους εύρεσης εργασίας με τη χρήση τεχνικών απόξεσης δεδομένων από το διαδίκτυο (web scraping).

2. Εξαγωγή πληροφορίας από διάφορους ιστοτόπους εύρεσης εργασίας με τη χρήση απόξεσης δεδομένων. Αυτή η διαδικασία συνέβαλλε σημαντικά στην αναγνώριση των αναγκών της αγοράς εργασίας για όλους τους τομείς που αναφέρθηκαν.
3. Ανάλυση του μέσου επιπέδου γνώσεων και δεξιοτήτων για έναν μέσο φοιτητή. Σε αυτό το πλαίσιο, ζητήθηκε από φοιτητές της σχολής να στείλουν τα βιογραφικά τους στην ομάδα έργου. Σημειώνεται ότι όλη η διαδικασία που ακολουθήθηκε έγινε νομότυπα και όλη η διαχείριση των δεδομένων ήταν GDPR compliant. Τελικά, η ομάδα έργου συγκέντρωσε 70 βιογραφικά από προπτυχιακούς και διδακτορικούς φοιτητές. Ο αριθμός αυτός θα μπορούσε να είναι μεγαλύτερος, ωστόσο η διαδικασία συγκέντρωσης βιογραφικών ξεκίνησε ακριβώς πριν το ξέσπασμα της πανδημίας, πράγμα που δυσχέρανε σημαντικά τη διαδικασία, καθώς έπρεπε οι φοιτητές να στείλουν μαζί με το βιογραφικό τους συμπληρωμένες και υπογεγραμμένες φόρμες στις οποίες παρέχουν τη συγκατάθεση τους για τη χρήση του βιογραφικού τους από το έργο. Επίσης, αξίζει να σημειωθεί ότι η υπογραφή των φοιτητών έπρεπε να είναι επικυρωμένη. Στην αρχή, οι φοιτητές έπρεπε να επισκεφτούν κάποια δημόσια αρχή για να την επικυρώσουν. Ωστόσο, στη συνέχεια, μπορούσαν να το κάνουν διαδικτυακά. Τα βιογραφικά που συγκεντρώθηκαν βοήθησαν τόσο στην ανάλυση του μέσου προφίλ γνώσεων και δεξιοτήτων ενός φοιτητή όσο και στη στρατηγική που έπρεπε να ακολουθηθεί στο νομικό κομμάτι του έργου, όπως ορίστηκε στα αντίστοιχα πακέτα εργασίας.
4. Ανάπτυξη μοντέλου γνώσης για το πρόγραμμα σπουδών της σχολής Η.Μ.Μ.Υ. που περιλαμβάνει πάνω από 140 μαθήματα εξειδίκευσης. Το μοντέλο αυτό συνέβαλλε σημαντικά στην εκπαίδευση και ανάπτυξη των συστημάτων συστάσεων της πλατφόρμας.
5. Μετά τη συγκέντρωση των βιογραφικών και την ανάλυση των απαιτήσεων της αγοράς εργασίας, πραγματοποιήθηκαν πολλές αναλύσεις για την εκτίμηση του κενού μεταξύ των γνώσεων και δεξιοτήτων που ζητούνται από την αγορά εργασίας και των γνώσεων και δεξιοτήτων που προσφέρει το δοθέν πρόγραμμα σπουδών. Το βήμα αυτό αποκάλυψε πληθώρα γνώσεων και δεξιοτήτων που είναι σε υψηλή ζήτηση από την αγορά εργασίας αλλά δε προσφέρονται σε κανένα από τα μαθήματα του δοθέντος προγράμματος σπουδών. Τέτοιες γνώσεις και δεξιότητες περιλαμβάνουν μεταξύ άλλων τις τεχνολογίες Docker, React, Android development, iOS development και JavaScript.
6. Οι υπηρεσίες συστάσεων και ανάλυσης δεδομένων της πλατφόρμας QualiChain λαμβάνουν υπόψη τη μοντελοποίηση του προγράμματος σπουδών και των δημοφιλών γνώσεων και δεξιοτήτων που δε προσφέρονται σε αυτό ώστε να προτείνουν μαθήματα στα οποία θα μπορούσαν να εισαχθούν. Η παρακάτω εικόνα παρουσιάζει την έξοδο της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών για μερικές από τις γνώσεις και δεξιότητες που πρέπει να εισαχθούν σε αυτό, μαζί με τα μαθήματα στα οποία προτείνεται να εισαχθεί κάθε γνώση ή

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

δεξιότητα. Κάποιες δεξιότητες προτείνονται σε πολλά μαθήματα, ενώ για άλλες δε φαίνεται να υπάρχουν πολύ σχετικά μαθήματα, και κατά συνέπεια προτείνεται η δημιουργία νέων μαθημάτων.

7. Υλοποίηση της υπηρεσίας προσωποποιημένων συστάσεων μαθημάτων και δεξιοτήτων στους φοιτητές και άλλους ενδιαφερόμενους.

The image shows two screenshots of a recommendation system interface. The top screenshot shows a list of programming languages on the left and a list of suggested courses on the right. The bottom screenshot shows a similar interface with a different set of programming languages and course suggestions.

Programming Language	Suggested Courses
css	Databases Internet Programming
json	
git	Distributed Systems Programming Languages II Digital Communications I Advanced Topics in Database Systems Operating Systems Operating Systems Laboratory Programming Languages I Algorithms and Complexity Artificial Intelligence Compilers
ios	
javascript	Databases Internet Programming Multimedia Technology Computer System Performance Distributed Systems Programming Languages I Software Engineering Compilers Programming Languages II
akka	Human- Computer Interaction
go	

**Εικόνα 38 - Παραδείγματα δεξιοτήτων προς εισαγωγή σε μαθήματα με βάση τις προτάσεις της υπηρεσίας ανασχεδιασμού προγράμματος σπουδών**

Τα 7 βήματα που παρουσιάστηκαν αποτελούν την προπαρασκευαστική φάση του πιλότου. Ακολούθως, η πλατφόρμα QualiChain και τα εργαλεία που αναπτύχθηκαν παρουσιάστηκαν εκτενώς σε φοιτητές και καθηγητές, και ξεκίνησε η υλοποίηση του πιλότου. Πριν την εισαγωγή νέων χρηστών στην πλατφόρμα, η προσέγγιση που θα ακολουθούνταν στο συγκεκριμένο πιλότο παρουσιάστηκε σε 6 εργαστήρια διαζώσης μέσα στην τάξη με πάνω από 70 φοιτητές κάθε φορά. Οι φοιτητές έδειξαν μεγάλο ενδιαφέρον για τις λειτουργίες της πλατφόρμας QualiChain και κατέληξαν ότι πραγματικά ταιριάζει με τις ανάγκες τους. Παράλληλα, η παρουσίαση του πιλότου σε επιστημονικά συνέδρια είχε ως αποτέλεσμα την πολύ θετική αντίδραση της επιστημονικής κοινότητας που περιλάμβανε μεταξύ άλλων διδακτορικούς φοιτητές και καθηγητές πανεπιστημίων από διεθνή πανεπιστήμια.

Για την αξιολόγηση της τεχνικής λύσης του έργου QualiChain, πραγματοποιήθηκε μια σειρά από εργαστήρια, ομάδες ενδιαφέροντος (focus groups) και επιδείξεις της πλατφόρμας σε προπτυχιακούς, και διδακτορικούς φοιτητές καθώς επίσης και καθηγητές της σχολής ώστε να καλυφθούν και τα τρία πιλοτικά σενάρια που έχουν παρουσιαστεί. Συγκεκριμένα, 50 προπτυχιακοί φοιτητές είχαν την ευκαιρία να γνωρίσουν την πλατφόρμα μέσα από τα μαθήματα του εργαστηρίου συστημάτων αποφάσεων. Επιπλέον, η πλατφόρμα παρουσιάστηκε σε 40 διδακτορικούς φοιτητές που περιλαμβάνουν τόσο φοιτητές του εργαστηρίου συστημάτων αποφάσεων όσο και άλλων ερευνητικών εργαστηρίων της

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

σχολής Η.Μ.Μ.Υ.. Τέλος, η πλατφόρμα παρουσιάστηκε σε 11 καθηγητές με τη χρήση email. Η πλατφόρμα παρουσιάστηκε σε κάθε μια από αυτές τις ομάδες ξεχωριστά. Τα εργαστήρια για προπτυχιακούς φοιτητές περιλάμβαναν 5-10 φοιτητές, ενώ τα εργαστήρια για τους διδακτορικούς περιλάμβαναν 10 φοιτητές. Τέλος, όσον αφορά τους καθηγητές, η πλατφόρμα παρουσιάστηκε στον καθένα ατομικά. Κάθε μια από τις συνεδρίες που αναφέρθηκαν κράτησε 1.5-2 ώρες.

Πιο συγκεκριμένα, τα πρώτα 15 λεπτά ήταν αφιερωμένα στη γνωριμία με τους συμμετέχοντες, και την εισαγωγή στην πλατφόρμα QualiChain και στο συγκεκριμένο πιλότο, καθώς επίσης και με τους στόχους του έργου και τους βασικούς ενδιαφερόμενους (stakeholders). Ακολούθως, οι συμμετέχοντες λάμβαναν ένα link για την πλατφόρμα και τους ζητούνταν να δημιουργήσουν έναν λογαριασμό. Στη συνέχεια, παρουσιαζόταν στους συμμετέχοντες η πλατφόρμα QualiChain επεξηγώντας τις διάφορες λειτουργικότητες που θα χρησιμοποιούσαν οι συμμετέχοντες.

Για τους προπτυχιακούς φοιτητές το επόμενο βήμα περιλάμβανε τη δημιουργία του προφίλ τους και την συμπλήρωση πληροφοριών για το βιογραφικό τους. Στη συνέχεια, οι φοιτητές καλούνταν να επιλέξουν μαθήματα από την αντίστοιχη ταμπέλα της πλατφόρμας, η οποία περιλάμβανε τόσο τα διαθέσιμα μαθήματα, όσο και προσωποποιημένες προτάσεις για να επιλέξουν οι φοιτητές. Επιπλέον, οι φοιτητές καλούνταν να συμπληρώσουν τα μαθήματα που έχουν ολοκληρώσει. Ακολούθως, παρουσιαζόταν στους φοιτητές η υπηρεσία πολυκριτηριακής υποστήριξης αποφάσεων, και καλούνταν να πραγματοποιήσουν μια ανάλυση σχετικά με την επιλογή ενός μαθήματος μεταξύ διαφορετικών επιλογών που προτείνονταν σε αυτούς με βάση το προφίλ τους. Στη συνέχεια, μοιράζονταν ερωτηματολόγια στους συμμετέχοντες και τους ζητούνταν να τα συμπληρώσουν όταν μπορέσουν, αφού περιηγηθούν ξανά στην πλατφόρμα και πειραματιστούν με τις λειτουργίες που τους παρέχονται.

Όμοια διαδικασία ακολουθήθηκε και για τους διδακτορικούς φοιτητές, με τη διαφορά ότι έπρεπε να δηλώσουν ότι είναι διδακτορικοί φοιτητές, και αντί για μαθήματα τους προτείνονταν σχετικές αγγελίες εργασίας ανάλογα με τα ενδιαφέροντα τους. Επίσης, η επίδειξη είχε να κάνει και με τη δυνατότητα επιβράβευσης με smart badges, ωστόσο αυτή η λειτουργία δεν εμπίπτει στο αντικείμενο της παρούσας διατριβής και κατά συνέπεια δε θα παρουσιαστεί περαιτέρω.

Τέλος, όσον αφορά τους καθηγητές, η πλατφόρμα παρουσιαζόταν με παρόμοιο τρόπο. Η πρώτη διαδικασία που έπρεπε να ολοκληρώσουν ήταν να δηλώσουν μέσα από την πλατφόρμα τα μαθήματα που διδάσκουν. Στη συνέχεια, καλούνταν να επισκεφτούν τη σελίδα ανασχεδιασμού προγράμματος σπουδών, και να ελέγξουν εάν διδάσκουν κάποιο μάθημα για το οποίο προτείνεται η εισαγωγή συγκεκριμένων γνώσεων και δεξιοτήτων. Στις περιπτώσεις στις οποίες μια γνώση ή δεξιότητα προτάθηκε για εισαγωγή σε περισσότερα του ενός μαθήματα ενός καθηγητή, ο τελευταίος καλούνταν να χρησιμοποιήσει την υπηρεσία υποστήριξης αποφάσεων για να διαλέξει το καταλληλότερο μάθημα. Τέλος, παρουσιαζόταν στους καθηγητές και η δυνατότητα χορήγησης smart badges, ωστόσο όπως έχει ήδη αναφερθεί δε θα παρουσιαστούν περαιτέρω.

Μετά την περάτωση του κάθε εργαστηρίου (workshop) οι συμμετέχοντες καλούνταν να συνεχίσουν να χρησιμοποιούν την πλατφόρμα και να στέλνουν τα σχόλια και τις προτάσεις τους για το πώς θα μπορούσε να βελτιωθεί. Τα σχόλια και οι προτάσεις των συμμετεχόντων βοήθησαν σημαντικά στη βελτίωση της πλατφόρμας, ώστε να γίνει ευκολότερη στη χρήση.

### 8.3.2 Αποτελέσματα ερωτηματολογίων

Η ενότητα αυτή παρουσιάζει τα αποτελέσματα των ερωτηματολογίων που συμπλήρωσαν οι χρήστες της πλατφόρμας QualiChain. Για τους φοιτητές τα αποτελέσματα παρουσιάζονται παρακάτω:

1. Πόσο ικανοποιημένος/η μείνατε από τη χρήση της πλατφόρμας QualiChain;
  - a. 0/90 ήταν πολύ απογοητευμένοι (very dissatisfied)
  - b. 1/90 ήταν απογοητευμένοι
  - c. 7/90 ήταν ουδέτεροι
  - d. 58/90 ήταν ικανοποιημένοι
  - e. 24/90 ήταν πολύ ικανοποιημένοι
2. Πόσο ικανοποιημένος/η μείνατε από την επιλογή μαθημάτων και άλλων δραστηριοτήτων από την πλατφόρμα QualiChain;
  - a. 0/90 ήταν πολύ απογοητευμένοι (very dissatisfied)
  - b. 2/90 ήταν απογοητευμένοι
  - c. 20/90 ήταν ουδέτεροι
  - d. 53/90 ήταν ικανοποιημένοι
  - e. 15/90 ήταν πολύ ικανοποιημένοι
3. Πόσο ικανοποιημένος/η μείνατε από την αποτελεσματικότητα της διαδικασίας επιλογής μαθήματος;
  - a. 0/90 ήταν πολύ απογοητευμένοι (very dissatisfied)
  - b. 5/90 ήταν απογοητευμένοι
  - c. 24/90 ήταν ουδέτεροι
  - d. 49/90 ήταν ικανοποιημένοι
  - e. 12/90 ήταν πολύ ικανοποιημένοι
4. Πόσο σχετικές θεωρείτε τις προτάσεις που λάβατε από την πλατφόρμα QualiChain σε σχέση με το ακαδημαϊκό/εκπαιδευτικό σας προφίλ;
  - a. 0/90 απάντησαν ότι οι προτάσεις δεν ήταν καθόλου σχετικές
  - b. 0/90 απάντησαν ότι οι προτάσεις ήταν κατά βάση άσχετες
  - c. 12/90 ότι οι προτάσεις ήταν ουδέτερες
  - d. 57/90 απάντησαν ότι οι προτάσεις ήταν σχετικές
  - e. 21/90 απάντησαν ότι οι προτάσεις ήταν πολύ σχετικές
5. Ποιος ήταν ο αριθμός μαθημάτων και άλλων δραστηριοτήτων που επιλέξατε με βάση τις προτάσεις της πλατφόρμας QualiChain;
  - a. 9/90 δεν επέλεξαν κανένα μάθημα ή δραστηριότητα που τους προτάθηκε
  - b. 27/90 επέλεξαν 1-3 μαθήματα ή δραστηριότητες που τους προτάθηκαν
  - c. 20/90 επέλεξαν 4-7 μαθήματα ή δραστηριότητες που τους προτάθηκαν
  - d. 21/90 επέλεξαν 8-10 μαθήματα ή δραστηριότητες που τους προτάθηκαν

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

- e. 13/90 επέλεξαν πάνω από 10 μαθήματα ή δραστηριότητες που τους προτάθηκαν
- 6. Σε τι βαθμό συμφωνείτε με την πρόταση «Το QualiChain με βοήθησε να δομήσω καλύτερα το ακαδημαϊκό/επαγγελματικό μου προφίλ και να διαλέξω μαθήματα και δραστηριότητες για τη βελτίωση του»;
  - a. 0/90 διαφωνούσαν έντονα με την πρόταση
  - b. 1/90 διαφωνούσαν με την πρόταση
  - c. 8/90 ούτε συμφωνούσαν ούτε διαφωνούσαν με την πρόταση
  - d. 48/90 συμφωνούσαν με την πρόταση
  - e. 33/90 συμφωνούσαν έντονα με την πρόταση
- 7. Σε τι βαθμό συμφωνείτε με την πρόταση «Πιστεύω ότι το QualiChain θα με βοηθήσει να βρω μια θέση εργασίας στο πεδίο ενδιαφέροντος μου γρηγορότερα μετά την αποφοίτησή μου»;
  - a. 0/90 διαφωνούσαν έντονα με την πρόταση
  - b. 1/90 διαφωνούσαν με την πρόταση
  - c. 16/90 ούτε συμφωνούσαν ούτε διαφωνούσαν με την πρόταση
  - d. 54/90 συμφωνούσαν με την πρόταση
  - e. 19/90 συμφωνούσαν έντονα με την πρόταση

Όσον αφορά τους καθηγητές, τα αποτελέσματα από τις απαντήσεις τους παρουσιάζονται παρακάτω:

- 1. Πόσο ικανοποιημένος/η μείνατε από τη χρήση της πλατφόρμας QualiChain;
  - a. 0/11 ήταν πολύ απογοητευμένοι (very dissatisfied)
  - b. 0/11 ήταν απογοητευμένοι
  - c. 1/11 ήταν ουδέτεροι
  - d. 8/11 ήταν ικανοποιημένοι
  - e. 2/11 ήταν πολύ ικανοποιημένοι
- 2. Πόσο ικανοποιημένος/η μείνατε κατά τη διαδικασία ανανέωσης των μαθημάτων σας με τη χρήση των υπηρεσιών της πλατφόρμας QualiChain;
  - a. 0/11 ήταν πολύ απογοητευμένοι (very dissatisfied)
  - b. 0/11 ήταν απογοητευμένοι
  - c. 3/11 ήταν ουδέτεροι
  - d. 5/11 ήταν ικανοποιημένοι
  - e. 3/11 ήταν πολύ ικανοποιημένοι
- 3. Πόσο σχετικές θεωρείτε τις προτάσεις που λάβατε από την πλατφόρμα QualiChain για την ενημέρωση των μαθημάτων σας;
  - a. 0/11 απάντησαν ότι οι προτάσεις δεν ήταν καθόλου σχετικές
  - b. 0/11 απάντησαν ότι οι προτάσεις ήταν κατά βάση άσχετες
  - c. 3/11 ότι οι προτάσεις ήταν ουδέτερες
  - d. 3/11 απάντησαν ότι οι προτάσεις ήταν σχετικές



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

- e. 5/11 απάντησαν ότι οι προτάσεις ήταν πολύ σχετικές
- 4. Σε τι βαθμό συμφωνείτε με την πρόταση «Το QualiChain με βοήθησε να δομήσω καλύτερα τα μαθήματα μου με νέες δεξιότητες και γνώσεις»;
  - a. 0/11 διαφωνούσαν έντονα με την πρόταση
  - b. 0/11 διαφωνούσαν με την πρόταση
  - c. 5/11 ούτε συμφωνούσαν ούτε διαφωνούσαν με την πρόταση
  - d. 5/11 συμφωνούσαν με την πρόταση
  - e. 1/11 συμφωνούσαν έντονα με την πρόταση
- 5. Σε τι βαθμό συμφωνείτε με την πρόταση «Πιστεύω ότι το QualiChain θα με βοηθήσει να προετοιμάσω καλύτερα τους φοιτητές μου για την αγορά εργασίας μέσω καλύτερα δομημένων μαθημάτων και προγράμματος σπουδών όσον αφορά τις διδασκόμενες δεξιότητες και γνώσεις»;
  - a. 0/11 διαφωνούσαν έντονα με την πρόταση
  - b. 0/11 διαφωνούσαν με την πρόταση
  - c. 3/11 ούτε συμφωνούσαν ούτε διαφωνούσαν με την πρόταση
  - d. 6/11 συμφωνούσαν με την πρόταση
  - e. 2/11 συμφωνούσαν έντονα με την πρόταση

### 8.3.3 Ανάλυση των KPIs

Τα αποτελέσματα των ερωτηματολογίων σε συνδυασμό με τις πληροφορίες χρήσης της πλατφόρμας βοηθούν να απαντήσουμε στην ερώτηση κατά πόσο ο πιλότος είναι επιτυχημένος.

#### 8.3.3.1 Περίπτωση χρήσης Α

Στην συγκεκριμένη περίπτωση χρήσης, οι φοιτητές δημιουργούν το προφίλ τους και χρησιμοποιούν τις υπηρεσίες ανάλυσης δεδομένων και υποστήριξης αποφάσεων για τη βελτίωση των γνώσεων και των δεξιοτήτων επιλέγοντας κατάλληλα μαθήματα και άλλες δραστηριότητες.

KPI 1 – Χρόνος για την πρόταση μαθημάτων: Το συγκεκριμένο KPI μετρήθηκε μέσα από την πλατφόρμα QualiChain. Συγκεκριμένα, πραγματοποιήθηκαν περισσότερες από 1397 προτάσεις μαθημάτων για τους διάφορους χρήστες της πλατφόρμας, συμπεριλαμβανομένων των αρχικών προτάσεων σε κάθε χρήστη όσο και τις προτάσεις που λάμβανε κάθε φορά που το προφίλ του ανανεωνόταν (με κάθε ανανέωση του προφίλ του, ο χρήστης λάμβανε και καινούριες προτάσεις). Ο μέσος χρόνος για την πρόταση μαθημάτων, όπως μετρήθηκε από την πλατφόρμα ήταν 0.236 δευτερόλεπτα. Ο μέσος χρόνος περιλαμβάνει όλες τις προτάσεις που γίνονταν στους χρήστες με τη συμπλήρωση του προφίλ τους και την υποβολή του βιογραφικού τους. Επίσης περιλαμβάνει και τις νέες προτάσεις που λαμβάνει ένας χρήστης μετά την ανανέωση του προφίλ ή του βιογραφικού του. Με βάση το γεγονός ότι η διαδικασία επιλογής κατάλληλων μαθημάτων, που είναι συμβατά με τις ανάγκες της αγοράς εργασίας, γίνεται από τους φοιτητές ψάχνοντας στο διαδίκτυο για ώρες, ώστε να είναι όσο το δυνατόν πιο σίγουροι για την επιλογή τους, συμπεραίνουμε ότι ένας μέσος χρόνος μικρότερος του ενός δευτερολέπτου είναι σημαντική

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

βελτίωση. Επιπλέον, όπως φαίνεται παρακάτω, στη μέτρηση του KPI 2 (ποσοστό επιτυχημένων προτάσεων), οι περισσότεροι φοιτητές ήταν ικανοποιημένοι από τις προτάσεις που λάμβαναν. Οπότε, συμπεραίνουμε ότι συνολικά η συγκεκριμένη λειτουργικότητα της πλατφόρμας κρίνεται επιτυχημένη από τους χρήστες που την χρησιμοποίησαν και την αξιολόγησαν.

KPI 2 – Ποσοστό επιτυχημένων προτάσεων μαθημάτων: Το συγκεκριμένο KPI μετρήθηκε μέσα από τις απαντήσεις σε ερωτηματολόγια. Στην ερώτηση «Πόσο σχετικές θεωρείτε τις προτάσεις μαθημάτων σε σχέση με το ακαδημαϊκό/επαγγελματικό σας προφίλ;» από τους 50 προπτυχιακούς φοιτητές οι 5 στους 50 (10%) απάντησαν βρήκαν τις προτάσεις ουδέτερες, 37 στους 50 (74%) απάντησαν ότι τις βρήκαν σχετικές ενώ 8 στους 50 (16%) απάντησαν ότι τις βρήκαν πολύ σχετικές. Επιπροσθέτως, στην ερώτηση, «Ποιος ήταν ο αριθμός μαθημάτων και άλλων δραστηριοτήτων που επιλέξατε με βάση τις προτάσεις της πλατφόρμας QualiChain;» 5 από τους 50 προπτυχιακούς φοιτητές (10%) απάντησαν ότι δεν επέλεξαν κανένα μάθημα με βάση τις προτάσεις της πλατφόρμας, 10 στους 50 (20%) επέλεξαν από 1 έως 3 μαθήματα, 10 στους 50 (20%) επέλεξαν από 4 έως 7 μαθήματα, 14 στους 50 (28%) επέλεξαν από 8 έως 10 μαθήματα και 11 στους 50 (22%) επέλεξαν περισσότερα από 10 μαθήματα με βάση τις προτάσεις του QualiChain. Πέρα από το 10% των ερωτηθέντων, οι οποίοι δεν επέλεξαν κανένα από τα προτεινόμενα μαθήματα ή δραστηριότητες, οι υπόλοιποι βρήκαν τις προτάσεις της πλατφόρμας πολύ σχετικές, με ποικιλία στις απαντήσεις σε ό,τι αφορά τον αριθμό των μαθημάτων και δραστηριοτήτων που επέλεξαν. Συνολικά, μπορούμε να πούμε ότι το σύστημα συστάσεων μαθημάτων και δεξιοτήτων του QualiChain αξιολογήθηκε θετικά από τους φοιτητές, και ότι βασίστηκαν στις προτάσεις του για να επιλέξουν μαθήματα.

KPI 3 – Ικανοποίηση των χρηστών με τη διαδικασία επιλογής: Το συγκεκριμένο KPI μετρήθηκε από τις απαντήσεις στα ερωτηματολόγια. Στην ερώτηση «Πόσο ικανοποιημένος/η μείνατε από την αποτελεσματικότητα της διαδικασίας επιλογής μαθήματος», 1 στους 50 (2%) απάντησε ότι δεν έμεινε ικανοποιημένος, 6 στους 50 (12%) δεν έμειναν ούτε ικανοποιημένοι αλλά ούτε απογοητεύτηκαν ενώ 35 στους 50 ερωτηθέντες (70%) έμειναν ικανοποιημένοι, και 8 στους 50 (16%) έμειναν πολύ ικανοποιημένοι με τη διαδικασία επιλογής μαθημάτων. Επιπροσθέτως, στην ερώτηση «Πόσο ικανοποιημένος/η μείνατε από την αποτελεσματικότητα της διαδικασίας επιλογής μαθήματος», 4 στους 50 (8%) ερωτηθέντες δεν έμειναν ικανοποιημένοι, 6 στους 50 (12%) δήλωσαν ουδέτεροι, ενώ 33 στους 50 (66%) έμειναν ικανοποιημένοι και 7 στους 50 ερωτηθέντες (14%) έμειναν πολύ ικανοποιημένοι από τη χρήση της πλατφόρμας. Το μικρό ποσοστό ανικανοποίητων χρηστών προέρχεται κυρίως από τη δυσκολία στη χρήση κάποιων υπηρεσιών υποστήριξης αποφάσεων. Επιπλέον, κάποιιοι χρήστες θα ήθελαν να συνδέεται η πλατφόρμα QualiChain με τη βάση δεδομένων της σχολής, ώστε να μην ξανά-δηλώνουν τα μαθήματα που έχουν επιλέξει και στο site της σχολής. Φυσικά, η διασύνδεση με την βάση δεδομένων της σχολής δεν ήταν εφικτή στο πλαίσιο του συγκεκριμένου έργου, λόγω πολύπλοκων διαχειριστικών διαδικασιών που χρειάζεται να γίνουν από τη γραμματεία της σχολής. Παρ' όλα αυτά, είναι ξεκάθαρο ότι οι περισσότεροι φοιτητές έμειναν ικανοποιημένοι με τις λειτουργικότητες της πλατφόρμας QualiChain και ενθουσιασμένοι με το γεγονός ότι παρείχε προσωποποιημένες υπηρεσίες με βάση τις ανάγκες τους.

### **8.3.3.2 Περίπτωση χρήσης Β**

KPI 1 – Χρόνος για τη διερεύνηση ανανέωσης ενός μαθήματος με βάση τις ανάγκες της αγοράς εργασίας: Το συγκεκριμένο KPI μετρήθηκε μέσω συνεντεύξεων και παρουσιάσεων σε καθηγητές της σχολής. Ο χρόνος διερεύνησης για ένα μάθημα περιλαμβάνει το χρόνο που χρειάζεται ένας καθηγητής να αναλύσει τις προτεινόμενες αλλαγές που παρέχει το σύστημα ανασχεδιασμού προγράμματος σπουδών καθώς επίσης και το χρόνο που χρειάζεται για να κάνει τις απαραίτητες αλλαγές μέσω της λειτουργίας ενημέρωσης του μαθήματος που προσφέρει η πλατφόρμα. Στα εργαστήρια που πραγματοποιήθηκαν για τους καθηγητές, η διαδικασία έπαιρνε περίπου 5 με 10 λεπτά, ενώ περιλάμβανε και επεξηγήσεις των λειτουργιών στους καθηγητές. Οπότε, μπορούμε να συμπεράνουμε ότι η όλη διαδικασία θα έπαιρνε λιγότερο από 5 λεπτά για ένα χρήστη που ξέρει πώς να χρησιμοποιήσει την πλατφόρμα. Επίσης, εφόσον ο περισσότερος χρόνος αφιερώνεται στην ανάλυση των προτεινόμενων αλλαγών, μπορούμε να συμπεράνουμε ότι ο χρόνος για την ανανέωση ενός μαθήματος με βάση τις ανάγκες της αγοράς εργασίας είναι σημαντικά μειωμένος σε σχέση με τη διαδικασία που ακολουθείται παραδοσιακά, στην οποία οι καθηγητές πρέπει να διερευνήσουν μόνοι τους τις ανάγκες της αγοράς εργασίας και να ανακαλύψουν μόνοι τους τις γνώσεις και δεξιότητες που ταιριάζουν στο μάθημα τους.

KPI 2 – Ικανοποίηση των καθηγητών από την διαδικασία ανανέωσης μαθημάτων του QualiChain: Το συγκεκριμένο KPI μετρήθηκε από τις απαντήσεις στα ερωτηματολόγια. Συγκεκριμένα, έντεκα συμμετέχοντες καθηγητές χρησιμοποίησαν και αξιολόγησαν τη λειτουργία ανανέωσης μαθήματος του QualiChain και τις προτάσεις για το πρόγραμμα σπουδών. Στην ερώτηση «Πόσο ικανοποιημένος/η μείνατε κατά τη διαδικασία ανανέωσης των μαθημάτων σας με τη χρήση των υπηρεσιών της πλατφόρμας QualiChain;» 3 από τους 11 (27%) απάντησαν ότι δεν ήταν ούτε ευχαριστημένοι ούτε δυσαρεστημένοι, 5 στους 11 (46%) απάντησαν ότι ήταν ικανοποιημένοι και 3 στους 11 (27%) απάντησαν ότι ήταν πολύ ικανοποιημένοι. Επιπλέον, στην ερώτηση «Πόσο σχετικές θεωρείτε τις προτάσεις που λάβατε από την πλατφόρμα QualiChain για την ενημέρωση των μαθημάτων σας;» 3 στους 11 (27%) απάντησαν ότι οι προτάσεις ήταν ουδέτερες, 3 στους 11 (27%) απάντησαν ότι ήταν σχετικές, και 5/11 (46%) απάντησαν ότι οι προτάσεις του συστήματος ήταν πολύ σχετικές. Με βάση τις απαντήσεις των καθηγητών, καταλήγουμε ότι οι περισσότεροι καθηγητές έμειναν αρκετά ικανοποιημένοι από τις προτάσεις που λάμβαναν για τη διαδικασία ανανέωσης μαθήματος. Το μικρό ποσοστό ουδέτερων απαντήσεων μπορεί να συσχετιστεί με καθηγητές που στα μαθήματα τους δεν προτεινόταν αλλαγές από το σύστημα.

## **8.4 Πολύτιμα μαθήματα και επεκτάσεις**

### **8.4.1 Συγκέντρωση βιογραφικών από φοιτητές και προστασία δεδομένων**

Κατά την αρχική περίοδο του συγκεκριμένου πιλότου, ζητήθηκαν βιογραφικά από τους προπτυχιακούς και τους διδακτορικούς φοιτητές, ώστε να αναλυθούν από την ομάδα έργου και μέσα από την ανάλυση να εξαχθούν συμπεράσματα ως προς το μέσο επίπεδο γνώσεων και δεξιοτήτων των φοιτητών. Μιας και το βιογραφικά που ζητήθηκαν αποτελούν προσωπικά δεδομένα, σύμφωνα με τη νομική στρατηγική του έργου που είναι συμβατή με το GDPR, έπρεπε να συνοδεύονται από μια φόρμα συναίνεσης (consent form). Η συγκεκριμένη φόρμα έπρεπε να είναι υπογεγραμμένη και επικυρωμένη από κάποια δημόσια

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

αρχή (π.χ. από την αστυνομία) πράγμα το οποίο ήταν σημαντικός επιπλέον φόρτος για τους φοιτητές. Για το λόγο αυτό, στην αρχική φάση του πιλότου, η συμμετοχή των φοιτητών ήταν χαμηλή όσον αφορά τα βιογραφικά που συγκεντρώθηκαν. Κατά το δεύτερο χρόνο εκτέλεσης του πιλότου, και κατά το πρώτο κύμα της πανδημίας Covid-19, έγιναν διαθέσιμες νέες ψηφιακές υπηρεσίες από την ελληνική κυβέρνηση που κατέστησαν δυνατό να πραγματοποιείται η διαδικασία επικύρωσης διαδικτυακά. Η συγκεκριμένη αλλαγή συνέβαλλε αποφασιστικά στην αύξηση των βιογραφικών που συγκεντρώθηκαν. Ενδεικτικά, πριν τη συγκεκριμένη αλλαγή, καταφέραμε να μαζέψουμε μόνο 10 βιογραφικά, ενώ μετά την εισαγωγή της διαδικτυακής πλατφόρμας επικύρωσης, συγκεντρώθηκαν ακόμα 60 βιογραφικά, σε πολύ μικρό χρονικό διάστημα. Είναι προφανές ότι στις μοντέρνες κοινωνίες, που οι τεχνολογίες πληροφορίας και επικοινωνιών παρέχουν καινοτόμες λύσεις για διάφορα καθημερινά προβλήματα και εργασίες, **θα πρέπει να γίνει σημαντική προσπάθεια για τη μετάβαση από τις πολύωρες διαχειριστικές διαδικασίες που απαιτούν φυσική παρουσία για να πραγματοποιηθούν σε ψηφιακές.**

#### **8.4.2 Ενημέρωση και βελτιστοποίηση προγράμματος σπουδών**

Ένας από τους σημαντικότερους στόχους του συγκεκριμένου πιλότου ήταν η ανάλυση του προγράμματος σπουδών της σχολής Η.Μ.Μ.Υ., η σύγκριση με τις γνώσεις και δεξιότητες που είναι σε υψηλή ζήτηση από την αγορά εργασίας και η πρόταση αλλαγών στο πρόγραμμα σπουδών ώστε να καλυφθεί το κενό μεταξύ των γνώσεων και δεξιοτήτων που προσφέρονται και των γνώσεων και δεξιοτήτων που είναι σε υψηλή ζήτηση από την αγορά εργασίας για τα επαγγέλματα που μπορούν να ακολουθήσουν οι απόφοιτοι της σχολής Η.Μ.Μ.Υ. Η διαδικασία πραγματοποιήθηκε για το αντικείμενο της τεχνολογίας της πληροφορίας και κυρίως την ανάπτυξη λογισμικού που αποτελεί αντικείμενο πολλών εκ των μαθημάτων της σχολής. Τα εργαλεία που αναπτύχθηκαν κατάφεραν να εντοπίσουν τις γνώσεις και δεξιότητες που δε προσφέρονται στο πρόγραμμα σπουδών και να παρέχουν προτάσεις για τα μαθήματα στα οποία θα μπορούσαν να εισαχθούν. Οι καθηγητές που αξιολόγησαν την πλατφόρμα στα πλαίσια του πιλότου θεώρησαν σε μεγάλο βαθμό κατάλληλες και χρήσιμες τις γνώσεις και δεξιότητες που προτείνονται για την ανάπτυξη των φοιτητών τους, ενώ ήταν θετικοί στην εισαγωγή τους στα μαθήματα τους, όπως και στη μελλοντική χρήση του QualiChain για την ανανέωση των μαθημάτων τους. Ωστόσο, η εισαγωγή νέων γνώσεων και δεξιοτήτων σε ένα μάθημα και γενικότερα στο πρόγραμμα σπουδών είναι πολύπλοκη διαδικασία. Συγκεκριμένα, οι καθηγητές θα πρέπει να λάβουν υπόψη και τις υπόλοιπες γνώσεις και δεξιότητες που προσφέρονται στο μάθημα τους, το συνολικό φόρτο εργασίας του μαθήματος και τις εργασίες του, καθώς επίσης και το βαθμό στον οποίο το μάθημα θα αλλάξει με την προσθήκη μιας γνώσης ή δεξιότητας. Ο συγκεκριμένος πιλότος, έκανε εμφανές ότι η διαδικασία ενημέρωσης ενός προγράμματος σπουδών σε ένα πανεπιστήμιο δεν είναι απλή και απαιτεί σημαντική προσπάθεια από όλους τους εμπλεκόμενους φορείς, και ειδικά από τους καθηγητές και το διοικητικό προσωπικό της σχολής. Από τα συμπεράσματα του συγκεκριμένου πιλότου καταλήγουμε ότι **μια τέτοια διαδικασία θα πρέπει να ξεκινάει με την αναγνώριση γνώσεων και δεξιοτήτων που προσφέρονται αλλά θεωρούνται παρωχημένες όσον αφορά τις απαιτήσεις της αγοράς εργασίας.** Αξίζει επίσης να σημειωθεί ότι είναι σημαντική η διάκριση μεταξύ παρωχημένων γνώσεων και δεξιοτήτων και εισαγωγικών γνώσεων και δεξιοτήτων, μιας και οι φοιτητές θα πρέπει πρώτα να μάθουν βασικές γνώσεις και δεξιότητες που μπορεί να μη ζητούνται από την αγορά εργασίας και στη συνέχεια να προχωρήσουν στην εκμάθηση πιο προχωρημένων γνώσεων και δεξιοτήτων.

### 8.4.3 Προσωποποιημένες υπηρεσίες για τους φοιτητές

Η πλατφόρμα QualiChain προσέφερε μεταξύ άλλων μια σειρά από λειτουργικότητες που ήταν διαμορφωμένες με βάση τις ατομικές ανάγκες του κάθε φοιτητή της σχολής H.M.M.Y. όπως η πρόταση δεξιοτήτων, μαθημάτων και θέσεων εργασίας. Παρότι οι περισσότεροι φοιτητές εξέφρασαν θετική άποψη για τις προσωποποιημένες αυτές υπηρεσίες, υπήρξαν και κάποια αρνητικά σχόλια σχετικά με τις προσδοκίες τους από την πλατφόρμα QualiChain. Σημαντική μερίδα φοιτητών απογοητεύτηκαν από το γεγονός ότι η πλατφόρμα QualiChain δεν ήταν συνδεδεμένη με την πλατφόρμα της γραμματείας και έπρεπε να δηλώσουν τα μαθήματα που επιλέγουν και στις δύο πλατφόρμες. Επιπλέον, δήλωσαν ότι θα ήθελαν και κάποιες επιπλέον λειτουργίες όπως για παράδειγμα ένα ημερολόγιο που να δείχνει τις υποχρεώσεις τους (μαθήματα και εργασίες), σύνδεση με το forum της σχολής, και ενημερώσεις για ανακοινώσεις της σχολής. Επιπροσθέτως, κάποιοι φοιτητές δήλωσαν ότι θα ήταν χρήσιμη η διασύνδεση με άλλες πλατφόρμες όπως είναι το LinkedIn, ώστε να μεταφέρονται αυτόματα τα βιογραφικά τους από τη μια πλατφόρμα στην άλλη. Τέλος, υπήρχαν κάποιες παραινέσεις για τη βελτίωση της διεπαφής χρήστη. Οπότε, παρότι συνολικά οι φοιτητές έμειναν ευχαριστημένοι από την πλατφόρμα, **η έλλειψη διασύνδεσης με άλλες πλατφόρμες, καθώς και τα περιθώρια βελτίωσης της διεπαφής χρήστη μπορούν να συμβάλλουν στη μείωση του ενδιαφέροντος για την πλατφόρμα.** Φυσικά, η διασύνδεση με την πλατφόρμα της σχολής παρουσιάζει σημαντικές δυσκολίες. Για να αρθεί ο συγκεκριμένος περιορισμός θα πρέπει να πραγματοποιηθούν σημαντικές αλλαγές στις διαδικασίες της σχολής, και να χρησιμοποιηθεί η πλατφόρμα QualiChain ή κάποια άλλη αντίστοιχη πλατφόρμα, ως επίσημη λύση από τη σχολή.

### 8.4.4 Πολυκριτηριακή υποστήριξη αποφάσεων

Η υπηρεσία πολυκριτηριακής υποστήριξης αποφάσεων παρείχε στους χρήστες ένα εργαλείο που βοηθάει στη λήψη αποφάσεων με μεγαλύτερη σιγουριά λαμβάνοντας υπόψη τη διαθέσιμη πληροφορία με την κατάταξη των εναλλακτικών επιλογών με βάση όλα τα κριτήρια που έχουν σημασία για αυτούς. Στα εργαστήρια που πραγματοποιήθηκαν για την παρουσίαση της υπηρεσίας στα πλαίσια του συγκεκριμένου πιλότου, παρήχθησαν εξηγήσεις για τη λειτουργικότητα της συγκεκριμένης υπηρεσίας. Η επίδειξη της υπηρεσίας παράλληλα με την παροχή επεξηγήσεων χρειάστηκε αρκετό χρόνο, καθώς οι περισσότεροι χρήστες δεν ήταν εξοικειωμένοι με την πολυκριτηριακή υποστήριξη αποφάσεων. Στο τέλος, οι χρήστες αντιλήφθηκαν τόσο πώς δουλεύει η υπηρεσία, όσο και την χρησιμότητα της. Επιπλέον, παρείχαν προτάσεις για το πώς θα μπορούσε να γίνει ευκολότερη στη χρήση, ειδικά για τους νέους χρήστες. Συγκεκριμένα, πολλοί φοιτητές πιστεύουν ότι εάν η υπηρεσία αυτή συνοδευόταν από κάποια παραδείγματα για το πώς να τη χρησιμοποιήσουν θα ήταν ευκολότερο να το κάνουν, χωρίς την ανάγκη για περαιτέρω επεξηγήσεις. Επιπροσθέτως, προτάθηκε η παροχή της επιλογής για αυτόματη εισαγωγή κριτηρίων από τους χρήστες, ανάλογα με τον τύπο της ανάλυσης που θα ήθελαν να πραγματοποιήσουν. Όταν αναφέρονται κριτήρια όπως η δυσκολία ή ο φόρτος εργασίας, παρότι είναι σχετικά υποκειμενικά, η υπηρεσία θα μπορούσε να παρέχει μια εκτίμηση μέσω της γνώσης που αποκτά από προηγούμενες αναλύσεις. Τέτοιες αλλαγές θα μείωναν σημαντικά το χρόνο που χρειάζονται οι χρήστες για να υπολογίσουν τη σημασία για κάθε κριτήριο. Στην πραγματικότητα, **η αξιοποίηση παρελθοντικών**

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

**αναλύσεων για τον υπολογισμό μέσων τιμών για κάποια κριτήρια θα σήμαινε ότι όσο χρησιμοποιείται το εργαλείο τόσο οι μέσες τιμές για τα κριτήρια θα γίνονταν και πιο ρεαλιστικές.**

#### **8.4.5 Η πλατφόρμα QualiChain**

Συνολικά, μπορούμε να συμπεράνουμε ότι η υλοποίηση της πλατφόρμας QualiChain υπό το πρίσμα του συγκεκριμένου πιλότου ήταν επιτυχής, όπως φαίνεται και από τις θετικές αντιδράσεις που εισπράξαμε από τους χρήστες μέσα από τα ερωτηματολόγια που συμπλήρωσαν. Το QualiChain παρουσίασε στους ενδιαφερόμενους της ανώτατης εκπαίδευσης έναν συνδυασμό από τεχνολογίες που δεν έχουν παρουσιαστεί μαζί στο αντικείμενο της εκπαίδευσης. Σε αυτό το πλαίσιο, μπορούμε να πούμε ότι η αξιοποίηση καινοτόμων τεχνολογιών μπορεί να οδηγήσει στην ανάπτυξη νέων πλατφορμών και εργαλείων οι οποίες καλύπτουν πραγματικές ανάγκες και παρέχουν πρωτότυπες λύσεις. Επιπλέον, ερευνητικά έργα όπως το QualiChain είναι κρίσιμα ώστε να φέρουν τους ανθρώπους σε επαφή με τις διάφορες καινοτόμες τεχνολογίες και να αυξήσουν την εμπιστοσύνη στην τεχνολογία και τις αποφάσεις τους. Για παράδειγμα, παρότι η τεχνολογία blockchain θεωρείται συνώνυμη με τα κρυπτονομίσματα, το QualiChain απέδειξε ότι μπορεί να εφαρμοστεί και σε άλλους τομείς όπως η εκπαίδευση ως λύση όταν υπάρχει ανάγκη για προστασία προσωπικών ή ευαίσθητων δεδομένων. Ειδικά στην εποχή της πανδημίας και μετά από αυτήν, **λειτουργικότητες όπως η αυτόματη επαλήθευση προσόντων, οι υπηρεσίες που βασίζονται στη συγκέντρωση και ανάλυση δεδομένων για την λήψη σημαντικών αποφάσεων και η υποστήριξη αποφάσεων μπορούν να βρεθούν στο προσκήνιο των μελλοντικών λύσεων.**

### **8.5 Συμπεράσματα**

Η συγκεκριμένη ενότητα παρουσιάζει τα βασικά συμπεράσματα που προέκυψαν από τη λειτουργία και εκτέλεση της περίπτωσης χρήσης που παρουσιάστηκε. Συνολικά, όπως παρατηρούμε από τις προηγούμενες ενότητες, ο πιλότος πέτυχε το σκοπό του όσον αφορά στα τεχνικά αποτελέσματα της πλατφόρμας QualiChain σε κρίσιμες ομάδες ενδιαφερόμενων ενός πανεπιστημίου. Επιπροσθέτως, αποδείχτηκε ότι μια πλατφόρμα σαν το QualiChain, με το συνδυασμό τεχνικών και τεχνολογιών που χρησιμοποιεί μπορεί να ενσωματωθεί στην καθημερινή λειτουργία ενός πανεπιστημίου, καθώς προσφέρει αρκετές λειτουργίες και καινοτομίες που βοηθούν διαφορετικές ομάδες χρηστών να ολοκληρώσουν συντομότερα και αποτελεσματικότερα τις εργασίες τους.

Η πρώτη ερευνητική ερώτηση αφορά τη χρήση υπηρεσιών αναλυτικής δεδομένων και υποστήριξης αποφάσεων για την επιλογή μαθημάτων και δεξιοτήτων για τη βελτίωση του ακαδημαϊκού/επαγγελματικού προφίλ του χρήστη, όσον αφορά τις ανάγκες της αγοράς εργασίας. Από προηγούμενη υποενότητα συμπεραίνουμε ότι οι φοιτητές που συμμετείχαν στο πιλοτικό πρόγραμμα ήταν συνολικά θετικοί και ικανοποιημένοι από τις παρεχόμενες από την πλατφόρμα QualiChain λειτουργίες. Ωστόσο, υπήρχε και χαμηλό ποσοστό απογοήτευσης από κάποιους φοιτητές, το οποίο έγκειται στο γεγονός ότι κάποιοι φοιτητές περίμεναν να είναι το σύστημα ενοποιημένο με τη βάση δεδομένων της σχολής H.M.M.Y., ώστε τα μαθήματα που επιλέγουν στο QualiChain να εμφανίζονται αυτόματα και στον διαδικτυακό τόπο της σχολής H.M.M.Y. και στα φοιτητικά τους στοιχεία, χωρίς να χρειάζεται να τα ξανά-επιλέξουν, ωστόσο κάτι τέτοιο δεν ήταν εφικτό στα πλαίσια ενός πιλοτικού

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

προγράμματος. Παρ' όλα αυτά, οι περισσότεροι φοιτητές έμειναν ικανοποιημένοι τόσο από τις προτάσεις που λάμβαναν για τη βελτίωση του προφίλ τους, όσο και για το γεγονός ότι μπορούσαν να επιλέξουν μαθήματα της σχολής τους μέσα από την πλατφόρμα QualiChain. Είναι προφανές ότι τέτοιες προσωποποιημένες υπηρεσίες θα εκλαμβάνονταν θετικά από τους φοιτητές σε περίπτωση που περισσότερα πανεπιστήμια αποφασίσουν να χρησιμοποιήσουν την πλατφόρμα QualiChain ή κάποια παρόμοια πλατφόρμα.

Η δεύτερη ερευνητική ερώτηση αφορά στη χρήση υπηρεσιών αναλυτικής δεδομένων και υποστήριξης αποφάσεων για την ανανέωση του προγράμματος σπουδών μιας σχολής ώστε να ανταποκρίνεται καλύτερα στις ανάγκες της αγοράς εργασίας. Όπως έχει ήδη αναφερθεί σε προηγούμενες ενότητες, οι καθηγητές που χρησιμοποίησαν το σύστημα έμειναν πολύ ικανοποιημένοι από τις προτάσεις που τους παρείχε η πλατφόρμα για τα μαθήματα τους όσο και για το πρόγραμμα σπουδών γενικότερα. Επιπροσθέτως, οι καθηγητές έμειναν πολύ ικανοποιημένοι και από τη διαδικασία ανανέωσης των μαθημάτων τους η οποία ήταν καλά σχεδιασμένη και εύκολη στη χρήση. Επίσης, αξίζει να σημειωθεί ότι η επέκταση της πλατφόρμας και σε άλλους τομείς πέρα από την τεχνολογία πληροφορίας και επικοινωνιών, αναμένεται να φέρει σημαντική προστιθέμενη αξία και για άλλους ακαδημαϊκούς οργανισμούς, επαγγέλματα και επαγγελματικούς τομείς.

Στο σημείο αυτό, πρέπει να αναφερθεί ότι η συμμετοχή χρηστών της σχολής Η.Μ.Μ.Υ. στην εκτέλεση του πιλότου ήταν πολύ ικανοποιητική με 90 φοιτητές και 11 καθηγητές. Διοργανώθηκε μια σειρά από εργαστήρια (workshops) ώστε το έργο QualiChain να γίνει γνωστό για τις λειτουργικότητες που παρέχει. Με βάση τις απαντήσεις στα ερωτηματολόγια, η πλειοψηφία των χρηστών έμειναν ικανοποιημένοι με τις λειτουργίες που τους προσφέρθηκαν. Το ξέσπασμα της πανδημίας Covid-19 δυσχέρανε σημαντικά την προσέγγιση φοιτητών και καθηγητών, ωστόσο, ο πιλότος κατάφερε να προσελκύσει σημαντική μερίδα χρηστών και να αντλήσει σημαντικά συμπεράσματα για την πλατφόρμα, όσο και σχόλια και προτάσεις για τη βελτίωση της.

Τέλος, αξίζει να σημειωθεί ότι οι αλλαγές στο εργασιακό περιβάλλον που προέκυψαν λόγω της πανδημίας επιδεικνύουν ακόμα περισσότερο τη δυναμική της πλατφόρμας QualiChain. Συγκεκριμένα, η απομακρυσμένη εργασία δημιουργεί μια παγκόσμια αγορά εργασίας, η οποία είναι πιο ανταγωνιστική από ποτέ. Για το λόγο αυτό, είναι σημαντικότερο από ποτέ οι φοιτητές να επιλέγουν μαθήματα που θα τους βοηθήσουν να ανταποκριθούν καλύτερα στον ανταγωνισμό. Επίσης, τα διαδικτυακά μαθήματα κερδίζουν όλο και μεγαλύτερη δημοφιλία, κατά συνέπεια τα πανεπιστήμια θα πρέπει και αυτά με τη σειρά τους να παρέχουν τις πιο προηγμένες γνώσεις και δεξιότητες στους φοιτητές μέσω των μαθημάτων που διδάσκουν. Για όλους αυτούς τους λόγους θεωρούμε ότι το έργο QualiChain, μέσω του συγκεκριμένου πιλότου πέτυχε την ανάδειξη των αναγκών της ανώτατης εκπαίδευσης και παρείχε μια καλή λύση για την κάλυψη των αναγκών αυτών.

## 8.6 Αναφορές

[1] “Matomo Analytics - The Google Analytics alternative that protects your data.”

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

<https://matomo.org/> (accessed May 03, 2023).

- [2] “Analytics.” <https://analytics.google.com/analytics/web/provision/#/provision> (accessed May 03, 2023).
- [3] “Feed | LinkedIn.” <https://www.linkedin.com/feed/> (accessed May 03, 2023).
- [4] “About SEAL | Seal.” <https://project-seal.eu/about> (accessed May 01, 2023).
- [5] “κανονισμός eIDAS | Shaping Europe’s digital future.” <https://digital-strategy.ec.europa.eu/el/policies/eidas-regulation> (accessed May 01, 2023).
- [6] “eduGAIN – enabling worldwide access.” <https://edugain.org/> (accessed May 01, 2023).



## ΚΕΦΑΛΑΙΟ 9: ΔΙΚΑΙΟΣΥΝΗ ΜΕΤΑΞΥ ΤΩΝ ΠΑΡΟΧΩΝ ΣΤΑ ΣΥΣΤΗΜΑΤΑ ΣΥΣΤΑΣΕΩΝ

### 9.1 Εισαγωγή

Στο παρελθόν, οι προτάσεις αντικειμένων στα συστήματα συστάσεων γινόταν στις περισσότερες περιπτώσεις με βάση τις υψηλότερες προβλεπόμενες βαθμολογίες για κάθε χρήστη. Φυσικά, πρόκειται για σημαντικό κριτήριο για την επιτυχία ενός συστήματος συστάσεων, καθώς προτείνει αντικείμενα τα οποία οι χρήστες φαίνεται να βρίσκουν ενδιαφέροντα. Ωστόσο, το συγκεκριμένο κριτήριο μπορεί να μην είναι επαρκές όταν είναι το μοναδικό που λαμβάνεται υπόψη. Αυτό συμβαίνει καθώς ένα σύστημα που λαμβάνει υπόψη μόνο αυτό το κριτήριο αποτυγχάνει να βοηθήσει το χρήστη να εξερευνήσει νέο περιεχόμενο, ενώ οι προτάσεις για κάθε χρήστη είναι αρκετά παρόμοιες μεταξύ τους. Έτσι, σχετικά άγνωστα αλλά καινοτόμα αντικείμενα μπορεί να μην ανακαλυφθούν ποτέ από πολλούς χρήστες. Επιπροσθέτως, οι περισσότεροι χρήστες λαμβάνουν προτάσεις για συγκεκριμένο τύπο αντικειμένων, ενώ το πρόβλημα αυτό μεγεθύνεται για συστήματα συστάσεων που βασίζονται στο περιεχόμενο (content-based recommender systems).

Οι συγκεκριμένες αδυναμίες έχουν εντοπιστεί από τους επιστήμονες εδώ και χρόνια και έχει προταθεί μια σειρά από εναλλακτικές μετρικές για την αξιολόγηση των συστημάτων συστάσεων, μεταξύ των οποίων η κάλυψη (coverage), η διαφοροποίηση (diversity), η δικαιοσύνη (fairness), η ευχάριστη έκπληξη (serendipity) και η καινοτομία (novelty). Όσον αφορά την κάλυψη, σύμφωνα με την εργασία [1] ένα αντικείμενο θεωρείται καλυμμένο αν προτείνεται τουλάχιστον σε ένα συγκεκριμένο αριθμό χρηστών. Από την άλλη πλευρά, η διαφοροποίηση ορίζεται ως η μέση ανομοιότητα μεταξύ όλων των ζευγών χρηστών ή αντικειμένων στο σύνολο αποτελεσμάτων (διαφοροποίηση αντικειμένου και χρήστη αντίστοιχα) [2]. Η ευχάριστη έκπληξη, είναι η θετική εμπειρία που αποκομίζει ο χρήστης από μη προσδοκώμενες αλλά καλές προτάσεις [3]. Η δικαιοσύνη αφορά κυρίως ομάδες ενδιαφερομένων και η διαισθητική ερμηνεία της είναι ότι ένα σύνολο αντικειμένων πρέπει να προτείνεται σε ομάδες χρηστών με τέτοιο τρόπο ώστε όλοι οι χρήστες της ομάδας να ικανοποιούνται με δίκαιο τρόπο [1]. Τέλος, η καινοτομία σε ένα σύστημα συστάσεων ορίζεται ως η αναλογία μεταξύ άγνωστων και γνωστών αντικειμένων που προτείνονται στο χρήστη [4].

Αξίζει να σημειωθεί ότι, τα τελευταία χρόνια, δίνεται μεγάλη σημασία, πέρα από τους χρήστες που λαμβάνουν προτάσεις αντικειμένων, και σε άλλους ενδιαφερόμενους που εμπλέκονται στη λειτουργία ενός συστήματος συστάσεων. Συγκεκριμένα, οι βασικοί ενδιαφερόμενοι ενός Συστήματος Συστάσεων είναι οι πάροχοι αντικειμένων, οι χρήστες, και οι υπεύθυνοι λειτουργίας του συστήματος (recommender system operators). Αυτή η ολιστική οπτική στα συστήματα συστάσεων κατά την οποία όλα τα ενδιαφερόμενα μέρη λαμβάνονται υπόψη είναι ευρέως γνωστή ως Πολυμετοχικά Συστήματα Συστάσεων (Multi-Stakeholder Recommender Systems - MSRS) [5].

Το συγκεκριμένο κεφάλαιο ασχολείται με τη δικαιοσύνη μεταξύ των παρόχων αντικειμένων, και συγκεκριμένα με τη δίκαιη κατανομή των αντικειμένων σε χρήστες, όσον αφορά την κάλυψη και τη διαφοροποίηση σε χρήστες των αντικειμένων για κάθε πάροχο αντικειμένων. Η ιδέα της δικαιοσύνης στην προκειμένη περίπτωση υπονοεί ότι τα αντικείμενα κάθε παρόχου προτείνονται σε παρόμοιο

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

αριθμό χρηστών όπως και τα αντικείμενα άλλων παρόχων, ενώ παράλληλα οι χρήστες στους οποίους προτείνονται τα αντικείμενα ενός παρόχου είναι το ίδιο διαφορετικοί μεταξύ τους όσο και οι χρήστες στους οποίους προτείνονται αντικείμενα από άλλους παρόχους αντικειμένων. Συγκεκριμένα, η δικαιοσύνη για τους παρόχους (provider fairness – P-fairness) [6] όσον αφορά την κάλυψη και τη διαφοροποίηση εισάγεται στο πρόβλημα προς επίλυση (πρόβλημα βελτιστοποίησης) ως περιορισμοί που πρέπει να ικανοποιούνται για κάθε πάροχο αντικειμένων. Έτσι, η μέση κάλυψη σε χρήστες και η διαφοροποίηση χρηστών ανά αντικείμενο για κάθε πάροχο πρέπει να ξεπερνάνε τουλάχιστον ένα συγκεκριμένο στόχο, ενώ οι παρεχόμενες προτάσεις αντικειμένων στους χρήστες θα πρέπει να είναι κατά το δυνατόν πιο ικανοποιητικές.

Σε αυτό το πλαίσιο, το σύστημα συστάσεων που παρουσιάζεται στο παρόν κεφάλαιο στοχεύει στη μεγιστοποίηση της συνολικής προβλεπόμενης βαθμολογίας των λιστών συστάσεων, ενώ παράλληλα ικανοποιούνται οι περιορισμοί για δίκαιη κάλυψη χρηστών και διαφοροποίηση χρηστών ανά αντικείμενο για κάθε πάροχο αντικειμένων. Συγκεκριμένα, το πρόβλημα μοντελοποιείται ως ένα πρόβλημα μαθηματικής βελτιστοποίησης όπως προτείνεται και στη βιβλιογραφία [7]. Επιπλέον, ακολουθείται η προτεινόμενη μεθοδολογία για την επίλυση του. Εφ' όσον το πρόβλημα δεν είναι γραμμικό και δεν είναι υπολογιστικά δυνατό να επιλυθεί αναλυτικά σε αποδεκτό χρόνο (ανήκει στην κατηγορία προβλημάτων NP-complete), και η μεθοδολογία που προτείνεται για την ικανοποίηση των περιορισμών για τη διαφοροποίηση δεν περιγράφεται επαρκώς, στο παρόν κεφάλαιο προτείνεται ένας ευριστικός αλγόριθμος για την επίλυση του προβλήματος. Ο συγκεκριμένος αλγόριθμος βρίσκει λύσεις που είναι πολύ κοντά στη θεωρητικά βέλτιστη λύση ενώ ικανοποιούν τους περιορισμούς για διαφοροποίηση και κάλυψη. Επιπλέον, τα αποτελέσματα του αλγορίθμου παρουσιάζονται αναλυτικά, ενώ συζητούνται επιγραμματικά και οι περιορισμοί στη λύση που αναπτύχθηκε αλλά και στη μοντελοποίηση του προβλήματος όπως προτείνεται στη βιβλιογραφία. Σε αυτό το πλαίσιο, αναγνωρίζονται κάποιες αδυναμίες στην προτεινόμενη μοντελοποίηση και προτείνονται δυο νέοι μαθηματικοί ορισμοί για τη διαφοροποίηση, οι οποίοι αντιμετωπίζουν τις αδυναμίες της προτεινόμενης μοντελοποίησης και προτείνεται μια μεθοδολογία για την επίλυση του νέου προβλήματος. Τέλος παρουσιάζονται κάποιοι περιορισμοί της νέας προτεινόμενης μεθοδολογίας παράλληλα με μελλοντικές επεκτάσεις για αυτήν.

### **Ερευνητικά ερωτήματα**

Το συγκεκριμένο κεφάλαιο στοχεύει στην απάντηση των ακόλουθων ερευνητικών ερωτημάτων:

1. Πώς μπορεί να επιτευχθεί δικαιοσύνη μεταξύ των παρόχων αντικειμένων σε ένα σύστημα συστάσεων όσον αφορά την κάλυψη των αντικειμένων και τη διαφοροποίηση των χρηστών ανά αντικείμενο;
2. Είναι ικανοποιητικά τα αποτελέσματα των προτεινόμενων μεθοδολογιών για το συγκεκριμένο πρόβλημα;
3. Ποιοι είναι οι περιορισμοί των προτεινόμενων λύσεων;
4. Πώς οι συγκεκριμένοι περιορισμοί μπορούν να αρθούν;

### **Συνεισφορά**

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Ο στόχος του παρόντος κεφαλαίου είναι η ανάπτυξη ενός συστήματος συστάσεων που θα παρέχει δίκαιες συστάσεις μεταξύ των παρόχων αντικειμένων όσον αφορά την κάλυψη και τη διαφοροποίηση των χρηστών στους οποίους προτείνονται τα αντικείμενα των παρόχων. Έτσι, η συνεισφορά του είναι πολλαπλή και περιλαμβάνει:

- Ανασκόπηση της πρόσφατης βιβλιογραφίας όσον αφορά τη δικαιοσύνη των παρόχων, την κάλυψη και τη διαφοροποίηση.
- Ορισμός και προτυποποίηση του προβλήματος.
- Εφαρμογή και επέκταση της προτεινόμενης από τη βιβλιογραφία μεθοδολογίας.
- Εφ' όσον η προτεινόμενη μεθοδολογία για την επίλυση των περιορισμών διαφοροποίησης είναι ασαφής, παρουσιάζεται ένας ευριστικός αλγόριθμος για την επίλυση των περιορισμών αυτών, ο οποίος αναπτύχθηκε στα πλαίσια της παρούσας διδακτορικής διατριβής.
- Αξιολόγηση των αποτελεσμάτων και συζήτηση για την προτεινόμενη μεθοδολογία.
- Πρόταση νέων ορισμών για την ποσοτικοποίηση της διαφοροποίησης, μιας και ο προτεινόμενος ορισμός παρουσιάζει σημαντικούς περιορισμούς.
- Πρόταση μια νέας υψηλού επιπέδου (high level) ευριστικής προσέγγισης για την επίλυση των περιορισμών για τη διαφοροποίηση με βάση τους νέους ορισμούς, καθώς το πρόβλημα παραμένει NP-Hard.

## 9.2 Υπόβαθρο και σχετικές εργασίες

Υπάρχει πληθώρα από επιστημονικές δημοσιεύσεις που αναγνωρίζουν την ανάγκη να προχωρήσουμε πέρα από την ακρίβεια των συστάσεων στα συστήματα συστάσεων. Ωστόσο, δεν υπάρχει κάποια κοινά αποδεκτή προσέγγιση για την ποσοτικοποίηση των εναλλακτικών μετρικών αξιολόγησης. Σε αυτό το πλαίσιο, στην ενότητα αυτή παρουσιάζεται η πρόσφατη βιβλιογραφία στα πολυμετοχικά συστήματα συστάσεων, και τις μετρικές αξιολόγησης για τη διαφοροποίηση, την κάλυψη και τη δικαιοσύνη, όσον αφορά τους ορισμούς και τη χρήση τους. Ακολουθως, παρουσιάζεται μια εργασία για παροχή συστάσεων σε πολυμετοχικά συστήματα συστάσεων (Multi-Stakeholder Recommender Systems - MSRS), με περιορισμούς για δικαιοσύνη μεταξύ των παρόχων σχετικά με την κάλυψη και τη διαφοροποίηση σε χρήστες που απολαμβάνουν τα αντικείμενα που παρέχουν από ένα σύστημα συστάσεων, καθώς επίσης και την προτεινόμενη λύση.

Τέλος, συζητούνται οι λόγοι που η προτεινόμενη προσέγγιση του προβλήματος πηγαίνει πέρα από την τρέχουσα τεχνολογική στάθμιση (state of the art).

### 9.2.1 Πολυμετοχικά Συστήματα Συστάσεων (MSRS)

Η πολυμετοχική οπτική (MSRS paradigm) στα συστήματα συστάσεων έχει γίνει ιδιαίτερα δημοφιλής τα τελευταία χρόνια. Το συγκεκριμένο υπόδειγμα, σε αντίθεση με το παραδοσιακό που τοποθετεί τον χρήστη (που λαμβάνει προτάσεις) στο επίκεντρο της προσοχής, λαμβάνει υπόψη πέρα από το χρήστη και τους άλλους μετόχους ενός συστήματος συστάσεων, όπως τους παρόχους αντικειμένων, το λειτουργό του συστήματος συστάσεων (recommender system operator) και την κοινωνία γενικότερα [5], [8]. Συγκεκριμένα, οι εργασίες [5], [8] αναγνωρίζουν την ανάγκη για εξέταση των συστημάτων

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

συστάσεων υπό το πρίσμα των διαφόρων μετόχων και προτείνουν το συγκεκριμένο υπόδειγμα. Από την άλλη πλευρά οι Abdollahrougi et al. [9], διερευνούν την δικαιοσύνη από πολλές πλευρές σε πολυμετοχικά συστήματα συστάσεων, όπως επίσης και πώς διαφορετικές οπτικές της δικαιοσύνης μπορούν να εισαχθούν σε τέτοια συστήματα. Επιπροσθέτως οι Milano et al. [10] ανέλυσαν τις ηθικές εκφάνσεις στα πολυμετοχικά συστήματα συστάσεων. Ακόμη, οι εργασίες [11], [12] επιχειρούν μια βιβλιογραφική ανασκόπηση στις υπάρχουσες περιπτώσεις χρήσης, μεθόδους και προκλήσεις και προτείνουν νέες κατευθύνσεις έρευνας για τα συστήματα συστάσεων. Τέλος, οι εργασίες [13], [14] παρουσιάζουν εφαρμογές των πολυμετοχικών συστημάτων συστάσεων. Η πρώτη εισάγει περιορισμούς για τους παρόχους στο πρόβλημα πολυμετοχικών συστάσεων, και το μοντελοποιεί ως ένα πρόβλημα βελτιστοποίησης ακέрайου προγραμματισμού, το οποίο επιλύεται με προσεγγίσεις και επιτυγχάνει ικανοποιητικά αποτελέσματα σε πραγματικές περιπτώσεις χρήσης. Στη δεύτερη εργασία, ένα μοντέλο δυαδικού ακέрайου προγραμματισμού πολλών στόχων προτείνεται για την πρόταση χορηγούμενων αντικειμένων. Ακολούθως, η ίδια εργασία προτείνει έναν αλγόριθμο για την επίλυση του προβλήματος με υπολογιστικά αποδοτικό τρόπο. Η προτεινόμενη προσέγγιση στη συγκεκριμένη εργασία εφαρμόστηκε σε μια πραγματική περίπτωση χρήσης με ικανοποιητικά αποτελέσματα και είναι εύκολα εφαρμόσιμη σε ήδη υπάρχοντα συστήματα συστάσεων, μιας και εφαρμόζεται μετά την αρχική επεξεργασία (postprocessing).

Φυσικά, στο παρόν κεφάλαιο δεν εξετάζονται γενικά τα πολυμετοχικά συστήματα συστάσεων. Αντίθετα, επικεντρώνεται στις μετρικές για κάλυψη και διαφοροποίηση μεταξύ των παρόχων αντικειμένων και αντιμετωπίζει το πρόβλημα της δικαιοσύνης μεταξύ αυτών όσον αφορά τις προαναφερθείσες μετρικές. Συνεπώς, στις επόμενες ενότητες παρουσιάζονται οι διάφοροι ορισμοί για την κάλυψη, τη διαφοροποίηση και τη δικαιοσύνη στα συστήματα συστάσεων.

### 9.2.2 Κάλυψη

Στην εργασία [15], ως κάλυψη αντικειμένων, ορίζεται η αναλογία από αντικείμενα για τα οποία το σύστημα μπορεί να δημιουργήσει προτάσεις (κάλυψη πρόβλεψης – prediction coverage), ή η αναλογία διαθέσιμων αντικειμένων τα οποία προτείνονται στους χρήστες (κάλυψη καταλόγου – catalog coverage).

Αξίζει να σημειωθεί ότι οι προηγούμενοι ορισμοί αναφέρονται σε κάλυψη αντικειμένων. Ωστόσο, υπάρχει και ένας λιγότερο δημοφιλής ορισμός για την κάλυψη που αφορά τους χρήστες και ορίζεται ως η αναλογία χρηστών στους οποίους έχουν προταθεί συγκεκριμένα αντικείμενα [16]. Η κάλυψη χρηστών παρουσιάζει σημαντικό ενδιαφέρον για λίστες αντικειμένων, με την έννοια ότι είναι σημαντικό για έναν πάροχο αντικειμένων να γνωρίζει σε τι ποσοστό των χρηστών προτείνονται τα αντικείμενα του. Στο συγκεκριμένο κεφάλαιο, ένας εκ των σημαντικότερων στόχων του συστήματος που αναπτύχθηκε είναι να ικανοποιήσει την ανάγκη κάλυψης χρηστών για τους διάφορους παρόχους.

Παρότι η κάλυψη αναφέρεται σε μια σειρά από ερευνητικές εργασίες, στις περισσότερες περιπτώσεις υπολογίζεται εκ των υστέρων μετά τον υπολογισμό της λίστας προτάσεων και χρησιμοποιείται μόνο για αξιολόγηση του συστήματος συστάσεων, χωρίς να λαμβάνεται υπόψη κατά τη διαδικασία υπολογισμού των συστάσεων [15], [16], [17], [18]. Η εργασία [1] είναι από τις πρώτες που εισάγει περιορισμούς για κάλυψη χρηστών κατά τον ορισμό του προβλήματος συστάσεων. Συγκεκριμένα, το πρόβλημα

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

μοντελοποιείται ως πρόγραμμα μαθηματικής βελτιστοποίησης με περιορισμούς για την κάλυψη. Η συγκεκριμένη προσέγγιση είναι αρκετά κοντά στην προσέγγιση που παρουσιάζεται στο παρόν κεφάλαιο.

### 9.2.3 Διαφοροποίηση

Όσον αφορά τη διαφοροποίηση, μια σειρά από ορισμούς έχει προταθεί στη βιβλιογραφία [16]. Συγκεκριμένα, έχει οριστεί ως η μέση απόσταση ανά ζεύγη μεταξύ των προτεινόμενων αντικειμένων, ή η συνολική ανά ζεύγη απόσταση μεταξύ των προτεινόμενων αντικειμένων [19]. Ο πρώτος ορισμός είναι ο πιο διαδεδομένος στη βιβλιογραφία. Ένα ενδιαφέρον θέμα σε ό,τι αφορά τη διαφοροποίηση είναι πώς υπολογίζεται η απόσταση (ή ανομοιότητα). Για παράδειγμα, όταν τα αντικείμενα είναι μοντελοποιημένα με περιγραφητές κειμένου (content descriptors), η απόσταση υπολογίζεται με τη χρήση ταξινομιών [19], ενώ όταν μοντελοποιούνται ως διανύσματα όρων η απόσταση υπολογίζεται ως το συμπλήρωμα της ομοιότητας Jaccard [20] ή της συνημιτονικής ομοιότητας [21]. Από την άλλη πλευρά όταν τα αντικείμενα αναπαρίστανται από διανύσματα βαθμολογιών, οι καταλληλότερες μετρικές είναι η απόσταση Hamming, η ομοιότητα Pearson [22] ή το συμπλήρωμα της συνημιτονικής ομοιότητας [21]. Στην προκειμένη περίπτωση έχει χρησιμοποιηθεί το συμπλήρωμα της ομοιότητας Pearson και τα αντικείμενα αναπαρίστανται ως διανύσματα βαθμολογιών. Φυσικά, υπάρχουν περισσότερες προσεγγίσεις για τον υπολογισμό της απόστασης μεταξύ αντικειμένων [20], [23], αλλά δε θα παρουσιαστούν στο παρόν κεφάλαιο.

Όμοια με τη διαφοροποίηση για τα αντικείμενα μπορεί να οριστεί και η διαφοροποίηση για τους χρήστες. Η ιδέα είναι ότι ένα αντικείμενο θα πρέπει να προτείνεται σε κατά το δυνατόν πιο διαφορετικούς μεταξύ τους χρήστες. Σε αυτό το πλαίσιο, η διαφοροποίηση για τους χρήστες ορίζεται ως η μέση ανά ζεύγη απόσταση μεταξύ των χρηστών στους οποίους προτείνεται ένα αντικείμενο. Ο συγκεκριμένος ορισμός μπορεί να μην είναι σημαντικός για συγκεκριμένα αντικείμενα, αλλά είναι πολύ σημαντικός για λίστες αντικειμένων και κατά συνέπεια για παρόχους αντικειμένων που στοχεύουν στη μεγιστοποίηση της διαφοροποίησης για τους χρήστες στους οποίους προτείνονται τα αντικείμενα τους. Ο τελευταίος ορισμός αποτελεί βασικό ζητούμενο του παρόντος κεφαλαίου.

### 9.2.4 Δικαιοσύνη

Σχετικά με τη δικαιοσύνη στα συστήματα συστάσεων, στις περισσότερες περιπτώσεις στη βιβλιογραφία αναφέρεται στην αντιμετώπιση ή μείωση της προκατάληψης από ένα μοντέλο μηχανικής μάθησης ή από ένα σύστημα συστάσεων [6]. Ωστόσο, τα συστήματα συστάσεων παρέχουν προσωποποιημένες προτάσεις, και κατά συνέπεια, είναι πολύ δύσκολο να αντιμετωπιστεί η προκατάληψη. Επιπροσθέτως, συνήθως τα συστήματα συστάσεων χρησιμοποιούνται (με διαφορετικό τρόπο) από διαφορετικούς ενδιαφερόμενους και ως εκ τούτου εγείρονται θέματα δικαιοσύνης για διαφορετικές ομάδες χρηστών. Έτσι ένα σύστημα συστάσεων δεν αρκεί μόνο να διασφαλίσει δικαιοσύνη ανάμεσα στους πελάτες αλλά είναι εξίσου σημαντικό να το κάνει και για τους παρόχους αντικειμένων. Για το λόγο αυτό η εργασία [6] διαχωρίζει τη δικαιοσύνη σε δικαιοσύνη πελάτη (customer fairness – C-fairness) και δικαιοσύνη παρόχου (provider fairness – P-fairness). Στο παρόν κεφάλαιο ασχολούμαστε με τη δικαιοσύνη παρόχου και συγκεκριμένα ο στόχος είναι τα αντικείμενα κάθε παρόχου να προτείνονται σε παρόμοιο αριθμό χρηστών όπως τα αντικείμενα άλλων χρηστών, και οι χρήστες στους οποίους προτείνονται τα αντικείμενα

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

κάθε παρόχου προτείνονται σε ομάδες χρηστών που είναι όσο διαφορετικοί μεταξύ τους είναι και οι χρήστες στους οποίους προτείνονται τα αντικείμενα των άλλων παρόχων.

Υπάρχει πληθώρα από πρακτικές εφαρμογές για την παροχή δίκαιων συστάσεων ως προς τους παρόχους. Συγκεκριμένα, η εργασία [24] αξιολογεί την δικαιοσύνη παρόχου όσον αφορά τις ανισότητες σε σχετικότητα, ορατότητα και έκθεση για τις ομάδες μειονοτήτων και προτείνει μια μεθοδολογία που συνδυάζει υπερδειγματοληψία παρατηρήσεων (observation upsampling) και τακτοποίηση απωλειών (loss regularisation) για το βαθμό σχετικότητας μεταξύ χρηστών και αντικειμένων, με ικανοποιητικά πειραματικά αποτελέσματα.

Σε σύγκριση με το παρόν κεφάλαιο, η προηγούμενη ερευνητική εργασία στοχεύει στη δικαιοσύνη για τις μειονότητες, ενώ το παρόν κεφάλαιο αντιμετωπίζει τους παρόχους ως ίσους και παρέχει στρατηγικές συστάσεων που επιτυγχάνουν τουλάχιστον μια ελάχιστη μέση κάλυψη και διαφοροποίηση ανά αντικείμενο για κάθε πάροχο, αν φυσικά υπάρχει μια τέτοια στρατηγική. Επιπλέον, η προηγούμενη έρευνα επιτρέπει τη σύνδεση ενός αντικειμένου με παραπάνω από έναν πάροχο, πράγμα που δεν ισχύει στο παρόν κεφάλαιο, καθώς θα έκανε πολύ πιο περίπλοκο το πρόβλημα, ενώ δεν ήταν και στις απαιτήσεις του προβλήματος.

Επιπλέον, στην εργασία [25], αναγνωρίζεται ότι οι προσωπικές προτιμήσεις κάθε χρήστη μπορεί να περιορίσουν την ικανότητα του συστήματος συστάσεων να παρέχει δίκαιες συστάσεις. Στην έρευνα αυτή παρουσιάζεται μια μέθοδος αναβαθμολόγησης (re-ranking) για προτάσεις που είναι ενήμερες για την δικαιοσύνη (fairness-aware recommendations) και μαθαίνουν τις προτιμήσεις του χρήστη αντί για ευαίσθητα χαρακτηριστικά όπως η φυλή. Η συγκεκριμένη προσέγγιση πέτυχε καλύτερα αποτελέσματα από άλλες αντίστοιχες μεθοδολογίες που προτείνονται στη βιβλιογραφία. Σε αντίθεση με το παρόν κεφάλαιο, η συγκεκριμένα μεθοδολογία βασίζεται στην αναβαθμολόγηση, ενώ ασχολείται επίσης με την διαφοροποίηση των αντικειμένων για κάθε χρήστη, ενώ το παρόν κεφάλαιο ασχολείται με τη διαφοροποίηση χρηστών για κάθε αντικείμενο ενός παρόχου αντικειμένων.

Επιπλέον, στην έρευνα [26] αναγνωρίζεται η ανάγκη για δικαιοσύνη παρόχου όσον αφορά την γεωγραφική κάλυψη, καθώς υπάρχει σημαντική ανισορροπία στα εκπαιδευτικά συστήματα συστάσεων, που σχετίζεται με τη γεωγραφική θέση των παρόχων. Η ανάλυση βασίζεται σε δεδομένα του πραγματικού κόσμου, και τα αποτελέσματα δείχνουν ότι υπάρχει σημαντική γεωγραφική ανισορροπία προς όφελος συστημάτων συστάσεων από τις Η.Π.Α. όσων αφορά ανοιχτά μαθήματα και αλληλεπιδράσεις. Μιας και τα παραδοσιακά συστήματα συστάσεων τείνουν να ενισχύουν τις χώρες που εκπροσωπούνται περισσότερο (ή τους παρόχους γενικότερα), η συγκεκριμένη εργασία παρουσιάζει μια μεθοδολογία που ρυθμίζει (regulates) το μερίδιο των συστάσεων για κάθε χώρα και τη θέση τους στη λίστα συστάσεων. Ο ορισμός της δικαιοσύνης που χρησιμοποιεί η συγκεκριμένη έρευνα είναι ο κοντινότερος με τον ορισμό που χρησιμοποιείται στο παρόν κεφάλαιο, καθώς αναγνωρίζει την έλλειψη ισότητας στις συστάσεις μεταξύ διαφορετικών χωρών και την έκθεση κάθε χώρας, ενώ το παρόν κεφάλαιο στοχεύει στην ρύθμιση της ανισορροπίας (έλλειψη ισότητας) σε κάλυψη και διαφοροποίηση για τα αντικείμενα κάθε παρόχου αντικειμένων.

Άλλη μια αξιοσημείωτη εργασία σχετικά με τη δικαιοσύνη στα συστήματα συστάσεων είναι η [27]. Η συγκεκριμένη εργασία αξιολογεί αλγοριθμικά τη δικαιοσύνη σε ένα σύστημα συστάσεων του

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

πραγματικού κόσμου και καταδεικνύει ότι η μέτρηση της δικαιοσύνης μέσω συγκρίσεων ανά ζεύγη από τυχαία πειράματα είναι ένας ασφαλής τρόπος για την ερμηνεία της δικαιοσύνης σε βαθμολογίες και προτείνει μια νέα μέθοδο ρύθμισης που συμβάλλει στην βελτίωση της δικαιοσύνης κατά ζεύγη σε παραγωγικά συστήματα συστάσεων μεγάλης κλίμακας. Η προτεινόμενη μέθοδος μπορεί να αναγνωρίσει και να βελτιώσει συστηματικές λάθος προβλέψεις για τις βαθμολογήσεις μιας συγκεκριμένης ομάδας. Από την άλλη πλευρά, η μέθοδος που παρουσιάζεται στο παρόν κεφάλαιο παράγει προτάσεις που βασίζονται σε προκαθορισμένα κατώφλια για τη μέση ελάχιστη κάλυψη και διαφοροποίηση χρηστών κάθε αντικείμενου για κάθε πάροχο και δε λαμβάνει υπόψη τη θέση στην οποία προτείνεται ένα αντικείμενο στη λίστα συστάσεων.

Τέλος άλλες πρακτικές εφαρμογές για την επίτευξη δικαιοσύνης σε συστήματα συστάσεων παρουσιάζονται στις εργασίες [28], [29].

### 9.2.5 Ορισμός προβλήματος

Έχοντας ήδη παρουσιάσει τα πολυμετοχικά συστήματα συστάσεων, όσο και τις βασικές μετρικές αξιολόγησης συστημάτων συστάσεων παράλληλα με διάφορες εφαρμογές τους στα συστήματα συστάσεων, είναι σημαντικό να παρουσιάσουμε την εργασία [7], η οποία αποτελεί τη βάση στην οποία στηρίχθηκε η δουλειά του συγκεκριμένου κεφαλαίου. Στην εργασία αυτοί οι συγγραφείς παρουσιάζουν μια μοντελοποίηση γραμμικού προγραμματισμού για το πρόβλημα της δικαιοσύνης μεταξύ των παρόχων αντικειμένων όσον αφορά την κάλυψη και τη διαφοροποίηση στα συστήματα συστάσεων, παράλληλα με την προτεινόμενη λύση του προβλήματος.

Στην εργασία αυτή, ως  $I$  ορίζεται το σύνολο των αντικειμένων, ενώ ως  $U$  ορίζεται το σύνολο των χρηστών. Ως βασικό (baseline) σύστημα συστάσεων ορίζεται ένα σύστημα item-item collaborative filtering που δημιουργεί λίστες προτάσεων  $L_u$  για κάθε χρήστη  $u$ , λαμβάνοντας υπόψη μόνο τις προτεινόμενες βαθμολογίες και χωρίς κανέναν άλλον περιορισμό. Επιπλέον, θεωρούμε ότι  $C$  είναι ο αριθμός των παρόχων, και ότι κάθε αντικείμενο ανήκει σε έναν ακριβώς πάροχο, ενώ  $I_c$  είναι το σύνολο αντικειμένων του παρόχου  $c$ , όπου  $c = 1, \dots, C$  για κάθε χρήστη  $u$  και αντικείμενο  $i$  ενός παρόχου. Τέλος, το  $r_{iu}$  δείχνει την προβλεπόμενη βαθμολογία με το βασικό αλγόριθμο συστάσεων.

Αξίζει να σημειωθεί ότι η προτεινόμενη μεθοδολογία μπορεί να επεκταθεί για κατηγορίες ή κλάσεις από αντικείμενα με την προϋπόθεση ότι κάθε αντικείμενο ανήκει σε μια ακριβώς κατηγορία. Τότε το  $C$  μπορεί να σημαίνει τις διαφορετικές κατηγορίες/κλάσεις και όχι απαραίτητα τους διαφορετικούς παρόχους. Για το λόγο αυτό, από εδώ και πέρα οι όροι πάροχος ή κατηγορία θα σημαίνουν το ίδιο πράγμα στα πλαίσια του παρόντος κεφαλαίου.

Η έξοδος του αλγορίθμου συστάσεων είναι μια λίστα από προτεινόμενα αντικείμενα για κάθε χρήστη. Κάθε λίστα συστάσεων θα έχει μέγεθος  $L$ . Επιπλέον, το  $L_u$  δείχνει το σύνολο προτεινόμενων αντικειμένων για το χρήστη  $u$ , ενώ το  $L_u'$  δείχνει τη νέα λίστα προτεινόμενων αντικειμένων για το χρήστη  $u$  η οποία ικανοποιεί τους περιορισμούς για κάλυψη και διαφοροποίηση.

Επίσης, έστω  $x = (x_{iu}: i \in I, u \in U)$ , δείχνει τη νέα δυαδική πολιτική συστάσεων (αν  $x_{iu} = 0$ , το αντικείμενο  $i$  δεν προτείνεται στο χρήστη  $u$ , αν  $x_{iu} = 1$ , το αντικείμενο  $i$  προτείνεται στο χρήστη  $u$ ). Η πολιτική αυτή

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

είναι το ζητούμενο που πρέπει να αποφασιστεί με τη χρήση του αλγορίθμου. Η συνολική απόκλιση μεταξύ της συνολικής βαθμολογίας της βασικής λίστας συστάσεων με την νέα ορίζεται ως:

$$Cost(x) = \frac{1}{L|U|} (\sum_{u \in U} \sum_{i \in Lu} riu - \sum_{i \in I} \sum_{u \in U} riu * xiu), \quad (1)$$

Σχετικά με την κάλυψη χρήστη, υπολογίζεται ως το άθροισμα της κάλυψης όλων των αντικειμένων για κάθε πάροχο. Έτσι η μέση κάλυψη χρήστη για τα αντικείμενα ενός παρόχου  $c$  είναι:

$$Cov(c,x) = \frac{1}{|c|} \sum_{i \in Ic} \sum_{u \in U} xiu, \quad (2)$$

Όσον αφορά τη διαφοροποίηση ένα αντικείμενο σε χρήστες για τον πάροχο  $c$ , υπολογίζεται για κάθε πάροχο ως εξής:

$$Div(c,x) = \frac{1}{|c|} \sum_{i \in Ic} \sum_{u \in U} \sum_{v \in U, v \neq u} duv * xiu * xiv, \quad (3)$$

Όπου  $duv$  είναι η ανομοιότητα μεταξύ των χρηστών  $u$  και  $v$ .

Πέρα από την απόλυτη διαφοροποίηση για κάθε κατηγορία/ πάροχο, η μέση κανονικοποιημένη διαφοροποίηση (για κάθε αντικείμενο και χρήστη), που θα χρησιμοποιηθεί ορίζεται ως εξής:

$$\overline{Div}(c,x) = \frac{2 \sum_{i \in Ic} \sum_{u \in U} \sum_{v \in U, v \neq u} duv * xiu * xiv}{|c| (\sum_{i \in Ic} \sum_{u \in U} xiu) * (\sum_{i \in Ic} \sum_{u \in U} xiu - 1)}, \quad (4)$$

Ή απλούστερα:

$$\overline{Div}(c,x) = \frac{2 \sum_{i \in Ic} \sum_{u \in U} \sum_{v \in U, v \neq u} duv * xiu * xiv}{|c| Kc |c| * (Kc |c| - 1)}, \quad (5)$$

Και οι περιορισμοί για την κάλυψη των παρόχων γίνονται:

$$Cov(c,x) \geq Kc, \quad (6)$$

Και αντίστοιχα για τη διαφοροποίηση:

$$Div(c,x) \geq Dc, \quad (7)$$

Όπου  $Kc$  η ελάχιστη κάλυψη χρηστών για τα αντικείμενα ενός παρόχου  $c$ , και  $Dc$  το ελάχιστο κατώφλι για την μέση κανονικοποιημένη διαφοροποίηση για αντικείμενα και χρήστες του παρόχου  $c$ .

Παράλληλα με τον ορισμό του προβλήματος βελτιστοποίησης στη συγκεκριμένη εργασία ([7]), παρέχεται μια λύση για την επίλυση των περιορισμών κάλυψης σε πολυωνυμικό χρόνο, μαζί με μια ευριστική λύση χαμηλής πολυπλοκότητας για την επίλυση των περιορισμών διαφοροποίησης. Τα αποτελέσματα της προτεινόμενης μεθόδου δείχνουν ικανοποιητική απόδοση σύμφωνα με τους συγγραφείς.

Στο πλαίσιο του παρόντος κεφαλαίου, ακολουθείται η προτεινόμενη μεθοδολογία και προτείνεται νέος ευριστικός αλγόριθμος για την επίλυση των περιορισμών διαφοροποίησης. Η νέα ευριστική λύση για την επίλυση των περιορισμών διαφοροποίησης ήταν απαραίτητη διότι η προτεινόμενη μέθοδος ήταν πολύ γενική, προτείνοντας την σύσταση αντικειμένων σε χρήστες που αυξάνουν κατά το μέγιστο τη διαφοροποίηση ενός παρόχου, αντικαθιστώντας προτάσεις αντικειμένων χαμηλής βαθμολογίας με την προϋπόθεση ότι δεν παραβιάζουν τους περιορισμούς κάλυψης. Λαμβάνοντας υπόψη ότι δεν είναι ούτε



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

εύκολο ούτε προφανές να υπολογιστεί ποιες προτάσεις πρέπει να αντικατασταθούν, στο παρόν κεφάλαιο παρέχεται μια πιο συγκεκριμένη προσέγγιση που παρουσιάζεται και αξιολογείται μαζί με τα αποτελέσματα της. Ακολούθως, εντοπίζονται και συζητούνται εκτενώς κάποιες αδυναμίες που εντοπίστηκαν στον προτεινόμενο ορισμό της διαφοροποίησης και προτείνονται δύο νέοι εναλλακτικοί ορισμοί, οι οποίοι αντιμετωπίζουν τις αδυναμίες που εντοπίστηκαν. Τέλος, προτείνεται μια νέα ευριστική μεθοδολογία για την επίλυση του νέου προβλήματος με τους νέους ορισμούς για τη διαφοροποίηση.

### 9.3 Η ευριστική λύση

Στην ενότητα αυτή παρουσιάζεται αναλυτικά η μεθοδολογία που χρησιμοποιήθηκε για την επίλυση του προβλήματος συστάσεων. Αποτελείται από τρία βήματα: (i) εύρεση της βέλτιστης λύσης για το βασικό πρόβλημα (χωρίς τους περιορισμούς για κάλυψη και διαφοροποίηση) με βάση τις καλύτερες προβλεπόμενες βαθμολογίες, (ii) εύρεση μιας λύσης που ικανοποιεί τους περιορισμούς κάλυψης και απέχει το λιγότερο δυνατό από τη βασική λύση στο πρόβλημα χωρίς τους περιορισμούς, και (iii) με βάση τη λύση που ικανοποιεί τους περιορισμούς της κάλυψης, εύρεση μιας λύσης που να ικανοποιεί τους περιορισμούς για διαφοροποίηση, αλλάζοντας τις λίστες, προτείνοντας περισσότερα αντικείμενα από τους παρόχους για τους οποίους δεν ικανοποιείται ο περιορισμός για διαφοροποίηση, αντικαθιστώντας αντικείμενα άλλων παρόχων τα οποία έχουν χαμηλή προβλεπόμενη βαθμολογία και οι πάροχοι στους οποίους ανήκουν παρουσιάζουν ικανοποιητική διαφοροποίηση. Όλα αυτά με την προϋπόθεση ότι οι περιορισμοί για κάλυψη και διαφοροποίηση στους παρόχους των οποίων τα αντικείμενα αντικαταστάθηκαν, συνεχίζουν να ικανοποιούνται.

#### 9.3.1 Η βασική λύση (πρόβλημα χωρίς περιορισμούς)

Όπως έχει αναφερθεί, το πρώτο βήμα της μεθοδολογίας είναι η εύρεση μιας λύσης στο πρόβλημα χωρίς τους περιορισμούς για κάλυψη και διαφοροποίηση. Για τη λύση σε αυτό το πρόβλημα εφαρμόστηκε η μέθοδος collaborative filtering με βάση τα αντικείμενα (item-based collaborative filtering), μιας και θεωρήθηκε η καταλληλότερη μέθοδος. Συγκεκριμένα, δημιουργήθηκε ο πίνακας χρηστών – αντικειμένων και προβλέφθηκαν οι βαθμολογίες των χρηστών για αντικείμενα τα οποία δεν είχαν βαθμολογήσει ακόμα. Για το σκοπό αυτό, χρησιμοποιήθηκε η μέθοδος των Κ-πλησιέστερων γειτόνων (K-Nearest Neighbors - KNN) [30]. Τα πειράματα με βάση τα διαθέσιμα δεδομένα έδειξαν ότι οι καλύτερες προβλέψεις προκύπταν για  $K=20$ . Όσον αφορά τη μέθοδο για τον υπολογισμό της ομοιότητας χρησιμοποιήθηκε η ομοιότητα Pearson [22].

Το αποτέλεσμα της συγκεκριμένης μεθοδολογίας είναι μια λίστα από αντικείμενα με τις καλύτερες προβλεπόμενες βαθμολογίες για κάθε χρήστη. Όπως προαναφέρθηκε, τα αποτελέσματα της μεθόδου αυτής θα χρησιμοποιηθούν για τον ορισμό του προβλήματος βελτιστοποίησης που περιγράφεται στις επόμενες ενότητες.

#### 9.3.2 Η λύση για την κάλυψη

Έχοντας διαθέσιμη τη λύση από το πρόβλημα χωρίς περιορισμούς, μπορούμε να λύσουμε το πρόβλημα υπό τους περιορισμούς για την κάλυψη, καθώς προκύπτει ένα πρόβλημα γραμμικού προγραμματισμού. Σύμφωνα με την εργασία [7] το πρόβλημα βελτιστοποίησης περιγράφεται από τις ακόλουθες εξισώσεις:

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

$$\max \left( \frac{1}{L|U|} \sum_{i \in I} \sum_{u \in U} r_{iu} * x_{iu} \right), \quad (8)$$

Έτσι ώστε:

$$Cov(c, x) = \frac{1}{|Ic|} \sum_{i \in Ic} \sum_{u \in U} x_{iu} \geq Kc, x_{iu} \in \{0,1\}, \quad (9)$$

$$\sum_{i \in I} x_{iu} = L, \text{ for all } u \in U \quad (10)$$

Όλα τα μεγέθη έχουν παρουσιαστεί αναλυτικά στο προηγούμενο κεφάλαιο.

Ο στόχος είναι η μεγιστοποίηση της συνολικής προβλεπόμενης βαθμολογίας για την νέα λίστα συστάσεων, ενώ η μέση κάλυψη για κάθε πάροχο πρέπει να είναι τουλάχιστον  $Kc$ , δηλαδή τα αντικείμενα του παρόχου  $c$  πρέπει να προτείνονται κατά μέσο όρο σε  $Kc$  χρήστες, ενώ ο τελευταίος περιορισμός δηλώνει ότι κάθε χρήστης λαμβάνει ακριβώς  $L$  προτάσεις αντικειμένων.

Όπως περιγράφεται επίσης στην εργασία [7], το πρόβλημα βελτιστοποίησης με τους περιορισμούς κάλυψης είναι γραμμικό και συνεπώς επιλύσιμο σε πολυωνυμικό χρόνο. Τα αποτελέσματα της λύσης στο πρόβλημα με περιορισμούς κάλυψης είναι μια λίστα από προτεινόμενα αντικείμενα, με τη μέγιστη προβλεπόμενη βαθμολογία, που επιτρέπουν την ικανοποίηση των περιορισμών κάλυψης. Οι λίστες αυτές θα χρησιμοποιηθούν ως βάση για την εύρεση λύσεων στο πρόβλημα της διαφοροποίησης.

### 9.3.3 Η γενική λύση

Όσον αφορά το στόχο της διαφοροποίησης, οι ακόλουθοι περιορισμοί πρέπει να προστεθούν στο μοντέλο που παρουσιάστηκε:

$$Div(c, x) = \frac{2 \sum_{i \in Ic} \sum_{u \in U} \sum_{v \in U, v \neq u} d_{uv} * x_{iu} * x_{iv}}{|Ic| * (Kc|Ic| - 1)} \geq D, \quad (11)$$

Όπου όλα τα μεγέθη έχουν παρουσιαστεί σε προηγούμενη ενότητα. Συγκεκριμένα, οι περιορισμοί για τη διαφοροποίηση δηλώνουν ότι η μέση διαφοροποίηση χρηστών ανά αντικείμενο για κάθε πάροχο πρέπει να είναι τουλάχιστον  $D$ .

Μιας και ο τύπος για τη διαφοροποίηση είναι σε τετραγωνική μορφή, το πρόβλημα γίνεται πολύ πιο πολύπλοκο, σε σχέση με τα προβλήματα που παρουσιάστηκαν μέχρι τώρα. Συγκεκριμένα, το πρόβλημα αυτό είναι ένα πρόβλημα ακέραιου προγραμματισμού τετραγωνικών περιορισμών, που ανήκει στην κατηγορία προβλημάτων NP-Hard, όπως αναγνωρίζεται και στην εργασία [7]. Αυτό σημαίνει ότι δεν μπορεί να επιλυθεί αναλυτικά σε αποδεκτό χρόνο. Κατά συνέπεια, αντί να ψάχνουμε τη βέλτιστη λύση του προβλήματος, είναι καλύτερο να ακολουθηθεί μια ευριστική προσέγγιση, ώστε να βρεθεί μια λύση κοντά στη βέλτιστη που ικανοποιεί τους περιορισμούς, όπως προτείνεται και στην εργασία [7]. Στην εργασία αυτή, η μέθοδος που προτείνεται είναι η εναλλαγή προτεινόμενων αντικειμένων μεταξύ των παρόχων με προτεραιότητα την βελτίωση της διαφοροποίησης, με την προϋπόθεση ότι εξακολουθούν να ισχύουν οι περιορισμοί για την κάλυψη. Ωστόσο, η συγκεκριμένη προσέγγιση είναι πολύ γενική, καθώς είναι πολύπλοκος ο υπολογισμός της διαφοροποίησης και της κάλυψης σε κάθε βήμα του αλγορίθμου, και ακόμα πιο πολύπλοκη η αναγνώριση αντικειμένων που αυξάνουν τη διαφοροποίηση στο μέγιστο, όπως και η αναγνώριση των αντικειμένων προς αντικατάσταση. Για τους λόγους αυτούς, ο

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

ορισμός της διαφοροποίησής διερευνήθηκε σε βάθος και στο παρόν κεφάλαιο προτείνεται μια πιο συγκεκριμένη ευριστική μέθοδος.

Συγκεκριμένα για την ανάπτυξη μιας ευριστικής λύσης, πραγματοποιήθηκε μια εξονυχιστική ερευνητική ανάλυση (exploratory analysis) στα δεδομένα για την αναγνώριση και ερμηνεία του τρόπου με τον οποίο αυξάνεται η διαφοροποίηση και πώς επηρεάζεται από άλλα μεγέθη όπως η διαφορετικότητα των χρηστών. Η αρχική διαίσθηση ήταν ότι η διαφοροποίηση θα αυξανόταν τα μέγιστα αν ένα αντικείμενο προτεινόταν σε όσο το δυνατόν πιο διαφορετικούς χρήστες. Ωστόσο, αποδείχτηκε εσφαλμένη.

Αντίθετα, παρατηρήθηκε ότι ο πιο σημαντικός παράγοντας για την αύξηση ή μείωση της διαφοροποίησης ενός αντικειμένου ήταν ο αριθμός χρηστών στους οποίους προτείνεται ένα αντικείμενο. Συγκεκριμένα, όσο μεγαλύτερος είναι αυτός ο αριθμός, τόσο μεγαλύτερη και η διαφοροποίηση για το συγκεκριμένο αντικείμενο και κατά συνέπεια και για τον πάροχο στον οποίο ανήκει.

Για το λόγο αυτό, σχεδιάστηκε ένας νέος ευριστικός αλγόριθμος για την ικανοποίηση των περιορισμών διαφοροποίησης. Χρησιμοποιεί ως βάση τη λύση με τους περιορισμούς κάλυψης, όπως περιγράφηκε στην προηγούμενη ενότητα. Στη συνέχεια, υπολογίζει τη διαφοροποίηση για κάθε πάροχο και βρίσκει τους παρόχους για τους οποίους η διαφοροποίηση δεν είναι επαρκής. Για κάθε έναν από αυτούς τους παρόχους βρίσκει το αντικείμενο που έχει προταθεί περισσότερες φορές και τους χρήστες στους οποίους δεν έχει προταθεί, ταξινομημένους με βάση την προβλεπόμενη βαθμολογία στο συγκεκριμένο αντικείμενο, και σε περίπτωση που έχουν προτάσεις από αντικείμενα παρόχων που έχουν επαρκή διαφοροποίηση, το αντικείμενο με τη χαμηλότερη προβλεπόμενη βαθμολογία από παρόχους με επαρκή διαφοροποίηση, αντικαθίσταται από μια πρόταση του πιο δημοφιλούς αντικειμένου του παρόχου του οποίου η διαφοροποίηση είναι προς αύξηση, με την προϋπόθεση φυσικά ότι η διαφοροποίηση της κατηγορίας που ανήκει το αντικείμενο που αντικαταστάθηκε εξακολουθεί να είναι επαρκής. Με αυτό το τρόπο η διαφοροποίηση του παρόχου αυξάνεται και η διαδικασία επαναλαμβάνεται μέχρι να ικανοποιείται ο περιορισμός της διαφοροποίησης για το συγκεκριμένο πάροχο. Επιπλέον, η ίδια διαδικασία επαναλαμβάνεται για κάθε πάροχο με ανεπαρκή διαφοροποίηση.

Μια πιο λεπτομερής εκδοχή του ευριστικού αλγορίθμου που αναπτύχθηκε παρουσιάζεται στον παρακάτω πίνακα:

Πίνακας 27 - Ο ευριστικός αλγόριθμος που αναπτύχθηκε

**Algorithm 1.** Heuristic algorithm for addressing diversity constraints.

```
1:  $X_{cov} = \text{calculate\_solution\_for\_Coverage}(A\_pred, Kc)$  // as a linear programming problem through cvxpy
2:  $X_{new} = X_{cov}$  // copy Coverage solution as a starting point for the final solution
3: foreach category  $c$ :
4:    $Div[c] = \text{calculate\_Diversity\_for\_category}(X_{new}, c)$  // as described in formula (11)
5:  $categories\_ordered\_by\_Diversity = \text{argsort}(Div)$ 
6:  $categories\_ordered\_by\_Diversity\_desc = \text{reverse}(categories\_ordered\_by\_Diversity)$ 
7:  $categories\ to\ change, changeable\_categories = [], categories\_ordered\_by\_Diversity$ 
8: foreach category  $c$  in  $categories\_ordered\_by\_Diversity$ :
9:   If  $Div[c] < D$ :
10:     $Categories\_to\_change.add(c)$ 
11:     $changeable\_categories.remove(c)$ 
12: foreach category  $c$  in  $categories\_to\_change$ :
13:    $Most\_recommended\_item[c] = \text{find\_most\_recommended\_item\_of\_category}(c)$ 
14:    $Users\_to\_recommend = \text{argsort}(A\_pred[most\_recommended\_item[c]])$ 
15:    $Users\_to\_recommend\_except\_recommended = users\ not\ recommended\ with\ the\ most\ recommended\ item$ 
16:   foreach user  $u$  in  $users\_to\_recommend\_except\_recommended$ :
17:     foreach category  $c2$  in  $changeable\_categories$ :
18:        $rec\_items\_of\_categ\_to\_usr = \text{find\_rec\_items\_of\_category\_to\_user}(c2, u)$ 
19:       foreach item  $i$  in  $rec\_items\_of\_categ\_to\_usr$ :
20:          $X_{new}[u, i] = 0$ 
21:          $X_{new}[u, most\_recommended\_item[c]] = 1$ 
// Calculate new Coverage for category  $c2$  according to formula (9)
22:    $Cov\_xnew[c2] = \text{calculate\_Coverage\_for\_category}(X\_new, c2)$ 
// Calculate new Diversity for categories  $c2$  and  $c$  according to formula (11)
23:    $Div\_xnew[c2] = \text{calculate\_Diversity\_for\_category}(X_{new}, c2)$ 
24:    $Div\_xnew[c] = \text{calculate\_Diversity\_for\_category}(X_{new}, c)$ 
// If Coverage or Diversity constraints for  $c2$  are violated rollback
25:   If  $Cov\_xnew[c2] < K[c2]$  or  $Div\_xnew[c2] < D$ :
26:      $X_{new}[u, i] = 1$ 
27:      $X_{new}[u, most\_recommended\_item[c2]] = 0$ 
28:   Else if  $Div\_xnew[c] > D$ :
29:     Break (line 12)
```

Περισσότερες πληροφορίες, όπως και η υλοποίηση του αλγορίθμου στη γλώσσα προγραμματισμού Python μπορούν να βρεθούν στο αποθετήριο Github [31] το οποίο παρέχεται μαζί με τα άλλα code repositories της παρούσας διδακτορικής διατριβής στο Παράρτημα II.

Ο αλγόριθμος που αναπτύχθηκε και παρουσιάστηκε ακολουθεί μια άπληστη προσέγγιση και αποφεύγει την οπισθοδρόμηση (backtracking) για την εύρεση μιας καλύτερης λύσης σε περίπτωση που έχει βρεθεί μια επαρκής λύση σε προηγούμενο βήμα. Με τον τρόπο αυτό το πρόβλημα επιλύεται αποδοτικά όσον αφορά τον απαιτούμενο χρόνο και τη μνήμη. Η πολυπλοκότητα του αλγορίθμου στη χειρότερη περίπτωση (worst case scenario) είναι  $O(L*U*Ctc* Ccc)$ , όπου  $Ctc$  είναι ο αριθμός παρόχων για τους οποίους οι περιορισμοί διαφοροποίησης δεν ικανοποιούνται και  $Ccc$  ο αριθμός παρόχων για τους οποίους οι περιορισμοί διαφοροποίησης ικανοποιούνται.

Απλούστερα, η πολυπλοκότητα στη χειρότερη περίπτωση είναι  $O(L*U*C^2)$ . Αυτό διότι στη χειρότερη περίπτωση, όλα τα αντικείμενα πρέπει να προσπελαστούν για όλους τους χρήστες, για όλους τους παρόχους που ικανοποιούνται από άποψη διαφοροποίησης, για κάθε έναν πάροχο για τον οποίο δεν ικανοποιούνται οι περιορισμοί διαφοροποίησης. Ωστόσο, αξίζει να σημειωθεί ότι ο αλγόριθμος ολοκληρώνεται σε πολύ λιγότερα βήματα στις περισσότερες περιπτώσεις, καθώς όταν ο περιορισμός της διαφοροποίησης ικανοποιηθεί σε κάποιο βήμα για έναν πάροχο, προχωράει με τον επόμενο πάροχο, χωρίς να προσπελάσει όλα τα αντικείμενα και όλους τους χρήστες.

## 9.4 Πειραματική διαδικασία και αποτελέσματα

Στην ενότητα αυτή παρουσιάζεται το σύνολο δεδομένων (dataset) που χρησιμοποιήθηκε, και η διαδικασία της προεπεξεργασίας που διενεργήθηκε για την επίλυση του προβλήματος. Στη συνέχεια, παρουσιάζεται το πείραμα και τα αποτελέσματα της ευριστικής λύσης που αναπτύχθηκε.

### 9.4.1 Το Dataset

Το dataset που χρησιμοποιήθηκε σε αυτό το κεφάλαιο αφορά συστήματα συστάσεων ταινιών και είναι διαθέσιμο στον ιστότοπο MovieLens [32]. Παρέχονται δεδομένα από ακέριες βαθμολογίες (1, η χαμηλότερη και 5 η υψηλότερη) για 9724 αντικείμενα (ταινίες) και 610 χρήστες. Καθώς το dataset είναι πολύ αραιό και οι διαθέσιμοι υπολογιστικοί πόροι ήταν περιορισμένοι, το πείραμα εκτελέστηκε για ένα υποσύνολο του αρχικού dataset. Συγκεκριμένα, αντικείμενα που είχαν δέκα ή λιγότερες βαθμολογίες από χρήστες δε χρησιμοποιήθηκαν. Κατά συνέπεια το dataset που χρησιμοποιήθηκε τελικά αποτελείται από 2121 ταινίες.

### 9.4.2 Η βασική λύση (χωρίς περιορισμούς για κάλυψη και διαφοροποίηση)

Μετά τη φάση της προεπεξεργασίας, δημιουργήθηκε ο πίνακας χρήστη αντικειμένου με τις βαθμολογίες, με βάση τον οποίο πραγματοποιήθηκε η ανάλυση και η μοντελοποίηση. Για την βασική λύση, ακολουθήθηκε η μεθοδολογία που περιγράφεται στην εργασία [7].

Συγκεκριμένα, μετά την εφαρμογή 5-fold cross validation στον πίνακα βαθμολογιών χρήστη αντικειμένου (80% των δεδομένων χρησιμοποιήθηκαν για εκπαίδευση και 20% για επαλήθευση), υπολογίστηκε η ομοιότητα Pearson μεταξύ των αντικειμένων στο σύνολο εκπαίδευσης και κατόπιν πειραματισμού παρατηρήθηκε ότι για τη μέθοδο των K κοντινότερων γειτόνων (K-nearest neighbors) ο αριθμός K=20 έδινε τις καλύτερες προβλέψεις στο σύνολο εκπαίδευσης. Συγκεκριμένα, η μέθοδος των 20 κοντινότερων γειτόνων οδήγησε στο χαμηλότερο μέσο τετραγωνικό σφάλμα (Mean Squared Error - MSE) στις προβλεπόμενες βαθμολογίες στο σύνολο εκπαίδευσης. Στη συνέχεια, με βάση το μέτρο ομοιότητας, εκτιμήθηκαν οι βαθμολογίες για τα αντικείμενα του συνόλου επαλήθευσης. Ακολούθως, αξιολογήθηκε το αποτέλεσμα της μεθόδου item based collaborative filtering, με το μέσο τετραγωνικό σφάλμα να είναι 0.8605 για το σύνολο επαλήθευσης. Αυτό σημαίνει ότι το σύστημα έχει σχετικά ικανοποιητική απόδοση, καθώς οι προβλεπόμενες βαθμολογίες αποκλίνουν γύρω στους 0.86 βαθμούς από τις πραγματικές.

Επομένως, χρησιμοποιώντας τη συγκεκριμένη μέθοδο ως βασική λύση, προβλέπονται οι βαθμολογίες που θα έδιναν οι χρήστες για αντικείμενα που δεν έχουν βαθμολογήσει (που έχουν την τιμή null στον αρχικό πίνακα βαθμολογιών). Η έξοδος του αλγορίθμου είναι μια λίστα προτάσεων που αποτελείται από τα δέκα αντικείμενα με τις υψηλότερες προτεινόμενες βαθμολογίες για κάθε χρήστη. Η μέση βαθμολογία για τα προτεινόμενα αντικείμενα είναι 4.938.

### 9.4.3 Πάροχοι αντικειμένων ή κατηγορίες

Το αποτέλεσμα της μεθόδου item based collaborative filtering μεγιστοποιεί το άθροισμα των βαθμολογιών που προκύπτει από τα προτεινόμενα αντικείμενα. Ωστόσο, η συνολική βαθμολογία δεν είναι η μόνη μετρική που πρέπει να ληφθεί υπόψη, ως στόχος του συστήματος. Συγκεκριμένα, ο στόχος

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

είναι να ικανοποιούνται επίσης οι περιορισμοί για κάλυψη και διαφοροποίηση για κάθε πάροχο αντικειμένων, και αυτό επιτυγχάνεται με τη χρήση κατωφλίων (thresholds) ανά πάροχο για τις παραπάνω μετρικές.

Καθώς η κάλυψη και η διαφοροποίηση στην προκειμένη περίπτωση είναι προσανατολισμένες στους παρόχους αντικειμένων και κάθε αντικείμενο συσχετίζεται με έναν πάροχο, το πείραμα πραγματοποιήθηκε για διαφορετικούς αριθμούς παρόχων με όμοια αποτελέσματα. Για το λόγο αυτό, παρουσιάζονται μόνο τα αποτελέσματα για είκοσι παρόχους αντικειμένων. Συγκεκριμένα, για το πείραμα δημιουργήθηκαν είκοσι πάροχοι και κάθε αντικείμενο ανατέθηκε σε έναν ακριβώς πάροχο, τυχαία χρησιμοποιώντας ομοιόμορφη κατανομή. Κατά συνέπεια, η ανάθεση ενός αντικειμένου σε έναν συγκεκριμένο πάροχο είναι ισοπίθανη με την ανάθεση σε οποιονδήποτε άλλο πάροχο. Έτσι, η κατανομή αντικειμένων στους διάφορους παρόχους είναι σχεδόν ομοιόμορφη, όπως φαίνεται και στον παρακάτω πίνακα.

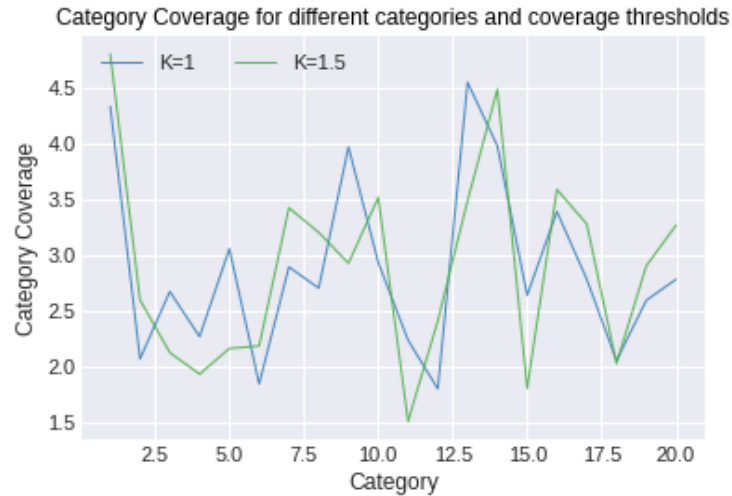
Πίνακας 28 - Κατανομή αντικειμένων σε παρόχους (ή κατηγορίες)

Category	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Items	101	91	105	115	105	107	90	121	101	101	90	114	94	120	115	104	114	108	107	118

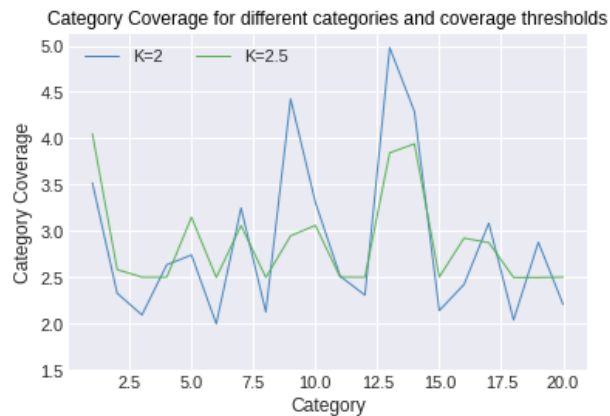
#### 9.4.4 Η λύση για την κάλυψη

Το πρόβλημα με τους περιορισμούς για κάλυψη επιλύθηκε για διακριτές τιμές του  $x$  (συγκεκριμένα 0 και 1), καθώς ένα αντικείμενο είτε προτείνεται είτε όχι σε ένα χρήστη. Για τη λύση του προβλήματος χρησιμοποιήθηκε η βιβλιοθήκη `cnxry` [33] της γλώσσας προγραμματισμού `Rython`. Παρότι οι νέες λίστες για τις διαφορετικές τιμές του  $K$  είναι διαφορετικές από τη βασική λίστα (που προκύπτει από τη λύση του βασικού προβλήματος), το άθροισμα των βαθμολογιών για τις διάφορες τιμές του  $K$  είναι το ίδιο με το βασικό πρόβλημα. Συγκεκριμένα, το συνολικό άθροισμα των βαθμολογιών είναι το ίδιο για  $K=1, 1.5, 2, 2.5$ , ενώ για  $K \geq 3$  δεν υπάρχει λύση στο πρόβλημα. Αυτό συμβαίνει διότι στο `dataset` που μελετάμε υπάρχουν πολλές τέλειες βαθμολογίες (βαθμός=5), όπως και προβλεπόμενες βαθμολογίες και τα αντικείμενα των νέων λιστών  $Lu'$  είναι διαφορετικά από τις λίστες της βασικής λύσης  $Lu$ . Ωστόσο, και τα νέα αντικείμενα έχουν τέλειες προβλεπόμενες βαθμολογίες όπως και τα παλιά. Ωστόσο, η κατανομή της κάλυψης διαφέρει ανά πάροχο αντικειμένων, όπως φαίνεται και στις παρακάτω εικόνες:

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 39 - Κάλυψη ανά πάροχο (ή κατηγορία) για K=1 και K=1.5



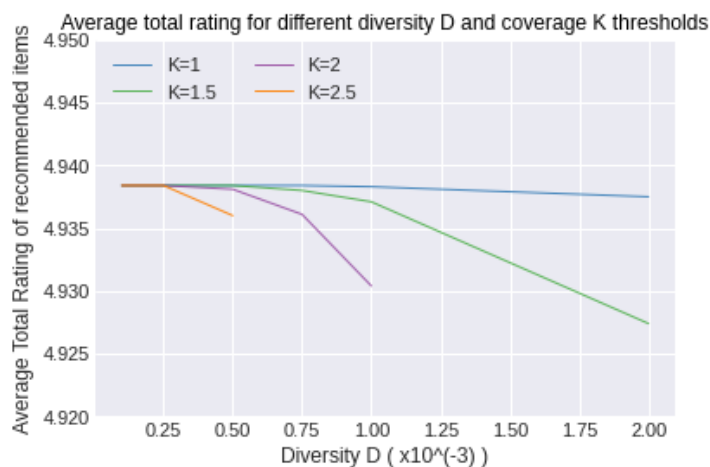
Εικόνα 40 - Κάλυψη ανά πάροχο (ή κατηγορία) για K=2 και K=2.5

Στις παραπάνω εικόνες είναι εύκολα αντιληπτό ότι όσο το K αυξάνεται, τόσο η κάλυψη τείνει να γίνεται πιο ομοιόμορφη ανά πάροχο αντικειμένων.

#### 9.4.5 Η γενική λύση

Όπως έχει ήδη αναφερθεί, η λύση του προβλήματος με τους περιορισμούς για κάλυψη χρησιμοποιείται ως βάση για την εύρεση μιας ικανοποιητικής λύσης που ικανοποιεί τους περιορισμούς για διαφοροποίηση ανά πάροχο. Σε αυτό το πλαίσιο, χρησιμοποιώντας τα αποτελέσματα που παρουσιάστηκαν στην προηγούμενη ενότητα, ακολουθήθηκε ο ευριστικός αλγόριθμος (Πίνακας 28) που παρουσιάστηκε προηγουμένως. Το αποτέλεσμα της προτεινόμενης λύσης φαίνεται στην παρακάτω εικόνα:

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 41 - Μέση βαθμολογία για τις διάφορες τιμές διαφοροποίησης και κάλυψης

Όπως φαίνεται στην παραπάνω εικόνα, για κάθε τιμή της κάλυψης, όσο το κατώφλι της διαφοροποίησης μεγαλώνει, τόσο η μέση βαθμολογία των προτεινόμενων αντικειμένων πέφτει ή παραμένει η ίδια. Επιπροσθέτως, για μεγαλύτερα κατώφλια κάλυψης το πρόβλημα γίνεται πιο δύσκολο και η μέση βαθμολογία πέφτει με πιο ραγδαίους ρυθμούς με την αύξηση του κατωφλίου για τη διαφοροποίηση. Πιο συγκεκριμένα, για κατώφλι κάλυψης  $K_c = 2.5$  το πρόβλημα δεν έχει λύση για κατώφλι διαφοροποίησης μεγαλύτερο του 0.0005 ( $D > 0.0005$ ). Στον παρακάτω πίνακα παρουσιάζεται η συνολική και η μέση βαθμολογία των προτεινόμενων αντικειμένων για τις διαφορετικές τιμές του κατωφλίου διαφοροποίησης ( $D$ ) και για κατώφλια κάλυψης  $K_c=1$  και  $K_c=2.5$ .

Πίνακας 29 - Μέση και συνολική βαθμολογία για τις διάφορες τιμές διαφοροποίησης και κάλυψης

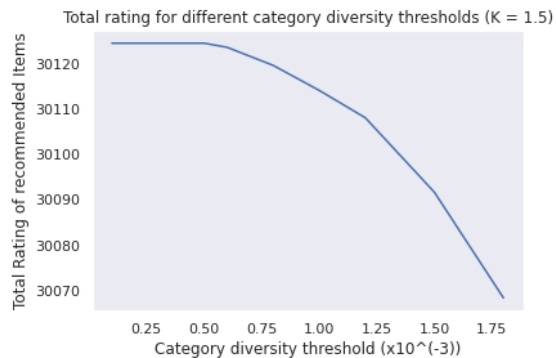
Diversity Threshold $D$	Total Rating ( $K_c = 1$ )	Average Rating ( $K_c = 1$ )	Total Rating ( $K_c = 2.5$ )	Average Rating ( $K_c = 2.5$ )
0.0001	30,124.411	4.9384	30,124.411	4.9384
0.0002	30,124.411	4.9384	30,124.411	4.9384
0.00025	30,124.411	4.9384	30,124.411	4.9384
0.0005	30,124.411	4.9384	30,110.984	4.9362
0.00075	30,124.411	4.9384	No solution	No solution
0.001	30,124.120	4.9383	No solution	No solution
0.002	30,119.159	4.9375	No solution	No solution

Όπως φαίνεται στον παραπάνω πίνακα και στην παραπάνω εικόνα, η συνολική βαθμολογία όλων των προτεινόμενων αντικειμένων για όλους τους χρήστες είναι πολύ ψηλή και στις περισσότερες περιπτώσεις είναι ίδια με τη βασική λύση. Αυτό δείχνει ότι ενώ ο ευριστικός αλγόριθμος μπορεί να μην εντοπίζει πάντα τη βέλτιστη λύση, βρίσκει λύσεις οι οποίες είναι πολύ κοντά στη βέλτιστη.

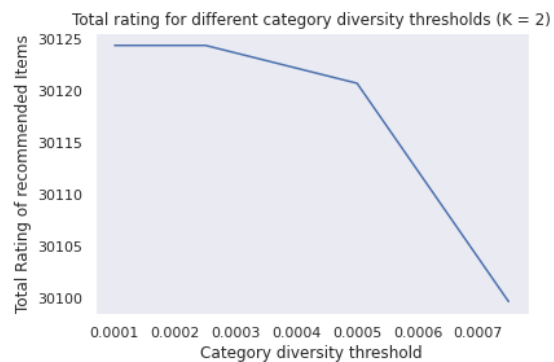


Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Τα παρακάτω γραφήματα δείχνουν με περισσότερη λεπτομέρεια πώς η διαφοροποίηση επηρεάζει τη συνολική βαθμολογία. Συγκεκριμένα, τα παρακάτω γραφήματα παρουσιάζουν τη συνολική βαθμολογία για τα διάφορα κατώφλια της διαφοροποίησης για κατώφλια κάλυψης  $K=1.5$  και  $K=2$  αντίστοιχα:



Εικόνα 42 - Συνολική βαθμολογία για διαφορετικά κατώφλια διαφοροποίησης για  $K=1.5$



Εικόνα 43 - Συνολική βαθμολογία για διαφορετικά κατώφλια διαφοροποίησης για  $K=2$

Στα παραπάνω διαγράμματα, φαίνεται ότι όσο το κατώφλι διαφοροποίησης αυξάνεται, τόσο μειώνεται η συνολική βαθμολογία, ενώ παράλληλα η συνολική βαθμολογία μειώνεται περισσότερο για μεγαλύτερες τιμές του κατωφλίου διαφοροποίησης. Σχετικά με τη σχέση μεταξύ διαφοροποίησης και κάλυψης, φαίνεται ότι για μεγαλύτερο κατώφλι κάλυψης, η συνολική βαθμολογία πέφτει ταχύτερα με την αύξηση του κατωφλίου διαφοροποίησης. Το φαινόμενο αυτό είναι λογικό καθώς μεγαλύτερο κατώφλι κάλυψης σημαίνει πιο αυστηρούς περιορισμούς καθώς επίσης και λιγότερες επιτρεπόμενες αλλαγές για αντικείμενα διαφορετικών παρόχων στις προτεινόμενες λίστες αντικειμένων.

Επιπλέον, αξίζει να σημειωθεί ότι τα αποτελέσματα (όπως φαίνεται στον πίνακα 2) όσον αφορά τη μέση και τη συνολική βαθμολογία είναι πολύ κοντά με τα αποτελέσματα της βασικής λύσης, τα οποία αποτελούν και την θεωρητικά βέλτιστη λύση του προβλήματος, καθώς το περιορισμένο πρόβλημα ως προς την κάλυψη και τη διαφοροποίηση, δεν είναι δυνατόν να παράγει καλύτερες συστάσεις από το βασικό πρόβλημα χωρίς τους περιορισμούς.

Όπως έχει ήδη αναφερθεί, ο προτεινόμενος αλγόριθμος είναι ευριστικός και δεν βρίσκει πάντα τη βέλτιστη λύση στο πρόβλημα βελτιστοποίησης. Ωστόσο, όπως φαίνεται από τα αποτελέσματα παρέχει πάντα μια λύση πολύ κοντά στη βέλτιστη. Συγκεκριμένα, ακολουθεί μια άπληστη προσέγγιση και

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

αποφεύγει την οπισθοδρόμηση και κατά συνέπεια, αν εντοπίσει μια λύση που είναι αρκετά καλή δε θα προσπαθήσει να εντοπίσει μια καλύτερη.

Οι περιπτώσεις στις οποίες αποτυγχάνει να εντοπίσει τη βέλτιστη λύση προκύπτουν από το γεγονός ότι όταν οι περιορισμοί διαφοροποίησης ικανοποιηθούν για έναν πάροχο, τα προτεινόμενα αντικείμενα για το συγκεκριμένο πάροχο θα αλλάξουν μόνο με την προϋπόθεση ότι σε μια συγκεκριμένη αντικατάσταση ενός αντικειμένου οι περιορισμοί για κάλυψη και διαφοροποίηση θα συνεχίσουν να ικανοποιούνται. Επίσης, θα είναι στο πλαίσιο ικανοποίησης των περιορισμών διαφοροποίησης για κάποιον άλλον πάροχο. Συνεπώς, είναι πιθανό να υπάρχει κάποιο καλύτερο πλάνο συστάσεων, με μεγαλύτερη μέση βαθμολογία από το αποτέλεσμα που προέκυψε από την ευριστική λύση. Για παράδειγμα, ο ευριστικός αλγόριθμος που παρουσιάστηκε μπορεί να αποτύχει σε περίπτωση που το αντικείμενο του παρόχου που προτάθηκε περισσότερες φορές έχει χαμηλότερη βαθμολογία στον χρήστη που προτείνεται από το δεύτερο πιο προτεινόμενο. Αν και τα δύο αντικείμενα ικανοποιούν τον περιορισμό της διαφοροποίησης τότε θα επιλεγεί το πρώτο παρότι το δεύτερο είναι καλύτερη λύση. Ακόμη, αν σε μια συγκεκριμένη εναλλαγή προτεινόμενων αντικειμένων ο περιορισμός για τη διαφοροποίηση κάποιας άλλης κατηγορίας παραβιάζεται δε θα πραγματοποιηθεί η εναλλαγή, ακόμα και αν υπάρχουν επιλογές που ικανοποιούν τους περιορισμούς. Φυσικά, τα αποτελέσματα είναι πολύ ικανοποιητικά τουλάχιστον για το συγκεκριμένο dataset, καθώς είναι πολύ κοντά στη θεωρητικά βέλτιστη λύση.

## 9.5 Αξιολόγηση και επεκτάσεις

### 9.5.1 Αξιολόγηση

Η μεθοδολογία που ακολουθήθηκε για την επίλυση του προβλήματος βελτιστοποίησης για την επίτευξη δικαιοσύνης μεταξύ των παρόχων αντικειμένων όσον αφορά την κάλυψη και τη διαφοροποίηση σε χρήστες των αντικειμένων τους, βρίσκει μια ευριστική λύση στο πρόβλημα, η οποία, σύμφωνα με τα αποτελέσματα είναι πολύ κοντά στη θεωρητικά βέλτιστη λύση η οποία προκύπτει από τη βασική λύση του προβλήματος χωρίς περιορισμούς, και κατά συνέπεια δρα ως άνω όριο για την προτεινόμενη λύση στο πρόβλημα που μελετήθηκε.

Όσον αφορά τη σύγκριση με αποτελέσματα άλλων αντίστοιχων ερευνητικών εργασιών, μόνο η εργασία [7] ασχολείται με το πρόβλημα της δικαιοσύνης μεταξύ των παρόχων ως προς την κάλυψη και τη διαφοροποίηση σε χρήστες των αντικειμένων τους. Τα αποτελέσματα του αλγορίθμου που προτείνεται σε αυτό το κεφάλαιο είναι παρόμοια με τα πειραματικά αποτελέσματα της προαναφερθείσας εργασίας, με την έννοια ότι όσο οι απαιτήσεις για κάλυψη και διαφοροποίηση ανά πάροχο αυξάνονται, η συνολική προβλεπόμενη βαθμολογία των προτεινόμενων αντικειμένων μειώνεται, ενώ οι περιορισμοί διαφοροποίησης έχουν μεγαλύτερο αντίκτυπο στη συνολική βαθμολογία των προτεινόμενων αντικειμένων. Ωστόσο, η μέθοδος που ακολουθείται στην παρούσα διδακτορική διατριβή (και το παρόν κεφάλαιο) δείχνει μικρότερη απόκλιση από τη βασική λύση, που αποτελεί το άνω όριο για τη συνολική προβλεπόμενη βαθμολογία των προτεινόμενων αντικειμένων. Φυσικά, αυτό δεν αποτελεί ασφαλές συμπέρασμα, καθώς η εργασία [7] δεν περιγράφει λεπτομερώς ούτε τη διαδικασία της προεπεξεργασίας αλλά ούτε και τις υπερπαραμέτρους που χρησιμοποιήθηκαν.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Σε γενικές γραμμές, το πλεονέκτημα της προσέγγισης του προβλήματος συστάσεων ως ένα πρόβλημα μαθηματικής βελτιστοποίησης με περιορισμούς είναι ότι εγγυάται ότι αν υπάρχει λύση τα κριτήρια που τίθενται ως περιορισμοί θα ικανοποιούνται. Φυσικά, οι περιορισμοί θα πρέπει να ορίζονται λαμβάνοντας υπόψη και τον αντίκτυπο που θα έχουν στις προτιμήσεις του χρήστη. Αυτό σημαίνει ότι αν οι περιορισμοί είναι πολύ αυστηροί, τα προτεινόμενα αντικείμενα μπορεί να μην είναι επαρκώς σχετικά με τις προτιμήσεις του χρήστη. Επομένως, οι διαχειριστές των συστημάτων συστάσεων θα πρέπει συνεχώς να αξιολογούν και να ενημερώνουν τις απαιτήσεις τους.

Κατά συνέπεια, η προτεινόμενη μεθοδολογία ή κάποια παρόμοια προσέγγιση μαθηματικής βελτιστοποίησης με περιορισμούς για ίση αντιμετώπιση των παρόχων αντικειμένων ή άλλων ενδιαφερόμενων ομάδων μπορεί να χρησιμοποιηθεί για την άμβλυνση του προβλήματος «Ο νικητής τα παίρνει όλα» (The winner takes it all phenomenon) [26], σύμφωνα με το οποίο, οι πιο σημαντικοί πάροχοι ή ομάδες παρόχων έχουν μεγάλο πλεονέκτημα έναντι των άλλων. Ο λόγος που μια τέτοια μεθοδολογία είναι κατάλληλη για την αντιμετώπιση του προβλήματος αυτού είναι ότι εγγυάται έναν ελάχιστο βαθμό έκθεσης (π.χ. κάλυψης ή διαφοροποίησης) για όλους τους παρόχους, αν υπάρχει λύση στο πρόβλημα. Επίσης, η συγκεκριμένη προσέγγιση μπορεί να προστατεύσει τις μειονότητες καθώς εγγυάται τουλάχιστον την απαιτούμενη κάλυψη και διαφοροποίηση. Ενώ, με μια παρόμοια προσέγγιση, η απαιτούμενη έκθεση και ορατότητα των μειονοτήτων μπορεί να βελτιωθεί σημαντικά.

Ωστόσο, αξίζει να σημειωθεί ότι το πρόβλημα όπως ορίζεται στη βιβλιογραφία [7] παρουσιάζει και κάποιους σημαντικούς περιορισμούς. Συγκεκριμένα, δε λαμβάνει υπόψη τη θέση ενός αντικειμένου στη λίστα συστάσεων. Στην πραγματικότητα η θέση του αντικειμένου στη λίστα συστάσεων είναι πολύ σημαντική καθώς διαδραματίζει σημαντικό ρόλο στη συμπεριφορά των χρηστών, ειδικά όταν τα αντικείμενα που προτείνονται είναι πολλά. Για παράδειγμα, ένας χρήστης μπορεί να δώσει μεγάλη σημασία στην κορυφή μιας λίστας προτάσεων αλλά να αγνοήσει τις τελευταίες προτάσεις της λίστας.

Ένας ακόμη περιορισμός για το πρόβλημα που ορίστηκε είναι ο ορισμός της διαφοροποίησης χρηστών για ένα αντικείμενο. Συγκεκριμένα, αποτυγχάνει να συλλάβει την διαφορετικότητα των χρηστών στους οποίους προτείνεται ένα αντικείμενο, ως το σημαντικότερο παράγοντα διαφοροποίησης. Αντίθετα, παρατηρείται ότι η συχνότητα πρότασης των αντικειμένων ενός παρόχου επηρεάζει πολύ περισσότερο τη διαφοροποίηση από ό,τι η διαφορετικότητα των χρηστών στους οποίους προτάθηκε, και η λύση που αναπτύχθηκε εκμεταλλεύεται τη συγκεκριμένη αδυναμία.

Για να κάνουμε πιο εμφανή τη συγκεκριμένη αδυναμία, προσεγγίσαμε τον ορισμό της διαφοροποίησης χρηστών ενός αντικειμένου όπως ορίζεται στη βιβλιογραφία υποθέτοντας ότι η διαφορετικότητα μεταξύ δύο διαφορετικών χρηστών είναι ίση με τη μέση διαφορετικότητα δύο οποιωνδήποτε χρηστών στο dataset (έστω  $d$ ). Με τη συγκεκριμένη προσέγγιση, αν θεωρήσουμε το  $d$  ως τη μέση διαφοροποίηση μεταξύ δύο διαφορετικών χρηστών, ο τύπος υπολογισμού της διαφοροποίησης αντικειμένου γίνεται:

$$Div(c, x) = \frac{2}{|Ic|} \frac{\sum_{i \in Ic} \sum_{u \in U} \sum_{v \in U, v \neq u} d * x_{iu} * x_{iv}}{Kc |Ic| * (Kc |Ic| - 1)} \geq D, \quad (12)$$

Ωστόσο, αν αναλύσουμε παραπάνω τη συνολική διαφορετικότητα ενός αντικειμένου προκύπτει:

Αν το αντικείμενο προτείνεται σε έναν ή κανέναν χρήστη:

$$Div(item) = 0$$

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Αν προτείνεται σε 2 χρήστες τον  $u_1$  και τον  $u_2$ :

$$Div(item) = d(u_1, u_2) = d$$

Αν προτείνεται σε 3 χρήστες, τους  $u_1, u_2, u_3$ :

$$Div(item) = d(u_1, u_2) + d(u_1, u_3) + d(u_2, u_3) = 2d + d = 3d$$

Με την ίδια λογική αν το αντικείμενο προτείνεται σε  $n$  χρήστες ( $u_1, \dots, u_n$ ) προκύπτει:

$$Div(item) = (1 + 2 + 3 + \dots + (n - 1)) * d$$

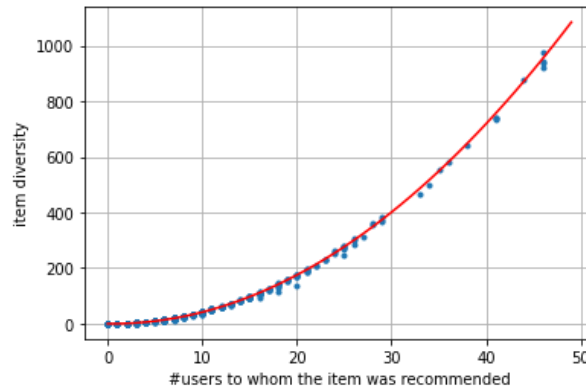
Και λαμβάνοντας υπόψη ότι  $(1 + 2 + 3 + \dots + (n-1))$  είναι ένα πολύ γνωστό άθροισμα που ισούται με  $(1/2) * n * (n-1)$ , προκύπτει:

$$Div(item) = \frac{1}{2} n(n - 1)d , \quad (13)$$

Με τον παραπάνω τύπο, προσεγγίστηκε η διαφοροποίηση όλων των αντικειμένων του dataset και τα αποτελέσματα φαίνονται στα παρακάτω διαγράμματα (Εικόνα 55 και 56):

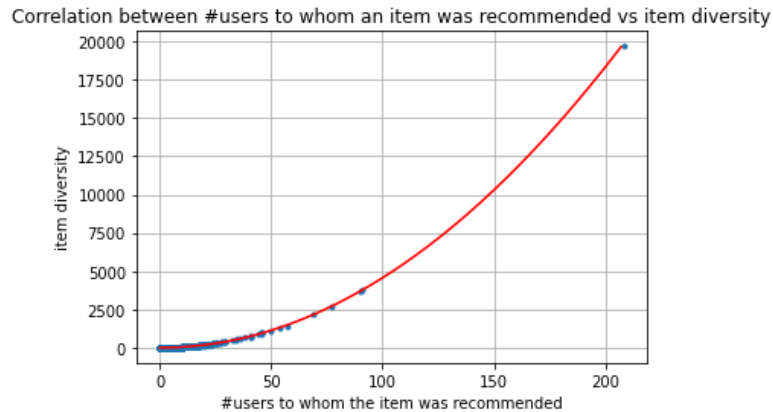
Συγκεκριμένα τα διαγράμματα αυτά δείχνουν την προσέγγιση της διαφοροποίησης για τις διαφορετικές τιμές του  $n$  (αριθμός χρηστών στους οποίους προτάθηκε ένα αντικείμενο). Συγκεκριμένα η κόκκινη γραμμή δείχνει την προσέγγιση που εφαρμόστηκε, ενώ τα μπλε σημεία δείχνουν την πραγματική διαφοροποίηση των αντικειμένων όπως υπολογίστηκε για το dataset και η συσχέτιση τους με το  $n$ . Η Εικόνα 44 δείχνει τα αποτελέσματα μετά την απομάκρυνση των ακραίων τιμών (outliers), ενώ η Εικόνα 45 δείχνει όλες τις διαφοροποιήσεις για κάθε αντικείμενο και κάθε τιμή του  $n$  που προέκυψε στα αποτελέσματα.

Correlation between #users to whom an item was recommended vs item diversity



Εικόνα 44 - Συσχέτιση μεταξύ διαφοροποίησης ανά αντικείμενο και αριθμού χρηστών στους οποίους προτείνεται (με απομάκρυνση ακραίων τιμών)

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 45 - Συσχέτιση μεταξύ διαφοροποίησης ανά αντικείμενο και αριθμού χρηστών στους οποίους προτείνεται

Όπως φαίνεται από τα παραπάνω διαγράμματα, η προσέγγιση δείχνει πολύ ακριβής σε σχέση με τα πραγματικά δεδομένα (τουλάχιστον για το συγκεκριμένο dataset), καθώς οι πραγματικές τιμές είναι πάρα πολύ κοντά με τις τιμές που προβλέφθηκαν, και αυτό ισχύει για κάθε αντικείμενο του dataset.

Στο σημείο αυτό, αξίζει να σημειωθεί ότι διενεργήθηκαν διάφοροι στατιστικοί έλεγχοι (όπως ο έλεγχος Kolmogorov-Smirnov τόσο για ομοιομορφία όσο και για κανονικότητα), για τον έλεγχο εάν η διαφορετικότητα μεταξύ των χρηστών είναι είτε ομοιόμορφη είτε ακολουθεί κανονική κατανομή γύρω από τη μέση τιμή  $d$ . Ωστόσο, στατιστικά η προσέγγιση δεν είναι έγκυρη για το συγκεκριμένο dataset. Παραταύτα, το αποτέλεσμα είναι εξαιρετικά ικανοποιητικό και δείχνει καθαρά ότι ο τύπος της διαφορετικότητας αντί να δίνει έμφαση στην διαφορετικότητα μεταξύ των χρηστών, δίνει έμφαση στο πιο συχνά προτεινόμενο αντικείμενο.

Φυσικά, οι πάροχοι αντικειμένων δε θα ήταν ικανοποιημένοι εάν το πιο συχνά προτεινόμενο αντικείμενο τους προτεινόταν σε περισσότερους χρήστες σε βάρος των λιγότερο προτεινόμενων αντικειμένων, ακόμα και αν η τιμή της μέσης διαφοροποίησης αυξανόταν για αυτούς. Αυτό διότι ένας από τους βασικούς στόχους της διαφοροποίησης είναι η ανακάλυψη περισσότερων αντικειμένων από περισσότερους χρήστες.

Επομένως, ο ορισμός της διαφοροποίησης χρηστών για ένα αντικείμενο και κατά συνέπεια ο ορισμός της διαφοροποίησης για τους παρόχους πρέπει να επανεξεταστεί, ενώ παράλληλα νέοι αλγόριθμοι για την αποδοτική λύση του προβλήματος θα πρέπει να σχεδιαστούν.

### 9.5.2 Νέος ορισμός για τη διαφοροποίηση και νέα ευριστική προσέγγιση

Για τον εκ νέου ορισμό της διαφοροποίησης είναι αναγκαίο να γίνουν αντιληπτά τα προβλήματα του αρχικού ορισμού. Σε αυτό το πλαίσιο, θα αναλύσουμε από την αρχή τον προτεινόμενο ορισμό από τη βιβλιογραφία [7]:

Πρωτίστως, η διαφοροποίηση χρηστών για ένα αντικείμενο ορίζεται ως το άθροισμα των ανομοιοτήτων μεταξύ των χρηστών στους οποίους προτείνεται ένα αντικείμενο, όπως φαίνεται παρακάτω:

$$Div(i) = \sum_{u \in U} \sum_{v \in U, v \neq u} d_{uv} * x_{iu} * x_{iv} , \quad (14)$$

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Ενώ η διαφοροποίηση για έναν πάροχο αντικειμένων ορίζεται ως το άθροισμα των επιμέρους διαφοροποιήσεων των αντικειμένων που ανήκουν στο συγκεκριμένο πάροχο.

Ακολούθως, η διαφοροποίηση παρόχου κανονικοποιείται διαιρώντας με τον αριθμό χρηστών στους οποίους προτείνονται τα αντικείμενα του συγκεκριμένου παρόχου. Κατά συνέπεια, η διαφοροποίηση παρόχου ανά αντικείμενο και ζεύγος χρηστών ορίζεται στον τύπο (4) και απλοποιείται (χρησιμοποιώντας τον περιορισμό της κάλυψης) στον τύπο (5).

Διαισθητικά μπορεί να οριστεί ως εξής:

$$\overline{Div}(c) = \frac{2}{|Ic|} \frac{\text{sum of all diversities for items in the specific category}}{\text{all pairs of users to whom items of class } c \text{ are assigned}}, \quad (15)$$

Παρότι ο ορισμός κανονικοποιεί την διαφοροποίηση παρόχου με βάση τα ζεύγη χρηστών στους οποίους προτείνονται τα αντικείμενα του, αποτυγχάνει να λάβει υπόψη τη δυναμική των πιο συχνά προτεινόμενων αντικειμένων. Συνεπώς, εάν ένας πάροχος διαθέτει  $x$  αντικείμενα και προτείνονται σε  $y$  χρήστες, η διαφοροποίηση μπορεί να αυξηθεί σημαντικά αν σε όλους τους χρήστες προτεινόταν το πιο συχνά προτεινόμενο αντικείμενο. Αυτή η συμπεριφορά δεν είναι επιθυμητή για το σύστημα συστάσεων.

Για το λόγο αυτό θα ήταν καλύτερο ο ορισμός της διαφοροποίησης παρόχου ανά αντικείμενο και ζεύγος χρηστών είτε να κανονικοποιούσε τις διαφοροποιήσεις ανά αντικείμενο είτε να οριζόταν μια νέα (υπερ)διαφοροποίηση για όλο το σύνολο αντικειμένων του κάθε παρόχου. Στην πρώτη περίπτωση ο ορισμός για την διαφοροποίηση παρόχου γίνεται:

$$\overline{Div}'(c) = \frac{2}{|Ic|} \sum_{i \in Ic} \frac{\sum_{u \in U} \sum_{v \in V: u \neq v} duv * xiu * xiv}{(\sum_{u \in U} xiu) * (\sum_{u \in U} xiu - 1)}, \quad (16)$$

Φυσικά, χρησιμοποιώντας τον παραπάνω ορισμό για τη διαφοροποίηση, το πρόβλημα εξακολουθεί να ανήκει στην κατηγορία NP-Hard και κατά συνέπεια δε λύνεται αποδοτικά σε πραγματικό χρόνο. Επιπλέον, ο ευριστικός αλγόριθμος που ορίστηκε σε προηγούμενη ενότητα δεν μπορεί να εφαρμοστεί πλέον.

Στη δεύτερη περίπτωση, η διαφοροποίηση παρόχου, αντί για το άθροισμα των διαφοροποιήσεων των αντικειμένων του, μπορεί να υπολογιστεί για κάθε πάροχο θεωρώντας ένα υπεραντικείμενο για κάθε πάροχο, το οποίο περιλαμβάνει όλα τα μεμονωμένα αντικείμενα του συγκεκριμένου παρόχου. Στη περίπτωση αυτή ένας νέος πίνακας  $Y(c \in C, u \in U)$  θα πρέπει να δημιουργηθεί. Ο πίνακας αυτός δείχνει τη δυαδική πολιτική συστάσεων για τα αντικείμενα του χρήστη (αν  $y(c,u)=0$  τότε κανένα αντικείμενο του παρόχου  $c$  δεν έχει προταθεί στο χρήστη  $u$ , ενώ αν  $y(c,u)=1$ , τότε ένα ή παραπάνω αντικείμενα του συγκεκριμένου παρόχου έχουν προταθεί στον χρήστη  $u$ ).

Έτσι η νέα διαφοροποίηση παρόχου μπορεί να οριστεί ως εξής:

$$\overline{Div}''(c) = \frac{2}{|Ic|} \frac{\sum_{u \in U} \sum_{v \in V: u \neq v} duv * ycu * ycv}{(\sum_{u \in U} ycu) * (\sum_{u \in U} ycu - 1)}, \quad (17)$$

Στον προηγούμενο ορισμό, ο παρονομαστής μπορεί να απλοποιηθεί όπως ακριβώς και στην εργασία [7] ως εξής:

$$\overline{Div}''(c) = \frac{2}{|Ic|} \frac{\sum_{u \in U} \sum_{v \in V: u \neq v} duv * ycu * ycv}{Kc|Ic| * (Kc|Ic| - 1)}, \quad (18)$$

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Ο τελευταίος ορισμός είναι πολύ απλούστερος (παρότι ακόμα παρουσιάζει τετραγωνικούς περιορισμούς στο πρόβλημα, που παραμένει NP-Hard), καθώς ο αριθμός των παρόχων είναι πολύ μικρότερος από τον αριθμό των αντικειμένων. Ωστόσο, δεν είναι κατάλληλος για μικρά συστήματα συστάσεων για τα οποία οι χρήστες έχουν βαθμολογήσει μεγάλο αριθμό αντικειμένων. Αντίθετα, για συστήματα συστάσεων μεγάλης κλίμακας και μεγάλων διαστάσεων αλλά αραιούς πίνακες βαθμολογιών, ο ορισμός αυτός μπορεί να αποδειχθεί πολύ χρήσιμος.

Όσον αφορά τη λύση του προβλήματος με τον νέο ορισμό της διαφοροποίησης και στις δύο περιπτώσεις μια ευριστική προσέγγιση μπορεί να λύσει το πρόβλημα. Για παράδειγμα, αυτό μπορεί να γίνει, όσον αφορά τον πρώτο ορισμό, με την πρόταση του αντικειμένου με τη μεγαλύτερη βαθμολογία σε χρήστες που έχουν τη μεγαλύτερη απόσταση από το κεντροειδές (centroid) των υπόλοιπων χρηστών στους οποίους προτάθηκε το αντικείμενο, ανταλλάσσοντας το αντικείμενο με τη χαμηλότερη βαθμολογία που τους έχει προταθεί, με την προϋπόθεση ότι εξακολουθούν να ικανοποιούνται οι σχετικοί περιορισμοί κάλυψης και διαφοροποίησης για τους παρόχους των οποίων τα αντικείμενα ανταλλάχθηκαν. Παρόμοια μεθοδολογία μπορεί να εφαρμοστεί και για τον δεύτερο ορισμό, με την πρόταση των αντικειμένων με τις καλύτερες βαθμολογίες σε χρήστες που έχουν τη μεγαλύτερη απόσταση από το κεντροειδές των χρηστών στους οποίους προτείνονται τα αντικείμενα του συγκεκριμένου παρόχου. Επίσης, σε περίπτωση που ο αριθμός των χρηστών είναι πολύ μεγάλος, μπορούν να χρησιμοποιηθούν και τεχνικές συσταδοποίησης (clustering), για την ομαδοποίηση των χρηστών με βάση την ομοιότητα που παρουσιάζουν μεταξύ τους, ενώ οι εναλλαγές αντικειμένων, θα πραγματοποιούνται μόνο για χρήστες από την ομάδα με τη μεγαλύτερη απόσταση από το κεντροειδές των χρηστών που τους έχουν προταθεί αντικείμενα του συγκεκριμένου παρόχου αντικειμένων.

### 9.5.3 Απαντήσεις στα ερευνητικά ερωτήματα που τέθηκαν

Έχοντας παρουσιάσει το θεωρητικό υπόβαθρο, τη σχετική βιβλιογραφία, τη μεθοδολογία και τα αποτελέσματα και αναγνωρίζοντας τους περιορισμούς της προτεινόμενης μεθοδολογίας, στην ενότητα αυτή, επιχειρούμε να απαντήσουμε στα ερευνητικά ερωτήματα που τέθηκαν στην αρχή του κεφαλαίου.

1. Πώς μπορεί να επιτευχθεί δικαιοσύνη μεταξύ των παρόχων αντικειμένων σε ένα σύστημα συστάσεων όσον αφορά την κάλυψη των αντικειμένων και τη διαφοροποίηση των χρηστών ανά αντικείμενο;

Η βιβλιογραφία παρέχει μια προτυποποίηση του προβλήματος μαζί με μια ενδεικτική ευριστική λύση. Συγκεκριμένα μοντελοποιεί το πρόβλημα ως ένα πρόβλημα μαθηματικού προγραμματισμού με τετραγωνικούς περιορισμούς, που αποτελεί πρόβλημα της κατηγορίας NP-Hard, που σημαίνει ότι δεν επιδέχεται αναλυτικής λύσης σε αποδεκτό χρόνο. Κατά συνέπεια, στο ίδιο άρθρο, προτείνεται μια ευριστική λύση για την επίλυση των περιορισμών διαφοροποίησης. Ωστόσο, εφόσον η λύση αυτή δεν περιγράφεται επαρκώς, στο παρόν κεφάλαιο προτείνεται μια νέα ευριστική λύση για αυτό το σκοπό.

2. Είναι ικανοποιητικά τα αποτελέσματα των προτεινόμενων μεθοδολογιών για το συγκεκριμένο πρόβλημα;

Τα αποτελέσματα της λύσης που προτείνεται σε αυτό το κεφάλαιο, είναι ικανοποιητικά για την επίτευξη επαρκούς κάλυψης και διαφοροποίησης χρηστών για κάθε πάροχο αντικειμένων.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Μάλιστα, τα αποτελέσματα είναι πολύ κοντά και στη θεωρητικά βέλτιστη λύση που προκύπτει από τη λύση του βασικού προβλήματος χωρίς τους περιορισμούς.

3. Ποιοι είναι οι περιορισμοί των προτεινόμενων λύσεων;

Υπάρχουν κάποιοι περιορισμοί στην προτεινόμενη μεθοδολογία οι οποίοι προκύπτουν από τη μοντελοποίηση του προβλήματος. Συγκεκριμένα, ο ορισμός της διαφοροποίησης αδυνατεί να λάβει σοβαρά υπόψη την ανομοιότητα μεταξύ των χρηστών στους οποίους προτείνεται ένα αντικείμενο. Αντίθετα, παρατηρείται ότι η πρόταση σε νέους χρήστες του πιο συχνά προτεινόμενου αντικειμένου είναι σημαντικότερος παράγοντας από τη διαφοροτικότητα των χρηστών που προτείνεται ένα αντικείμενο και το γεγονός αυτό χρησιμοποιείται για τη σχεδίαση της προτεινόμενης ευριστικής λύσης. Για το λόγο αυτό το πρόβλημα και συγκεκριμένα ο ορισμός της διαφοροποίησης πρέπει να οριστεί εκ νέου.

Επιπροσθέτως, η προτεινόμενη μεθοδολογία δε λαμβάνει υπόψη τη θέση ενός αντικειμένου στη λίστα συστάσεων ενός χρήστη, πράγμα το οποίο είναι πολύ σημαντικό στα σύγχρονα συστήματα συστάσεων.

4. Πώς οι συγκεκριμένοι περιορισμοί μπορούν να αρθούν;

Για την αντιμετώπιση του προβλήματος με τον ορισμό της διαφοροποίησης, στο παρόν κεφάλαιο προτείνονται δύο νέοι ορισμοί για τη διαφοροποίηση των παρόχων σε χρήστες και συζητούνται επιγραμματικά τα πλεονεκτήματα και τα μειονεκτήματά τους. Επίσης, προτείνονται ευριστικές μέθοδοι για την επίλυση των εκ νέου ορισμένων περιορισμών για διαφοροποίηση.

## 9.6 Συμπεράσματα και μελλοντικές επεκτάσεις

Η συνεισφορά του παρόντος κεφαλαίου είναι πολλαπλή. Συγκεκριμένα, η μεθοδολογία που προτείνεται στη βιβλιογραφία εφαρμόζεται για την επίλυση του προβλήματος της δικαιοσύνης μεταξύ των παρόχων για κάλυψη και διαφοροποίηση χρηστών σε πολυμετοχικά συστήματα συστάσεων. Επίσης, μιας και το πρόβλημα ανήκει στην κατηγορία NP-Hard και δεν επιδέχεται αναλυτικής λύσης σε αποδεκτό χρόνο, και καθώς η προτεινόμενη από τη βιβλιογραφία ευριστική λύση είναι πολύ γενική και δεν εξηγείται επαρκώς, στο παρόν κεφάλαιο αναπτύχθηκε νέος ευριστικός αλγόριθμος για την επίλυση των περιορισμών διαφοροποίησης που εξηγείται με λεπτομέρεια. Ακόμη, στο κεφάλαιο αυτό παρουσιάζεται η υπολογιστική πολυπλοκότητα του προτεινόμενου αλγορίθμου, παράλληλα με τους περιορισμούς που παρουσιάζει. Επιπλέον, παρουσιάζονται εκτενώς τα αποτελέσματα της προτεινόμενης λύσης και είναι πολύ κοντά στα θεωρητικά βέλτιστα αποτελέσματα για το συγκεκριμένο dataset.

Στη συνέχεια, στο παρόν κεφάλαιο αξιολογούνται και συζητούνται σε βάθος τα αποτελέσματα, και αναγνωρίζονται κάποιες αδυναμίες στον ορισμό του προβλήματος. Η σημαντικότερη αδυναμία έγκειται στον ορισμό της διαφοροποίησης και στη σχέση της με την ανομοιότητα μεταξύ των χρηστών. Προς αυτή την κατεύθυνση, παρέχεται μια ανάλυση για να ανακαλύψουμε τι επηρεάζει την διαφοροποίηση και ανακαλύπτουμε ότι ο σημαντικότερος παράγοντας είναι η συχνότητα που προτείνονται τα πιο συχνά προτεινόμενα αντικείμενα. Για το λόγο αυτό, προτείνονται δύο νέοι ορισμοί για την διαφοροποίηση καθώς επίσης και μια νέα ευριστική μεθοδολογία για την εύρεση της λύσης.

Όσον αφορά τις μελλοντικές επεκτάσεις του παρόντος κεφαλαίου και της προτεινόμενης μεθοδολογίας, προτεραιότητα αποτελεί η ένταξη της σειράς εμφάνισης ενός αντικειμένου στη λίστα προτάσεων ενός



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

χρήστη στο μαθηματικό μοντέλο του συστήματος συστάσεων, καθώς είναι πολύ σημαντική η θέση ενός αντικειμένου στη λίστα συστάσεων ενός χρήστη. Αυτό θα γίνει χρησιμοποιώντας και άλλες μετρικές αξιολόγησης όπως είναι το Normalised Discounted Cumulative Gain (NDCG), που μετράει την ποιότητα της θέσης των προτεινόμενων αντικειμένων για κάθε χρήστη.

Επιπλέον, η αντιπροσώπευση των μειονοτήτων θα εξεταστεί σε σχέση με την κάλυψη και τη διαφοροποίηση, και θα συγκριθεί με άλλες αντίστοιχες εργασίες (π.χ. την [24]). Επίσης, διαφορετικές μετρικές θα εξεταστούν για να μοντελοποιηθούν ως περιορισμοί του προβλήματος βελτιστοποίησης.

Τέλος, οι προτεινόμενες μέθοδοι θα ήταν χρήσιμο εφαρμοστούν σε πραγματικά δεδομένα και διαφορετικά datasets, όπως επίσης και σε συστήματα συστάσεων του πραγματικού κόσμου. Επίσης, πέρα από την αξιολόγηση της απόδοσης του συστήματος με βάση τις προτεινόμενες μετρικές, θα ήταν χρήσιμο το σύστημα να αξιολογηθεί από πραγματικούς χρήστες, όσον αφορά την κάλυψη και τη διαφοροποίηση όσο και την ικανοποίηση τους από το σύστημα, συμπεριλαμβανομένων απλών χρηστών όσο και παρόχων αντικειμένων.

## 9.7 Αναφορές

- [1] “Efficient and Fair Item Coverage in Recommender Systems | IEEE Conference Publication | IEEE Xplore.” <https://ieeexplore.ieee.org/abstract/document/8511997> (accessed May 03, 2023).
- [2] M. Kunaver and T. Požrl, “Diversity in recommender systems – A survey,” *Knowl Based Syst*, vol. 123, pp. 154–162, May 2017, doi: 10.1016/J.KNOSYS.2017.02.009.
- [3] S. M. McNee, J. Riedl, and J. A. Konstan, “Being accurate is not enough: How accuracy metrics have hurt recommender systems,” *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1097–1101, 2006, doi: 10.1145/1125451.1125659.
- [4] K. G. Saranya and G. Sudha Sadasivam, “Personalized News Article Recommendation with Novelty Using Collaborative Filtering Based Rough Set Theory,” *Mobile Networks and Applications*, vol. 22, no. 4, pp. 719–729, Aug. 2017, doi: 10.1007/S11036-017-0842-9/FIGURES/9.
- [5] H. Abdollahpouri, R. Burke, and B. Mobasher, “Recommender systems as multistakeholder environments,” *UMAP 2017 - Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, pp. 347–348, Jul. 2017, doi: 10.1145/3079628.3079657.
- [6] R. Burke, N. Sonboli, and A. Ordonez-Gauger, “Balanced Neighborhoods for Multi-sided Fairness in Recommendation,” *Proceedings of Machine Learning Research*, vol. 81. PMLR, pp. 202–214, Jan. 21, 2018. Accessed: May 03, 2023. [Online]. Available: <https://proceedings.mlr.press/v81/burke18a.html>
- [7] I. Koutsopoulos and M. Halkidi, “Optimization of Multi-stakeholder Recommender Systems for Diversity and Coverage,” *IFIP Adv Inf Commun Technol*, vol. 627, pp. 703–714, 2021, doi: 10.1007/978-3-030-79150-6\_55/FIGURES/2.

- [8] R. Burke, H. Abdollahpouri, E. C. Malthouse, K. P. Thai, and Y. Zhang, “Recommendation in multistakeholder environments,” *RecSys 2019 - 13th ACM Conference on Recommender Systems*, pp. 566–567, Sep. 2019, doi: 10.1145/3298689.3346973.
- [9] H. Abdollahpouri and R. Burke, “Multi-stakeholder Recommendation and its Connection to Multi-sided Fairness,” *CEUR Workshop Proc*, vol. 2440, Jul. 2019, Accessed: May 03, 2023. [Online]. Available: <https://arxiv.org/abs/1907.13158v1>
- [10] S. Milano, M. Taddeo, and L. Floridi, “Ethical aspects of multi-stakeholder recommendation systems,” <https://doi.org/10.1080/01972243.2020.1832636>, vol. 37, no. 1, pp. 35–45, 2020, doi: 10.1080/01972243.2020.1832636.
- [11] H. Abdollahpouri et al., “Multistakeholder recommendation: Survey and research directions,” *User Model User-adapt Interact*, vol. 30, no. 1, pp. 127–158, Mar. 2020, doi: 10.1007/S11257-019-09256-1/TABLES/2.
- [12] Y. Zheng, “Multi-Stakeholder Recommendations: Case Studies, Methods and Challenges,” *Proceedings of the 13th ACM Conference on Recommender Systems*, pp. 578–579, Jul. 2019, doi: 10.1145/3298689.
- [13] Ö. Sürer, R. Burke, and E. C. Malthouse, “Multistakeholder recommendation with provider constraints,” *RecSys 2018 - 12th ACM Conference on Recommender Systems*, vol. 9, pp. 54–62, Sep. 2018, doi: 10.1145/3240323.3240350.
- [14] E. C. Malthouse, K. A. Vakeel, Y. K. Hessary, R. Burke, and M. Fuduric, “A Multistakeholder Recommender Systems Algorithm for Allocating Sponsored Recommendations,” in *Proceedings of the CEUR Workshop, Lviv, Ukraine, 2019*.
- [15] M. Ge, C. Delgado-Battenfeld, and D. Jannach, “Beyond accuracy: Evaluating recommender systems by coverage and serendipity,” *RecSys’10 - Proceedings of the 4th ACM Conference on Recommender Systems*, pp. 257–260, 2010, doi: 10.1145/1864708.1864761.
- [16] M. Kaminskis and D. Bridge, “Diversity, Serendipity, Novelty, and Coverage,” *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 7, no. 1, Dec. 2016, doi: 10.1145/2926720.
- [17] M. Rahman and J. C. Oh, “Graph bandit for diverse user coverage in online recommendation,” *Applied Intelligence 2017 48:8*, vol. 48, no. 8, pp. 1979–1995, Jun. 2017, doi: 10.1007/S10489-017-0977-1.
- [18] M. Hammar, R. Karlsson, and B. J. Nilsson, “Using maximum coverage to optimize recommendation systems in e-commerce,” *Proceedings of the 7th ACM conference on Recommender systems*, pp. 265–272, Oct. 2013, doi: 10.1145/2507157.2507169.
- [19] C. N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen, “Improving recommendation lists through topic diversification | Proceedings of the 14th international conference on World Wide Web,” in *WWW ’05: Proceedings of the 14th international conference on World Wide Web, 2005*, pp. 22–32. Accessed: May 03, 2023. [Online]. Available: <https://dl.acm.org/doi/10.1145/1060745.1060754>
- [20] S. Vargas, P. Castells, and D. Vallet, “Intent-oriented diversity in recommender systems,” *SIGIR’11 - Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1211–1212, 2011, doi: 10.1145/2009916.2010124.

- [21] J. Ben Schafer, D. Frankowski, J. Herlocker, and S. Sen, “Collaborative filtering recommender systems,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4321 LNCS, pp. 291–324, 2007, doi: 10.1007/978-3-540-72079-9\_9/COVER.
- [22] J. P. Kelly and D. Bridge, “Enhancing the diversity of conversational collaborative recommendations: A comparison,” *Artif Intell Rev*, vol. 25, no. 1–2, pp. 79–95, Apr. 2006, doi: 10.1007/S10462-007-9023-8/METRICS.
- [23] C. Yu, L. V. S. Lakshmanan, and S. Amer-Yahia, “Recommendation diversification using explanations,” *Proc Int Conf Data Eng*, pp. 1299–1302, 2009, doi: 10.1109/ICDE.2009.225.
- [24] L. Boratto, G. Fenu, and M. Marras, “Interplay between upsampling and regularization for provider fairness in recommender systems,” *User Model User-adapt Interact*, vol. 31, no. 3, pp. 421–455, Jul. 2021, doi: 10.1007/S11257-021-09294-8/TABLES/4.
- [25] N. Sonboli, F. Eskandarian, R. Burke, W. Liu, and B. Mobasher, “Opportunistic Multi-aspect Fairness through Personalized Re-ranking,” *UMAP 2020 - Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, pp. 239–247, Jul. 2020, doi: 10.1145/3340631.3394846.
- [26] E. Gómez, C. Shui Zhang, L. Boratto, M. Salamó, and M. Marras, “The Winner Takes it All: Geographic Imbalance and Provider (Un)fairness in Educational Recommender Systems,” *SIGIR 2021 - Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1808–1812, Jul. 2021, doi: 10.1145/3404835.3463235.
- [27] A. Beutel et al., “Fairness in recommendation ranking through pairwise comparisons,” *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 2212–2220, Jul. 2019, doi: 10.1145/3292500.3330745.
- [28] X. Lin, M. Zhang, Y. Zhang, Z. Gu, Y. Liu, and S. Ma, “Fairness-aware group recommendation with pareto-efficiency,” *RecSys 2017 - Proceedings of the 11th ACM Conference on Recommender Systems*, pp. 107–115, Aug. 2017, doi: 10.1145/3109859.3109887.
- [29] D. Serbos, S. Qi, N. Mamoulis, E. Pitoura, and P. Tsaparas, “Fairness in package-to-group recommendations,” *26th International World Wide Web Conference, WWW 2017*, pp. 371–379, 2017, doi: 10.1145/3038912.3052612.
- [30] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, “Item-based collaborative filtering recommendation algorithms,” *Proceedings of the 10th International Conference on World Wide Web, WWW 2001*, pp. 285–295, Apr. 2001, doi: 10.1145/371920.372071.
- [31] “Recommender systems optimisation for provider fairness” [https://github.com/vkarakolis-epu/recsys\\_provider\\_fairness\\_optimization](https://github.com/vkarakolis-epu/recsys_provider_fairness_optimization) (accessed May 03, 2023).
- [32] “Index of /datasets/movielens.” <http://files.grouplens.org/datasets/movielens/> (accessed May 03, 2023).
- [33] “Welcome to CVXPY 1.3 — CVXPY 1.3 documentation.” <https://www.cvxpy.org/> (accessed May 03, 2023).

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## ΚΕΦΑΛΑΙΟ 10: ΗΘΙΚΕΣ ΕΠΙΠΤΩΣΕΙΣ ΣΤΑ ΣΥΣΤΗΜΑΤΑ ΣΥΣΤΑΣΕΩΝ

### 10.1 Εισαγωγή

Στις μέρες μας τα συστήματα συστάσεων παίζουν σημαντικό ρόλο σε πολλές πτυχές της ανθρώπινης ζωής, από τη διασκέδαση μέχρι τα ψώνια και την κοινωνικοποίηση. Αυτό συμβαίνει διότι δρουν ως προσωποποιημένο φίλτρο για κάθε χρήστη, προστατεύοντας από την υπερφόρτωση πληροφορίας που παρέχει το σημερινό διαδίκτυο. Η υπερφόρτωση πληροφορίας είναι τόσο εκτεταμένη στις μέρες μας που χωρίς τα συστήματα συστάσεων να φιλτράρουν την πληροφορία που φτάνει στους τελικούς χρήστες, η πλοήγηση στο διαδίκτυο θα ήταν αδύνατη [1]. Έτσι, τα συστήματα συστάσεων πρέπει να έχουν ως κύριο σκοπό την παροχή προτάσεων αντικειμένων που είναι επιθυμητά από τους χρήστες. Στη βιβλιογραφία, στις περισσότερες περιπτώσεις, αυτό επιτυγχάνεται με τη σύσταση αντικειμένων που προβλέπεται ότι θα ικανοποιήσουν τον χρήστη στο μέγιστο βαθμό. Ωστόσο, στην πραγματικότητα, τα σύγχρονα συστήματα συστάσεων οδηγούνται από το ψηφιακό μάρκετινγκ και δίνουν έμφαση σε εμπορικούς στόχους αντί για τον τελικό χρήστη [2].

Σε αυτό το πλαίσιο, τα τελευταία χρόνια η στόχευση έχει μετατοπιστεί από την ικανοποίηση του χρήστη στη μεγιστοποίηση του κέρδους. Έτσι, ο στόχος των μοντέρνων συστημάτων συστάσεων είναι να κρατήσουν τους χρήστες απορροφημένους και προσκολλημένους στην πλατφόρμα και όχι να αποκλείσουν πληροφορίες για τις οποίες οι χρήστες δεν ενδιαφέρονται. Ένα παράδειγμα της μετατόπισης αυτής αποτελεί το γεγονός ότι ο αλγόριθμος που κέρδισε το διαγωνισμό του Netflix δε χρησιμοποιήθηκε ποτέ. Στο διαγωνισμό αυτό το Netflix προσέφερε το βραβείο του ενός εκατομμυρίου δολαρίων σε περίπτωση που επιτύγχανε 10% αύξηση στην ακρίβεια πρόβλεψης. Ο λόγος που δε χρησιμοποιήθηκε ποτέ ο συγκεκριμένος αλγόριθμος είναι ότι το Netflix άλλαξε από εταιρεία ενοικίασης DVD σε εταιρεία ροής περιεχομένου στο διαδίκτυο (internet content streaming company). Επομένως, ο στόχος δεν ήταν πλέον να μαντέψει τις προτιμήσεις των χρηστών αλλά να τους κάνει να συνεχίσουν να παρακολουθούν το περιεχόμενο του Netflix και να τους κρατήσει εγγεγραμμένους και προσκολλημένους στην πλατφόρμα [3].

Φυσικά, αυτή η μετατόπιση στους στόχους των συστημάτων συστάσεων μπορεί να έχει αρνητικές συνέπειες για τους χρήστες και την κοινωνία γενικότερα. Για παράδειγμα, η Frances Haugen, μια πρώην εργαζόμενη στο Facebook, κατηγόρησε την εταιρεία ότι δίνει προτεραιότητα στα κέρδη έναντι των ανθρώπων [4]. Συγκεκριμένα, αποκάλυψε, μεταξύ άλλων, ότι η χρήση του Instagram έχει αρνητικό αντίκτυπο στην ψυχική υγεία των έφηβων κοριτσιών, και ότι το Facebook είναι υπεύθυνο για την αυξανόμενη βία στις αναπτυσσόμενες χώρες, ενώ ο αλγόριθμος συστάσεων του Facebook φέρεται να δίνει προτεραιότητα στο διχαστικό περιεχόμενο έναντι του ουδέτερου περιεχομένου.

Ένα άλλο παράδειγμα του αρνητικού αντίκτυπου που έχει το Facebook στους χρήστες του παρουσιάζεται στην εργασία [5]. Συγκεκριμένα, αναφέρεται ότι ένας ερευνητής ανακάλυψε ότι πάνω από το 33% των ομάδων του Facebook για τη γερμανική πολιτική παρουσίασαν αυξανόμενο όγκο σε εξτρεμιστικό περιεχόμενο, ενώ το 64% των συμμετεχόντων σε εξτρεμιστικές ομάδες στο ίδιο μέσο, μπήκαν σε αυτές τις ομάδες λόγω του αλγορίθμου συστάσεων του Facebook.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Επιπροσθέτως, μια έρευνα για τα βιβλία της Amazon που σχετίζονται με τα εμβόλια έδειξε ότι ο αριθμός των βιβλίων που ήταν επιφυλακτικά έναντι στον εμβολιασμό ήταν πάνω από διπλάσιος σε σχέση με τα βιβλία που τον υποστήριζαν στο ηλεκτρονικό κατάστημα της Amazon, ενώ το σύστημα συστάσεων πρότεινε βιβλία τα οποία ενίσχυαν τις πεποιθήσεις του χρήστη. Οπότε, αν ένας χρήστης έψαχνε ένα αντιεμβολιαστικό βιβλίο, το σύστημα θα του πρότεινε και άλλα αντιεμβολιαστικά βιβλία [6]. Επιπλέον, στην εργασία [7], οι συγγραφείς δηλώνουν ότι ο αλγόριθμος συστάσεων της Amazon προωθεί θεωρίες συνωμοσίας. Έτσι, ένας χρήστης που κοιτάει για βιβλία που υποστηρίζουν μια συγκεκριμένη θεωρία συνωμοσίας θα πάρει προτάσεις βιβλίων που υποστηρίζουν και άλλες τέτοιες θεωρίες, ενώ παρόμοια είναι η κατάσταση για βιβλία που υποστηρίζουν ρατσιστικές ή εξτρεμιστικές ιδεολογίες.

Σε αυτό το πλαίσιο, στο παρόν κεφάλαιο, διερευνάται ο βαθμός στον οποίο ένα σύστημα συστάσεων μπορεί να θεωρηθεί ηθικά προβληματικό και πώς τα προβλήματα ηθικής που δημιουργούνται μπορούν να περιοριστούν.

## 10.2 Θεωρητικό υπόβαθρο και σχετικές εργασίες

Όπως έχει ήδη αναφερθεί, στα παραδοσιακά συστήματα συστάσεων και στην πλειοψηφία της βιβλιογραφίας, η σημαντικότερη μετρική για την αξιολόγηση της απόδοσης ενός συστήματος συστάσεων είναι η ακρίβεια των προβλεπόμενων βαθμολογιών για τα διαθέσιμα αντικείμενα από όλους τους χρήστες. Η συγκεκριμένη προσέγγιση επικεντρώνεται κυρίως στην ικανοποίηση των χρηστών. Από την άλλη πλευρά, τα μοντέρνα συστήματα συστάσεων δίνουν περισσότερη έμφαση στην αφοσίωση των χρηστών (user engagement) και στη μεγιστοποίηση του κέρδους [3]. Ωστόσο, συνήθως οι αλγόριθμοι που αποφασίζουν τις πολιτικές συστάσεων δεν αποκαλύπτονται και κατά συνέπεια οι χρήστες αντιλαμβάνονται τα συστήματα συστάσεων ως «μαύρα κουτιά» [8]. Συνεπώς, είναι κοινά αποδεκτό ότι περισσότερες μετρικές αξιολόγησης πρέπει να λαμβάνονται υπόψη για την αξιολόγηση των συστημάτων συστάσεων [9], και θα πρέπει να επικεντρώνονται όχι μόνο στους τελικούς χρήστες αλλά σε όλους τους ενδιαφερόμενους [10].

Πιο συγκεκριμένα, οι βασικοί ενδιαφερόμενοι ενός συστήματος συστάσεων είναι οι χρήστες, οι πάροχοι αντικειμένων, και οι επιχειρηματίες και διαχειριστές (operators) του συστήματος. Σε αυτό το πλαίσιο, έχει προταθεί στη βιβλιογραφία ένα νέο υπόδειγμα για τη μελέτη και αξιολόγηση των συστημάτων συστάσεων που ονομάζεται Πολυμετοχικά Συστήματα Συστάσεων (Multi-Stakeholder Recommender System paradigm). Το συγκεκριμένο υπόδειγμα δίνει έμφαση σε όλους τους μετόχους ενός συστήματος συστάσεων και τα τελευταία χρόνια κερδίζει σταθερά σε δημοφιλία [11].

Από την άλλη πλευρά, οι σημαντικότερες μετρικές αξιολόγησης των συστημάτων συστάσεων που πηγαινούν πέρα από την ακρίβεια των προβλεπόμενων βαθμολογιών είναι η κάλυψη, η διαφοροποίηση, η ευχάριστη έκπληξη και η δικαιοσύνη, οι οποίες έχουν παρουσιαστεί σε προηγούμενα κεφάλαια της παρούσας διδακτορικής διατριβής. Για λόγους πληρότητας και πολύ συνοπτικά, η κάλυψη (αντικειμένου) είναι ο αριθμός χρηστών στους οποίους ένα αντικείμενο προτείνεται [12]. Η διαφοροποίηση ορίζεται ως η μέση απόσταση όλων των ζευγών αντικειμένων ή χρηστών σε μια λίστα προτάσεων [13]. Η ευχάριστη έκπληξη αξιολογεί σε ποιο βαθμό προτείνονται στο χρήστη μη αναμενόμενα αλλά ευχάριστα αντικείμενα [9]. Και τέλος, η δικαιοσύνη έχει πολλαπλούς στόχους με βάση τις ανάγκες των ενδιαφερόμενων για τους οποίους εξετάζεται. Μπορεί να διακριθεί σε δικαιοσύνη

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

πελάτη (customer fairness – C-fairness) που εξετάζει τη δικαιοσύνη στους χρήστες του συστήματος συστάσεων και σε δικαιοσύνη παρόχου (provider fairness – P-fairness) που εξετάζει τη δικαιοσύνη μεταξύ των παρόχων [14].

Παρότι έχει γίνει αρκετή έρευνα για τη βελτίωση της απόδοσης των συστημάτων συστάσεων [15], μόνο λίγες εργασίες εξετάζουν τον αντίκτυπο των συστημάτων συστάσεων στους χρήστες. Το ίδιο ισχύει και για τις ηθικές προκλήσεις που παρουσιάζονται από την ευρεία χρήση των συστημάτων συστάσεων. Συγκεκριμένα, οι περισσότερες έρευνες καταπιάνονται με ένα συγκεκριμένο σύστημα συστάσεων και τον αντίκτυπο που έχει στους χρήστες [4] ή με ένα συγκεκριμένο φαινόμενο όπως είναι οι κάμαρες ηχούς (echo chambers) [16], [17] και η πόλωση [18] που σχετίζεται με τη λειτουργία των συστημάτων συστάσεων. Όσον αφορά τις ηθικές εκφάνσεις των συστημάτων συστάσεων γενικά, η εργασία [19] παρουσιάζει περιπτώσεις στις οποίες πραγματικά συστήματα συστάσεων τροποποιήθηκαν για να υπηρετήσουν διάφορες ανθρώπινες αξίες, και προτείνει μεθόδους που αξιοποιούν τη μεγαλύτερη συμμετοχή από όλους τους ενδιαφερόμενους, ώστε να επιτευχθεί μεγαλύτερη σύμπλευση μεταξύ των συστημάτων συστάσεων και των ανθρώπινων αξιών. Επιπλέον, η εργασία [2] παρουσιάζει μια ανάλυση των ηθικών προκλήσεων που ανακύπτουν από τη λειτουργία των συστημάτων συστάσεων και διακρίνει τα διαφορετικά είδη ηθικού αντίκτυπου. Επίσης, οι ίδιοι συγγραφείς στην εργασία [20] εξετάζουν τις ηθικές εκφάνσεις των συστημάτων συστάσεων υπό το πρίσμα του υποδείγματος των πολυμετοχικών συστημάτων συστάσεων. Σε αυτό το πλαίσιο, καταλήγουν ότι το συγκεκριμένο υπόδειγμα έχει σημαντικά πλεονεκτήματα έναντι της παραδοσιακής εστιασμένης στο χρήστη προσέγγισης.

Κατά τη συγγραφή του παρόντος κεφαλαίου, απουσίαζε από τη βιβλιογραφία ένα πλαίσιο αξιολόγησης για την εκτίμηση της κρισιμότητας των ηθικών διαστάσεων των συστημάτων συστάσεων, όσο και ένα ρυθμιστικό πλαίσιο για την αντιμετώπιση τέτοιων προβλημάτων. Συνεπώς, το παρόν κεφάλαιο ανασκοπεί τις ηθικές εκφάνσεις των συστημάτων συστάσεων όπως παρουσιάζονται στη βιβλιογραφία και προτείνει ένα πλαίσιο αξιολόγησης για την ηθική κρισιμότητα ενός συστήματος συστάσεων. Σε αυτό το πλαίσιο, το πλαίσιο αξιολόγησης που προτείνεται εφαρμόζεται σε μια σειρά από γνωστά συστήματα συστάσεων και τα αποτελέσματα παρουσιάζονται και αναλύονται. Τέλος, προτείνεται ένα ρυθμιστικό πλαίσιο για την βελτίωση των συστημάτων συστάσεων όσον αφορά τα ηθικά προβλήματα που παρουσιάζουν.

### **10.3 Ηθικές προκλήσεις στα συστήματα συστάσεων**

Η απόφαση για το τι θεωρείται ηθικά αποδεκτό στις υπολογιστικές και τεχνολογικές πρακτικές είναι άκρως υποκειμενικό ή ακόμα και φιλοσοφικό θέμα. Το παρόν κεφάλαιο επικεντρώνεται σε περιπτώσεις των οποίων οι ηθικές επιπτώσεις είναι αναμφισβήτητες και κοινά αποδεκτές, λαμβάνοντας υπόψη μόνο τη συμπεριφορά και τα αποτελέσματα ενός συστήματος συστάσεων ώστε να αναγνωριστούν τυχόντα προβλήματα ηθικής. Για το λόγο αυτό, στο παρόν κεφάλαιο προτείνεται ο εξής κανόνας: Ένα σύστημα συστάσεων θεωρείται ως ηθικά προβληματικό όταν [2], [21]:

- A) Βλάπτει την ωφέλεια κάποιων από τους μετόχους του (stakeholders) ή μπορεί να το κάνει στο μέλλον.
- B) Παραβιάζει τα δικαιώματα τους ή μπορεί να το κάνει στο μέλλον.

Σε αυτό το πλαίσιο, αναγνωρίζουμε τις ακόλουθες ηθικές προκλήσεις στα Συστήματα Συστάσεων:

- 1. Ιδιωτικότητα:** Λαμβάνουμε υπόψη ότι οι παράμετροι που χρησιμοποιούνται για τη δημιουργία προσωποποιημένων συστάσεων δημιουργούν ένα προφίλ δεδομένων για κάθε χρήστη [21], ενώ η μέθοδος collaborative filtering καθιστά αναπόφευκτη τη μαζική συλλογή προσωπικών δεδομένων [2]. Η ιδιωτικότητα ενός χρήστη κινδυνεύει όταν τα δεδομένα του συλλέγονται ή μοιράζονται χωρίς την πλήρη συναίνεση του. Είναι γνωστό ότι τα συστήματα συστάσεων συλλέγουν δεδομένα από άλλες πλατφόρμες μέσω της χρήσης cookies και ότι οι όροι χρήσης (terms and conditions) τους οποίους ο χρήστης πολύ συχνά αποδέχεται χωρίς να διαβάσει είναι συχνά ασαφείς και γενικοί, καθιστώντας αδύνατο για ένα χρήστη να έχει πλήρη γνώση σε τι συναίνεσε [21]. Επιπλέον, υπάρχει ο κίνδυνος να διαρρεύσουν και να γίνουν αντικείμενο εκμετάλλευσης τα ιδιωτικά δεδομένα ενός χρήστη [2]. Επίσης, πέρα από τα δεδομένα που ένα σύστημα συστάσεων συλλέγει από έναν χρήστη υπάρχουν και τα δεδομένα που συμπεραίνονται από το σύστημα για το χρήστη μέσα από τα δεδομένα που έχει συλλέξει [2]. Έτσι, ο χρήστης είναι πολύ πιθανόν να μη γνωρίζει για τα συμπεράσματα που έχει δημιουργήσει το σύστημα, και στη χρήση των οποίων μπορεί να μην έδιναν τη συγκατάθεση τους αν γνώριζαν για την ύπαρξη τους. Τέλος, η τεχνική collaborative filtering μπορεί από μόνη της να εγείρει ανησυχίες ιδιωτικότητας, καθώς ένα σύστημα συστάσεων collaborative filtering μπορεί να δημιουργήσει προφίλ χρηστών με τη χρήση άλλων προφίλ χρηστών, χρησιμοποιώντας πολύ λίγη πληροφορία από τον χρήστη.
- 2. Ακατάλληλο περιεχόμενο:** Οι προτάσεις ενός συστήματος συστάσεων μπορεί να είναι αποτελεσματικά στοχευμένες και παράλληλα ακατάλληλες [22], σύμφωνα με τον κανόνα που προτάθηκε προηγουμένως. Αυτό μπορεί να αποτελεί υποκειμενικό θέμα, ωστόσο υπάρχουν αναμφισβήτητες περιπτώσεις ακατάλληλου περιεχομένου όπως είναι οι περιπτώσεις προπαγάνδας, ρατσιστικού περιεχομένου, ή περιεχομένου που προωθεί το μίσος, και σε κάποιες περιπτώσεις ακατάλληλο περιεχόμενο για ανήλικους χρήστες [22].
- 3. Προσωπική ταυτότητα:** Τα συστήματα συστάσεων μπορεί να παραβιάζουν την αυτονομία και την προσωπική ταυτότητα (ή την ατομικότητα) ενός χρήστη με διάφορους τρόπους. Η αίσθηση του εαυτού μας καλλιεργείται μέσα από διάφορες κατηγορίες στις οποίες θεωρούμε ότι ανήκουμε [23]. Εφόσον το προφίλ ενός χρήστη δημιουργείται δυναμικά με τη χρήση τόσο των δεδομένων του όσο και δεδομένων άλλων χρηστών και δε δημιουργείται από τον ίδιο το χρήστη, η αίσθηση ταυτότητας, οι προτιμήσεις και η αυτόνομη κίνηση του χρήστη μπορεί να παραβιάζονται, ενώ οι κατηγορίες μέσα από τις οποίες το σύστημα συστάσεων βλέπει το χρήστη και του προτείνει περιεχόμενο μπορεί να διαφέρουν από αυτά που ο χρήστης πιστεύει για τον εαυτό του [2]. Τέλος, τα εξελιγμένα συστημάτων συστάσεων, των οποίων ο στόχος έχει αλλάξει προς την κατεύθυνση του να κρατάνε το χρήστη προσκολλημένο στη πλατφόρμα [3], μπορεί να απομονώσουν συγκεκριμένες προτιμήσεις του χρήστη και πρακτικά να τον κάνουν να εθιστεί, εκμεταλλευόμενα την προσωπική του ταυτότητα [2],[3].
- 4. Διαφάνεια:** Η λειτουργία ενός συστήματος συστάσεων, που αποτελείται από τον τρόπο που συλλέγει, επεξεργάζεται και χρησιμοποιεί τα δεδομένα των χρηστών για να κάνει προτάσεις, ευρέως γνωστή και ως «ο αλγόριθμος», δεν είναι σχεδόν ποτέ διαθέσιμη στο κοινό. Αντίθετα, κρατιέται κρυφή, ώστε να διασφαλιστεί το ανταγωνιστικό πλεονέκτημα των επιχειρήσεων που επιτυγχάνεται μέσω του αλγορίθμου [22]. Το γεγονός αυτό, με τη σειρά του, καθιστά αδύνατο για τους χρήστες να γνωρίζουν τι συμβαίνει πίσω από τη διεπαφή με το σύστημα και κατά συνέπεια να γνωρίζουν σε τι συναινούν.



5. **Δικαιοσύνη:** Τα συστήματα συστάσεων μπορούν να κατηγορηθούν για την ενίσχυση της προκατάληψης. Είναι γνωστό ότι οι πραγματικές προτιμήσεις των χρηστών μπορεί να διαφέρουν από τις βαθμολογίες που δίνουν σε αντικείμενα διαδικτυακά, μιας και οι βαθμολογίες μπορεί να επηρεάζονται από κοινωνικούς παράγοντες που μπορεί να εμπεριέχουν προκατάληψη [24]. Οι βρόχοι ανατροφοδότησης (feedback loops) μπορούν με τη σειρά τους να αποτελέσουν πηγή προκατάληψης, όταν ένα σύστημα συνεχίζει να δημιουργεί προτάσεις για μια συγκεκριμένη ομάδα χρηστών. Επίσης τα συστήματα συστάσεων μπορούν να προκαλέσουν ανισορροπία στον πληθυσμό, που συμβαίνει όταν τα διαθέσιμα δεδομένα στο σύστημα απεικονίζουν υπάρχουσες κοινωνικές ανισότητες, ενισχύοντας την κοινωνική προκατάληψη έναντι των μειονοτήτων, του φύλου κλπ. [2].
6. **Κοινωνικές επιπτώσεις:** Πρόκειται για μεγάλης κλίμακας αρνητικές συνέπειες της λειτουργίας των συστημάτων συστάσεων. Με την αποτροπή έκθεσης ενός χρήστη σε διαφορετικές απόψεις, οι χρήστες καταλήγουν αποκομμένοι μέσα στις λεγόμενες «φούσκες φίλτρου» (filter bubbles) [21]. Σε μεγαλύτερη κλίμακα πρόκειται για τις «κάμαρες ηχούς» (echo chambers) στις οποίες ομάδες ανθρώπων επικαιροποιούν και ενισχύουν οι ίδιοι τις ίδιες απόψεις για ένα θέμα, απομονωμένοι από κάθε αντίθετη άποψη. Το φαινόμενο αυτό οδηγεί στη διάβρωση του δημόσιου διαλόγου [2]. Επιπλέον, τα συστήματα συστάσεων μπορούν να γίνουν αντικείμενο εκμετάλλευσης ενεργών ομάδων που προωθούν περιεχόμενο, φαινόμενο που οδηγεί στη διάδοση προπαγανδιστικού περιεχομένου μέσω του αλγορίθμου συστάσεων [2]. Τέλος, κυρίως στα μέσα κοινωνικής δικτύωσης, έχουν παρατηρηθεί φαινόμενα μαζικού επηρεασμού της κοινωνίας. Ένα παράδειγμα είναι το σκάνδαλο της Cambridge Analytica το 2018 [21], [25].

## 10.4 Το προτεινόμενο πλαίσιο αξιολόγησης

Με βάση όσα έχουν παρουσιαστεί μέχρι τώρα στο παρόν κεφάλαιο, είναι απαραίτητη η ανάπτυξη ενός πλαισίου για τη μέτρηση της κρισιμότητας ενός συστήματος συστάσεων, έτσι ώστε οι αρμόδιοι φορείς να μπορούν να αποφανθούν πόσο ηθικά προβληματική είναι η λειτουργία ενός συστήματος συστάσεων.

Στην προηγούμενη ενότητα καταλήξαμε σε 6 ηθικές προκλήσεις που συνδέονται με τη χρήση των συστημάτων συστάσεων. Θεωρούμε ότι οι πέντε πρώτες προκλήσεις (ιδιωτικότητα, ακατάλληλο περιεχόμενο, προσωπική ταυτότητα, διαφάνεια και δικαιοσύνη) έχουν την ίδια βαρύτητα, ενώ οι κοινωνικές επιπτώσεις είναι πολύ πιο σημαντικές μιας και επηρεάζουν μεγάλο μέρος της κοινωνίας και μπορούν να επηρεάσουν σημαντικά την ωφέλεια των μετόχων και άλλων ενδιαφερομένων.

Έτσι, προτείνουμε μια κλίμακα μέτρησης της κρισιμότητας ενός συστήματος συστάσεων, ως προς τις ηθικές του επιπτώσεις. Συγκεκριμένα, ορίζουμε τρία διαφορετικά επίπεδα: σύστημα χαμηλής κρισιμότητας, κρίσιμο σύστημα και σύστημα υψηλής κρισιμότητας.

Για την ένταξη ενός συστήματος συστάσεων σε κάποια από τις παραπάνω κατηγορίες, χρησιμοποιούμε τα δύο ακόλουθα κριτήρια:

**Κριτήριο Α:** Αν αναγνωρίζονται κοινωνικές επιπτώσεις στο σύστημα συστάσεων που μελετάται τότε θεωρείται σύστημα υψηλής κρισιμότητας. Αν αυτό δε συμβαίνει προχωράμε με το Κριτήριο Β.

**Κριτήριο Β:** αν κατά την εξέταση ενός συστήματος συστάσεων αναγνωριστούν:

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

- 0-1 προβληματικές εκφάνσεις από τις εναπομείνουσες 5, το σύστημα συστάσεων είναι **σύστημα χαμηλής κρισιμότητας**.
- 2-3 προβληματικές εκφάνσεις από τις εναπομείνουσες 5, το σύστημα συστάσεων χαρακτηρίζεται ως **κρίσιμο σύστημα**.
- 4-5 προβληματικές εκφάνσεις από τις εναπομείνουσες 5, το σύστημα συστάσεων χαρακτηρίζεται ως **σύστημα υψηλής κρισιμότητας**.

## 10.5 Αξιολόγηση δημοφιλών συστημάτων συστάσεων

Στην ενότητα αυτή θα αναλυθούν κάποιες μελέτες περιπτώσεων για κάποια πραγματικά συστήματα συστάσεων και θα αξιολογηθούν ως προς την κρισιμότητα, όπως παρουσιάστηκε στην προηγούμενη ενότητα. Συγκεκριμένα, θα εξεταστούν οι περιπτώσεις συστημάτων συστάσεων του Facebook, της Amazon και του TikTok και θα αξιολογηθεί η κρισιμότητα σε καθένα από αυτά.

### Η περίπτωση του Facebook

Το **Facebook** είναι το δημοφιλέστερο μέσο κοινωνικής δικτύωσης στον κόσμο, και αριθμούσε περισσότερους από 2.6 δισεκατομμύρια ενεργούς χρήστες κατά τη συγγραφή της παρούσας διατριβής [26]. Έχει αποδειχθεί ότι οι χρήστες του Facebook είναι αρκετά πολωμένοι [16], με την έννοια ότι φαινόμενα κάμαρας ηχούς παρατηρούνται πολύ συχνά στο συγκεκριμένο κοινωνικό δίκτυο. Οι χρήστες γενικά τείνουν να προτιμούν πληροφορίες οι οποίες επιβεβαιώνουν τα πιστεύω τους και την άποψη τους. Το σύστημα συστάσεων του Facebook έχει κατηγορηθεί ότι ενισχύει τα φαινόμενα κάμαρας ηχούς [17]. Συγκεκριμένα το σύστημα συστάσεων ανθρώπων, δηλαδή το σύστημα που προτείνει νέες συνδέσεις με άλλους ανθρώπους στο χρήστη, επηρεάζει ευθέως την πληροφορία και τις απόψεις στις οποίες θα εκτεθεί ένας χρήστης. Συνεπώς, άτομα που υποστηρίζουν την ίδια περιοχή στο πολιτικό φάσμα, ή την ίδια θεωρία συνωμοσίας, είναι πολύ πιο πιθανό να έρθουν σε επαφή μέσω του συγκεκριμένου συστήματος συστάσεων από ό,τι δύο άτομα με διαφορετικές απόψεις πάνω σε ένα συγκεκριμένο θέμα [16],[17]. Ενδεικτικά, ένας αναλυτής του Facebook, ανακάλυψε ότι το 2016, περισσότερες από το ένα τρίτο (33%) όλων των ομάδων του Facebook για τη γερμανική πολιτική παρουσίαζαν μια αυξητική τάση για εξτρεμιστικό περιεχόμενο και το 64% όσων συμμετείχαν σε μια εξτρεμιστική ομάδα, το έκαναν λόγω του συστήματος συστάσεων του Facebook [5]. Στις 17 Μαρτίου του 2018, ο Guardian αποκάλυψε μια διαρροή 50 εκατομμυρίων προφίλ στο Facebook και την εκμετάλλευσή τους από την εταιρεία Cambridge Analytica [27]. Ο Christopher Wiley, πρώην εργαζόμενος της συγκεκριμένης εταιρείας δήλωσε ότι εκμεταλλεύτηκαν το Facebook για την εξόρυξη εκατομμυρίων προφίλ χρηστών, και ανέπτυξαν μοντέλα για την εκμετάλλευση όσων γνώριζαν για αυτούς τους χρήστες στοχεύοντας τους εσωτερικούς τους δαίμονες [27]. Πολλά από τα προφίλ που διέρρευσαν ανήκαν σε Αμερικανούς πολίτες και η Cambridge Analytica χρησιμοποίησε τα δεδομένα τους για να αναγνωρίσει χαρακτηριστικά που θα αποκάλυπταν ποιες ήταν οι πολιτικές τους προτιμήσεις. Με βάση τα ευρήματα, η εταιρεία σχεδίασε ισχυρή προπαγάνδα που θα έφτανε στους χρήστες που στόχευε μέσω του συστήματος συστάσεων του Facebook, προσπαθώντας να τους πείσει να ψηφίσουν Donald Trump στις επόμενες εκλογές. Σε μια παρουσίαση για έμπειρα στελέχη το 2018, κάποιοι από τους υπαλλήλους του Facebook έγραψαν ότι οι αλγόριθμοι του Facebook εκμεταλλεύονται την έλξη του ανθρώπινου μυαλού

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

προς τη διχόνοια. Αν αφηθεί ανεξέλεγκτο, το Facebook θα συνεχίσει να προτείνει στους χρήστες όλο και περισσότερο διχαστικό περιεχόμενο για να αυξήσει το χρόνο τους στην πλατφόρμα [5].

Για την κρισιμότητα του συστήματος συστάσεων του Facebook, με βάση την κλίμακα μέτρησης κρισιμότητας που προτάθηκε, θα εφαρμόσουμε τα Κριτήρια Α και Β που παρουσιάστηκαν σε προηγούμενη ενότητα. Θα ξεκινήσουμε με το δεύτερο κριτήριο ώστε να παρουσιαστεί με μεγαλύτερη σαφήνεια η μέθοδος που προτάθηκε.

Σύμφωνα με την υπόθεση της Cambridge Analytica, οι χρήστες συμφώνησαν με τη συλλογή των δεδομένων τους χωρίς να πληροφορηθούν επαρκώς ποια δεδομένα θα μοιράζονταν, και πώς θα υπόκειντο σε επεξεργασία. Επίσης τα δεδομένα τελικά διέρρευσαν και έγιναν αντικείμενο εκμετάλλευσης από μια τρίτη εταιρεία. Η περίπτωση αυτή αποτελεί χαρακτηριστική περίπτωση παραβίασης **ιδιωτικότητας**.

Επιπροσθέτως, εύκολα εντοπίζεται **ακατάλληλο περιεχόμενο**, μιας και ψεύτικες ειδήσεις και προπαγάνδα μοιράζονται καθημερινά μέσω του συστήματος συστάσεων του Facebook.

Ακόμη, η ταυτότητα διαφόρων χρηστών χρησιμοποιείται ώστε να τους εθίσει σε συγκεκριμένου τύπου περιεχόμενο. Παράδειγμα αποτελούν οι εξτρεμιστικές ομάδες του Facebook. Κατά συνέπεια, προκύπτει **πρόβλημα προσωπικής ταυτότητας** για τους χρήστες.

Εφόσον οι χρήστες δεν έχουν κάποιο τρόπο για να ενημερωθούν με ακρίβεια για το πώς τα δεδομένα τους συλλέγονται και επεξεργάζονται από το σύστημα συστάσεων του Facebook, προκύπτει θέμα **διαφάνειας**.

Επομένως, στο σύστημα συστάσεων του Facebook αναγνωρίζονται τέσσερις από τις πέντε ηθικά προβληματικές εκφάνσεις των συστημάτων συστάσεων (παραβίαση ιδιωτικότητας, ακατάλληλο περιεχόμενο, θέματα προσωπικής ταυτότητας και θέματα διαφάνειας). Αυτό σημαίνει ότι είναι ένα σύστημα υψηλής κρισιμότητας. Επίσης, όσον αφορά το κριτήριο Α (με το οποίο θα έπρεπε να είχε ξεκινήσει η ανάλυση), προκύπτουν επίσης θέματα με σοβαρές κοινωνικές επιπτώσεις στο Facebook. Πιο συγκεκριμένα, παρατηρείται απομόνωση χρηστών, προώθηση προπαγάνδας και φαινόμενα μαζικού επηρεασμού. Συνεπώς, σύμφωνα και με τα δύο κριτήρια, το Facebook θεωρείται ένα **σύστημα υψηλής κρισιμότητας**.

### **Η περίπτωση της Amazon**

Η **Amazon** είναι μια πολυεθνική τεχνολογική εταιρεία. Οι βασικές της υπηρεσίες περιλαμβάνουν το ηλεκτρονικό εμπόριο, την υπολογιστική νέφους (cloud computing), την ροή περιεχομένου (content streaming) και την τεχνητή νοημοσύνη. Είναι μία από τις μεγαλύτερες εταιρείες παγκοσμίως. Στην παρούσα ενότητα, θα ασχοληθούμε με την υπηρεσία ηλεκτρονικού εμπορίου της Amazon.

Το σύστημα συστάσεων της Amazon βασίζεται στην μέθοδο collaborative filtering, ωστόσο δεν υπάρχουν διαθέσιμες πληροφορίες για τη λειτουργία του για το κοινό και κατά συνέπεια, ενεργεί ως μαύρο κουτί (black box) [6], [28]. Η εργασία [6] ασχολείται με τα βιβλία για τον εμβολιασμό που είναι διαθέσιμα προς αγορά στην Amazon. Παρατηρείται ότι τα βιβλία που υποστηρίζουν αντιεμβολιαστικές θέσεις ήταν παραπάνω από διπλάσια σε σχέση με τα βιβλία που ήταν υπέρ του εμβολιασμού. Επιπλέον, αποδείχτηκε ότι βιβλία που μοιράζονταν την ίδια οπτική στο θέμα των εμβολιασμών προτεινόνταν μαζί πολύ

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

συχνότερα από βιβλία που εξέφραζαν διαφορετικές απόψεις [6]. Αυτό σημαίνει ότι ένας χρήστης που έψαχνε ένα βιβλίο για τον εμβολιασμό που εξέφραζε μια συγκεκριμένη άποψη, πιθανότατα θα λάμβανε προτάσεις για βιβλία που υποστήριζαν την ίδια άποψη.

Επιπροσθέτως, το σύστημα συστάσεων της Amazon προτείνει διάφορες θεωρίες συνωμοσίας όταν ένας χρήστης ασχολούνταν με κάποια τέτοια θεωρία [7]. Επίσης, η λειτουργία αυτόματης συμπλήρωσης (autocomplete) της Amazon στην υπηρεσία αναζήτησης, παρατηρήθηκε ότι έδινε προβάδισμα σε θεωρίες συνωμοσίας. Έτσι, μπορεί ένας χρήστης να μην ενδιαφερόταν να ψάξει για κάποια συγκεκριμένη θεωρία συνωμοσίας, αλλά η λειτουργία αυτόματης συμπλήρωσης μπορεί να την πρότεινε [7]. Για παράδειγμα, με την συμπλήρωση της λέξης “vaccine” που σημαίνει εμβόλιο, η αυτόματη συμπλήρωση πρότεινε αποτελέσματα όπως “vaccines are dangerous” (τα εμβόλια είναι επικίνδυνα) ή “vaccines are the biggest pharmaceutical fraud in history” (τα εμβόλια είναι η μεγαλύτερη φαρμακευτική απάτη στην ιστορία).

Για όλους αυτούς τους λόγους, στο σύστημα της Amazon παρατηρούνται περιπτώσεις **ακατάλληλου περιεχομένου**.

Επίσης, παρατηρείται ότι το σύστημα συστάσεων της Amazon ενισχύει την προκατάληψη. Αυτό συμβαίνει, για παράδειγμα μέσω των βρόχων ανατροφοδότησης, που εξακολουθούν να προτείνουν περιεχόμενο πάνω σε μια συγκεκριμένη θεωρία. Ταυτόχρονα, οι προτάσεις του συστήματος απεικονίζουν στερεότυπα που συναντώνται στην κοινωνία και προωθεί προκατειλημμένες απόψεις. Έτσι προκύπτει ξεκάθαρο θέμα **δικαιοσύνης**.

Επίσης εντοπίζονται περιπτώσεις πρόληψης της έκθεσης ενός χρήστη σε διαφορετικές απόψεις, όπως για παράδειγμα στην περίπτωση προτάσεων βιβλίων σχετικών με τον εμβολιασμό. Τα παραπάνω προβλήματα ανήκουν στις **κοινωνικές επιπτώσεις**. Αυτό σημαίνει ότι ικανοποιείται το κριτήριο Α. Έτσι το σύστημα συστάσεων της Amazon θεωρείται **σύστημα υψηλής κρισιμότητας**.

### **Η περίπτωση του TikTok**

Με τη χρήση παρόμοιας μεθοδολογίας καταλήγουμε ότι το **TikTok** θεωρείται **κρίσιμο σύστημα**, καθώς εντοπίζονται περιπτώσεις **ακατάλληλου περιεχομένου** (εξτρεμιστικό και ρατσιστικό περιεχόμενο), και **προσωπικής ταυτότητας** (οι χρήστες εθίζονται στην κατανάλωση συγκεκριμένου τύπου περιεχομένου, που επιλέγεται από τον αλγόριθμο με βάση τις υποθέσεις που κάνει για την ταυτότητα του χρήστη). Αξίζει να σημειωθεί ότι η έρευνα που έχει γίνει για τις κοινωνικές επιπτώσεις του TikTok δεν είναι επαρκής, οπότε θα μπορούσε να αποτελεί και σύστημα υψηλής κρισιμότητας [29], [30], [31], [32], [33], [34].

## **10.6 Ρυθμιστικό πλαίσιο για την εύρυθμη λειτουργία των Συστημάτων Συστάσεων**

Στην παρούσα ενότητα προτείνεται ένα ρυθμιστικό πλαίσιο που στοχεύει στην βελτίωση των συστημάτων συστάσεων, όσον αφορά τις ηθικές τους επιπτώσεις. Αποτελείται από διάφορες προτάσεις που θα μπορούσαν να εφαρμοστούν και ανεξάρτητα μεταξύ τους. Ο στόχος είναι τα συστήματα

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

συστάσεων να συμβαδίσουν με την ανθρώπινη και κοινωνική ευημερία, χωρίς να επιβραδύνουν την πρόοδο και την εξέλιξη της.

1. **Σύσταση σώματος ελέγχου:** Σε συμφωνία με το ερευνητικό πρόγραμμα που προτείνεται στην εργασία [21], προτείνουμε την ίδρυση ενός μόνιμου σώματος ελέγχου της λειτουργίας των συστημάτων συστάσεων. Ο ρόλος του θα είναι να ελέγχει πανευρωπαϊκά και παγκόσμια τη λειτουργία των συστημάτων συστάσεων ώστε η λειτουργία των τελευταίων να είναι σύμφωνη με τις αρχές της υπεύθυνης έρευνας και καινοτομίας της Ευρωπαϊκής Επιτροπής (European Commission's Responsible Research and Innovation - RRI). Σύμφωνα με την RRI, «οι κοινωνικοί φορείς (ερευνητές, πολίτες, υπεύθυνοι χάραξης πολιτικών, επιχειρήσεις, οργανισμού, κλπ.) δουλεύουν μαζί καθ' όλη τη διάρκεια της διαδικασίας έρευνας και καινοτομίας για την καλύτερη ευθυγράμμιση τόσο της διαδικασίας όσο και των αποτελεσμάτων, με τις αξίες, τις ανάγκες και τις προσδοκίες της κοινωνίας» [35]. Η αποστολή του σώματος ελεγκτών θα είναι η έρευνα σε πραγματικά συστήματα συστάσεων, έτσι ώστε να παρακολουθείται από κοντά η εξέλιξη τους όσο και τα αποτελέσματα της στους χρήστες και την κοινωνία ως σύνολο. Έτσι, το σώμα ελεγκτών θα πρέπει να αποτελείται από μια διεπιστημονική ομάδα επιστημόνων και αναλυτών που θα μπορούν να μελετήσουν όλες τις εκφάνσεις του θέματος. Όπως έχει ήδη αναφερθεί παραπάνω, είναι αναμενόμενο ότι η βελτίωση των συστημάτων συστάσεων σε θέματα ηθικής, μπορεί να έρθει με κάποιο κόστος όσον αφορά την αποδοτικότητα τους. Για το λόγο αυτό, το σώμα ελεγκτών θα είναι σε συνεχή επικοινωνία με την τεχνολογική βιομηχανία, ενθαρρύνοντας την ανάπτυξη συστημάτων συστάσεων που θα ισορροπήσουν μεταξύ της ανάπτυξης και της ηθικής ακεραιότητας. Επιπλέον, το σώμα αυτό θα χρειάζεται μια νομική υπόσταση ώστε να παίρνει μέρος σε σημαντικές εργασίες όπως η χάραξη πολιτικής, η συμβουλευτική σε πολιτικούς φορείς, αναλαμβάνοντας στοχευμένη δράση και προτείνοντας τεχνολογικά ακριβείς λύσεις, και παράλληλα ενθαρρύνοντας πολιτικές παρεμβάσεις σε κρίσιμα θέματα. Η κλίμακα μέτρησης κρισιμότητας που προτείνεται στο παρόν κεφάλαιο μπορεί επίσης να χρησιμοποιηθεί ώστε να αποφασιστεί η καλύτερη δυνατή δράση. Για παράδειγμα, αν ένα σύστημα συστάσεων θεωρείται χαμηλής ή μεσαίας κρισιμότητας, μπορεί να τεθεί υπό παρακολούθηση από το σώμα ελεγκτών και οι τελευταίοι να προτείνουν βελτιώσεις. Αντίθετα, οι ελεγκτές θα χρειάζοταν να εφαρμόσουν πιο αυστηρά μέτρα σε ένα σύστημα υψηλής κρισιμότητας.
2. **Πολυμετοχικά Συστήματα Συστάσεων:** Η παραδοσιακή αντιμετώπιση των συστημάτων συστάσεων που επικεντρώνεται στο χρήστη πλέον θεωρείται από πολλούς ερευνητές απαρχαιωμένη, λαμβάνοντας υπόψη την παρούσα κατάσταση της τεχνολογίας και της συνέπειες της. Σε αυτό το πλαίσιο, προτείνεται η υιοθέτηση του υποδείγματος των πολυμετοχικών συστημάτων συστάσεων. Σύμφωνα με την ανάλυση του συγκεκριμένου υποδείγματος [20], τα συστήματα συστάσεων είναι πολυμετοχικά περιβάλλοντα στα οποία συμμετέχουν διαφορετικές ομάδες ενδιαφερομένων. Οι ομάδες αυτές έχουν ενδιαφέροντα που συνήθως δεν ταυτίζονται. Θεωρώντας ένα σύστημα συστάσεων ως πολυμετοχικό σύστημα, και παρατηρώντας τον αντίκτυπο που έχει στις διαφορετικές ομάδες ενδιαφερομένων, μπορούμε να συμπεράνουμε τους τρόπους με τους οποίους βλέπει ή ωφελεί την κοινωνία και κατά πόσο υπερασπίζεται ή καταπατά τα δικαιώματα των πολιτών και κατά συνέπεια, τα συμφέροντα της κάθε ομάδας μπορούν να έρθουν σε ισορροπία. Οι μέτοχοι ενός συστήματος συστάσεων με μια αφαιρετική ματιά είναι [20]:
  - a. Οι χρήστες

- b. Οι πάροχοι αντικειμένων, οι οποίοι κάνουν διαθέσιμες τις προτάσεις του συστήματος συστάσεων.
- c. Το σύστημα, δηλαδή η πλατφόρμα
- d. Η κοινωνία

Για παράδειγμα, στο σύστημα συστάσεων του Spotify, οι ομάδες ενδιαφερομένων είναι οι χρήστες, οι καλλιτέχνες των οποίων η δουλειά είναι διαθέσιμη στην πλατφόρμα, το ίδιο το Spotify και η κοινωνία. Το επίπεδο αφαίρεσης που χρησιμοποιείται είναι σημαντικός παράγοντας για την προσέγγιση αυτή. Αυτό διότι τοποθετεί το πλαίσιο στο οποίο μπορούν να εντοπιστούν τυχόν προβληματικές συνέπειες από τη λειτουργία ενός συστήματος συστάσεων [20]. Στην περίπτωση μιας πλατφόρμας ηλεκτρονικού εμπορίου με ένα σύστημα που προτείνει αντικείμενα για αγορά, συνήθως, δε θα ασχολούνταν κανείς με τις περιβαλλοντικές συνέπειες που προκαλούνται από τις προτάσεις του συστήματος. Ωστόσο, προσθέτοντας το περιβάλλον ως μέτοχο του συστήματος, με την πολυμετοχική προσέγγιση, η συσκευασία των προτεινόμενων αντικειμένων θα είχε σημασία, π.χ. αν το προτεινόμενο αντικείμενο περιέχει πλαστικό. Η πολυμετοχική προσέγγιση μπορεί να χρησιμοποιηθεί ως πυξίδα στη λήψη αποφάσεων όταν αναπτύσσεται ένα σύστημα συστάσεων, οδηγώντας στις καλύτερες δυνατές επιλογές αφού λαμβάνει υπόψη όλους τους ενδιαφερόμενους.

3. **Αντιμετώπιση κοινωνικής πόλωσης:** Η πόλωση συμβάλλει στη διάβρωση της δημοκρατίας και στη μετατόπιση προς τη βία [36]. Οι κοινωνικές επιπτώσεις από τη λειτουργία των συστημάτων συστάσεων συνδέονται σε μεγάλο βαθμό με την πόλωση. Η βελτίωση της διαφοροποίησης και της διαφορετικότητας (diversity) φαίνεται να βοηθάει στη μείωση των επιπέδων πόλωσης [36]. Επιπλέον, προτείνεται η ανάπτυξη αλγορίθμων συστάσεων οι οποίοι θα ενθαρρύνουν τον πολιτισμένο διάλογο, προωθώντας το πολιτισμένο περιεχόμενο έναντι εκφράσεων μίσους. Αυτή η προσέγγιση βέβαια μπορεί και η ίδια να εγείρει θέματα ηθικής, ωστόσο η έρευνα έχει δείξει ότι με την ενθάρρυνση διαλόγου με βάση το σεβασμό, ο οποίος δέχεται τη διαφορετικότητα των απόψεων μέσω της προώθησης πολιτισμένων και ποιοτικών σχολίων, έχει θετικές επιπτώσεις στο διάλογο που θα ακολουθήσει [36]. Για το λόγο αυτό, θα ήταν πολύ χρήσιμο να αναπτυχθούν συστήματα συστάσεων που θα αναγνωρίζουν τότε μια συγκεκριμένη αντίδραση ενός χρήστη είναι μια χρήσιμη ανταλλαγή απόψεων και τότε μια διχαστική διαφωνία. Αυτό μπορεί να οδηγήσει μελλοντικά, μέσω της ενισχυτικής μάθησης, σε συστήματα συστάσεων τα οποία θα μπορούν να αναγνωρίσουν μοτίβα συστάσεων που θα οδηγήσουν σε επιθυμητά μακροπρόθεσμα αποτελέσματα στην κοινωνία. Επιπροσθέτως, αλγόριθμοι βελτιστοποίησης για την εξασφάλιση επαρκούς διαφορετικότητας και διαφοροποίησης, όπως αυτοί που προτείνονται στο προηγούμενο κεφάλαιο μπορούν να αποδειχθούν ιδιαίτερα χρήσιμοι [37]. Σήμερα, τα συστήματα συστάσεων στα κοινωνικά δίκτυα έχουν ως στόχο να καταπολεμήσουν τις συγκρούσεις μέσω του μετριασμού του διχαστικού περιεχομένου (ρητορική μίσους, προπαγάνδα, ρατσιστικό περιεχόμενο, κλπ.), αλλά αυτό συνήθως γίνεται εκτός πλαισίου και το περιεχόμενο πολλές φορές διαγράφεται λανθασμένα. Σε αντίθεση με τη συγκεκριμένη πρακτική, τα συστήματα συστάσεων θα πρέπει να στοχεύουν στην πιο παραγωγική επίλυση των συγκρούσεων.
4. **Ο έλεγχος στα χέρια των χρηστών:** Προτείνουμε τον εκδημοκρατισμό των συστημάτων συστάσεων, με την έννοια του σεβασμού στα δικαιώματα των χρηστών και την παροχή πλήρους διαφάνειας σχετικά με τη λειτουργία των συστημάτων συστάσεων. Οι χρήστες θα πρέπει να

ενημερώνονται, κατά το δυνατόν, γιατί τους προτείνεται το περιεχόμενο που βλέπουν και τα κριτήρια που αποφασίζουν τι τους προτείνεται. Πρέπει επίσης να γίνεται προσεκτική διαχείριση των προσωπικών δεδομένων με πλήρη διαφάνεια, ώστε οι χρήστες να μπορούν να δώσουν πάντα την συγκατάθεση τους όντας απολύτως ενημερωμένοι. Ιδανικά, οι χρήστες θα ήταν χρήσιμο να έχουν την επιλογή για τροποποίηση των κριτηρίων μέσω των οποίων τους προτείνονται αντικείμενα, με τη χρήση μιας διεπαφής χρήστη, ώστε να ανταποκρίνονται καλύτερα στις ανάγκες τους. Ένα τέτοιο παράδειγμα είναι το σύστημα συστάσεων PeerChooser, που βασίζεται στη μέθοδο collaborative filtering και παρέχει μια διαδραστική παρουσίαση για τη διαδικασία συστάσεων [38]. Περιλαμβάνει επίσης την επιλογή για τον χρήστη να αφήσει τα σχόλια του και να αλλάξει τις παραμέτρους του συστήματος ώστε να αντανακλούν καλύτερα τις ανάγκες και τη διάθεση του κάθε στιγμή [38].

## 10.7 Συμπεράσματα και μελλοντικές επεκτάσεις

Στο κεφάλαιο αυτό εξετάζονται οι ηθικές επιπτώσεις των συστημάτων συστάσεων, όπως έχουν παρουσιαστεί στη βιβλιογραφία. Στη συνέχεια, προτείνεται ένα ποιοτικό πλαίσιο αξιολόγησης για την αξιολόγηση της ηθικής κρισιμότητας ενός συστήματος συστάσεων. Το πλαίσιο αξιολόγησης εφαρμόζεται σε μια σειρά από δημοφιλή συστήματα συστάσεων και τα αποτελέσματα δείχνουν ότι πολλά από αυτά είναι συστήματα υψηλής κρισιμότητας όσον αφορά τους ηθικούς προβληματισμούς που εγείρουν. Τέλος, προτείνεται μια σειρά από ρυθμιστικές ενέργειες για την αντιμετώπιση των ηθικών προβλημάτων που εγείρονται, και καθεμία από αυτές επεξηγείται αναλυτικά.

Όσον αφορά τις μελλοντικές επεκτάσεις, το επόμενο βήμα είναι η εφαρμογή του προτεινόμενου πλαισίου αξιολόγησης σε συστήματα συστάσεων μικρότερης κλίμακας, και η εφαρμογή διορθωτικών κινήσεων με βάση τις ρυθμιστικές ενέργειες που προτάθηκαν. Στη συνέχεια, τα συστήματα θα αξιολογηθούν από πραγματικούς χρήστες τόσο με τις διορθωτικές κινήσεις που έγιναν όσο και χωρίς αυτές ώστε να διαπιστωθεί ο αντίκτυπος που είχαν οι αλλαγές σε διάφορους τομείς όπως είναι η ικανοποίηση των χρηστών, και η διαφορετικότητα. Επιπλέον, η αυτοματοποίηση κάποιων από τις ρυθμιστικές ενέργειες που προτάθηκαν θα εξεταστεί εκτενώς. Τέλος, θα εξεταστεί ενδελεχώς πώς το προτεινόμενο πλαίσιο αξιολόγησης, όσο και οι προτεινόμενες ρυθμιστικές ενέργειες θα μπορούσαν να βελτιωθούν.

## 10.8 Αναφορές

- [1] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, "Recommender systems: An introduction," vol. 9780521493369. 2010.  
[https://books.google.gr/books?hl=el&lr=&id=eygTJBd\\_U2cC&oi=fnd&pg=PR5&dq=%5B1%5D%09D.+Jannach,+M.+Zanker,+A.+Felfernig,+and+G.+Friedrich,+Recommender+systems:+An+introduction,+vol.+9780521493369.+2010.&ots=mWB050FUvJ&sig=WsgX3Lw0ONygsfjwmcjrjvhvGYhns&redir\\_esc=y#v=onepage&q&f=false](https://books.google.gr/books?hl=el&lr=&id=eygTJBd_U2cC&oi=fnd&pg=PR5&dq=%5B1%5D%09D.+Jannach,+M.+Zanker,+A.+Felfernig,+and+G.+Friedrich,+Recommender+systems:+An+introduction,+vol.+9780521493369.+2010.&ots=mWB050FUvJ&sig=WsgX3Lw0ONygsfjwmcjrjvhvGYhns&redir_esc=y#v=onepage&q&f=false) (accessed May 03, 2023).
- [2] S. Milano, M. Taddeo, and L. Floridi, "Recommender systems and their ethical challenges," *AI Soc*, vol. 35, no. 4, pp. 957–967, Dec. 2020, doi: 10.1007/S00146-020-00950-Y/TABLES/4.
- [3] N. Seaver, "Captivating algorithms: Recommender systems as traps," <https://doi.org/10.1177/1359183518820366>, vol. 24, no. 4, pp. 421–436, Dec. 2018, doi:

10.1177/1359183518820366.

- [4] “Facebook revelations: what is in cache of internal documents? | Facebook | The Guardian.” <https://www.theguardian.com/technology/2021/oct/25/facebook-revelations-from-misinformation-to-mental-health> (accessed May 03, 2023).
- [5] “Facebook Executives Shut Down Efforts to Make the Site Less Divisive - WSJ.” <https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499> (accessed May 03, 2023).
- [6] J. Shin and T. Valente, “Algorithms and Health Misinformation: A Case Study of Vaccine Books on Amazon,” <https://doi.org/10.1080/10810730.2020.1776423>, vol. 25, no. 5, pp. 394–401, May 2020, doi: 10.1080/10810730.2020.1776423.
- [7] “Recommended Reading: Amazon’s algorithms, conspiracy theories and extremist literature - ISD.” <https://www.isdglobal.org/isd-publications/recommended-reading-amazons-algorithms-conspiracy-theories-and-extremist-literature/> (accessed May 03, 2023).
- [8] M. Millicamp, N. N. Htun, Y. Jin, and K. Verbert, “Effects of personal characteristics on music recommender user interfaces,” UMAP 2018 - Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization, vol. 18, pp. 101–109, Jul. 2018, doi: 10.1145/3209219.3209223.
- [9] S. M. McNee, J. Riedl, and J. A. Konstan, “Being accurate is not enough: How accuracy metrics have hurt recommender systems,” Conference on Human Factors in Computing Systems - Proceedings, pp. 1097–1101, 2006, doi: 10.1145/1125451.1125659.
- [10] H. Abdollahpouri, R. Burke, and B. Mobasher, “Recommender systems as multistakeholder environments,” UMAP 2017 - Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization, pp. 347–348, Jul. 2017, doi: 10.1145/3079628.3079657.
- [11] R. Burke, H. Abdollahpouri, E. C. Malthouse, K. P. Thai, and Y. Zhang, “Recommendation in multistakeholder environments,” RecSys 2019 - 13th ACM Conference on Recommender Systems, pp. 566–567, Sep. 2019, doi: 10.1145/3298689.3346973.
- [12] I. Koutsopoulos and M. Halkidi, “Optimization of Multi-stakeholder Recommender Systems for Diversity and Coverage,” IFIP Adv Inf Commun Technol, vol. 627, pp. 703–714, 2021, doi: 10.1007/978-3-030-79150-6\_55/FIGURES/2.
- [13] M. Kunaver and T. Požrl, “Diversity in recommender systems – A survey,” Knowl Based Syst, vol. 123, pp. 154–162, May 2017, doi: 10.1016/J.KNOSYS.2017.02.009.
- [14] R. Burke, N. Sonboli, and A. Ordonez-Gauger, “Balanced Neighborhoods for Multi-sided Fairness in Recommendation,” Proceedings of Machine Learning Research, vol. 81. PMLR, pp. 202–214, Jan. 21, 2018. Accessed: May 03, 2023. [Online]. Available: <https://proceedings.mlr.press/v81/burke18a.html>
- [15] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions,” IEEE Trans Knowl Data Eng, vol. 17, no. 6, pp. 734–749, Jun. 2005, doi: 10.1109/TKDE.2005.99.
- [16] W. Quattrociocchi, A. Scala, and C. R. Sunstein, “Echo Chambers on Facebook,” SSRN Electronic Journal, Jun. 2016, doi: 10.2139/SSRN.2795110.



- [17] F. Cinus, M. Minici, C. Monti, and F. Bonchi, “The Effect of People Recommenders on Echo Chambers and Polarization,” *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 16, pp. 90–101, May 2022, doi: 10.1609/ICWSM.V16I1.19275.
- [18] M. Badami, O. Nasraoui, W. Sun, and P. Shafto, “Detecting polarization in ratings: An automated pipeline and a preliminary quantification on several benchmark data sets,” *Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017*, vol. 2018-January, pp. 2682–2690, Jul. 2017, doi: 10.1109/BIGDATA.2017.8258231.
- [19] J. Stray, I. Vendrov, J. Nixon, S. Adler, and D. Hadfield-Menell, “What are you optimizing for? Aligning Recommender Systems with Human Values,” Jul. 2021, Accessed: May 03, 2023. [Online]. Available: <https://arxiv.org/abs/2107.10939v1>
- [20] S. Milano, M. Taddeo, and L. Floridi, “Ethical aspects of multi-stakeholder recommendation systems,” <https://doi.org/10.1080/01972243.2020.1832636>, vol. 37, no. 1, pp. 35–45, 2020, doi: 10.1080/01972243.2020.1832636.
- [21] A. Koene et al., “Ethics of personalized information filtering,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9089, pp. 123–132, 2015, doi: 10.1007/978-3-319-18609-2\_10/COVER.
- [22] T. Y. Tang and P. Winoto, “I should not recommend it to you even if you will like it: the ethics of recommender systems,” <https://doi.org/10.1080/13614568.2015.1052099>, vol. 22, no. 1–2, pp. 111–138, Jan. 2015, doi: 10.1080/13614568.2015.1052099.
- [23] K. de Vries, “Identity, profiling algorithms and a world of ambient intelligence,” *Ethics and Information Technology 2010 12:1*, vol. 12, no. 1, pp. 71–85, Jan. 2010, doi: 10.1007/S10676-009-9215-9.
- [24] S. Yao and B. Huang, “Beyond Parity: Fairness Objectives for Collaborative Filtering,” *Adv Neural Inf Process Syst*, 2017.
- [25] A. D. I. Kramer, J. E. Guillory, and J. T. Hancock, “Experimental evidence of massive-scale emotional contagion through social networks,” *Proc Natl Acad Sci U S A*, vol. 111, no. 24, pp. 8788–8790, Jun. 2014, doi: 10.1073/PNAS.1320040111/ASSET/E6DFA794-CF07-443A-8FD7-0DFD58FEF380/ASSETS/GRAPHIC/PNAS.1320040111FIG01.JPEG.
- [26] “Facebook users by country 2023 | Statista.” <https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/> (accessed May 03, 2023).
- [27] “Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach | Cambridge Analytica | The Guardian.” <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> (accessed May 03, 2023).
- [28] B. Smith and G. Linden, “Two Decades of Recommender Systems at Amazon.com,” *IEEE Internet Comput*, vol. 21, no. 3, pp. 12–18, May 2017, doi: 10.1109/MIC.2017.72.
- [29] J. C. Medina Serrano, O. Papakyriakopoulos, and S. Hegelich, “Dancing to the Partisan Beat: A First Analysis of Political Communication on TikTok,” *WebSci 2020 - Proceedings of the 12th ACM Conference on Web Science*, pp. 157–166, Jul. 2020, doi: 10.1145/3394231.3397916.
- [30] Y. Zhao, “Analysis of TikTok’s Success Based on Its Algorithm Mechanism,” *Proceedings - 2020*

- International Conference on Big Data and Social Sciences, ICBDS 2020, pp. 19–23, Aug. 2020, doi: 10.1109/ICBDSS51270.2020.00012.
- [31] K. E. Anderson, “Getting acquainted with social networks and apps: it is time to talk about TikTok,” *Library Hi Tech News*, vol. 37, no. 4, pp. 7–12, May 2020, doi: 10.1108/LHTN-01-2020-0001/FULL/PDF.
- [32] J. Cook, “Far-right activists are taking their message to gen Z on TikTok,” *HuffPost*, 2019.
- [33] “Tiktok has a Nazi problem,” *VICE*, 2018.
- [34] “Dance and sing while spreading a hoax — this is what TikTok looks like now - Poynter.” <https://www.poynter.org/fact-checking/2019/dance-and-sing-while-spreading-a-hoax-this-is-what-tiktok-looks-like-now/> (accessed May 03, 2023).
- [35] European Commission, *HORIZON 2020 - Responsible research & innovation*. 2020.
- [36] J. Stray, “Designing Recommender Systems to Depolarize,” *First Monday*, Jul. 2021, doi: 10.5210/fm.v27i5.12604.
- [37] E. Karakolis, P. Kokkinakos, and D. Askounis, “Provider Fairness for Diversity and Coverage in Multi-Stakeholder Recommender Systems,” *Applied Sciences* 2022, Vol. 12, Page 4984, vol. 12, no. 10, p. 4984, May 2022, doi: 10.3390/APP12104984.
- [38] J. O’Donovan, B. Smyth, B. Gretarsson, S. Bostandjiev, and T. Höllerer, “PeerChooser: Visual interactive recommendation,” *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1085–1088, 2008, doi: 10.1145/1357054.1357222.

## ΚΕΦΑΛΑΙΟ 11: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ

### 11.1 Σύνοψη

Η συγκεκριμένη διδακτορική διατριβή στοχεύει στη γεφύρωση του χάσματος μεταξύ των γνώσεων και δεξιοτήτων που κατέχει το εργατικό δυναμικό και των γνώσεων και δεξιοτήτων που ζητούνται από την αγορά εργασίας στον τομέα της τεχνολογίας. Οι γνώσεις και δεξιότητες που διαθέτουν οι εργαζόμενοι συνήθως καλλιεργούνται στα εκπαιδευτικά ιδρύματα της χώρας. Ωστόσο, δε συμβαδίζουν πάντα με τις ανάγκες της αγοράς εργασίας, η οποία εξελίσσεται ραγδαία, ειδικά στον τομέα της τεχνολογίας (Πρόβλημα 1 κεφαλαίου 1).

Το πρόβλημα αυτό αποτελεί ένα από τα αντικείμενα μελέτης του ευρωπαϊκού ερευνητικού έργου QualiChain, το οποίο περιλαμβάνει μια σειρά από εκπαιδευτικούς οργανισμούς, οργανισμούς πρόσληψης ανθρώπινου δυναμικού, καθώς επίσης και οργανισμούς ανάπτυξης τεχνολογίας και λογισμικού. Η έρευνα που πραγματοποιήθηκε στο πλαίσιο της συγκεκριμένης διδακτορικής διατριβής χρησιμοποιήθηκε από το έργο QualiChain.

Στο πλαίσιο της γεφύρωσης του χάσματος μεταξύ των γνώσεων και δεξιοτήτων που κατέχει το εργατικό δυναμικό και των γνώσεων και δεξιοτήτων που απαιτούνται από την αγορά εργασίας, αναπτύχθηκε μια σειρά από υπηρεσίες για την κάλυψη της απόστασης μεταξύ των διδασκόμενων γνώσεων και δεξιοτήτων στις σχολές ανώτατης εκπαίδευσης και στα ανοιχτά μαθήματα και των γνώσεων και δεξιοτήτων που είναι σε υψηλή ζήτηση από την αγορά εργασίας. Συγκεκριμένα, αναπτύχθηκε μια υπηρεσία με το όνομα JobCrawler για τη συγκέντρωση δεδομένων για την αγορά εργασίας μέσω της εφαρμογής τεχνικών απόξεσης δεδομένων από το διαδίκτυο (web scraping) σε κάποιους ιστοτόπους αγγελιών εργασίας. Επιπλέον, αναπτύχθηκε ένα εργαλείο εξαγωγής πληροφορίας μέσα από τα δεδομένα από τις αγγελίες εργασίας που εξάχθηκαν μέσω της υπηρεσίας JobCrawler με τη χρήση τεχνικών επεξεργασίας φυσικής γλώσσας. Με τη χρήση του εργαλείου αυτού, εξάγονται αυτόματα πληροφορίες για τις γνώσεις δεξιότητες που απαιτεί μια θέση εργασίας. Επιπλέον, χρησιμοποιήθηκε για την εξαγωγή πληροφορίας για τις γνώσεις και δεξιότητες που προσφέρονται σε ένα μάθημα μιας σχολής δίνοντας του ως είσοδο το πρόγραμμα σπουδών της σχολής ή μια λίστα μαθημάτων με τις περιγραφές τους.

Με τη χρήση των παραπάνω εργαλείων αναπτύχθηκε μια σημαντική βάση γνώσης, η οποία αποτελείται από θέσεις εργασίας, μαθήματα καθώς επίσης και σχετικές γνώσεις και δεξιότητες για κάθε ένα από αυτά. Πάνω στη συγκεκριμένη βάση γνώσης αναπτύχθηκαν δύο συστήματα συστάσεων. Το πρώτο (υπηρεσία σύστασης μαθημάτων και δεξιοτήτων - Course and Skill Recommender) παρέχει προσωποποιημένες προτάσεις δεξιοτήτων και μαθημάτων σε φοιτητές και δια βίου μαθητές, ώστε να βελτιώσουν το προφίλ τους με βάση τις ανάγκες της αγοράς εργασίας, αλλά και τις προσωπικές τους ανάγκες και προτιμήσεις.

Το δεύτερο (υπηρεσία ανασχεδιασμού προγράμματος σπουδών - Curriculum Designer) αναλύει ένα πρόγραμμα σπουδών ή μια λίστα από μαθήματα με τις περιγραφές τους και βρίσκει τις γνώσεις και δεξιότητες που δε προσφέρονται σε αυτό αλλά είναι σε υψηλή ζήτηση από την αγορά εργασίας. Επιπλέον, προτείνει την εισαγωγή των γνώσεων και δεξιοτήτων που απουσιάζουν από το πρόγραμμα σπουδών σε συγκεκριμένα μαθήματα, τα οποία είναι ταιριαστά με τις γνώσεις και δεξιότητες αυτές.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Αξίζει να σημειωθεί ότι οι προαναφερθείσες υπηρεσίες ενημερώνονται με νέα δεδομένα ανά τακτά χρονικά διαστήματα, ώστε να παρέχουν επίκαιρη πληροφορία κάθε στιγμή. Για το λόγο αυτό, αποτελούν πολύτιμο σύμμαχο τόσο για τους δια βίου μαθητές όσο και για καθηγητές ή διοικητικά συμβούλια σχολών ανώτατης εκπαίδευσης ή οργανισμών παροχής ανοιχτών μαθημάτων και άλλων εκπαιδευτικών οργανισμών. Φυσικά, ο ρόλος των εργαλείων αυτών είναι καθαρά συμβουλευτικός, με την έννοια ότι παρέχουν προτάσεις και τη λογική πίσω από καθεμιά από αυτές, και ο χρήστης μπορεί να τις αξιοποιήσει κατά το δοκούν.

Επιπροσθέτως, καθώς τα συστήματα συστάσεων που αναπτύχθηκαν έχουν ως στόχο την παροχή προτάσεων για τη λήψη ενημερωμένων αποφάσεων από τους χρήστες, κρίθηκε σκόπιμη η ανάπτυξη μιας υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων, η οποία μπορεί να χρησιμοποιηθεί συμπληρωματικά με τα συστήματα συστάσεων που αναπτύχθηκαν. Με την υπηρεσία αυτή, ο χρήστης μπορεί να είναι πιο σίγουρος για την απόφαση του λαμβάνοντας υπόψη όλα τα απαραίτητα κριτήρια, συμπεριλαμβανομένων και των υποκειμενικών κριτηρίων που είναι σημαντικά για αυτόν.

Όλες οι προαναφερθείσες υπηρεσίες αξιοποιήθηκαν στο πλαίσιο ενός πιλοτικού προγράμματος που εφαρμόστηκε στη σχολή Η.Μ.Μ.Υ. του Ε.Μ.Π.. Συγκεκριμένα, οι εφαρμογές αυτές παρουσιάστηκαν σε 90 φοιτητές (προπτυχιακούς και διδακτορικούς) και σε 11 καθηγητές. Οι συγκεκριμένοι χρήστες κλήθηκαν να χρησιμοποιήσουν και να αξιολογήσουν τις υπηρεσίες που τους αφορούσαν και να παρέχουν την άποψη τους με τη μορφή συμπλήρωσης μιας σειράς ερωτηματολογίων. Σε γενικές γραμμές, οι χρήστες έμειναν αρκετά ευχαριστημένοι από τις υπηρεσίες που τους παρουσιάστηκαν. Ωστόσο, ένα μικρό μέρος των χρηστών εξέφρασαν ουδέτερες ή αρνητικές απόψεις. Οι χρήστες αυτοί κλήθηκαν να παρέχουν τις προτάσεις τους για τη βελτίωση των υπηρεσιών που αναπτύχθηκαν, και μέσα από τις απαντήσεις τους αντλήθηκαν σημαντικά συμπεράσματα ως προς τις μελλοντικές επεκτάσεις και βελτιώσεις των υπηρεσιών αυτών.

Αξίζει να σημειωθεί ότι κατά την ενασχόληση του συγγραφέα με το πρόβλημα των εκπαιδευτικών συστημάτων συστάσεων για την κάλυψη της απόστασης της εκπαίδευσης με τις ανάγκες της αγοράς εργασίας, προέκυψαν κάποια επιπρόσθετα προβλήματα σχετικά με τα Συστήματα Συστάσεων, με τα οποία επίσης καταπιάνεται η παρούσα διδακτορική διατριβή. Συγκεκριμένα, ένα πρόβλημα που προέκυψε στα αρχικά στάδια του έργου QualiChain ήταν η δικαιοσύνη μεταξύ των παρόχων των μαθημάτων, με την έννοια της κατά το δυνατόν δίκαιης πρότασης των μαθημάτων κάθε παρόχου τόσο σε αριθμό χρηστών όσο και στη διαφορετικότητα αυτών (Πρόβλημα 2, κεφαλαίου 1). Αξίζει να σημειωθεί ότι η υπάρχουσα βιβλιογραφία δεν παρείχε κάποια επαρκή λύση για το συγκεκριμένο πρόβλημα το οποίο ανήκει στην κατηγορία προβλημάτων NP-Hard. Ωστόσο, παρείχε μια μοντελοποίηση και μια υποτυπώδη ευριστική λύση υψηλού επιπέδου (high-level heuristic solution). Στο πλαίσιο της παρούσας διδακτορικής αναπτύχθηκε ευριστικός αλγόριθμος που βρίσκει λύσεις πολύ κοντά στη θεωρητικά βέλτιστη λύση για το πρόβλημα στο σύνολο δεδομένων (dataset) που χρησιμοποιήθηκε. Επιπλέον, στην παρούσα διδακτορική διατριβή αναγνωρίστηκαν κάποιες αδυναμίες στη μοντελοποίηση του προβλήματος, όπως παρέχεται από τη βιβλιογραφία. Για το λόγο αυτό, προτάθηκε μια εναλλακτική μοντελοποίηση για το πρόβλημα καθώς επίσης και μια νέα ευριστική προσέγγιση για την επίλυση του.

Τέλος, λόγω της σημαντικής ενασχόλησης με τα συστήματα συστάσεων, στην παρούσα διατριβή ήρθαμε αντιμέτωποι, στην υφιστάμενη βιβλιογραφία, με διάφορες ηθικά επιλήψιμες εκφάνσεις που συνδέονται

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

με τη λειτουργία των σύγχρονων συστημάτων συστάσεων. Τα θέματα αυτά είναι πολύ σημαντικά για τα συστήματα συστάσεων, και ο απώτερος στόχος για τα συστήματα συστάσεων που αναπτύχθηκαν είναι να χρησιμοποιηθούν ως προϊόν. Κατά συνέπεια, ήταν σημαντικό να μελετηθούν οι ηθικές προεκτάσεις από τη λειτουργία τους. Για το λόγο αυτό, ερευνήθηκαν σε βάθος οι ηθικά προβληματικές εκφάνσεις των συστημάτων συστάσεων, παράλληλα με τους τρόπους με τους οποίους μπορούν να αξιολογηθούν και να αντιμετωπιστούν (Πρόβλημα 3, κεφαλαίου 1). Σε αυτό το πλαίσιο, αναγνωρίστηκαν και κατηγοριοποιήθηκαν οι προβληματικές εκφάνσεις των συστημάτων συστάσεων, προτάθηκε μέτρο για την αξιολόγηση της κρισιμότητας για ένα σύστημα συστάσεων με βάση τις κατηγορίες που αναγνωρίστηκαν. Το μέτρο εφαρμόστηκε σε κάποια από τα πιο γνωστά συστήματα συστάσεων παγκοσμίως για την αξιολόγηση της κρισιμότητας τους. Τέλος, προτάθηκε μια σειρά από ανεξάρτητες ρυθμιστικές ενέργειες για την εξυγίανση των συστημάτων συστάσεων.

## **11.2 Συμπεράσματα**

Στην παρούσα ενότητα παρουσιάζονται τα βασικά συμπεράσματα της συγκεκριμένης διδακτορικής διατριβής. Συγκεκριμένα, παρουσιάζονται τα βασικά συμπεράσματα για τις υπηρεσίες που αναπτύχθηκαν στο πλαίσιο του προβλήματος 1, όπως παρουσιάζεται στο 1<sup>ο</sup> κεφάλαιο (γεφύρωση του χάσματος μεταξύ των γνώσεων και δεξιοτήτων που προσφέρονται από εκπαιδευτικούς οργανισμούς και αυτών που απαιτούνται από την αγορά εργασίας στον τομέα της τεχνολογίας). Στη συνέχεια, παρουσιάζονται τα συμπεράσματα της διατριβής όσον αφορά το πρόβλημα 2 (Δικαιοσύνη μεταξύ των παρόχων αντικειμένων στα Συστήματα Συστάσεων) και τέλος παρουσιάζονται συνοπτικά τα βασικά συμπεράσματα για το πρόβλημα 3 (Ηθικά προβληματικές εκφάνσεις Συστημάτων Συστάσεων).

### **11.2.1 Πρόβλημα 1: Το χάσμα δεξιοτήτων μεταξύ αγοράς εργασίας και τεχνολογικής εκπαίδευσης**

#### **11.2.1.1 Προτάσεις μαθημάτων και δεξιοτήτων**

Η υπηρεσία συστάσεων μαθημάτων και δεξιοτήτων παρέχει προσωποποιημένες προτάσεις δεξιοτήτων και μαθημάτων σε φοιτητές, δια βίου μαθητές και υποψήφιους εργαζομένους, με βάση δεδομένα που εξάγονται τόσο για την αγορά εργασίας όσο και για την ανώτατη εκπαίδευση. Στην παρούσα διδακτορική διατριβή παρουσιάστηκε η λειτουργία του συστήματος για την αγορά εργασίας στον επιστημονικό κλάδο της επιστήμης υπολογιστών και της τεχνολογίας της πληροφορίας, και τα μαθήματα που προτείνονται προέρχονται από το πρόγραμμα σπουδών της σχολής Η.Μ.Μ.Υ. του Ε.Μ.Π. Ωστόσο, η χρήση της υπηρεσίας μπορεί εύκολα να επεκταθεί με επιπλέον σχολές και πανεπιστήμια, ή και διαφορετικούς επιστημονικούς κλάδους με μικρές αλλαγές.

Τα αποτελέσματα της υπηρεσίας αυτής φαίνονται λογικά και ικανοποιητικά για τους χρήστες. Επιπλέον, οι εξειδικεύσεις που αναγνωρίστηκαν με την τεχνική της συσταδοποίησης (clustering) είναι αντιπροσωπευτικές και αντιστοιχίζονται με την αγορά εργασίας και οι παρεχόμενες προτάσεις θα μπορούσαν ρεαλιστικά να βοηθήσουν έναν χρήστη να βελτιώσει τις γνώσεις και τις δεξιότητες του κατά τέτοιο τρόπο που θα είναι πιο ελκυστικός στην αγορά εργασίας.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Επιπλέον, αξίζει να σημειωθεί ότι με τη χρήση της συγκεκριμένης υπηρεσίας συστάσεων, ο χρήστης μπορεί να εξοικονομήσει σημαντικό χρόνο, καθώς η ανάλυση των αναγκών της αγοράς εργασίας είναι μια ιδιαίτερα χρονοβόρα διαδικασία για έναν άνθρωπο, ενώ το σύστημα παρέχει την απαραίτητη πληροφορία σε λίγα δευτερόλεπτα. Ακόμη, η υπηρεσία αυτή βασίζεται πάντα σε επικαιροποιημένα δεδομένα, καθώς οι αναλύσεις δεδομένων εκτελούνται περιοδικά, πράγμα που είναι πολύ δύσκολο να πραγματοποιηθεί από κάποιον άνθρωπο.

Τα συμπεράσματα αυτά επαληθεύονται και από την πιλοτική χρήση και αξιολόγηση της συγκεκριμένης υπηρεσίας στη σχολή Η.Μ.Μ.Υ. του Ε.Μ.Π., καθώς αξιολογήθηκε θετικά από σημαντική μερίδα φοιτητών που τη χρησιμοποίησαν, τόσο ως προς τη σχετικότητα των προτάσεων με τα ενδιαφέροντα τους, όσο και με βάση τη συμβατότητα με τις ανάγκες της αγοράς εργασίας. Περισσότερες πληροφορίες σχετικά με την αξιολόγηση της υπηρεσίας παρουσιάζονται στο κεφάλαιο 8.

### **11.2.1.2 Ανασχεδιασμός προγράμματος σπουδών**

Η υπηρεσία συστάσεων για τον ανασχεδιασμό προγράμματος σπουδών παρέχει προτάσεις για την εισαγωγή γνώσεων και δεξιοτήτων σε ένα δοθέν πρόγραμμα σπουδών ή μια λίστα μαθημάτων. Οι συγκεκριμένες γνώσεις και δεξιότητες προτείνονται διότι αφ' ενός είναι σε μεγάλη ζήτηση από την αγορά εργασίας και αφ' ετέρου δεν προσφέρονται στο δοθέν πρόγραμμα σπουδών ή στη δοθείσα λίστα μαθημάτων. Στην παρούσα διδακτορική διατριβή παρουσιάστηκε η λειτουργία του συστήματος για την αγορά εργασίας στον επιστημονικό κλάδο της τεχνολογίας της πληροφορίας και των επικοινωνιών, ενώ το πρόγραμμα σπουδών στο οποίο προτείνονται οι αλλαγές είναι το πρόγραμμα σπουδών της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου. Ωστόσο, η εφαρμογή μπορεί εύκολα να επεκταθεί για επιπλέον εκπαιδευτικούς οργανισμούς και ιδρύματα, ή και για διαφορετικούς επιστημονικούς κλάδους με στοχευμένες αλλαγές.

Τα αποτελέσματα φαίνονται λογικά και ικανοποιητικά για τους χρήστες. Επιπλέον, οι κανόνες συσχέτισης που εξάχθηκαν φαίνονται λογικοί και παρόμοιοι με τις συσχετίσεις που θα φανταζόταν ένας πεπειραμένος επαγγελματίας στο αντικείμενο της ανάπτυξης λογισμικού.

Η υπηρεσία ανασχεδιασμού προγράμματος σπουδών μπορεί να χρησιμοποιηθεί συμβουλευτικά για την ενημέρωση ενός προγράμματος σπουδών κάποιου εκπαιδευτικού οργανισμού, ενώ ένας τέτοιος οργανισμός μπορεί να αντλήσει σημαντική προστιθέμενη αξία από τη χρήση της, καθώς το σύστημα είναι πάντα ενημερωμένο για τις τελευταίες εξελίξεις στην αγορά εργασίας.

Επιπλέον, αξίζει να σημειωθεί ότι με τη χρήση της συγκεκριμένης υπηρεσίας, ο χρήστης μπορεί να εξοικονομήσει σημαντικό χρόνο, καθώς η ανάλυση των αναγκών της αγοράς εργασίας είναι μια ιδιαίτερα χρονοβόρα διαδικασία για έναν άνθρωπο, ενώ το σύστημα παρέχει την απαραίτητη πληροφορία σε λίγα δευτερόλεπτα. Επιπροσθέτως, παρέχει αντίστοιχη πληροφορία και για τα δοθέντα μαθήματα, απαλλάσσοντας το χρήστη από την ανάγκη να ψάξει και στις περιγραφές των μαθημάτων, ώστε να βρει σε ποια μαθήματα θα μπορούσαν να ενταχθούν οι νέες γνώσεις και δεξιότητες. Ακόμη, η υπηρεσία αυτή βασίζεται πάντα σε επικαιροποιημένα δεδομένα, καθώς οι αναλύσεις δεδομένων εκτελούνται περιοδικά, πράγμα που είναι πολύ δύσκολο να πραγματοποιηθεί από κάποιον άνθρωπο.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Τα συμπεράσματα αυτά επαληθεύονται και από την πιλοτική χρήση και αξιολόγηση της συγκεκριμένης υπηρεσίας στη σχολή Η.Μ.Μ.Υ. του Ε.Μ.Π., καθώς αξιολογήθηκε θετικά από 11 καθηγητές της σχολής που τη χρησιμοποίησαν. Συγκεκριμένα, οι καθηγητές επιβεβαίωσαν ότι οι γνώσεις και δεξιότητες που πρότεινε το σύστημα, ήταν εύστοχες, ενώ θεώρησαν και αρκετά σχετικές τις προτάσεις για την εισαγωγή τους σε συγκεκριμένα μαθήματα. Περισσότερες πληροφορίες σχετικά με την αξιολόγηση της υπηρεσίας παρουσιάζονται στο κεφάλαιο 8.

### **11.2.1.3 Πολυκριτηριακή υποστήριξη αποφάσεων**

Η υπηρεσία πολυκριτηριακής υποστήριξης αποφάσεων παρέχει στους χρήστες της πλατφόρμας ένα εργαλείο που τους βοηθάει να είναι πιο σίγουροι για τις αποφάσεις που λαμβάνουν, ταξινομώντας τις εναλλακτικές τους με βάση όλα τα κριτήρια τα οποία οι ίδιοι θεωρούν σημαντικά. Συγκεκριμένα, στο πλαίσιο του έργου QualiChain χρησιμοποιήθηκε για την υποβοήθηση της απόφασης επιλογής υποψηφίου για μια θέση εργασίας, όσο και για την επιλογή μαθημάτων και δεξιοτήτων από φοιτητές και δια βίου μαθητές, όπως επίσης και για την υποστήριξη αποφάσεων για την ενημέρωση ενός προγράμματος σπουδών με την εισαγωγή νέων γνώσεων και δεξιοτήτων.

Ένας περιορισμός που παρατηρήθηκε είναι ότι πολλοί χρήστες της πλατφόρμας δεν ήταν εξοικειωμένοι με την πολυκριτηριακή υποστήριξη αποφάσεων και τις μεθόδους της και κατά συνέπεια χρειάστηκε να προστεθούν επεξηγήσεις στην διεπαφή χρήστη της εφαρμογής, όπως επίσης και πολλές παρουσιάσεις για το πώς μπορεί να χρησιμοποιηθεί αποτελεσματικά. Επιπλέον, αξίζει να σημειωθεί ότι είχε σχεδιαστεί η ανάπτυξη περισσότερων μεθόδων πολυκριτηριακής ανάλυσης αποφάσεων στην υπηρεσία, αλλά λόγω της έλλειψης εξοικείωσης από τους χρήστες, αποφασίστηκε να δοθεί περισσότερη έμφαση στην βελτίωση της διεπαφής και στην παροχή επεξηγήσεων και παραδειγμάτων, αντί για την ανάπτυξη νέων πιο πολύπλοκων μεθόδων.

Τα συμπεράσματα αυτά προέκυψαν από την πιλοτική χρήση και αξιολόγηση της συγκεκριμένης υπηρεσίας στη σχολή Η.Μ.Μ.Υ. του Ε.Μ.Π. Περισσότερες πληροφορίες σχετικά με την αξιολόγηση της υπηρεσίας παρουσιάζονται στο κεφάλαιο 8.

### **11.2.1.4 Πιλοτική εφαρμογή**

Οι υπηρεσίες και οι τεχνολογίες που αναπτύχθηκαν στα πλαίσια του Προβλήματος 1 εφαρμόστηκαν πιλοτικά στη σχολή Η.Μ.Μ.Υ. του Ε.Μ.Π. για το επιστημονικό πεδίο της τεχνολογίας της πληροφορίας και των επικοινωνιών. Ο πιλότος αποσκοπούσε στη βελτίωση του προγράμματος σπουδών με βάση τις ανάγκες της αγοράς εργασίας, όσο και στην υποβοήθηση των φοιτητών στην επιλογή μαθημάτων και δεξιοτήτων, ώστε να έχουν καλύτερες προοπτικές καριέρας, αναπτύσσοντας γνώσεις και δεξιότητες που είναι σε υψηλή ζήτηση από την αγορά εργασίας.

Συνολικά, θεωρούμε ότι ο πιλότος πέτυχε το σκοπό του όσον αφορά στα τεχνικά αποτελέσματα της πλατφόρμας QualiChain σε κρίσιμες ομάδες ενδιαφερόμενων ενός πανεπιστημίου. Επιπροσθέτως, αποδείχτηκε ότι μια πλατφόρμα σαν το QualiChain, με το συνδυασμό τεχνικών και τεχνολογιών που χρησιμοποιεί μπορεί να ενσωματωθεί στην καθημερινή λειτουργία ενός πανεπιστημίου, καθώς προσφέρει αρκετές λειτουργίες και καινοτομίες που υποστηρίζουν διαφορετικές ομάδες χρηστών να ολοκληρώσουν συντομότερα και αποτελεσματικότερα της εργασίες τους.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Όσον αφορά τη χρήση υπηρεσιών αναλυτικής δεδομένων και υποστήριξης αποφάσεων για την επιλογή μαθημάτων και δεξιοτήτων για τη βελτίωση του ακαδημαϊκού ή του επαγγελματικού προφίλ του χρήστη, με βάση τις ανάγκες της αγοράς εργασίας, συμπεραίνουμε ότι οι φοιτητές που συμμετείχαν στο πιλοτικό πρόγραμμα ήταν συνολικά θετικοί και ικανοποιημένοι από τις παρεχόμενες υπηρεσίες. Ωστόσο, υπήρχε ένα χαμηλό ποσοστό φοιτητών που δεν ήταν και τόσο ευχαριστημένοι. Η απογοήτευση των συγκεκριμένων φοιτητών έγκειται στο γεγονός ότι περίμεναν το σύστημα να είναι ενοποιημένο με τη βάση δεδομένων της σχολής Η.Μ.Μ.Υ., ώστε τα μαθήματα που επιλέγουν στο QualiChain να εμφανίζονται αυτόματα και στον διαδικτυακό τόπο της σχολής Η.Μ.Μ.Υ. και στα φοιτητικά τους στοιχεία, χωρίς να χρειάζεται να τα ξανά-επιλέξουν, ωστόσο κάτι τέτοιο δεν ήταν εφικτό στο πλαίσιο ενός ερευνητικού προγράμματος. Παρ' όλα αυτά, οι περισσότεροι φοιτητές έμειναν ικανοποιημένοι τόσο από τις προτάσεις που λάμβαναν για τη βελτίωση του προφίλ τους, όσο και για το γεγονός ότι μπορούσαν να επιλέξουν μαθήματα της σχολής τους μέσα από την πλατφόρμα. Είναι προφανές ότι τέτοιες προσωποποιημένες υπηρεσίες θα εκλαμβάνονταν θετικά από τους φοιτητές σε περίπτωση που περισσότερα πανεπιστήμια αποφασίσουν να χρησιμοποιήσουν την πλατφόρμα QualiChain.

Όσον αφορά τη χρήση υπηρεσιών αναλυτικής δεδομένων και υποστήριξης αποφάσεων για την ανανέωση του προγράμματος σπουδών μιας σχολής ώστε να ανταποκρίνεται καλύτερα στις ανάγκες της αγοράς εργασίας, οι καθηγητές που χρησιμοποίησαν το σύστημα έμειναν πολύ ικανοποιημένοι από της προτάσεις που τους παρείχε η πλατφόρμα για τα μαθήματα τους όσο και για το πρόγραμμα σπουδών γενικότερα. Επιπροσθέτως, οι καθηγητές έμειναν πολύ ικανοποιημένοι και από τη διαδικασία ανανέωσης των μαθημάτων τους η οποία ήταν καλά σχεδιασμένη και εύκολη στη χρήση. Επίσης, αξίζει να σημειωθεί ότι η επέκταση της πλατφόρμας και σε άλλους τομείς πέρα από την τεχνολογία πληροφορικής και επικοινωνιών, αναμένεται να φέρει σημαντική προστιθέμενη αξία και για άλλες σχολές και επαγγέλματα.

Τέλος, αξίζει να σημειωθεί ότι οι αλλαγές στο εργασιακό περιβάλλον που προέκυψαν λόγω της πανδημίας επιδεικνύουν ακόμα περισσότερο τη δυναμική των υπηρεσιών που αναπτύχθηκαν. Συγκεκριμένα, η απομακρυσμένη εργασία δημιουργεί μία παγκόσμια αγορά εργασίας, η οποία είναι πιο ανταγωνιστική από ποτέ. Για το λόγο αυτό, είναι σημαντικό οι φοιτητές να επιλέγουν μαθήματα που θα τους βοηθήσουν να ανταποκριθούν στον ανταγωνισμό. Επίσης, τα διαδικτυακά μαθήματα κερδίζουν όλο και μεγαλύτερη δημοφιλία, κατά συνέπεια τα πανεπιστήμια θα πρέπει και αυτά με τη σειρά τους να προσφέρουν τις πιο προηγμένες γνώσεις και δεξιότητες στους φοιτητές μέσω των μαθημάτων τους. Για όλους αυτούς τους λόγους, θεωρούμε ότι οι εφαρμογές που αναπτύχθηκαν και αξιολογήθηκαν μέσω του συγκεκριμένου πιλότου πέτυχαν την ανάδειξη των αναγκών της ανώτατης εκπαίδευσης και παρείχαν μια καλή λύση για την κάλυψη των αναγκών αυτών.

### **11.2.2 Πρόβλημα 2: Δικαιοσύνη μεταξύ των παρόχων αντικειμένων στα Συστήματα Συστάσεων**

Σχετικά με το πρόβλημα της δικαιοσύνης μεταξύ των παρόχων όσον αφορά την κάλυψη και τη διαφοροποίηση των αντικειμένων τους σε χρήστες, εφαρμόστηκε σε κάποιο βαθμό, η μεθοδολογία που προτείνεται στη βιβλιογραφία για την επίλυση του προβλήματος. Συγκεκριμένα, μιας και το πρόβλημα ανήκει στην κατηγορία NP-Hard και δεν επιδέχεται αναλυτικής λύσης σε αποδεκτό χρόνο, και καθώς η



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

προτεινόμενη ευριστική λύση στη βιβλιογραφία είναι πολύ γενική και δεν επεξηγείται επαρκώς, στην παρούσα διατριβή (κεφάλαιο 9) αναπτύχθηκε νέος ευριστικός αλγόριθμος για την επίλυση των περιορισμών διαφοροποίησης που εξηγείται με λεπτομέρεια. Ακόμη, παρουσιάστηκε η υπολογιστική πολυπλοκότητα του προτεινόμενου αλγόριθμου, παράλληλα με τους περιορισμούς που παρουσιάζει. Επιπλέον, παρουσιάστηκαν εκτενώς τα αποτελέσματα της προτεινόμενης λύσης και είναι πολύ κοντά στα θεωρητικά βέλτιστα αποτελέσματα για το συγκεκριμένο dataset.

Αξίζει να σημειωθεί ότι κατά την ενασχόληση με το συγκεκριμένο πρόβλημα, αναγνωρίστηκαν κάποιες αδυναμίες στον ορισμό του. Η σημαντικότερη αδυναμία έγκειται στον ορισμό της διαφοροποίησης και στη σχέση της με την ανομοιότητα μεταξύ των χρηστών. Προς αυτή την κατεύθυνση, παρουσιάστηκε μια ανάλυση ως προς το τι επηρεάζει την διαφορετικότητα. Τα αποτελέσματα κατέδειξαν ότι ο σημαντικότερος παράγοντας είναι η συχνότητα που προτείνονται τα πιο συχνά προτεινόμενα αντικείμενα πράγμα το οποίο δε θα έπρεπε να συμβαίνει. Για το λόγο αυτό προτάθηκαν δύο νέοι ορισμοί για την διαφοροποίηση καθώς επίσης και μια νέα ευριστική μεθοδολογία για την εύρεση της λύσης.

### **11.2.3 Πρόβλημα 3: Ηθικά προβληματικές εκφάνσεις Συστημάτων Συστάσεων**

Στην παρούσα διδακτορική διατριβή εξετάστηκαν οι ηθικές επιπτώσεις των συστημάτων συστάσεων, όπως έχουν παρουσιαστεί στη βιβλιογραφία. Στη συνέχεια, προτάθηκε ένα ποιοτικό πλαίσιο αξιολόγησης για την αξιολόγηση της ηθικής κρισιμότητας ενός συστήματος συστάσεων. Το πλαίσιο αξιολόγησης εφαρμόστηκε σε μια σειρά από δημοφιλή συστήματα συστάσεων και τα αποτελέσματα κατέδειξαν ότι πολλά από αυτά είναι συστήματα υψηλής κρισιμότητας όσον αφορά τους ηθικούς προβληματισμούς που εγείρουν. Τέλος, προτάθηκε μια σειρά από ρυθμιστικές ενέργειες για την αντιμετώπιση των ηθικών προβλημάτων που εγείρονται, και καθεμία από αυτές επεξηγείται αναλυτικά.

## **11.3 Μελλοντικές επεκτάσεις**

### **11.3.1 Πρόβλημα 1: Το χάσμα δεξιοτήτων μεταξύ αγοράς εργασίας και τεχνολογικής εκπαίδευσης**

Όσον αφορά τις μελλοντικές επεκτάσεις για τις υπηρεσίες που αναπτύχθηκαν, ο βασικός μας στόχος είναι η μετατροπή των υπηρεσιών προτάσεων μαθημάτων, ανασχεδιασμού προγράμματος σπουδών και πολυκριτηριακής υποστήριξης αποφάσεων, σε προϊόν. Για να γίνει αυτό, θα πρέπει να προστεθούν περισσότεροι επαγγελματικοί και επιστημονικοί κλάδοι στη βάση γνώσης των υπηρεσιών αυτών. Επιπλέον, θα χρειαστεί νέα ανάλυση δεδομένων για τα δεδομένα κάθε κλάδου, ώστε να βελτιστοποιηθούν τα αποτελέσματα των αλγορίθμων συστάσεων. Επιπροσθέτως, για να γίνει η πλατφόρμα κανονικό προϊόν θα πρέπει να βρεθούν επιπλέον εκπαιδευτικοί οργανισμοί οι οποίοι θα υιοθετήσουν την πλατφόρμα και θα παρέχουν τα δεδομένα τους. Αυτό θα έχει ως αποτέλεσμα την προσέλκυση περισσότερων χρηστών.

Η προσέλκυση περισσότερων χρηστών και η χρήση των υπηρεσιών που αναπτύχθηκαν από αυτούς μπορεί αφενός να συμβάλει στη συγκέντρωση σχολίων για τη βελτίωση των παρεχόμενων υπηρεσιών, αλλά και στη συγκέντρωση επιπλέον δεδομένων, για την παροχή καλύτερων και πιο προσωποποιημένων προτάσεων. Σε αυτό το πλαίσιο, θα μπορούσαν να εφαρμοστούν και επιπλέον αλγόριθμοι συστάσεων

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

όπως είναι ο αλγόριθμος collaborative filtering. Ακόμη, θα μπορούσαν να χρησιμοποιηθούν περισσότερες μετρικές για την αξιολόγηση των υπηρεσιών συστάσεων, όπως είναι η ακρίβεια (precision), η ανάκληση (recall), και το Normalised Discounted Cumulative Gain (NDCG). Η αξιολόγηση με τις προαναφερθείσες μετρικές είναι πολύ σημαντική για τη βελτίωση των προτάσεων σε βάθος χρόνου, μιας και μέχρι στιγμής η αξιολόγηση γίνεται με τη χρήση ερωτηματολογίων.

Μια άλλη επέκταση που μελετούμε για το σύστημα συστάσεων μαθημάτων και δεξιοτήτων είναι η παροχή προτάσεων με γνώμονα το μέλλον, όπως επίσης και για διαφορετικές περιοχές, τις οποίες θα ορίζει ο χρήστης. Με αυτό τον τρόπο, ο χρήστης θα μπορεί να λαμβάνει προσωποποιημένες και έγκυρες συστάσεις τόσο για την περίοδο που στοχεύει, όσο και για την περιοχή για την οποία ενδιαφέρεται.

Επιπλέον, θα είχε ενδιαφέρον και η παρακολούθηση της επαγγελματικής πορείας χρηστών της πλατφόρμας μακροχρόνια. Με αυτό τον τρόπο, θα μπορούσαμε να διαπιστώσουμε ουσιαστικά σε ποιο βαθμό οι παρεχόμενες υπηρεσίες όντως συμβάλλουν στη βελτίωση της καριέρας κάποιου χρήστη. Ωστόσο, δεν υπήρχε αρκετός χρόνος για να γίνει κάτι τέτοιο στα πλαίσια του συγκεκριμένου διδακτορικού και του ερευνητικού έργου QualiChain.

Άλλη μια μελλοντική κατεύθυνση για τη βελτίωση των παρεχόμενων υπηρεσιών είναι η δοκιμή αλγορίθμων για τη δικαιοσύνη μεταξύ των παρόχων αντικειμένων, όπως αυτοί που παρουσιάστηκαν στο κεφάλαιο 9, με την προϋπόθεση ότι θα προστεθούν νέοι πάροχοι μαθημάτων στις υπηρεσίες συστάσεων. Επίσης, θα πρέπει να διερευνηθεί και ο αντίκτυπος αυτών των αλγορίθμων στις προτιμήσεις του χρήστη και να γίνει η κατάλληλη παραμετροποίηση, ώστε να μην χαθεί το ενδιαφέρον και η εμπιστοσύνη των χρηστών στις υπηρεσίες συστάσεων που αναπτύχθηκαν.

Αντίστοιχα, με την ανάπτυξη και εξάπλωση των παρεχόμενων υπηρεσιών συστάσεων σε περισσότερους χρήστες, θα πρέπει να πραγματοποιηθεί και μια αξιολόγηση ως προς τον ηθικό αντίκτυπο που έχει στους χρήστες η λειτουργία τους, με βάση το πλαίσιο αξιολόγησης που προτείνεται στο κεφάλαιο 10. Σε περίπτωση που διαπιστωθούν ηθικά επιλήψιμα θέματα, θα πρέπει να εφαρμοστούν οι κατάλληλες ρυθμιστικές ενέργειες. Μερικές τέτοιες ενέργειες προτείνονται επίσης στο κεφάλαιο 10.

Τέλος, όσον αφορά τις μελλοντικές επεκτάσεις της υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων, θα δοθεί έμφαση στη βελτίωση της διεπαφής χρήστη αλλά και στην ενημέρωση των χρηστών πάνω στις μεθόδους της πολυκριτηριακής υποστήριξης αποφάσεων, καθώς διαπιστώθηκε ότι οι χρήστες δυσκολεύονταν να χρησιμοποιήσουν τη συγκεκριμένη υπηρεσία. Επιπλέον, με βάση την αξιολόγηση από τους χρήστες, θα εξεταστεί η αυτόματη προσθήκη κριτηρίων και η αυτόματη συμπλήρωση της σημαντικότητας καθενός από αυτά με βάση τα κριτήρια που έχουν θέσει άλλοι χρήστες στο παρελθόν.

### **11.3.2 Πρόβλημα 2: Δικαιοσύνη μεταξύ των παρόχων αντικειμένων στα Συστήματα Συστάσεων**

Όσον αφορά τις μελλοντικές επεκτάσεις της έρευνας σχετικά με τη δικαιοσύνη μεταξύ των παρόχων αντικειμένων, αυτές περιλαμβάνουν την ένταξη της σειράς εμφάνισης ενός αντικειμένου στη λίστα προτάσεων ενός χρήστη στο μαθηματικό μοντέλο του συστήματος συστάσεων, καθώς είναι πολύ σημαντική η θέση ενός αντικειμένου στη λίστα συστάσεων ενός χρήστη. Αυτό θα γίνει χρησιμοποιώντας

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

και άλλες μετρικές αξιολόγησης όπως είναι το Normalised Discounted Cumulative Gain (NDCG), που μετράει την ποιότητα της θέσης των προτεινόμενων αντικειμένων για κάθε χρήστη.

Επιπλέον, περιλαμβάνουν την αντιπροσώπευση των μειονοτήτων, η οποία θα εξεταστεί σε σχέση με την κάλυψη και τη διαφοροποίηση, και θα συγκριθεί με άλλες αντίστοιχες εργασίες, ενώ, διαφορετικές μετρικές θα εξεταστούν για να μοντελοποιηθούν ως περιορισμοί του προβλήματος βελτιστοποίησης.

Τέλος, οι προτεινόμενες μέθοδοι θα ήταν χρήσιμο εφαρμοστούν σε πραγματικά δεδομένα και διαφορετικά datasets, όπως επίσης και σε συστήματα συστάσεων του πραγματικού κόσμου. Επίσης, πέρα από την αξιολόγηση της απόδοσης του συστήματος με βάση τις προτεινόμενες μετρικές, θα ήταν χρήσιμο το σύστημα να αξιολογηθεί από πραγματικούς χρήστες, όσον αφορά την κάλυψη και τη διαφοροποίηση όσο και την ικανοποίηση τους από το σύστημα, συμπεριλαμβανομένων απλών χρηστών όσο και παρόχων αντικειμένων.

### **11.3.3 Πρόβλημα 3: Ηθικά προβληματικές εκφάνσεις Συστημάτων Συστάσεων**

Το επόμενο βήμα, όσον αφορά την έρευνα στα πλαίσια του συγκεκριμένου προβλήματος, είναι η εφαρμογή του προτεινόμενου πλαισίου αξιολόγησης σε συστήματα συστάσεων μικρότερης κλίμακας, και η εφαρμογή διορθωτικών κινήσεων με βάση τις ρυθμιστικές ενέργειες που προτάθηκαν. Επιπροσθέτως, τα συστήματα θα πρέπει να αξιολογηθούν από πραγματικούς χρήστες τόσο με τις διορθωτικές κινήσεις που έγιναν όσο και χωρίς αυτές ώστε να διαπιστωθεί ο αντίκτυπος των αλλαγών σε διάφορους τομείς όπως είναι η ικανοποίηση των χρηστών, και η διαφορετικότητα. Επιπλέον, η αυτοματοποίηση κάποιων από τις ρυθμιστικές ενέργειες που προτάθηκαν θα πρέπει να εξεταστεί εκτενώς. Τέλος, είναι σημαντικό να εξεταστεί ενδελεχώς πώς το προτεινόμενο πλαίσιο αξιολόγησης, όσο και οι προτεινόμενες ρυθμιστικές ενέργειες θα μπορούσαν να βελτιωθούν.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## ΠΑΡΑΡΤΗΜΑ Ι: ΔΗΜΟΣΙΕΥΜΕΝΟ ΕΡΓΟ

### I.1 Δημοσιεύσεις σε διεθνή επιστημονικά περιοδικά

1. **Karakolis, E.**, Kokkinakos, P., & Askounis, D. (2022). Provider Fairness for Diversity and Coverage in Multi-Stakeholder Recommender Systems. *Applied Sciences*, 12(10), 4984.  
<https://doi.org/10.3390/app12104984>
2. **Karakolis, E.**, Kapsalis, P., Skalidakis, S., Kontzinos, C., Kokkinakos, P., Markaki, O., & Askounis, D. (2022). Bridging the Gap between Technological Education and Job Market Requirements through Data Analytics and Decision Support Services. *Applied Sciences*, 12(14), 7139.  
<https://doi.org/10.3390/app12147139>
3. Pelekis, S., Pipergias, A., **Karakolis, E.**, Mouzakitits, S., Santori, F., Ghoreishi, M., & Askounis, D. (2023). Targeted demand response for flexible energy communities using clustering techniques. *Sustainable Energy, Grids and Networks*, 101134.  
<https://doi.org/10.1016/j.segan.2023.101134>
4. Pelekis, S., Seisopoulos, I. K., Spiliotis, E., Pountridis, T., **Karakolis, E.**, Mouzakitits, S., & Askounis, D. (2023). A comparative assessment of deep learning models for day-ahead load forecasting: Investigating key accuracy drivers. *Sustainable Energy, Grids and Networks (SEGAN)*, 101171.  
<https://doi.org/10.1016/j.segan.2023.101171>
5. Kontzinos, C., Kokkinakos, P., Kapsalis, P., Markaki, O., **Karakolis, V.**, & Psarras, J. (2020). Exploring Blockchain, Semantics and Decision Support to Optimise Qualification Certification, Recruitment and Competency Management: an Assessment of Challenges, Current Practices and Opportunities. *International Journal on Advances in Intelligent Systems Volume 13, Number 3 & 4*, 2020.
6. Kontzinos, C., Kokkinakos, P., Kapsalis, P., Markaki, O., **Karakolis, V.**, & Psarras, J. (2020). Leveraging Blockchain, Analytics and Decision Support to Facilitate Qualifications' Verification, Recruitment and Competency Management: The QualiChain Project and Initial Results. *International Journal on Advances in Intelligent Systems Volume 13, Number 3 & 4*, 2020.
7. Vlachou, V., Kontzinos, C., Markaki, O., Kokkinakos, P., **Karakolis, V.**, & Psarras, J. (2020). Leveraging Hyperledger Iroha for the Issuance and Verification of Higher-Education Certificates. *International Journal of Educational and Pedagogical Sciences*, 14(9), 755-763.
8. Kontzinos, C., **Karakolis, V.**, Skalidakis, S., Markaki, O., Androutsopoulou, M., & Psarras, J. (2019). Combining Blockchain, Semantics and Data Analytics for University Process Optimisation. *IADIS International Journal on Computer Science & Information Systems*, 14(2).

### I.2 Παρουσιάσεις σε διεθνή επιστημονικά συνέδρια

1. **Karakolis E.**, Oikonomidis F., Askounis D. Identifying and Addressing Ethical Challenges in Recommender Systems. *The Thirteenth International Conference on Information, Intelligence, Systems and Applications (IISA 2022)*. 2022

2. **Karakolis, E.**, Pelekis, S., Mouzakitis, S., Kormpakis, G., Michalakopoulos, V., & Psarras, J. The I-ENERGY reference architecture for the provision of next generation energy services through artificial intelligence. IADIS International Conference e-Society 2023 (e-Society 2023). 2023
3. Touloumis K., Kapsalis P., Sarmas E., Stamatopoulos S., **Karakolis E.**, Marinakis V. (2023). A unified framework for querying dynamic and semantic data sources. Υποβλήθηκε στο διεθνές επιστημονικό συνέδριο The Fourteenth International Conference on Information, Intelligence, Systems and Applications (IISA 2023).
4. Touloumis, K., **Karakolis, E.**, Kapsalis, P., Marinakis, V., & Psarras, J. BD4NRG Query Engine – Intuitive, efficient and federated querying on big data. IADIS International Conference e-Society 2023 (e-Society 2023). 2023
5. Pelekis, S., Sarmas, E., Georgiadou, A., **Karakolis, E.**, Ntanos, C., Dimitropoulos, N., & Doukas, H. TWINP2G: A digital twin architecture for optimal power-to-gas planning. IADIS International Conference e-Society 2023 (e-Society 2023). 2023
6. **Karakolis E.**, Alexakis K., Kapsalis P., Mouzakitis S., Psarras J. An end-to-end approach for scalable real time Anomaly detection in smart buildings. The Thirteenth International Conference on Information, Intelligence, Systems and Applications (IISA 2022). 2022
7. Pelekis S., **Karakolis E.**, Silva F., Schoinas V., Mouzakitis S., Kormpakis G., Amaro N., Psarras J. In Search of Deep Learning Architectures for Load Forecasting: A Comparative Analysis and the Impact of the Covid-19 Pandemic on Model Performance. The Thirteenth International Conference on Information, Intelligence, Systems and Applications (IISA 2022). 2022
8. **Karakolis E.**, Pelekis S., Mouzakitis S., Markaki O., Papapostolou A., Korbakis G., Psarras J. Artificial Intelligence for Next Generation Energy Services Across Europe - The I-ENERGY Project IADIS International Conference e-Society 2022 (e-Society 2022). 2022
9. Kapsalis P., Kormpakis G., Alexakis K., **Karakolis E.**, Mouzakitis S., Askounis D. A Reasoning Engine Architecture for Building Energy Metadata Management. The Thirteenth International Conference on Information, Intelligence, Systems and Applications (IISA 2022). 2022
10. Alexakis K., Kapsalis P., Mylona Z., Kormpakis G., **Karakolis E.**, Ntanos C., Askounis D. Intelligent Querying for Implementing Building Aggregation Pipelines. The Thirteenth International Conference on Information, Intelligence, Systems and Applications (IISA 2022). 2022
11. Kormpakis G, Kapsalis P., Alexakis K., Pelekis S., **Karakolis E.**, Doukas H. An Advanced Visualisation Engine with Role-Based Access Control for Building Energy Visual Analytics. The Thirteenth International Conference on Information, Intelligence, Systems and Applications (IISA 2022). 2022
12. **Karakolis E.**, Kapsalis P., Skalidakis S., Kontzinos C., Kokkinakos P., Markaki O., Vlachou S., Psarras J. (2020). QualiChain Curriculum Designer – Providing Intelligent Recommendations for Curriculum Restructuring Based on Labour Market Needs. 13th International Conference on Education, Research and Innovation (ICERI), Seville, Spain.
13. **Karakolis E.**, Kapsalis P., Skalidakis S., Kontzinos C., Kokkinakos P., Markaki O., Vlachou S., Psarras J. (2020). QualiChain Recommender – Providing Personalised Skill and Course Recommendations Based on Labour Market Needs. 13th International Conference on Education, Research and Innovation (ICERI), Seville, Spain.

14. Kontzinos C., Markaki O., Kokkinakos P., **Karakolis V.**, Skalidakis S., Psarras J. (2020). Using Blockchain, Semantics and Data Analytics to Optimise Qualification Certification, Recruitment and Competency Management: a Landscape Review, The 12th International Conference on Mobile, Hybrid, and On-line Learning (eLmL 2020), Valencia, Spain.
15. Kontzinos C., Markaki O., Kokkinakos P., **Karakolis V.**, Skalidakis S., Psarras J. (2020) Decentralised Qualifications' Verification and Management for Learner Empowerment, Education Reengineering and Public Sector Transformation: The QualiChain project, The 12th International Conference on Mobile, Hybrid, and On-line Learning (eLmL 2020), Valencia, Spain.
16. Kontzinos C., Markaki O., Kokkinakos P., **Karakolis V.**, Skalidakis S., Psarras J. (2019) Implementing Blockchain and Computer Intelligence for University Process Optimisation: the QualiChain case, 12th International Conference on Education, Research and Innovation (ICERI), Seville, Spain.
17. Kontzinos C., Markaki O., Kokkinakos P., **Karakolis V.**, Skalidakis S., Psarras, J. (2019) University process optimisation through smart curriculum design and blockchain-based student accreditation, 18th International Conference on WWW/Internet, Cagliari, Italy
18. Lampropoulos G., Pelekis S., **Karakolis V.**, Anastasatos G., Kormpakis G., Tsapelas I., Mouzakitis S., Marinakis V., Psarras J. I-ENERGY: Delivering the next-generation Artificial Intelligence services for the Energy Sector, 31st European Conference on Operational Research, University of West Attica, Athens, Greece (EURO 2021), 2021, pp. 278.
19. **Karakolis V.**, Kapsalis P., Kontzinos C., Skalidakis S., Vlachou V., Kokkinakos P., Markaki O., Psarras J. Combining recommender systems with multicriteria decision support to facilitate optimal course selection by learners - The QualiChain project's case, 31st European Conference on Operational Research, University of West Attica, Athens, Greece (EURO 2021), 2021, pp. 108
20. Anastasatos G., **Karakolis V.**, Marinakis V., Doukas H. AI models, Visual Analytics and Querying for Big Data in the Energy Sector, 31st European Conference on Operational Research, University of West Attica, Athens, Greece (EURO 2021), 2021, pp. 304
21. Mouzakitis, S., Markaki, O., Papapostolou, K., **Karakolis, E.**, Pelekis, S., & Psarras, J. (2023). Enhancing decision support systems for the energy sector with sustainable artificial intelligence solutions. Intelligent Systems Conference (IntelliSys) 2023.

### I.3 Υποβληθείσες εργασίες για τις οποίες εκκρεμεί η απόφαση

1. Tzortzis, A. M., Pelekis, S., Spiliotis, E., **Karakolis, E.**, Mouzakitis, S., Psarras, J. & Askounis, D. (2023). Transfer learning for day-ahead load forecasting: A case study for European national electricity demand time series. Υποβλήθηκε στο διεθνές επιστημονικό περιοδικό Mathematics.

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



## ΠΑΡΑΡΤΗΜΑ ΙΙ: ΥΠΗΡΕΣΙΕΣ ΤΗΣ ΠΛΑΤΦΟΡΜΑΣ QUALICHAIN ΣΤΟ GITHub ΚΑΙ ΑΛΛΕΣ ΥΠΗΡΕΣΙΕΣ ΛΟΓΙΣΜΙΚΟΥ ΠΟΥ ΑΝΑΠΤΥΧΘΗΚΑΝ ΣΤΗΝ ΠΑΡΟΥΣΑ ΔΙΑΤΡΙΒΗ

Υπηρεσία	GitHub Repository	Ενότητα
Identity and Access Control Management	<a href="https://github.com/QualiChain/identity-access-management">https://github.com/QualiChain/identity-access-management</a>	4.2.1
Knowledge Graph Engine – DeTrusty	<a href="https://github.com/QualiChain/knowledge-engine">https://github.com/QualiChain/knowledge-engine</a>	4.2.2 <b>Error! Reference source not found.</b>
Verification for the Blockchain-Based OpenBadges Platform	<a href="https://github.com/QualiChain/verification">https://github.com/QualiChain/verification</a>	4.2.3
Verification for the Blockchain-Based Consortium and Certificate Management Platform	<a href="https://github.com/QualiChain/consortium">https://github.com/QualiChain/consortium</a>	4.2.4
Recruiting Webapp	<a href="https://github.com/QualiChain/hei-did">https://github.com/QualiChain/hei-did</a>	4.2.6.2
Knowledge Extraction	<a href="https://github.com/QualiChain/knowledge-extraction">https://github.com/QualiChain/knowledge-extraction</a>	4.2.5
Job Post Extraction – JobWatch	<a href="https://github.com/QualiChain/knowledge-extraction/tree/master/jobwatchpy">https://github.com/QualiChain/knowledge-extraction/tree/master/jobwatchpy</a>	4.2.5.1
Job Post Extraction for Greek job portals Component – JobCrawler	<a href="https://github.com/QualiChain/JobCrawler">https://github.com/QualiChain/JobCrawler</a>	4.2.5.2
Domain Ontology-Based Information Extraction – DOB-IE	<a href="https://github.com/QualiChain/knowledge-extraction/tree/master/dobie">https://github.com/QualiChain/knowledge-extraction/tree/master/dobie</a>	4.2.5.3
Falcon	<a href="https://github.com/SDM-TIB/falcon2.0">https://github.com/SDM-TIB/falcon2.0</a>	4.2.5.4 <b>Error! Reference source not found.</b>
Course Recommendation	<a href="https://github.com/QualiChain/course_recommendation">https://github.com/QualiChain/course_recommendation</a>	4.2.6.1, 5
Curriculum Designer	<a href="https://github.com/QualiChain/curriculum_designer">https://github.com/QualiChain/curriculum_designer</a>	4.2.6.1, 6
Profiling and CV Customization	<a href="https://github.com/QualiChain/analytics">https://github.com/QualiChain/analytics</a>	4.2.6.2
Recruitment and Internal Allocation	<a href="https://github.com/QualiChain/analytics">https://github.com/QualiChain/analytics</a>	4.2.6.2
Career Advisor	<a href="https://github.com/QualiChain/analytics">https://github.com/QualiChain/analytics</a>	4.2.6.2
Degree comparison	<a href="https://github.com/QualiChain/degree-comparison">https://github.com/QualiChain/degree-comparison</a>	4.2.6.1

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

<b>Multi-Criteria Decision Support</b>	<a href="https://github.com/QualiChain/qualichain_mcdss">https://github.com/QualiChain/qualichain_mcdss</a>	4.2.6.3, 7
<b>Visualization Engine</b>	<a href="https://github.com/QualiChain/visualisation_engine">https://github.com/QualiChain/visualisation_engine</a>	4.2.7.1
<b>Front End</b>	<a href="https://github.com/QualiChain/fe">https://github.com/QualiChain/fe</a>	4.2.7.2
<b>Distributed file storage</b>	<a href="https://github.com/QualiChain/P2Cstore">https://github.com/QualiChain/P2Cstore</a>	
<b>Back end and APIs</b>	<a href="https://github.com/QualiChain/qualichain_backend">https://github.com/QualiChain/qualichain_backend</a>	
<b>QualiChain Mediator</b>	<a href="https://github.com/QualiChain/qualichain_mediator">https://github.com/QualiChain/qualichain_mediator</a>	
<b>Recommender systems optimisation for provider fairness</b>	<a href="https://github.com/vkarakolis-epu/recsys_provider_fairness_optimization">https://github.com/vkarakolis-epu/recsys_provider_fairness_optimization</a>	9

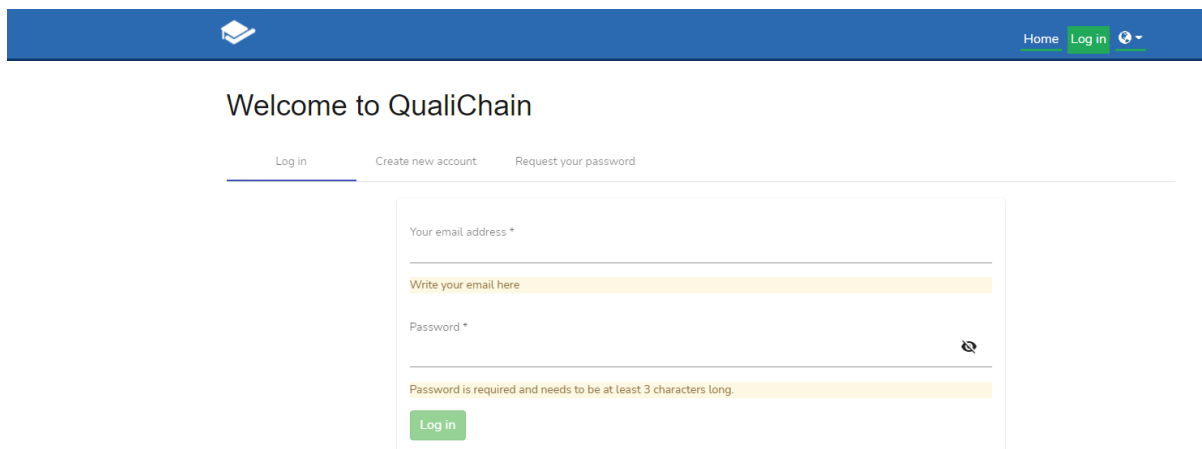
Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## ΠΑΡΑΡΤΗΜΑ ΙΙΙ: ΕΠΙΔΕΙΞΗ ΤΗΣ ΠΛΑΤΦΟΡΜΑΣ QUALICHAIN

Στο παράρτημα αυτό παρουσιάζονται παραδείγματα από όλες τις λειτουργίες της πλατφόρμας QualiChain. Σε πρώτη φάση παρουσιάζεται η αρχική σελίδα και η σελίδα εισόδου στην πλατφόρμα και η εγγραφή νέου χρήστη. Στη συνέχεια παρουσιάζονται οι λειτουργίες για τους φοιτητές, τους δια βίου μαθητές και αυτούς που αναζητούν εργασία. Ακολούθως παρουσιάζονται οι λειτουργίες για τους καθηγητές και τους διδάσκοντες μαθημάτων και τέλος παρουσιάζονται οι λειτουργίες για τους υπεύθυνους ανθρώπινου δυναμικού.



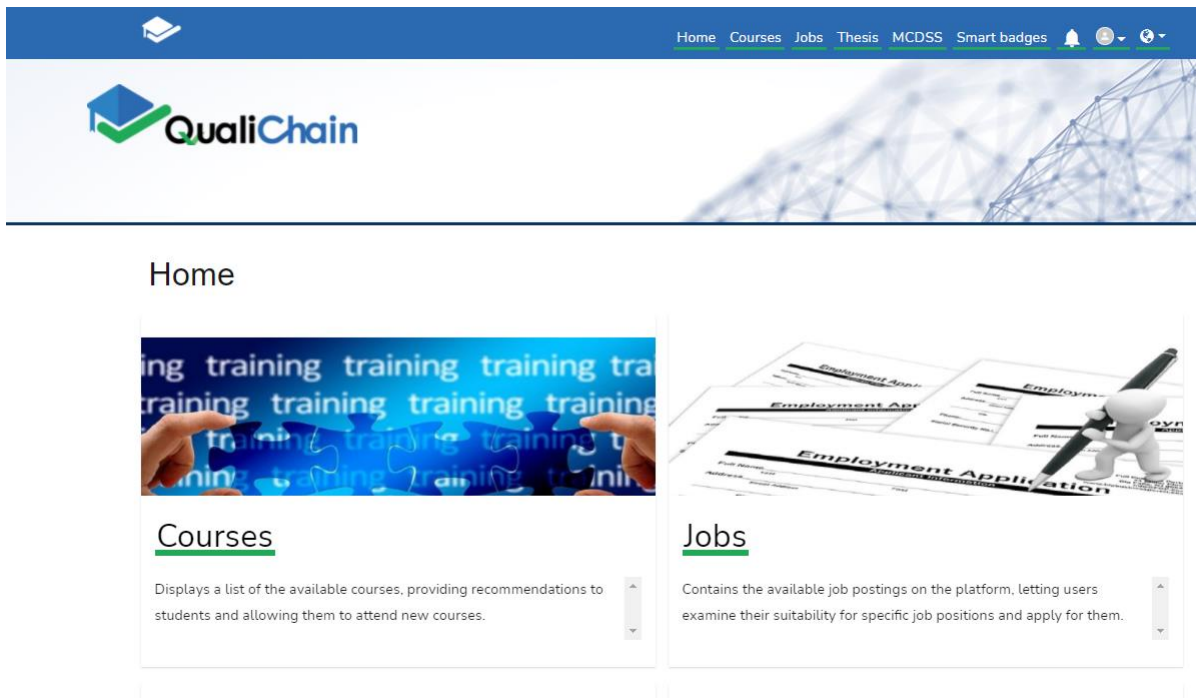
Εικόνα 46 - Αρχική σελίδα πλατφόρμας QualiChain



Εικόνα 47 - Σελίδα εισόδου ή εγγραφής στην πλατφόρμα

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

### III.1 Λειτουργίες για φοιτητές, δια βίου μαθητές και ανθρώπους που αναζητούν εργασία



Εικόνα 48 - Αρχική σελίδα για φοιτητές και δια βίου μαθητές

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

### III.1.1 Λειτουργίες επισκόπησης και επιλογής μαθήματος

#### Courses

Filter

Name	Semester	Academic Organization	Action
Forecasting Techniques	8		<a href="#">View</a>
Management of the Digital Enterprise	8		<a href="#">View</a>
Operating Systems	6		<a href="#">View</a>
Management and Decision Support Systems	7		<a href="#">View</a>
Human- Computer Interaction	7		<a href="#">View</a>

Items per page **5** 1 - 5 of 158 |< < > >|

Recommended courses

- Human- Computer Interaction
- Computer System Performance
- Digital VLSI Systems
- Management Game
- Operating Systems Laboratory

[View more](#) [MCDSS](#)

My enrolled courses

- Data Mining
- Forecasting Techniques

Course	Demand in Job Market (number of job postings)
Multimedia Technology	2.2
Computational Techniques for Information Transmission Systems	2.2
Programming Languages I	3.2
Algorithms and Complexity	3.2
Internet Programming	3.2
Databases	5.2
Data Mining	5.2

Εικόνα 49 - Επιλογή από λίστα μαθημάτων

## Course: Management of the Digital Enterprise

Enroll me to the course Set as completed

**Description:** The course covers a wide range of technical and managerial topics regarding e-commerce and e-government. The course examines topics like: e-business models and mechanisms; customer interface design; internet promotion methods and computational advertising algorithms; online search types and functional description of search engines; recommender system techniques and algorithms; business process modelling and management; Enterprise Resource planning (ERP) Systems architectures and their critical success factors; Customer Relationship Management (CRM) Systems; knowledge management approaches, systems and architectures; mobile commerce technologies and applications; e-government, e-democracy and e-participation concepts and frameworks; Web 2.0 and collective intelligence: approaches and mechanisms.

**Semester:** 8

**Related Skills:** [crm](#), [erp](#)

### Smart Badges

This course does not have any smart badge yet.

### Professors

No professors

Εικόνα 50 - Επισκόπηση μαθήματος Management of the Digital Enterprise

## Recommended courses

Filter

MCDSS

Name	Description	Score	Action
<input type="checkbox"/> Human-Computer Interaction	The course deals with the design, development and evaluation methodologies of computer systems that interact with users to a significant degree. The purpose is to review theoretical models, technologies, methods and tools for the design and development of interactive computing systems. The course includes (1) Human modeling of computer system and conceptual interaction framework, (2) Cognitive models, perception and representation, attention and memory representation and organization of knowledge(3) Interactive Systems Design Methodologies: Human-centric Design, Usability Requirements, Task Analysis, GOMS Models, Dialogue Methods, Interface Design, Usability and Accessibility of Internet Services, Interactive Systems Evaluation Techniques, (4) Overview of Interactive Technologies and Applications: Distributed Intelligence Systems, Ubiquitous Computer, Augmented Reality Systems, Synthetic Animation, Accessibility Interface Systems, Collaboration Applications.	1	<a href="#">View</a>
<input type="checkbox"/> Computer System Performance	Main approaches for modelling and analysis of computer system performance: analytical models, simulation and empirical techniques. Introduction to queueing theory, queueing networks (general product-form networks, convolution algorithms, mean value analysis), techniques based on queueing networks (operational laws, bounds, hierarchical decomposition, load-independent and load-dependent service models, non-product-form networks, exact and approximate solution algorithms, analysis of specific systems), discrete event simulation (random number generation, program construction, transient removal, statistical analysis of simulation results), measurement techniques and tools (workload, benchmarking, monitoring and capacity planning, design and analysis of experiments). Applications to the performance analysis of various computer systems with emphasis on currently established paradigms (client/server, intranets και Internet, Web services etc).	1	<a href="#">View</a>
<input type="checkbox"/> Digital VLSI Systems	Implementation of VLSI circuits for arithmetic operations. Implementation of digital filters. Circuits based on special numerical systems for high speed applications. VLSI design tools. Circuit Description Languages for automated Design. Design and implementation of digital systems.	1	<a href="#">View</a>
<input type="checkbox"/> Management Game	The scope of the course is to offer to the students a realistic picture of different real-world problems related to fields of the management science. This is achieved by utilizing a management simulation game where the students are called to apply selected	1	<a href="#">View</a>

Εικόνα 51 - Προτεινόμενα μαθήματα

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

Manual input      File input

+ Add criteria    + Add alternative    > Send data

Select method

- MAUT
- TOPSIS
- ELECTRE I
- PROMETHEE II

	Criteria name *	Weight *	Type *
	score	1	0
	difficulty	0.8	1
	state of the art	0.9	0

Alternative name *	Value *	Value *	Value *
Human- Computer Interaction	1	0.6	0.7
Computer System Performance	1	0.7	0.5
Digital VLSI Systems	1	0.9	0.8

Εικόνα 52 - Σελίδα πολυκριτηριακής υποστήριξης αποφάσεων

ⓘ

← Back    Compare

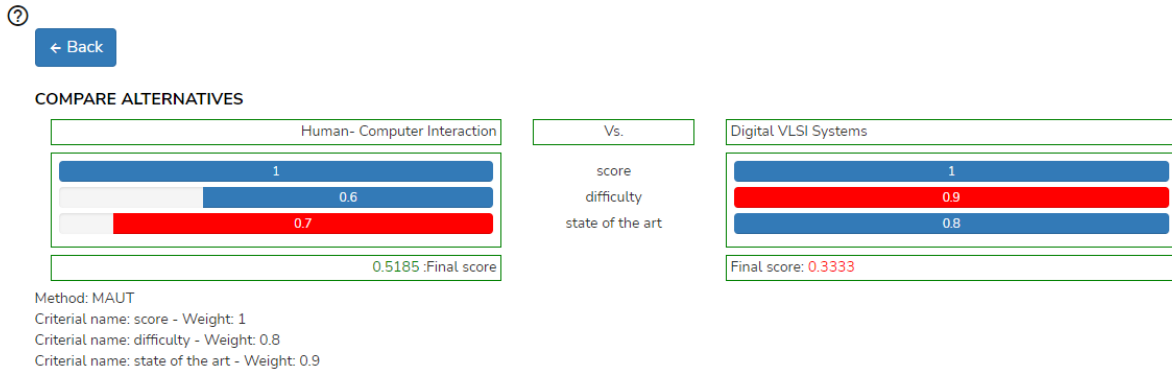
**RESULTS**

Alternative	Ranking	Score
<input type="checkbox"/> Human- Computer Interaction	1	0.5185
<input type="checkbox"/> Digital VLSI Systems	2	0.3333
<input type="checkbox"/> Computer System Performance	3	0.1975

Method: MAUT  
 Criterial name: score - Weight: 1  
 Criterial name: difficulty - Weight: 0.8  
 Criterial name: state of the art - Weight: 0.9

Εικόνα 53 - Αποτελέσματα υπηρεσίας πολυκριτηριακής υποστήριξης αποφάσεων

## Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 54 - Σελίδα σύγκρισης εναλλακτικών επιλογών

### III.1.2 Λειτουργίες επισκόπησης και αίτησης σε αγγελίες εργασίας

#### Jobs

by title      Submission start date (>=)

Filter: \_\_\_\_\_ by organization ▼ \_\_\_\_\_

Submission end date (<=) \_\_\_\_\_

Title	Employment type	Seniority Level	Recruitment organisation	Submission start date	Submission end date	Action
Java junior Back end software engineer	Part-time	Entry level	ASEP	May 1, 2021	Aug 1, 2023	<a href="#">View</a>
Senior Javascript Backend Software Engineer	Full-time	Director		May 10, 2021	Aug 1, 2023	<a href="#">View</a>
Senior Java Backend Software Engineer	Full-time	Director		Oct 6, 2021	Aug 1, 2023	<a href="#">View</a>
Java Back end software engineer	Full-time	Mid-Senior level		May 10, 2021	Aug 1, 2023	<a href="#">View</a>
Senior Frontend Developer	Full-time	Internship	ASEP			<a href="#">View</a>

Recommended jobs ^

- Hardware Engineer ▼
- Psicologia Clínica ▼
- 3E\_Test\_Nanouris ▼
- first Job ▼
- IT services ▼

[View more](#)

Job applications ^

- Java Back end software engineer ▼

Εικόνα 55 - Διαθέσιμες αγγελίες εργασίας



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## Job: EEP 3E/2017

[Apply for this job](#)

**Description:** COMPUTER SCIENCE  
**Specialization:** Web Developer  
**Submission start date:** Nov 25, 2021  
**Submission end date:** Dec 15, 2021  
**Employment type:** Full-time  
**Seniority Level:** Entry level  
**Recruitment organisation:** ASEP  
**Country:** GREECE  
**State:** ATTIKI  
**City:** ATHENS

### Required Skills

#	Name	Priority	Level
1	Java (computer programming)	High	Expert

### Required Degrees

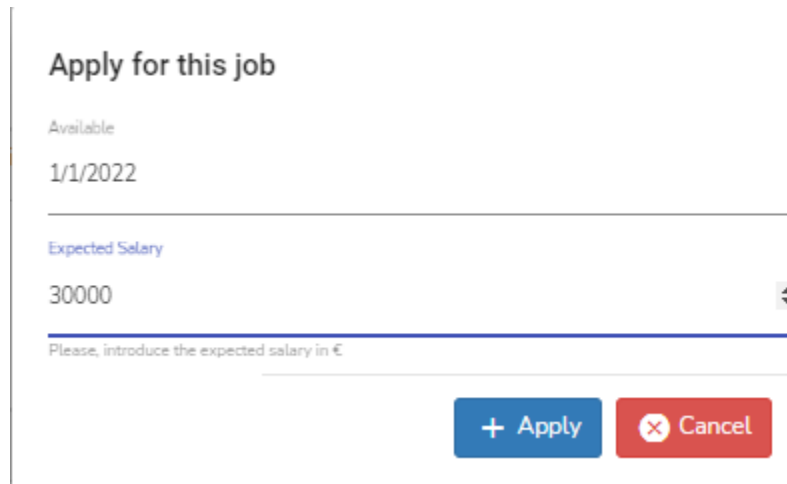
#	Title	Description
1	BSc	computer science
2	MSc	computer science
3	PhD	computer science
4	CERTIFICATE OF PROFICIENCY IN ENGLISH	C1 ENGLISH LANGUAGE

### Required Experience

#	Position	Description	Duration
1	PhD	JAVA frameworks (ADF etc) developer, and/or as web services developer and/or as an Oracle Weblogic Application Server Administrator and/or in IT Project Management with emphasis in E-government	ONE (1) YEAR
2	MSc	JAVA frameworks (ADF etc) developer, and/or as web services developer and/or as an Oracle Weblogic Application Server Administrator and/or in IT Project Management with emphasis in E-government	THREE (3) YEARS

Εικόνα 56 - Επισκόπηση συγκεκριμένης αγγελίας εργασίας

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



**Apply for this job**

Available  
1/1/2022

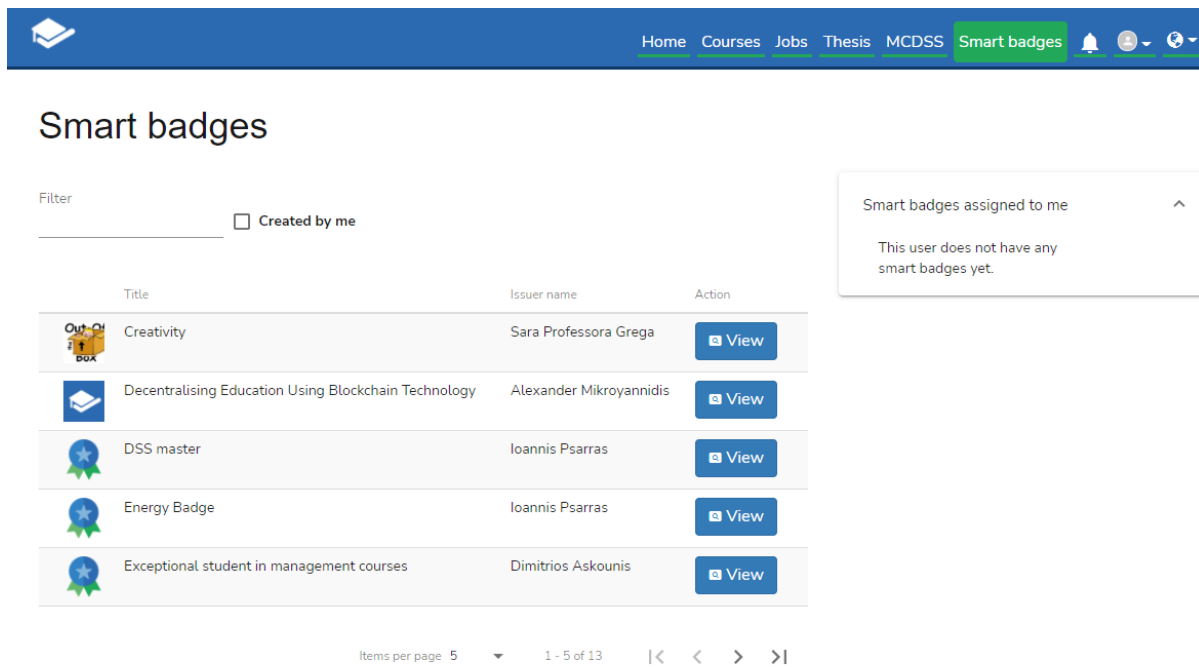
Expected Salary  
30000




Please, introduce the expected salary in €

[+ Apply](#) [Cancel](#)

Εικόνα 57 - Αίτηση για θέση εργασίας






### III.1.3 Επισκόπηση και επιδότηση με smart badges



Home Courses Jobs Thesis MCDSS **Smart badges**   

## Smart badges

Filter  Created by me

Title	Issuer name	Action
 Creativity	Sara Professora Grega	<a href="#">View</a>
 Decentralising Education Using Blockchain Technology	Alexander Mikroyannidis	<a href="#">View</a>
 DSS master	Ioannis Psarras	<a href="#">View</a>
 Energy Badge	Ioannis Psarras	<a href="#">View</a>
 Exceptional student in management courses	Dimitrios Askounis	<a href="#">View</a>

Items per page 5 1 - 5 of 13 |< < > >|




Smart badges assigned to me

This user does not have any smart badges yet.

Εικόνα 58 - Λίστα από διαθέσιμα smart badges

## My colleagues

Filter

Name	Surname	Organizations	Smart Badges	Action
Vagelis Karakolis	Vagelis Karakolis	NTUA	  	<a href="#">Award</a>
Dimitris Strotman	Dimitris Strotman	NTUA	This user does not have any smart badges yet.	<a href="#">Award</a>
Dimitris Strotopoulos	Dimitris Strotopoulos	NTUA	This user does not have any smart badges yet.	<a href="#">Award</a>
Deriziotis Panagiotis	Deriziotis Panagiotis	NTUA	This user does not have any smart badges yet.	<a href="#">Award</a>
Domingo Mendivil	Domingo Mendivil	NTUA	This user does not have any smart badges yet.	<a href="#">Award</a>

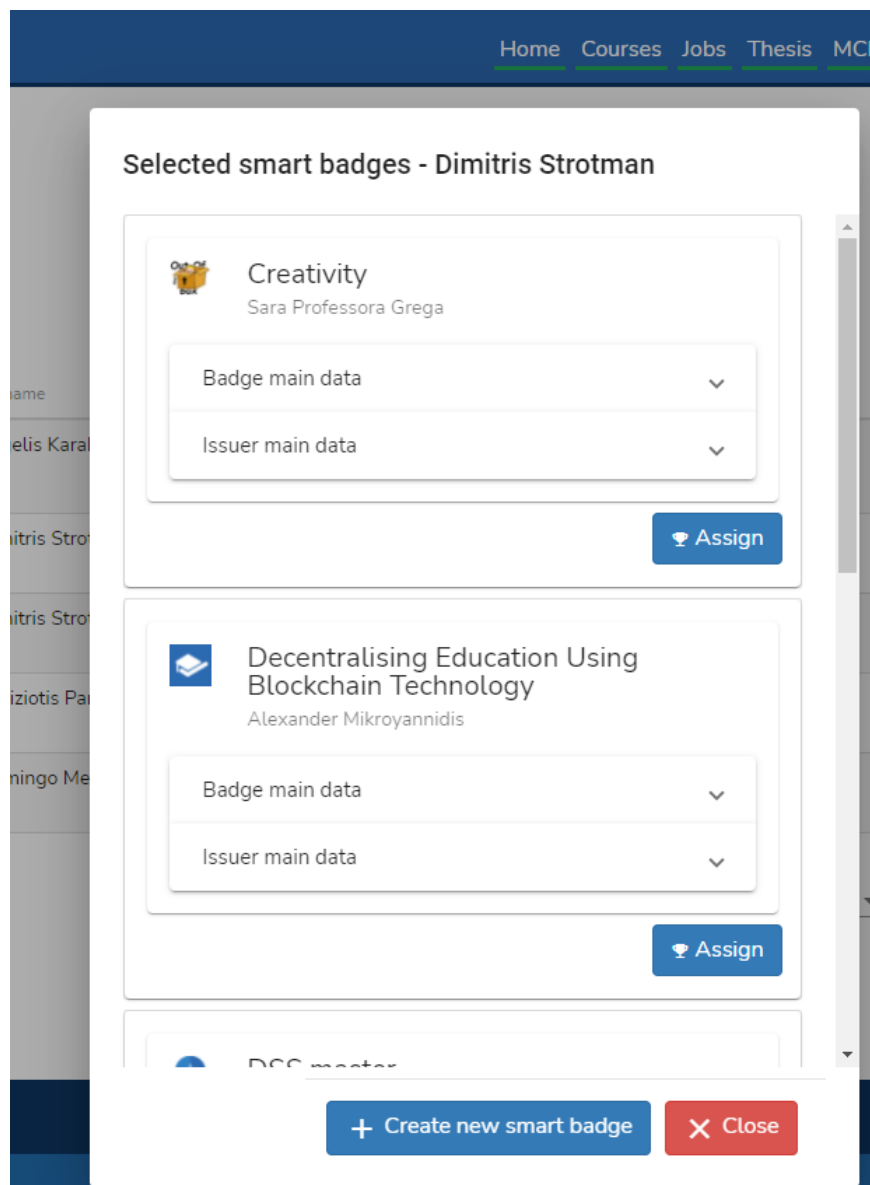
Items per page 5

1 - 5 of 84

[|<](#) [<](#) [>](#) [>|](#)

Εικόνα 59 - Σελίδα με συναδέλφους για επίδοση smart badge

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Εικόνα 60 - Φόρμα για επιδότηση smart badge

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

### III.1.4 Λειτουργία προφίλ

The screenshot displays a user profile page with a blue header containing navigation links: Home, Courses, Jobs, Thesis, MCDSS, Smart badges, and user icons. The main content area is titled "My profile" and includes several sections:

- Profile Information:** Username: Vagelis Karakolis Student, Name: Vagelis Karakolis Student, Surname: Vagelis Karakolis Student, Email: vk\_student@epu.ntua.gr, Role: Student, Internal relocation availability: No. Includes buttons for "Open PDF", "Edit", and "Delete own profile".
- CV Section:** A "CV" button and a "Best Career Options" button. Below are tabs for "Career path", "Completed Courses", "Smart Badges", and "Competency Development". A "Career Advisor" button is located below the tabs.
- CV Completeness:** A circular progress indicator and a "CV completeness" label with a question mark.
- My enrolled courses:** A list of courses including "Data Mining" and "Forecasting Techniques".
- My thesis:** A dropdown menu.
- Recommended jobs:** A dropdown menu.
- Recommended courses:** A list of courses with checkboxes: "Human- Computer Interaction", "Computer System Performance", "Digital VLSI Systems", and "Management Game".

Εικόνα 61 - Η σελίδα προφίλ για έναν χρήστη

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

CV - Vagelis Karakolis Student

Title  
Software Engineer

Description  
Senior Backend Software Engineer

Target sector  
Backend Developer

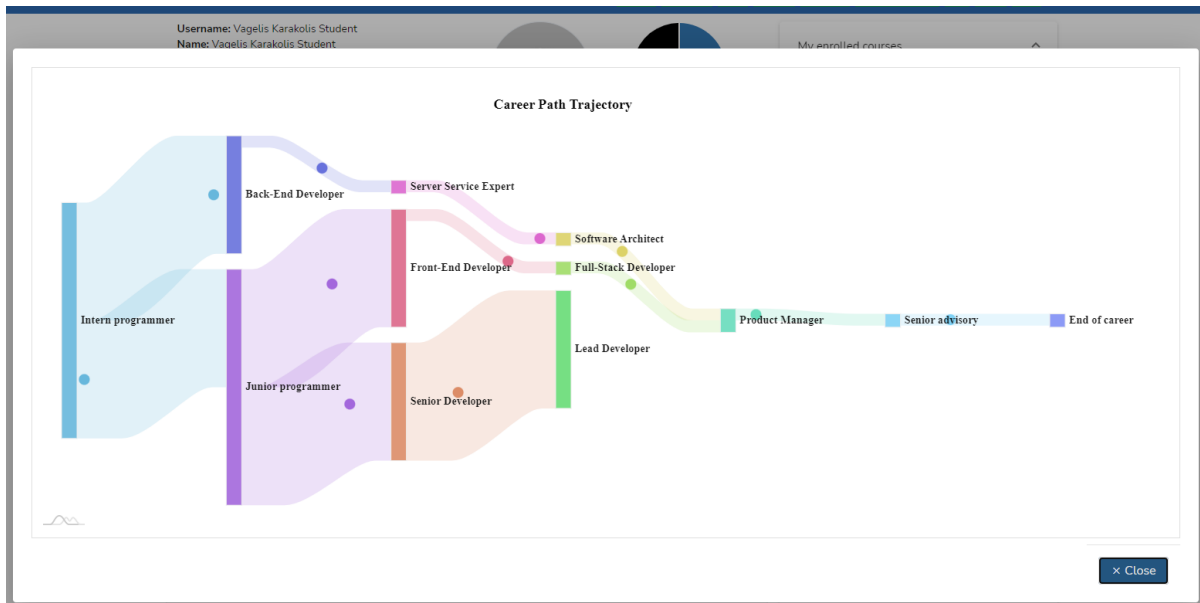
Skills

Skill 1	Delete
<p>Label + <b>Java (computer programming)</b></p> <p>Proficiency level Advanced</p> <p>Comment</p>	<input type="button" value="Delete"/>

Εικόνα 62 - Φόρμα δημιουργίας βιογραφικού

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

### III.1.5 Συμβουλευτική καριέρας

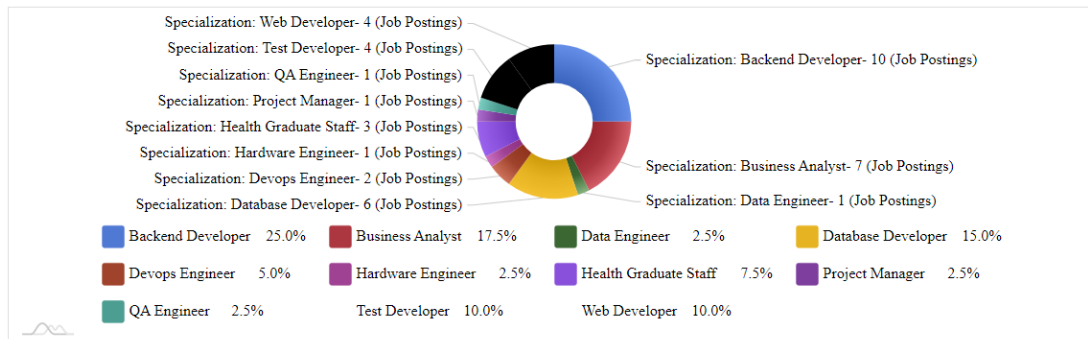


Εικόνα 63 - Σχεδιασμός προσωποποιημένου μονοπατιού καριέρας (career path)



### Career Advisor

[← Back to user profile](#)



### Targeted Job Titles

Please, type targeted job titles...

Εικόνα 64 - Σελίδα συμβουλευτικής καριέρας

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

[Competency Development](#)

[+ Create](#)

Filter

Competency	Level	Progress	Acquired date	Next evaluation	Action
Java (computer programming)	Advanced	0			<a href="#">Edit</a>
database management systems	Advanced	0			<a href="#">Edit</a>
JavaScript	Medium	0			<a href="#">Edit</a>

Items per page **5** 1 - 3 of 3 [|<](#) [<](#) [>](#) [>|](#)

Εικόνα 65 - Σελίδα ανάπτυξης δεξιοτήτων

## III.2 Λειτουργίες για καθηγητές και διδάσκοντες μαθημάτων

### III.2.1 Επισκόπηση και τροποποίηση μαθήματος

#### Courses

[+ Create Course](#) [Export to Excel](#)

Filter

Teaching Courses  
Data Mining

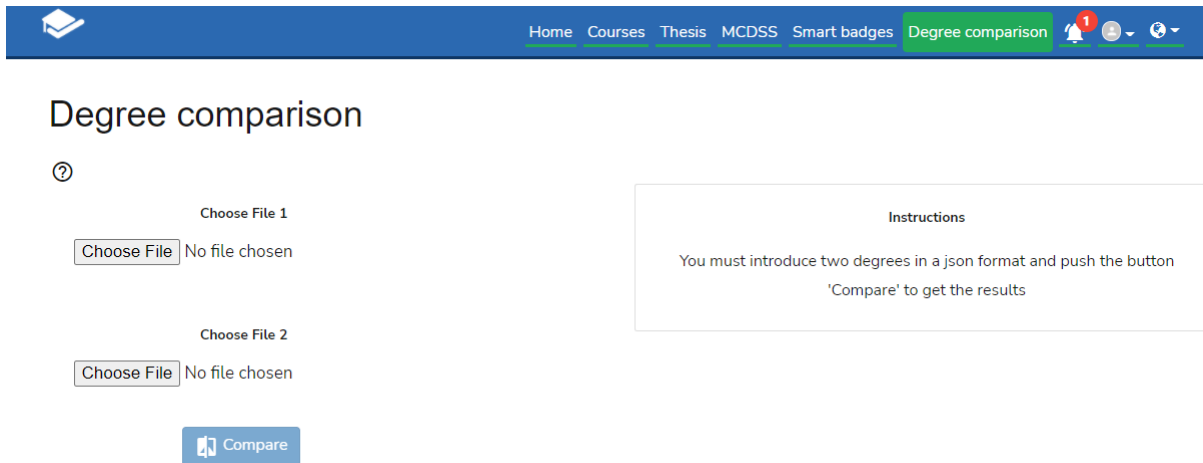
Name	Semester	Academic Organization	Action
ALTC 2021 Pre-conference Workshop			<a href="#">View</a>
Using the QualiChain platform			<a href="#">View</a>
Data Mining	10		<a href="#">View</a> <a href="#">Edit</a> <a href="#">Delete</a>
Production and Operations Management	7	NTUA	<a href="#">View</a>
Utilização de Testes Psicológicos na Gestão de Recursos Humanos			<a href="#">View</a>



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

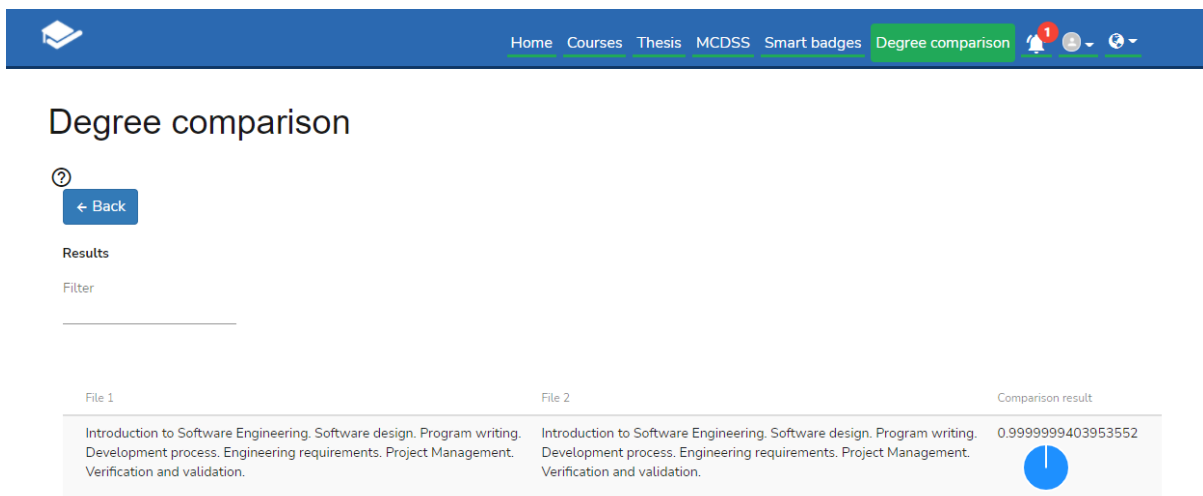
Εικόνα 66 - Σελίδα επισκόπησης μαθημάτων για τους καθηγητές

### III.2.2 Σύγκριση διπλωμάτων (Degree Comparison)



The screenshot shows the 'Degree comparison' page. At the top, there is a navigation bar with links for Home, Courses, Thesis, MCDSS, Smart badges, and Degree comparison. Below the navigation bar, the page title 'Degree comparison' is displayed. There are two file upload sections: 'Choose File 1' and 'Choose File 2', each with a 'Choose File' button and the text 'No file chosen'. A 'Compare' button is located below these sections. To the right, an 'Instructions' box states: 'You must introduce two degrees in a json format and push the button 'Compare' to get the results'.

Εικόνα 67 - Σελίδα σύγκρισης διπλωμάτων



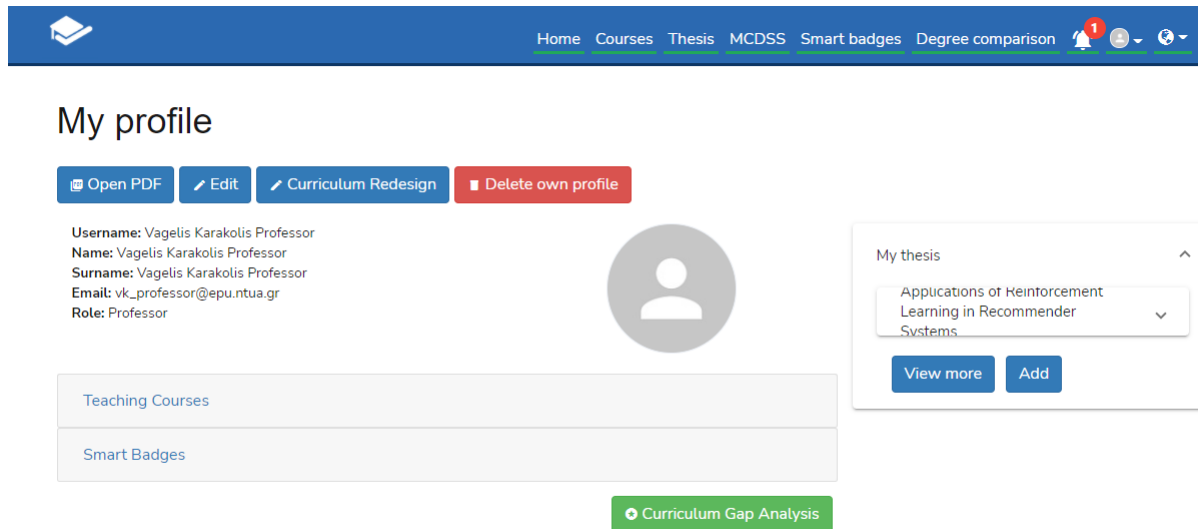
The screenshot shows the 'Degree comparison' page with the results of a comparison. A 'Back' button is visible. The 'Results' section shows a table with three columns: 'File 1', 'File 2', and 'Comparison result'. Both 'File 1' and 'File 2' contain the text: 'Introduction to Software Engineering. Software design. Program writing. Development process. Engineering requirements. Project Management. Verification and validation.' The 'Comparison result' column shows the value '0.9999999403953552' and a blue circular progress indicator.

File 1	File 2	Comparison result
Introduction to Software Engineering. Software design. Program writing. Development process. Engineering requirements. Project Management. Verification and validation.	Introduction to Software Engineering. Software design. Program writing. Development process. Engineering requirements. Project Management. Verification and validation.	0.9999999403953552

Εικόνα 68 - Αποτέλεσμα υπηρεσίας σύγκρισης διπλωμάτων

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

### III.2.3 Λειτουργία προφίλ



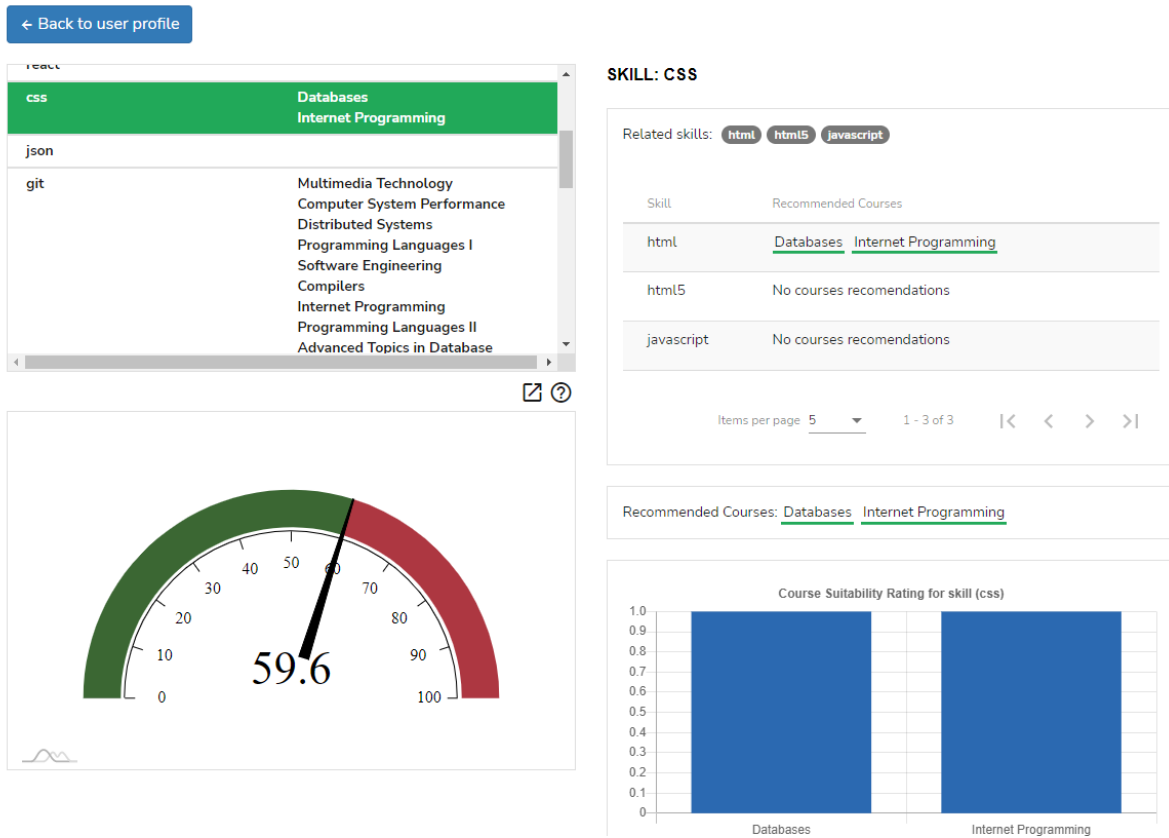
The screenshot shows a user profile page with a blue header. The header contains a logo on the left and navigation links: Home, Courses, Thesis, MCDSS, Smart badges, Degree comparison, and a notification bell with a red '1'. The main content area is titled 'My profile'. Below the title are four buttons: 'Open PDF', 'Edit', 'Curriculum Redesign', and 'Delete own profile'. The user's details are listed: Username: Vagelis Karakolis Professor, Name: Vagelis Karakolis Professor, Surname: Vagelis Karakolis Professor, Email: vk\_professor@epu.ntua.gr, and Role: Professor. A circular profile picture placeholder is shown. To the right, a 'My thesis' section displays a dropdown menu with 'Applications of Reinforcement Learning in Recommender Systems' selected, and 'View more' and 'Add' buttons. Below the profile details are two expandable sections: 'Teaching Courses' and 'Smart Badges'. At the bottom right, there is a green button labeled 'Curriculum Gap Analysis'.

Εικόνα 69 - Σελίδα για το προφίλ καθηγητή ή διδάσκοντα μαθημάτων

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

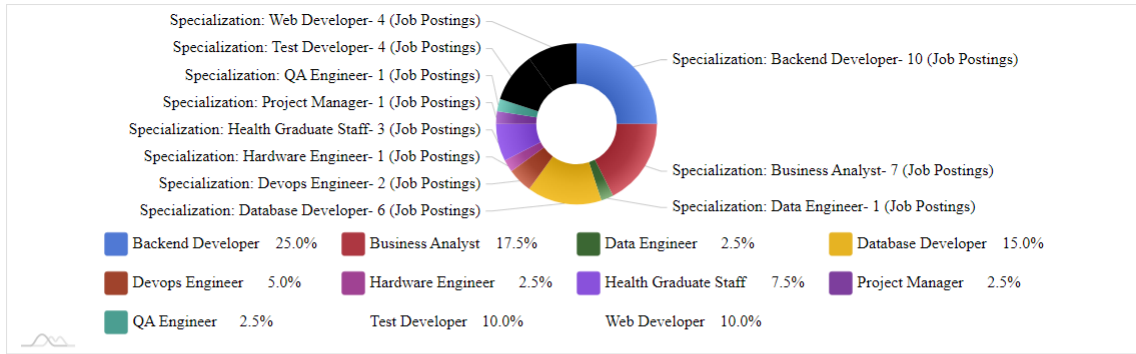
### III.2.4 Λειτουργία ανάλυσης και ανασχεδιασμού προγράμματος σπουδών

#### Curriculum Redesign

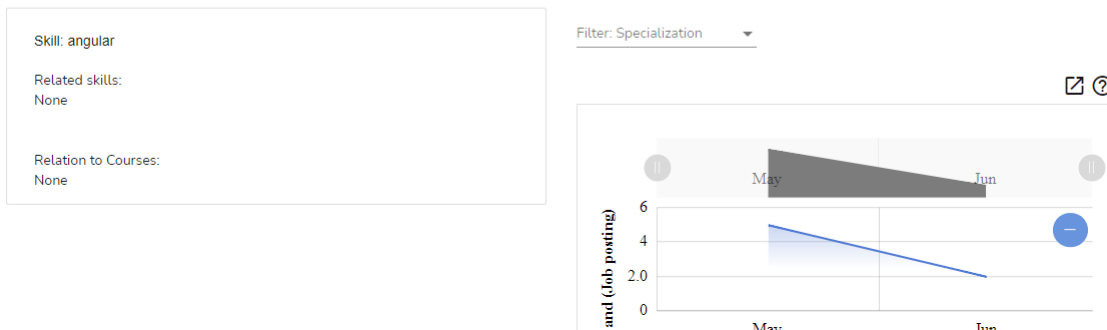


Εικόνα 70 - Η σελίδα ανασχεδιασμού προγράμματος σπουδών

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



### Currently Missing Skills

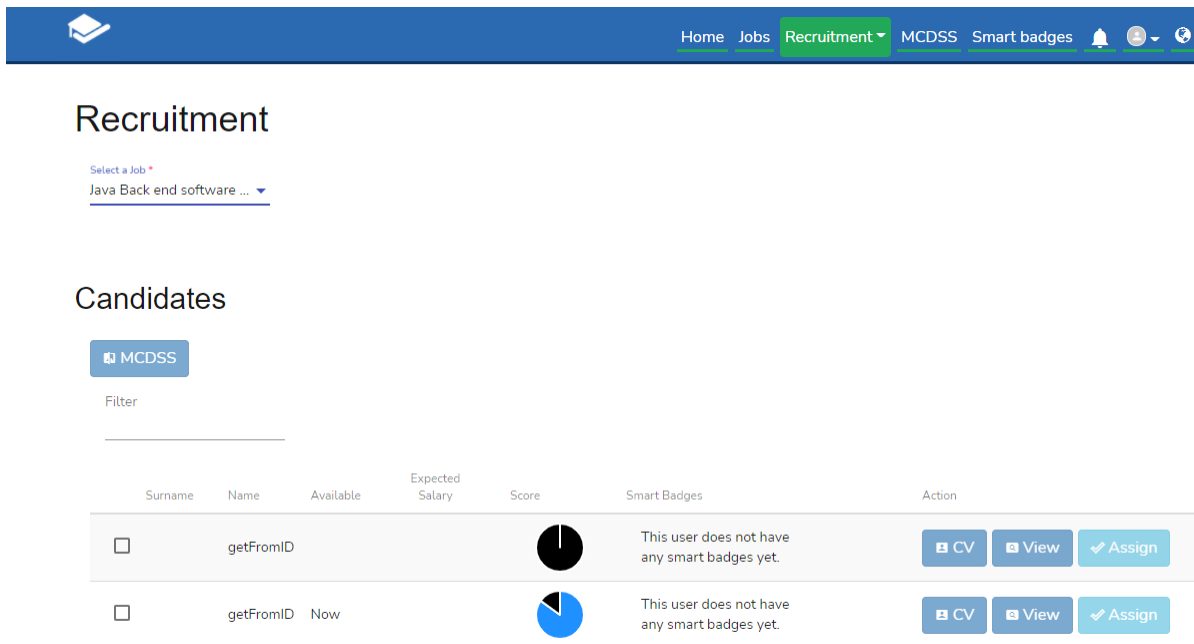


Εικόνα 71 - Η σελίδα ανάλυσης και εύρεσης δεξιοτήτων που δε προσφέρονται στο πρόγραμμα σπουδών

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

### III.3 Λειτουργίες για υπεύθυνους ανθρώπινου δυναμικού

#### III.3.1 Επιλογή υποψηφίου για θέση εργασίας

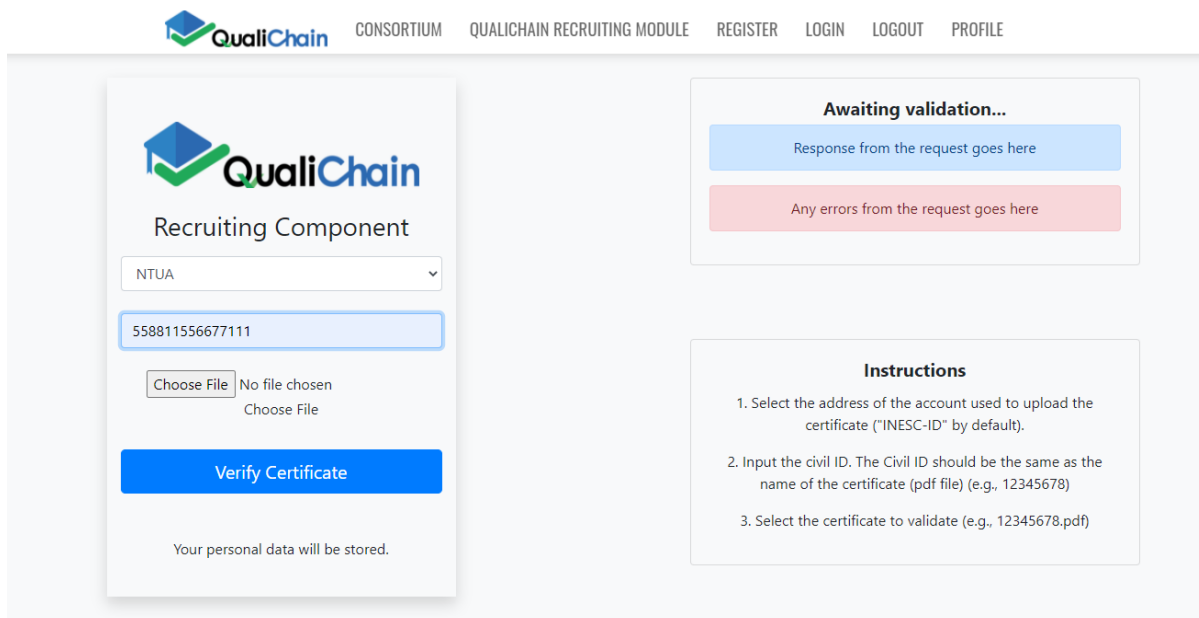


The screenshot shows a web application interface for recruitment. At the top, there is a navigation bar with a logo and menu items: Home, Jobs, Recruitment (highlighted), MCDSS, Smart badges, and user profile icons. Below the navigation bar, the page title is "Recruitment". A dropdown menu shows "Select a Job" with "Java Back end software ..." selected. The main section is titled "Candidates" and includes a filter button for "MCDSS". Below the filter, there is a table of candidates with columns: Surname, Name, Available, Expected Salary, Score, Smart Badges, and Action. Two candidates are listed, both with the name "getFromID". The first candidate has a score of 0 and no smart badges. The second candidate has a score of approximately 25% and no smart badges. Each candidate row has a checkbox, a CV icon, a View icon, and an Assign icon.

Surname	Name	Available	Expected Salary	Score	Smart Badges	Action
<input type="checkbox"/>	getFromID			0	This user does not have any smart badges yet.	<a href="#">CV</a> <a href="#">View</a> <a href="#">Assign</a>
<input type="checkbox"/>	getFromID	Now		25%	This user does not have any smart badges yet.	<a href="#">CV</a> <a href="#">View</a> <a href="#">Assign</a>

Εικόνα 72 - Σελίδα επιλογής υποψηφίου

#### III.3.2 Έλεγχος εγκυρότητας ακαδημαϊκού τίτλου



The screenshot shows the "QualiChain" web application interface. The top navigation bar includes the QualiChain logo, "CONSORTIUM", "QUALICHAIN RECRUITING MODULE", and links for "REGISTER", "LOGIN", "LOGOUT", and "PROFILE". The main content area is titled "Recruiting Component" and features a form for certificate verification. The form includes a dropdown menu for "NTUA", a text input field containing "558811556677111", a "Choose File" button, and a "Verify Certificate" button. Below the form, a message states "Your personal data will be stored." To the right of the form, there is a section titled "Awaiting validation..." with two colored boxes: a blue box for "Response from the request goes here" and a red box for "Any errors from the request goes here". Below this, there is an "Instructions" section with three numbered steps: 1. Select the address of the account used to upload the certificate ("INESC-ID" by default). 2. Input the civil ID. The Civil ID should be the same as the name of the certificate (pdf file) (e.g., 12345678). 3. Select the certificate to validate (e.g., 12345678.pdf).

Εικόνα 73 - Σελίδα ελέγχου εγκυρότητας ακαδημαϊκού τίτλου

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

The screenshot shows the QualiChain Recruiting Component interface. At the top, there is a navigation bar with the QualiChain logo and links for CONSORTIUM, QUALICHAIN RECRUITING MODULE, REGISTER, LOGIN, LOGOUT, and PROFILE. The main content area is divided into two columns. The left column contains the 'Recruiting Component' form, which includes a dropdown menu set to 'NTUA', a text input field containing '55881155667711', a file upload button labeled 'Choose File' with a file named '55881155667711.pdf', and a blue 'Verify Certificate' button. Below the button, it states 'Your personal data will be stored.' The right column contains two panels. The top panel, titled 'Done | NTUA | Error', displays a blue error message: 'Verification failure. The hash of the provided document does not correspond to the stored hash'. Below this is a red box containing the uploaded document hash and the stored hash, showing they differ. The bottom panel, titled 'Instructions', lists three steps: 1. Select the address of the account used to upload the certificate ('INESC-ID' by default). 2. Input the civil ID. The Civil ID should be the same as the name of the certificate (pdf file) (e.g., 12345678). 3. Select the certificate to validate (e.g., 12345678.pdf).

Εικόνα 74 - Παράδειγμα άκυρου ακαδημαϊκού τίτλου

The screenshot shows the QualiChain Recruiting Component interface. At the top, there is a navigation bar with the QualiChain logo and links for CONSORTIUM, QUALICHAIN RECRUITING MODULE, REGISTER, LOGIN, LOGOUT, and PROFILE. The main content area is divided into two columns. The left column contains the 'Recruiting Component' form, which includes a dropdown menu set to 'INESC-ID', a text input field containing '55881155667711', a file upload button labeled 'Choose File' with a file named '55881155667711.pdf', and a blue 'Verify Certificate' button. Below the button, it states 'Your personal data will be stored.' The right column contains two panels. The top panel, titled 'Done | INESC-ID | Success', displays a blue success message: 'Successful certificate verification - Certificate hash is: 0x11afca4b63beb5717820de60f95aea708b36ac6d3e2545c43d42ee2c4e12d8af'. Below this is a red box. The bottom panel, titled 'Instructions', lists three steps: 1. Select the address of the account used to upload the certificate ('INESC-ID' by default). 2. Input the civil ID. The Civil ID should be the same as the name of the certificate (pdf file) (e.g., 12345678). 3. Select the certificate to validate (e.g., 12345678.pdf).

Εικόνα 75 - Παράδειγμα έγκυρου ακαδημαϊκού τίτλου

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

The screenshot shows the QualiChain Consortium Component interface. At the top, there is a navigation bar with the QualiChain logo and menu items: CONSORTIUM, QUALICHAIN RECRUITING MODULE, REGISTER, LOGIN, LOGOUT, and PROFILE. The main content area is divided into two columns. The left column contains the QualiChain logo, the text 'Consortium Component', a text input field for 'CivillID (e.g., 12345678)', a 'Choose File' button, and a blue 'Upload Certificate' button. Below the button, it states 'Your personal data will be stored.' The right column contains two boxes. The top box is titled 'Awaiting uploading...' and has two sub-boxes: a light blue one for 'Response from the request goes here' and a light red one for 'Any errors from the request goes here'. The bottom box is titled 'Instructions' and contains two numbered steps: 1. Input the civil ID. The Civil ID should be the same as the name of the certificate (pdf file) (e.g., 12345678) and 2. Select the certificate to upload (e.g., 12345678.pdf).

Εικόνα 76 - Σελίδα για ανέβασμα ακαδημαϊκού τίτλου, ώστε να είναι επαληθεύσιμος

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας



Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## ΠΑΡΑΡΤΗΜΑ IV: ΔΕΞΙΟΤΗΤΕΣ ΠΟΥ ΕΞΑΧΘΗΚΑΝ ΚΑΙ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗΝ ΑΝΑΛΥΣΗ

A/A	string	kind	occurrences
1	Git	tool	419
2	C	tool	314
3	SQL	tool	311
4	JavaScript	tool	294
5	Linux	product	231
6	Java	tool	219
7	Python	tool	215
8	REST	tool	214
9	CSS	tool	179
10	HTML5	tool	159
11	HTML	tool	150
12	IOS	product	140
13	UX	tool	140
14	Docker	product	138
15	Akka	tool	138
16	Scripting	tool	130
17	JSON	product	125
18	PHP	tool	125
19	NoSQL	tool	122
20	Angular	product	110
21	API	tool	107
22	Jenkins	tool	105
23	Android	product	91
24	Windows	product	91
25	R	tool	88
26	React	tool	86
27	GO	product	85
28	CSS3	tool	85
29	AWS	tool	82
30	TCP	tool	77
31	Ruby	tool	76
32	Maven	tool	73
33	Ansible	tool	72
34	Azure	tool	70
35	Swift	tool	69
36	Jira	tool	68
37	Unix	product	67
38	Cassandra	product	66
39	Kafka	tool	65
40	Bash	tool	60
41	Scala	tool	58
42	Xcode	tool	57

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

---

43	HTTP	tool	56
44	Gradle	tool	56
45	TypeScript	tool	38
46	Perl	tool	36
47	PostgreSQL	tool	35
48	Selenium	tool	31
49	ERP	product	28
50	IIS	product	21
51	ETL	tool	16
52	Debian	product	12
53	Cucumber	tool	10
54	Protractor	tool	6
55	jMeter	tool	5
56	ABAP	tool	4

---

## ΠΑΡΑΡΤΗΜΑ V: ΟΙ ΔΕΞΙΟΤΗΤΕΣ ΠΟΥ ΔΕ ΔΙΔΑΣΚΟΝΤΑΙ ΣΤΑ ΠΛΑΙΣΙΑ ΤΟΥ ΔΟΘΕΝΤΟΣ ΠΡΟΓΡΑΜΜΑΤΟΣ ΣΠΟΥΔΩΝ ΑΛΛΑ ΕΙΝΑΙ ΣΕ ΥΨΗΛΗ ΖΗΤΗΣΗ ΑΠΟ ΤΗΝ ΑΓΟΡΑ ΕΡΓΑΣΙΑΣ

A/A	Δεξιότητα
1	Windows
2	Docker
3	R
4	Scripting
5	CSS3
6	API
7	React
8	CSS
9	JSON
10	Git
11	Android
12	Angular
13	TCP
14	REST
15	iOS
16	JavaScript
17	Akka
18	Go
19	HTML5
20	Linux
21	AWS
22	Jenkins
23	PHP

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

## ΠΑΡΑΡΤΗΜΑ VI: ΚΑΝΟΝΕΣ ΣΥΣΧΕΤΙΣΗΣ ΜΕΤΑΞΥ ΤΩΝ ΔΕΞΙΟΤΗΤΩΝ

Association	Antecedents	Consequents	Support	Confidence	Lift
1	CSS3	HTML5	0.046	0.965	10.78
2	HTML5, CSS	JavaScript	0.045	0.964	5.826
3	HTML, JavaScript	CSS	0.042	0.872	8.658
4	Jenkins	Git	0.051	0.857	3.635
5	CSS	JavaScript	0.083	0.821	4.963
6	Android	IOS	0.041	0.802	10.182
7	Docker	Git	0.062	0.797	3.381
8	HTML5	JavaScript	0.066	0.736	4.448
9	HTML, CSS	JavaScript	0.042	0.728	4.401
10	JSON	REST	0.051	0.728	6.045
11	HTML	CSS	0.058	0.687	6.817
12	HTML5, JavaScript	CSS	0.045	0.684	6.788
13	JSON	Git	0.048	0.68	2.884
14	Akka	UX	0.051	0.659	8.37
15	UX	Akka	0.051	0.65	8.37
16	REST	Git	0.078	0.65	2.755
17	NoSQL	Git	0.044	0.648	2.746
18	NoSQL	Python	0.041	0.598	4.945
19	CSS	HTML	0.058	0.575	6.817
20	HTML	JavaScript	0.048	0.573	3.465
21	JavaScript, CSS	HTML5	0.045	0.544	6.082
22	HTML5	CSS	0.047	0.522	5.182
23	IOS	Android	0.041	0.521	10.182
24	HTML5	CSS3	0.046	0.516	10.782
25	CSS, JavaScript	HTML	0.042	0.51	6.044
26	HTML5	JavaScript, CSS	0.045	0.503	6.082
27	JavaScript	CSS	0.083	0.5	4.963
28	HTML	CSS, JavaScript	0.042	0.5	6.044
29	CSS	HTML5	0.047	0.46	5.182
30	Python	C	0.056	0.46	2.606
31	HTML5	Git	0.041	0.459	1.947
32	Java	SQL	0.055	0.447	2.557
33	CSS	HTML5, JavaScript	0.045	0.447	6.788
34	Python	Git	0.052	0.433	1.834
35	REST	JSON	0.051	0.425	6.045
36	CSS	HTML, JavaScript	0.042	0.419	8.658
37	REST	JavaScript	0.049	0.407	2.457
38	CSS	Git	0.041	0.402	1.706
39	JavaScript	HTML5	0.066	0.398	4.448
40	JavaScript	Git	0.065	0.391	1.659
41	Java	Git	0.043	0.352	1.49
42	Python	NoSQL	0.041	0.34	4.945

Συστήματα Συστάσεων για τη γεφύρωση του χάσματος δεξιοτήτων μεταξύ αγοράς εργασίας και εκπαίδευσης στον τομέα της τεχνολογίας

43	Git	REST	0.078	0.332	2.755
44	Java	JavaScript	0.04	0.329	1.988
45	C	Python	0.056	0.315	2.606
46	SQL	Java	0.055	0.315	2.557
47	JavaScript	REST	0.049	0.296	2.457
48	SQL	C	0.052	0.296	1.674
49	C	SQL	0.052	0.293	1.674
50	JavaScript	HTML	0.048	0.293	3.465
51	Git	JavaScript	0.065	0.274	1.659
52	JavaScript	HTML5, CSS	0.045	0.272	5.826
53	JavaScript	SQL	0.044	0.269	1.535
54	C	Git	0.047	0.264	1.121
55	Git	Docker	0.062	0.263	3.381
56	JavaScript	HTML, CSS	0.042	0.255	4.401
57	SQL	JavaScript	0.044	0.254	1.535
58	JavaScript	Java	0.041	0.245	1.987
59	Git	Python	0.052	0.222	1.835
60	Git	Jenkins	0.051	0.215	3.635
61	Git	JSON	0.048	0.203	2.884