

ΣΗΜΑΣΙΟΛΟΓΙΚΗ ΑΝΑΖΗΤΗΣΗ ΣΤΙΣ ΒΙΒΛΙΟΘΗΚΕΣ: ΠΡΟΚΛΗΣΕΙΣ ΚΑΙ ΔΥΝΑΤΟΤΗΤΕΣ ΓΙΑ ΒΕΛΤΙΩΣΗ

Αλέξανδρος Ζ. Σπυρόπουλος^{1,2*}, Ανθούλα Μπάλιου², Χαράλαμπος Μπράτσας³, Εμμανουήλ Γαρουφάλλου², Βασίλειος Τσιάντος¹

¹ Τμήμα Φυσικής, Σχολή Θετικών Επιστημών, Πανεπιστημιούπολη Καβάλας, Διεθνές Πανεπιστήμιο της Ελλάδος (IHU), 57001, Θεσσαλονίκη, Ελλάς. daspyro@physics.ihu.gr, tsianto@physics.ihu.gr

² Τμήμα Βιβλιοθηκονομίας, Αρχειονομίας και Συστημάτων Πληροφόρησης, Σχολή Κοινωνικών Επιστημών, Αλεξάνδρεια Πανεπιστημιούπολη, Διεθνές Πανεπιστήμιο της Ελλάδος (IHU), 57400 Θεσσαλονίκη, Ελλάς. daspyro@physics.ihu.gr; anthi@uom.edu.gr; mgarou@ihu.gr

³ Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων, Αλεξάνδρεια Πανεπιστημιούπολη, Διεθνές Πανεπιστήμιο της Ελλάδος (IHU), 57400 Θεσσαλονίκη, Ελλάς. cbratsas@iee.ihu.gr

Εισαγωγή

Η σημασιολογική αναζήτηση αποτελεί ένα από τα πιο ενδιαφέροντα και δυναμικά πεδία της πληροφορικής και της τεχνητής νοημοσύνης, επικεντρώνοντας στην ανάπτυξη αλγορίθμων και τεχνολογιών που επιδιώκουν την ακριβέστερη και πιο σημασιολογική απεικόνιση των ερωτημάτων των χρηστών [1–3]. Αντίθετα με τις παραδοσιακές μεθόδους αναζήτησης που βασίζονται κυρίως στην αντιστοίχιση λέξεων και φράσεων, η σημασιολογική αναζήτηση φιλοδοξεί να κατανοήσει το "νόημα" πίσω από την πληροφορία, επιτρέποντας έτσι την ανάκτηση πιο σχετικών και ποιοτικά υψηλών αποτελεσμάτων [2,4–6]. Ενώ η πρόοδος σε αυτό το πεδίο έχει ήδη σημειώσει σημαντικές επιτυχίες, παραμένουν αδιευκρίνιστα ζητήματα και ανοιχτές προκλήσεις, ιδιαίτερα στο πλαίσιο των σύγχρονων βιβλιοθηκών, όπου η διαχείριση μεγάλων όγκων δεδομένων και η κατανόηση των αναγκών των χρηστών είναι ζωτικής σημασίας [7–11].

Οι σύγχρονες βιβλιοθήκες αντιμετωπίζουν μια σειρά από προκλήσεις που απαιτούν την αναθεώρηση των παραδοσιακών μεθόδων αναζήτησης πληροφοριών [12–14]. Η αύξηση του ψηφιακού περιεχομένου, η εισαγωγή της τεχνητής νοημοσύνης στα συστήματα πληροφόρησης και η πολυπλοκότητα των χρηστικών αναγκών, απαιτούν την εξεύρεση ενός ισορροπημένου συστήματος που να συνδυάζει την ακρίβεια με την προσβασιμότητα [15–19].

Επιπλέον, τα προβλήματα ποιότητας δεδομένων και η ανάγκη για πιο προσαρμοστικά συστήματα αναζήτησης είναι εξίσου κρίσιμα [13,16]. Οι βιβλιοθήκες αναζητούν λύσεις που θα επιτρέπουν την αποτελεσματική εξόρυξη γνώσης από τεκστολογικά δεδομένα, καθώς και τη βελτίωση των συστημάτων συστάσεων [14,17,19]. Ολοένα και περισσότερο, το ενδιαφέρον μετατοπίζεται προς την εξερεύνηση των δυνατοτήτων του σημασιολογικού ιστού στην ενίσχυση της αναζήτησης και της πλοήγησης σε πολύπλοκες πληροφοριακές συλλογές [7,15,19].

Το παρόν άρθρο βιβλιογραφικής ανασκόπησης έχει ως κεντρικό του στόχο τη συστηματική εξέταση της εξελισσόμενης πεδίου της σημασιολογικής αναζήτησης στις σύγχρονες βιβλιοθήκες. Στόχος είναι η ανάδειξη των κρίσιμων πτυχών, των μεθοδολογικών προκλήσεων και των ερευνητικών κενών που χαρακτηρίζουν το πεδίο αυτό. Ειδικότερα, το άρθρο στοχεύει στην αναγνώριση και καταγραφή των τεχνολογικών, των οργανωτικών και των ανθρωποκεντρικών διαστάσεων που επηρεάζουν την αποτελεσματικότητα των σημασιολογικών αλγορίθμων αναζήτησης και των συστημάτων συστάσεων σε βιβλιοθήκες. Δεδομένου ότι η σημασιολογική αναζήτηση είναι ένας ευρύς και πολυδιάστατος τομέας, η βιβλιογραφική ανασκόπηση αυτή θα προσπαθήσει να διευκρινίσει τις πιθανές διεπαφές μεταξύ της τεχνητής νοημοσύνης, της εξόρυξης γνώσης από τεκστολογικά δεδομένα, της ποιότητας δεδομένων και της προσαρμοστικής αναζήτησης πληροφοριών. Μέσα από αυτήν την πολυδιάστατη προσέγγιση, το άρθρο επιδιώκει να αναδείξει τα κενά στην υφιστάμενη ερευνητική βιβλιογραφία και να προτείνει πιθανές κατευθύνσεις για μελλοντική επιστημονική έρευνα. Στην εποχή της πληροφορίας, οι μηχανές

αναζήτησης στις βιβλιοθήκες έχουν καταστεί απαραίτητες για την εύρεση και την πρόσβαση σε πληροφορίες. Η αποτελεσματικότητα αυτών των συστημάτων διαδραστικής αναζήτησης καθίσταται όλο και πιο κρίσιμη καθώς οι χρήστες εξαρτώνται ολοένα περισσότερο από αυτά για την έρευνα και τη μάθηση. Σε αυτό το πλαίσιο, το παρόν έργο εξετάζει το εξής ερευνητικό ερώτημα: 'Πώς οι σύγχρονες τεχνικές βελτίωσης ερωτημάτων και η κατανόηση των συμπεριφορών των χρηστών επηρεάζουν την αποτελεσματικότητα των συστημάτων αναζήτησης σε βιβλιοθήκες;'. Μέσω αυτής της ερώτησης, το άρθρο αναδεικνύει τις πολυπλοκότητες και τις αδυναμίες των σημασιολογικών συστημάτων αναζήτησης, καθώς και τις δυνατότητες για βελτίωση των αποτελεσμάτων και της χρηστικότητας στις βιβλιοθήκες του μέλλοντος.

Ιστορική Αναδρομή

Η σημασιολογική αναζήτηση, παρά τη σχετικά πρόσφατη εμφάνιση της στο προσκήνιο της επιστημονικής κοινότητας, ενσωματώνει μια μακρά ιστορία που αποτελείται από εξελίξεις σε πολλά επιστημονικά πεδία. Οι ρίζες της βρίσκονται στη θεωρία της πληροφορίας και στη λεξικολογία, ενώ η εξέλιξη της τεχνολογίας και ιδιαίτερα του διαδικτύου, έχει συμβάλει σημαντικά στη διαμόρφωση των σημερινών της χαρακτηριστικών [20]. Στα τέλη της δεκαετίας του '90 και τα πρώτα χρόνια του 21ου αιώνα, οι πρώτες προσπάθειες για την εφαρμογή σημασιολογικών αλγορίθμων σε συστήματα αναζήτησης άρχισαν να γίνονται ορατές, συχνά υπό την ονομασία "σημασιολογικό διαδίκτυο" [21–24].

Με την περαιτέρω εξέλιξη των τεχνολογιών της τεχνητής νοημοσύνης και της εξόρυξης γνώσης από δεδομένα, η σημασιολογική αναζήτηση έχει αρχίσει να αποτελεί έναν ολοένα και πιο κρίσιμο τομέα για την εξατομίκευση της εμπειρίας του χρήστη και την αυτοματοποίηση της επεξεργασίας πληροφοριών. Επομένως, η σημασιολογική αναζήτηση είναι πλέον ένα πολυδιάστατο πεδίο που αφορά τόσο την ακαδημαϊκή έρευνα όσο και την εφαρμοσμένη τεχνολογία [21,24–26].

Η εξέλιξη των αλγορίθμων και των μεθόδων αναζήτησης είναι ένα δυναμικό και συνεχώς μεταβαλλόμενο πεδίο, ιδιαίτερα με την έξαρση των τεχνολογιών της τεχνητής νοημοσύνης και της εξόρυξης δεδομένων [27]. Από τους πρώτους αλγόριθμους ταξινόμησης βασισμένους στη συχνότητα εμφάνισης των λέξεων, μέχρι τις σύγχρονες τεχνικές της ανάλυσης συναισθημάτων και του Natural Language Processing, η διαδικασία της αναζήτησης έχει εξελιχθεί σε μια πολύπλοκη και εξειδικευμένη διαδικασία [28,29].

Ενδεικτικά, τεχνολογίες όπως τα νευρωνικά δίκτυα και οι μηχανές συστάσεων αρχίζουν να ενσωματώνονται ενεργά στα συστήματα αναζήτησης, προσφέροντας πιο εξειδικευμένες και εξατομικευμένες εμπειρίες στους χρήστες [30–32]. Η εξέλιξη αυτή ανοίγει νέα πεδία για την επιστημονική έρευνα, επιτρέποντας την περαιτέρω βελτίωση και εξειδίκευση των συστημάτων αναζήτησης, ειδικά στο πλαίσιο της σημασιολογικής ανάλυσης [32–34].

Θεωρητικό Πλαίσιο

Στην ενότητα αυτή, επικεντρωνόμαστε στην εξέταση των βασικών θεωριών και μοντέλων που υποστηρίζουν και καθοδηγούν τον τομέα της σημασιολογικής αναζήτησης. Ένα από τα πρώτα μοντέλα που αξίζει να αναφερθεί είναι το Vector Space Model (VSM) [35,36]. Αν και το VSM δημιουργήθηκε αρχικά για την αξιολόγηση εγγράφων, το μοντέλο έχει προσαρμοστεί επιτυχώς για την αξιολόγηση της ομοιότητας μεταξύ ενός ερωτήματος και ενός συνόλου εγγράφων σε μια βάση δεδομένων [35–37]. Το Term Frequency-Inverse Document Frequency (TF-IDF) είναι ένα άλλο κλασικό μοντέλο που έχει εφαρμοστεί ευρέως στη σημασιολογική αναζήτηση [38,38–40]. Το TF-IDF αξιολογεί τη σημασία ενός όρου μέσα σε ένα έγγραφο σε σχέση με ένα σύνολο εγγράφων, δίνοντας μεγαλύτερο βάρος στους όρους που είναι σπάνιοι αλλά σημαντικοί [35,38–42].

Με την εμφάνιση της Σημασιολογικής Ιστού, οι οντολογίες έχουν κερδίσει σημαντικό χώρο στο πεδίο [43]. Οι οντολογίες δίνουν τη δυνατότητα για πιο πολύπλοκες και εξατομικευμένες αναζητήσεις, καθώς επιτρέπουν την ερμηνεία του περιεχομένου με βάση ένα σύνολο προκαθορισμένων κανόνων και σχέσεων [35,38,39,41,43].

Η ανάλυση συναισθημάτων και το Latent Semantic Indexing (LSI) είναι επίσης σημαντικά στοιχεία. Το LSI προσπαθεί να εντοπίσει τις υποκείμενες διαστάσεις του περιεχομένου, ενώ η ανάλυση συναισθημάτων ασχολείται με την εξαγωγή αντικειμενικών και υποκειμενικών στοιχείων από το κείμενο [38,39]. Στη νεότερη ερευνητική βιβλιογραφία, τα νευρωνικά δίκτυα και οι τεχνικές της μηχανικής μάθησης εμφανίζονται ως δυναμικοί παράγοντες για τη βελτίωση της ακρίβειας και της ευελιξίας της σημασιολογικής αναζήτησης [37,40,42,43]. Επιπλέον, το Data Mining και οι τεχνικές εξόρυξης κανόνων συμπεριφοράς καθιστούν δυνατή την αναγνώριση πατρών και την ανάπτυξη προσαρμοστικών στρατηγικών αναζήτησης [44]. Η επεξεργασία φυσικής γλώσσας (Natural Language Processing - NLP) επιτρέπει την καλύτερη κατανόηση της γλωσσικής και σημασιολογικής δομής των κειμένων, διευκολύνοντας την αυτόματη οργάνωση και κατηγοριοποίηση των δεδομένων [45,46].

Τέλος, η έννοια της εξατομικευσης προσφέρει μια νέα διάσταση στη σημασιολογική αναζήτηση [37,40]. Συστήματα που λαμβάνουν υπόψη τις προτιμήσεις, το ιστορικό αναζήτησης και τη συμπεριφορά του χρήστη μπορούν να προσφέρουν πιο εξατομικευμένα και σχετικά αποτελέσματα [37,40].

Συνοπτικά, οι θεωρίες και τα μοντέλα που καθοδηγούν τη σημασιολογική αναζήτηση είναι ποικίλα και περιλαμβάνουν τόσο τις κλασικές προσεγγίσεις της αξιολόγησης εγγράφων όσο και τις πιο πρόσφατες τεχνικές που προκύπτουν από τις προόδους στην τεχνολογία και την επιστήμη των δεδομένων.

Εξέταση της Σύγχρονης Έρευνας

Στην ενότητα "Εξέταση της Σύγχρονης Έρευνας", πραγματοποιούμε μια κριτική αξιολόγηση των πρόσφατων επιστημονικών ερευνών που έχουν διεξαχθεί στον τομέα της σημασιολογικής αναζήτησης. Ενώ η αύξηση των δεδομένων και η εξέλιξη της τεχνολογίας έχουν επιτρέψει την εφαρμογή πιο προηγμένων μεθόδων, όπως τα νευρωνικά δίκτυα και το deep learning, υπάρχουν επίσης προκλήσεις και αδυναμίες που πρέπει να αντιμετωπιστούν [19,33].

Ενδεικτικά, οι επιστημονικές προσεγγίσεις που βασίζονται σε μηχανική μάθηση και νευρωνικά δίκτυα μπορεί να απαιτούν μεγάλους όγκους δεδομένων για την εκπαίδευση, πράγμα που θέτει ζητήματα σχετικά με την προσβασιμότητα και την ηθική των δεδομένων [31,47,48]. Επιπλέον, η εξατομικευση της αναζήτησης, παρόλο που αποτελεί ενδιαφέρον πεδίο έρευνας, εγείρει σημαντικές ανησυχίες σχετικά με την προστασία της ιδιωτικότητας και την διαφάνεια των αλγορίθμων [32–34]. Εν κατακλείδι, η σύγχρονη έρευνα στον τομέα παρουσιάζει ενδιαφέρον και προοπτικές, αλλά επίσης αντιμετωπίζει και ουσιαστικές προκλήσεις που απαιτούν περαιτέρω προσοχή [33,34].

Ας εξετάσουμε τώρα τα στοιχεία που επηρεάζουν την ακρίβεια της σημασιολογικής αναζήτησης. Ένας κρίσιμος παράγοντας είναι η ποιότητα και η δομή των δεδομένων που χρησιμοποιούνται. Η οργάνωση των δεδομένων, καθώς και η συνοχή και η συνέπεια των μεταδεδομένων, είναι κρίσιμες για την ακρίβεια των αποτελεσμάτων [49–53].

Επιπρόσθετα, η απόδοση των αλγορίθμων που χρησιμοποιούνται επηρεάζεται από την δυνατότητα της σύστασης και της αντιμετώπισης της ασάφειας στη φυσική γλώσσα [54–56]. Ειδικές τεχνικές όπως η αναγνώριση οντοτήτων και η διαχείριση συνώνυμων μπορούν να βελτιώσουν σημαντικά την ακρίβεια [54]. Τέλος, η διεπαφή και η εμπειρία του χρήστη έχουν άμεσο αντίκτυπο στην αποτελεσματικότητα της αναζήτησης, καθώς οι χρήστες πρέπει να μπορούν να ερμηνεύσουν και να εφαρμόσουν τα αποτελέσματα με ακρίβεια [51,57].

Τεχνικές Βελτίωσης Ερωτημάτων

Η βελτίωση των ερωτημάτων αναζήτησης αποτελεί κρίσιμο στοιχείο για την αύξηση της ακρίβειας και της σχετικότητας των αποτελεσμάτων αναζήτησης. Οι τεχνικές που εφαρμόζονται περιλαμβάνουν την ανάλυση και επεξεργασία του ερωτήματος με στόχο την αφαίρεση ασήμαντων λέξεων, την κανονικοποίηση των όρων και την ανάλυση των σημασιολογικών σχέσεων [58–60].

Η ανάπτυξη σημασιολογικών δικτύων και οντολογιών παίζει σημαντικό ρόλο στην βελτίωση των αναζητήσεων [32,34,53,58,59]. Με την εφαρμογή τους, μπορούμε να κατανοήσουμε τις σχέσεις μεταξύ των οντοτήτων και να διευκρινίσουμε τις αμφίσημες εκφράσεις. Επιπλέον, η χρήση αλγορίθμων μηχανικής μάθησης μπορεί να προσφέρει δυναμικές λύσεις προσαρμοσμένες στις ανάγκες του κάθε χρήστη [43,48,58].

Επίσης, η κατανόηση του προφίλ του χρήστη, των προτιμήσεων και του τρέχοντος πλαισίου μπορεί να οδηγήσει σε μια πιο στοχευμένη και εξατομικευμένη αναζήτηση. Η υποστήριξη των συστημάτων αναζήτησης με πληροφορίες για τις προτιμήσεις των χρηστών είναι καίριας σημασίας για την παροχή εξατομικευμένων λύσεων [60–62].

Η συνεχής εξέλιξη και προσαρμογή των τεχνικών βελτίωσης ερωτημάτων παρέχει τη δυνατότητα για πιο αποτελεσματικές και δυναμικές στρατηγικές αναζήτησης. Η ενσωμάτωση των τελευταίων εξελίξεων στην τεχνολογία και την έρευνα μπορεί να οδηγήσει σε μια ενισχυμένη εμπειρία αναζήτησης που ανταποκρίνεται ακόμη πιο αποτελεσματικά στις ανάγκες των χρηστών [37,39,45,61,62].

Στην πεδία της αναζήτησης πληροφοριών, η αξιολόγηση της αποτελεσματικότητας των τεχνικών βελτίωσης ερωτημάτων αποτελεί κεντρικό ζήτημα [63,64]. Ερευνητικά δεδομένα και περιπτώσεις μελέτης παρέχουν πολύτιμη υποστήριξη στην κατανόηση της αποτελεσματικότητας αυτών των τεχνικών [63].

Σε αρκετές περιπτώσεις, έχει διαπιστωθεί ότι η εφαρμογή οντολογιών και σημασιολογικών δικτύων ενισχύει την ακρίβεια των αποτελεσμάτων αναζήτησης, καθώς επιτρέπει μια πιο βαθιά κατανόηση των σχέσεων μεταξύ των δεδομένων [65,66]. Ταυτόχρονα, η χρήση των αλγορίθμων μηχανικής μάθησης για την προσαρμογή των αναζητήσεων στις ανάγκες του κάθε χρήστη έχει αποδειχθεί επιτυχής σε διάφορα ερευνητικά προγράμματα [66–68].

Επίσης, η εξατομίκευση της αναζήτησης με βάση το προφίλ και τις προτιμήσεις των χρηστών έχει εξεταστεί εκτενώς και έχει αποδειχθεί ότι αυξάνει την ικανοποίηση του χρήστη και την αποδοτικότητα της αναζήτησης. Αρκετές εμπειρικές μελέτες έχουν επιβεβαιώσει την αντικειμενική βελτίωση της ποιότητας των αποτελεσμάτων αναζήτησης με την εφαρμογή των παραπάνω τεχνικών [63,64,66].

Κατανόηση Συμπεριφορών Χρήστη

Η κατανόηση της συμπεριφοράς των χρηστών στην αναζήτηση αποτελεί πεδίο που έχει προσελκύσει έντονο ενδιαφέρον τα τελευταία χρόνια, κυρίως λόγω της πολυπλοκότητας και της δυναμικότητας των αναγκών και των συμπεριφορών των χρηστών [69,70]. Η εξατομίκευση των αποτελεσμάτων αναζήτησης βασίζεται σε μεγάλο βαθμό στην ανάλυση και την ερμηνεία των δράσεων και των προτιμήσεων των χρηστών [69–71].

Η συμπεριφορά των χρηστών μπορεί να αναλυθεί μέσα από διάφορους παράγοντες, όπως το ιστορικό αναζήτησης, την αλληλεπίδραση με τα αποτελέσματα, τον χρόνο παραμονής σε συγκεκριμένες ιστοσελίδες και άλλες μετρικές. Οι τεχνολογίες που εφαρμόζονται για την ανίχνευση και την ανάλυση της συμπεριφοράς των χρηστών καθορίζουν τον τρόπο με τον οποίο τα συστήματα αναζήτησης

προσαρμόζονται στις μεμονωμένες ανάγκες [64,69–73]. Πέρα από την εξατομίκευση, η κατανόηση της συμπεριφοράς των χρηστών συμβάλλει στην ανάπτυξη πιο αποτελεσματικών και προσανατολισμένων προς τον χρήστη υπηρεσιών [74–76]. Για παράδειγμα, η δυνατότητα προβλέψεων για τις πιθανές ερωτήσεις ή ανάγκες των χρηστών μπορεί να αποτελέσει βασικό εργαλείο για την βελτίωση της εμπειρίας αναζήτησης [73,75].

Στην άλλη πλευρά, η κατανόηση της συμπεριφοράς των χρηστών μπορεί να παραμερίσει ζητήματα προσωπικής ιδιωτικότητας και ηθικής. Η εξαιρετικά λεπτομερής παρακολούθηση και ανάλυση των δραστηριοτήτων των χρηστών απαιτεί προσεκτική ισορροπία μεταξύ της βελτίωσης της υπηρεσίας και του σεβασμού των δικαιωμάτων των χρηστών στην ιδιωτικότητα [13,14,31].

Συνοψίζοντας, η κατανόηση της συμπεριφοράς των χρηστών στην αναζήτηση είναι απαραίτητη για την παραγωγή δυναμικών, προσαρμοσμένων και αποτελεσματικών αποτελεσμάτων. Ταυτόχρονα, πρέπει να γίνεται με σεβασμό προς τα ηθικά και ιδιωτικά δικαιώματα των χρηστών, και με την εφαρμογή των κατάλληλων τεχνικών και μεθοδολογιών [33,34,69,73,76].

Συμπεράσματα

Συνοψίζοντας τα κύρια ευρήματα της παρούσας έρευνας, αναδεικνύεται μια πολύπλοκη κατανόηση των συμπεριφορών των χρηστών, της ιστορικής εξέλιξης των μηχανών αναζήτησης και των τεχνικών βελτίωσης ερωτημάτων στο πλαίσιο των βιβλιοθηκών. Από την εισαγωγή και την αρχική χρήση των μηχανών αναζήτησης, έως τη σύγχρονη κατανόηση των πολυδιάστατων και πολύπλοκων συμπεριφορών των χρηστών, εντοπίζουμε σύνθετες διαδρομές και διαλειτουργικότητες.

Η αναλυτική εξέταση της σύγχρονης έρευνας υπογράμμισε τη σπουδαία συνεισφορά της σημασιολογικής αποτύπωσης της υπάρχουσας γνώσης σχετικά με διαφορετικούς τομείς του ανθρώπινου επιστητού. Έτσι μπορεί να επιτευχθεί περαιτέρω εξέλιξη στον τομέα της επεξεργασίας φυσικής γλώσσας και της μηχανικής μάθησης, με σκοπό τη βελτίωση της εμπειρίας του χρήστη των φυσικών ή ψυφιακών βιβλιοθηκών. Επιπλέον, η εφαρμογή σύγχρονων τεχνικών βελτίωσης ερωτημάτων αναδείχθηκε ως ζωτικής σημασίας για την επίτευξη ακριβέστερων και πιο σχετικών αποτελεσμάτων.

Παρά την ανάλυση και την πρόοδο που έχει σημειωθεί στους τομείς της κατανόησης συμπεριφορών χρηστών, της ιστορικής εξέλιξης των μηχανών αναζήτησης, των τεχνικών βελτίωσης ερωτημάτων, καθώς και της σημασιολογικής προσέγγισης στην αποδελτίωση και ταξινόμηση της γνώσης, η παρούσα έρευνα αποκάλυψε κάποια σημαντικά κενά. Τα ζητήματα της ηθικής και της διαφάνειας στην παρακολούθηση των χρηστών δεν έχουν εξεταστεί επαρκώς, αφήνοντας περιθώρια για περαιτέρω έρευνα.

Τα θεωρητικά πλαίσια που αφορούν τις μηχανές αναζήτησης εξακολουθούν να είναι ασαφή και κατά κάποιο τρόπο αποσπασματικά. Αυτό καθιστά απαραίτητη την ανάπτυξη ενιαίων και ολοκληρωμένων θεωρητικών προσεγγίσεων που θα διευκολύνουν την κατανόηση και την εφαρμογή των νεότερων ερευνητικών εξελίξεων.

Η εξέταση των τεχνικών βελτίωσης ερωτημάτων πρέπει επίσης να επεκταθεί προς την κατεύθυνση της ενσωμάτωσης διαφορετικών γλωσσών και πολιτισμών. Η ικανότητα εξυπηρέτησης ενός παγκόσμιου κοινού απαιτεί την κατανόηση και την εφαρμογή των πολυπολιτισμικών αναγκών και προκλήσεων.

Επιπλέον, τα κενά στην κατανόηση των συμπεριφορών των χρηστών μπορούν να γεφυρωθούν μέσω της χρήσης ενισχυμένης προσωποποιημένης μάθησης και προσαρμογής στις αλλαγές των αναγκών και των προτιμήσεων του χρήστη. Η εξέλιξη αυτών των στρατηγικών μπορεί να προσφέρει μια πιο δυναμική και ανταποκριτική εμπειρία.

Τέλος, η εφαρμογή νέων τεχνολογιών και η εξερεύνηση καινοτόμων μεθόδων παρέχει ευκαιρίες για περαιτέρω ερευνητική ανάπτυξη και πρόοδο. Η επιδίωξη των προτεινόμενων μελλοντικών κατευθύνσεων μπορεί να συμβάλει στη δημιουργία μιας πιο ολοκληρωμένης και αποτελεσματικής εικόνας του πεδίου των μηχανών αναζήτησης και της αλληλεπίδρασης με τους χρήστες.

Βιβλιογραφία

1. Lassila, O. Web Metadata: A Matter of Semantics. *IEEE Internet Comput.* **1998**, *2*, 30–37.
2. Pérez-Agüera, J.R.; Arroyo, J.; Greenberg, J.; Iglesias, J.P.; Fresno, V. Using BM25F for Semantic Search. In Proceedings of the Proceedings of the 3rd international semantic search workshop; 2010; pp. 1–8.
3. Gasparini, A.A.; Kautonen, H. Understanding Artificial Intelligence in Research Libraries: An Extensive Literature Review. *Liber Q. Te J. Eur. Res. Libr.* **2022**, *32*, 1–36.
4. Eller, D.W. Transparency and the Future of Semantic Searching in Academic Libraries. *Inf. Serv. Use* **2022**, *42*, 453–461.
5. Al-Natsheh, H.T.; Martinet, L.; Muhlenbach, F.; Rico, F.; Zighed, D.A. Semantic Search-by-Examples for Scientific Topic Corpus Expansion in Digital Libraries. In Proceedings of the 2017 IEEE International Conference on Data Mining Workshops (ICDMW); IEEE, 2017; pp. 747–756.
6. Ding, H. Semantic Search in Peer-to-Peer Based Digital Libraries. **2005**.
7. Jamil, H. Semantic Querying of Knowledge Rich Legal Digital Libraries Using Prism 1. In *Legal Knowledge and Information Systems*; IOS Press, 2022; pp. 63–72.
8. Elizarov, A.; Kirillovich, A.; Lipachev, E.; Nevzorova, O. Semantic Formula Search in Digital Mathematical Libraries. In Proceedings of the 2017 Second Russia and Pacific Conference on Computer Technology and Applications (RPC); IEEE, 2017; pp. 39–43.
9. Garoufallou, E.; Fallucchi, F.; De Luca, E.W. *Metadata and Semantic Research: 13th International Conference, MTSR 2019, Rome, Italy, October 28–31, 2019, Revised Selected Papers*; Springer Nature, 2019; Vol. 1057; ISBN 3-030-36599-9.
10. Garoufallou, E.; Ovalle-Perandones, M.A. *Metadata and Semantic Research*; Springer, 2022; ISBN 3-030-98875-9.
11. Ai, Q.; Bai, T.; Cao, Z.; Chang, Y.; Chen, J.; Chen, Z.; Cheng, Z.; Dong, S.; Dou, Z.; Feng, F. Information Retrieval Meets Large Language Models: A Strategic Report from Chinese IR Community. *AI Open* **2023**.
12. Liu, J.; Liu, J.; Chen, Y. Application of Artificial Intelligence Technology in Information Retrieval of University Library. In Proceedings of the Innovative Computing: Proceedings of the 4th International Conference on Innovative Computing (IC 2021); Springer, 2022; pp. 221–228.
13. Albakour, D.; Ducatel, G.; Kruschwitz, U. The Role of Search for Field Force Knowledge Management. In *Transforming Field and Service Operations: Methodologies for Successful Technology-Driven Business Transformation*; Springer, 2013; pp. 117–132.
14. Losiewicz, P.; Oard, D.W.; Kostoff, R.N. Textual Data Mining to Support Science and Technology Management. *J. Intell. Inf. Syst.* **2000**, *15*, 99–119.
15. Virkus, S.; Garoufallou, E. Data Science from a Library and Information Science Perspective. *Data Technol. Appl.* **2019**, *53*, 422–441.

16. Linckels, S.; Meinel, C. *E-Librarian Service: User-Friendly Semantic Search in Digital Libraries*; Springer, 2011; ISBN 3-642-17742-5.
17. Badami, M.; Benatallah, B.; Baez, M. Adaptive Search Query Generation and Refinement in Systematic Literature Review. *Inf. Syst.* **2023**, 102231.
18. Pavlidis, G. From Digital Recording to Advanced AI Applications in Archaeology and Cultural Heritage. In *“And in Length of Days Understanding”(Job 12: 12) Essays on Archaeology in the Eastern Mediterranean and Beyond in Honor of Thomas E. Levy*; Springer, 2023; pp. 1627–1656.
19. Dave, D.; Celestino, A.; Varde, A.S.; Anu, V. Management of Implicit Requirements Data in Large SRS Documents: Taxonomy and Techniques. *ACM SIGMOD Rec.* **2022**, *51*, 18–29.
20. Melnykova, N. Semantic Search Personalized Data as Special Method of Processing Medical Information. In *Advances in Intelligent Systems and Computing: Selected Papers from the International Conference on Computer Science and Information Technologies, CSIT 2016, September 6-10 Lviv, Ukraine*; Shakhovska, N., Ed.; Advances in Intelligent Systems and Computing; Springer International Publishing: Cham, 2017; pp. 315–325 ISBN 978-3-319-45991-2.
21. Bevilacqua, V.; Santarcangelo, V.; Magarelli, A.; Bianco, A.; Mastronardi, G.; Cascini, E. A Semantic Search Framework for Document Retrievals (Literature, Art and History) Based on Thesaurus Multiwordnet Like. In *Proceedings of the Advanced Intelligent Computing*; Huang, D.-S., Gan, Y., Bevilacqua, V., Figueroa, J.C., Eds.; Springer: Berlin, Heidelberg, 2012; pp. 456–463.
22. Nagarajan, G.; Thyagarajan, K.K. A Machine Learning Technique for Semantic Search Engine. *Procedia Eng.* **2012**, *38*, 2164–2171, doi:10.1016/j.proeng.2012.06.260.
23. Li, L.; Dong, Z.; Xie, K. Ontology of General Concept for Semantic Searching. In *Proceedings of the 2010 Second International Conference on Computer Modeling and Simulation*; January 2010; Vol. 2, pp. 81–84.
24. Enser, P. The Evolution of Visual Information Retrieval. *J. Inf. Sci.* **2008**, *34*, 531–546, doi:10.1177/0165551508091013.
25. Kambau, R.A.; Hasibuan, Z.A. Evolution of Information Retrieval System: Critical Review of Multimedia Information Retrieval System Based on Content, Context, and Concept. In *Proceedings of the 2017 11th International Conference on Information & Communication Technology and System (ICTS)*; July 2017; pp. 91–98.
26. Gleason, P. Identifying Identity: A Semantic History. *J. Am. Hist.* **1983**, *69*, 910–931, doi:10.2307/1901196.
27. Aggarwal, C.C.; Wang, H. Text Mining in Social Networks. In *Social Network Data Analytics*; Aggarwal, C.C., Ed.; Springer US: Boston, MA, 2011; pp. 353–378 ISBN 978-1-4419-8462-3.
28. Langley, P.; Carbonell, J.G. Approaches to Machine Learning. *J. Am. Soc. Inf. Sci.* **1984**, *35*, 306–316, doi:10.1002/asi.4630350509.
29. Gordon, D.F.; Desjardins, M. Evaluation and Selection of Biases in Machine Learning. *Mach. Learn.* **1995**, *20*, 5–22, doi:10.1007/BF00993472.
30. Mosa, M.A. Real-Time Data Text Mining Based on Gravitational Search Algorithm. *Expert Syst. Appl.* **2019**, *137*, 117–129, doi:10.1016/j.eswa.2019.06.065.
31. Wang, Q.; Huang, Y.; Jasin, S.; Singh, P.V. Algorithmic Transparency with Strategic Users. *Manag. Sci.* **2023**, *69*, 2297–2317, doi:10.1287/mnsc.2022.4475.
32. Hoaihongthong, S.; Kwiecien, K. Semantic Search System for Common Cultural Traditions of the Greater Mekong Subregion. *Int. J. Innov. Res. Sci. Stud.* **2022**, *5*, 289–296, doi:10.53894/ijirss.v5i4.709.

33. Moataz, T.; Shikfa, A.; Cuppens-Boulahia, N.; Cuppens, F. Semantic Search over Encrypted Data. In Proceedings of the ICT 2013; February 2013; pp. 1–5.
34. Kats, Y. Semantic Search and NLP-Based Diagnostics. In Proceedings of the 2014 IEEE 27th International Symposium on Computer-Based Medical Systems; February 2014; pp. 277–280.
35. Guo, Y.; Chen, D.; Le, J. An Extended Vector Space Model for XML Information Retrieval. In Proceedings of the 2009 Second International Workshop on Knowledge Discovery and Data Mining; January 2009; pp. 797–800.
36. Abu-Salih, B. Applying Vector Space Model (VSM) Techniques in Information Retrieval for Arabic Language 2018.
37. Sabna, E. INFORMATION RETRIEVAL PADA DATA JUDUL SKRIPSI BERBASIS TEXT MENGGUNAKAN VECTOR SPACE MODEL. *J. Ilmu Komput.* **2021**, *10*, 132-1`35, doi:10.33060/JIK/2021/Vol10.Iss2.230.
38. Fang, W.-D.; Zhang, L.; Wang, Y.-X.; Dong, S.-B. Toward a Semantic Search Engine Based on Ontologies. In Proceedings of the 2005 International Conference on Machine Learning and Cybernetics; December 2005; Vol. 3, pp. 1913-1918 Vol. 3.
39. Fernandez, M.; Lopez, V.; Sabou, M.; Uren, V.; Vallet, D.; Motta, E.; Castells, P. Semantic Search Meets the Web. In Proceedings of the 2008 IEEE International Conference on Semantic Computing; December 2008; pp. 253–260.
40. Pol Gonçalves, F.J.; de Oliveira Carmo, V.C.; Toquetti de Melo, V.; Cunha, R. da S.; Santos, I.H.F.; Barreira, R.A.; Cugnasca, C.E.; Cozman, F.G.; Gomi, E.S. Semantic Search in Offshore Engineering With Linguistics And Neural Processing Pipelines.; American Society of Mechanical Engineers Digital Collection, October 11 2021.
41. Mohd Nafis, N.S.; Awang, S. An Enhanced Hybrid Feature Selection Technique Using Term Frequency-Inverse Document Frequency and Support Vector Machine-Recursive Feature Elimination for Sentiment Classification. *IEEE Access* **2021**, *9*, 52177–52192, doi:10.1109/ACCESS.2021.3069001.
42. Patel, A.; Debnath, N.C. *Data Science with Semantic Technologies: Deployment and Exploration*; CRC Press, 2023; ISBN 978-1-00-088129-5.
43. Spyropoulos, A.Z.; Kornilakis, A.; Makris, G.C.; Bratsas, C.; Tsiantos, V.; Antoniou, I. Semantic Representation of the Intersection of Criminal Law & Civil Tort. *Data* **2022**, *7*, 176, doi:10.3390/data7120176.
44. From Data Mining to Wisdom Mining - Salma Khan, Muhammad Shaheen, 2023 Available online: https://journals.sagepub.com/doi/full/10.1177/01655515211030872?casa_token=BPtvdQ1qe8gAAAAA%3AtcsWOhPPXVZRNvFfqU_pjZavldlkdov4INAmNjdVmdl5YNTwEX4Haf9tkW9J5zcax-siG_rNnVWw7Uc (accessed on 24 August 2023).
45. Khurana, D.; Koli, A.; Khatter, K.; Singh, S. Natural Language Processing: State of the Art, Current Trends and Challenges. *Multimed. Tools Appl.* **2023**, *82*, 3713–3744, doi:10.1007/s11042-022-13428-4.
46. Pre-Train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing | ACM Computing Surveys Available online: <https://dl.acm.org/doi/full/10.1145/3560815> (accessed on 24 August 2023).
47. Diakopoulos, N.; Koliska, M. Algorithmic Transparency in the News Media. *Digit. Journal.* **2017**, *5*, 809–828, doi:10.1080/21670811.2016.1208053.
48. Hanley, H.W.A.; Kumar, D.; Durumeric, Z. Happenstance: Utilizing Semantic Search to Track Russian State Media Narratives about the Russo-Ukrainian War on Reddit. *Proc. Int. AAAI Conf. Web Soc. Media* **2023**, *17*, 327–338, doi:10.1609/icwsm.v17i1.22149.

49. Greenberg, J.; Garoufallou, E. Change and a Future for Metadata. In Proceedings of the Metadata and Semantics Research; Garoufallou, E., Greenberg, J., Eds.; Springer International Publishing: Cham, 2013; pp. 1–5.
50. Sharma, S.; Jain, S. Comprehensive Review on Semantic Information Retrieval and Ontology Engineering 2023.
51. Megaw, E.D. Factors Affecting Visual Inspection Accuracy. *Appl. Ergon.* **1979**, *10*, 27–32, doi:10.1016/0003-6870(79)90006-1.
52. Ezzini, S.; Abualhaija, S.; Arora, C.; Sabetzadeh, M.; Briand, L.C. Using Domain-Specific Corpora for Improved Handling of Ambiguity in Requirements. In Proceedings of the 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE); February 2021; pp. 1485–1497.
53. Garoufallou, E.; Ovalle-Perandones, M.-A. *Metadata and Semantic Research: 14th International Conference, MTSR 2020, Madrid, Spain, December 2–4, 2020, Revised Selected Papers*; Springer Nature, 2021; ISBN 978-3-030-71903-6.
54. Natural Language Processing: State of the Art, Current Trends and Challenges | SpringerLink Available online: <https://link.springer.com/article/10.1007/s11042-022-13428-4> (accessed on 24 August 2023).
55. Schilhan, L.; Kaier, C.; Lackner, K. Increasing Visibility and Discoverability of Scholarly Publications with Academic Search Engine Optimization. **2021**, *34*, 6, doi:10.1629/uksg.534.
56. Setlur, V.; Kanyuka, A.; Srinivasan, A. Olio: A Semantic Search Interface for Data Repositories 2023.
57. Maulud, D.H.; Zeebaree, S.R.M.; Jacksi, K.; Sadeeq, M.A.M.; Sharif, K.H. State of Art for Semantic Analysis of Natural Language Processing. *Qubahan Acad. J.* **2021**, *1*, 21–28, doi:10.48161/qaj.v1n2a44.
58. Herodotou, H.; Borisov, N.; Babu, S. Query Optimization Techniques for Partitioned Tables. In Proceedings of the Proceedings of the 2011 ACM SIGMOD International Conference on Management of data; Association for Computing Machinery: New York, NY, USA, March 12 2011; pp. 49–60.
59. Khan, M.; Khan, M.N.A. Exploring Query Optimization Techniques in Relational Databases. *Int. J. Database Theory Appl.* **2013**, *6*, 11–20.
60. Azhir, E.; Jafari Navimipour, N.; Hosseinzadeh, M.; Sharifi, A.; Darwesh, A. Deterministic and Non-Deterministic Query Optimization Techniques in the Cloud Computing. *Concurr. Comput. Pract. Exp.* **2019**, *31*, e5240, doi:10.1002/cpe.5240.
61. A Large-Scale Evaluation and Analysis of Personalized Search Strategies | Proceedings of the 16th International Conference on World Wide Web Available online: https://dl.acm.org/doi/abs/10.1145/1242572.1242651?casa_token=ldcOX8DYUtAAAAA:nwf5EQ72kPUpc0Wc9CPhLRWXMh3zLCu37D82paMtxwrezDsmCjn0KnWLubKH-EOHyZT_1d_hXV1bKw (accessed on 24 August 2023).
62. Yang, Y. Personalized Search Strategies for Spatial Information on the Web. *IEEE Intell. Syst.* **2012**, *27*, 12–20, doi:10.1109/MIS.2010.108.
63. UPS | Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval Available online: https://dl.acm.org/doi/abs/10.1145/2009916.2009999?casa_token=S6WAZqz0jg4AAAAA:O8WAKJplG6ek_3chx1Enbe2HNIURC05sXALTwBBsVs-hNNdbbEqTAEEmVKLYIUhIQgQffQ4rT5wu4_Lc (accessed on 24 August 2023).
64. Shim, K.; Sellis, T.; Nau, D. Improvements on a Heuristic Algorithm for Multiple-Query Optimization. *Data Knowl. Eng.* **1994**, *12*, 197–222, doi:10.1016/0169-023X(94)90014-0.

65. Vallet, D.; Fernández, M.; Castells, P. An Ontology-Based Information Retrieval Model. In Proceedings of the The Semantic Web: Research and Applications; Gómez-Pérez, A., Euzenat, J., Eds.; Springer: Berlin, Heidelberg, 2005; pp. 455–470.
66. Padmanaban, R.; Valentine, P.; Nivetha, A.; Aruna, A. A Framework for User Customizable Privacy Preserving Search. *Adv. Nat. Appl. Sci.* **2015**, *9*, 187–192.
67. Graefe, G.; McKenna, W.J. The Volcano Optimizer Generator: Extensibility and Efficient Search. In Proceedings of the Proceedings of IEEE 9th International Conference on Data Engineering; April 1993; pp. 209–218.
68. Dang, D.V.; Nguyen, H.D.; Ngo, H.; Pham, V.T.; Nguyen, D. Information Retrieval from Legal Documents with Ontology and Graph Embeddings Approach. In Proceedings of the Advances and Trends in Artificial Intelligence. Theory and Applications; Fujita, H., Wang, Y., Xiao, Y., Moonis, A., Eds.; Springer Nature Switzerland: Cham, 2023; pp. 300–312.
69. Sfakakis, M.; Kapidakis, S. User Behavior Tendencies on Data Collections in a Digital Library. In Proceedings of the Research and Advanced Technology for Digital Libraries; Agosti, M., Thanos, C., Eds.; Springer: Berlin, Heidelberg, 2002; pp. 550–559.
70. Laiche, F.; Ben Letaifa, A.; Aguilí, T. QoE-Aware Traffic Monitoring Based on User Behavior in Video Streaming Services. *Concurr. Comput. Pract. Exp.* **2023**, *35*, e6678, doi:10.1002/cpe.6678.
71. Kuhlthau, C.C. Perceptions of the Information Search Process in Libraries: A Study of Changes from High School through College. *Inf. Process. Manag.* **1988**, *24*, 419–427, doi:10.1016/0306-4573(88)90045-3.
72. Zhou, Y.; Zhao, J.; Zhang, J. Prediction of Learners' Dropout in E-Learning Based on the Unusual Behaviors. *Interact. Learn. Environ.* **2023**, *31*, 1796–1820, doi:10.1080/10494820.2020.1857788.
73. Lown, C.; Sierra, T.; Boyer, J. How Users Search the Library from a Single Search Box | Lown | College & Research Libraries. **2017**, doi:https://doi.org/10.5860/crl-321
74. Sarker, I.H.; Colman, A.; Han, J.; Khan, A.I.; Abushark, Y.B.; Salah, K. BehavDT: A Behavioral Decision Tree Learning to Build User-Centric Context-Aware Predictive Model. *Mob. Netw. Appl.* **2020**, *25*, 1151–1161, doi:10.1007/s11036-019-01443-z.
75. Köse, A.; Tepljakov, A.; Petlenkov, E. Dynamic Predictive Modeling Approach of User Behavior in Virtual Reality Based Application. In Proceedings of the 2019 27th Mediterranean Conference on Control and Automation (MED); July 2019; pp. 57–62.
76. Günther, S.; Göttert, P.; Hagen, M. Exploring LSTMs for Simulating Search Sessions in Digital Libraries. In Proceedings of the Linking Theory and Practice of Digital Libraries; Silvello, G., Corcho, O., Manghi, P., Di Nunzio, G.M., Golub, K., Ferro, N., Poggi, A., Eds.; Springer International Publishing: Cham, 2022; pp. 469–473.