

Οι νέες υπηρεσίες ανακάλυψης περιεχομένου στο SearchCulture.gr

Λαγούδη Ε., Αγγελάκη Γ., Παπανότη Α., Γεωργιάδης Χ.

Εθνικό Κέντρο Τεκμηρίωσης και Ηλεκτρονικού Περιεχομένου

Λέξεις-κλειδιά: SearchCulture.gr, Υπηρεσίες Ανακάλυψης, σημασιολογικοί εμπλουτισμοί, γεωχωρικά δεδομένα

Οι υπηρεσίες ανακάλυψης του περιεχομένου στον Εθνικό Συσσωρευτή Πολιτιστικού Περιεχομένου SearchCulture.gr βασίζονται στις καινοτόμες μεθόδους σημασιολογικού εμπλουτισμού που πραγματοποιεί η ομάδα του SearchCulture.gr και συνεχώς επεκτείνονται με νέες ενδιαφέρουσες λειτουργικότητες. Στην παρούσα εργασία θα παρουσιάσουμε, καταρχήν, τις εκτεταμένες διαδικασίες αναδρομικού σημασιολογικού εμπλουτισμού με γεωχωρική πληροφορία. 550.000 τεκμήρια εμπλουτίστηκαν με το νέο πεδίο «ΕΚΤ Τόπος», με αναφορές σε όρους ενός δίγλωσσου ιεραρχικού λεξιλογίου τοποθεσιών που δημιουργήθηκε και που αποτελεί την ελληνική απόδοση επιλεγμένων όρων της γεωγραφικής βάσης δεδομένων GeoNames και ενός συμπληρωματικού δίγλωσσου λεξιλογίου με ιστορικές ή μη διοικητικές περιοχές, που, από κοινού περιλαμβάνουν πάνω από 12.000 όρους. Το νέο πεδίο χρησιμοποιείται στην προηγμένη αναζήτηση, ως φίλτρο επιμερισμού των αποτελεσμάτων, βοηθά στην αποσαφήνιση όρων διαφοροποιώντας π.χ. του όρου Τρίπολη στην Πελοπόννησο και Τρίπολη στην Λιβύη, και, τέλος, ο χρήστης μπορεί να πλοηγηθεί στα τεκμήρια με βάση το χάρτη. Άλλες επιμέρους νέες λειτουργικότητες που θα παρουσιαστούν αφορούν την αναζήτηση και πλοήγηση προσώπων με βάση το φύλο, την ιδιότητα και τον τόπο γέννησης ή θανάτου και τη δυνατότητα πλοήγησης σε χάρτη με βάση των τόπο γέννησης ή θανάτου. Τέλος, οι γεωχωρικοί εμπλουτισμοί αξιοποιήθηκαν στη δημιουργία θεματικών εκθέσεων, όπως για τους πολιτισμούς της Μικράς Ασίας. Συνδυαστικά μαζί και με τα άλλα σημασιολογικά «ΕΚΤ πεδία» “Δημιουργός”, “Αναφερόμενο Πρόσωπο”, “Θέμα”, “Ιστορική Περίοδος”, “Χρονολογία/Εύρος Χρονολογιών”, “Τύπος Τεκμηρίου” ανοίγουν συναρπαστικούς τρόπους πλοήγησης σε όλο το εύρος και το βάθος της ελληνικής πολιτιστικής κληρονομιάς καθιστώντας το SearchCulture.gr πρότυπο εθνικό συσσωρευτή ανάμεσα στους συσσωρευτές της Europeana.

1. Εισαγωγή

Το SearchCulture.gr είναι η εθνική πύλη συσσώρευσης της ψηφιακής πολιτιστικής κληρονομιάς και της σύγχρονης καλλιτεχνικής δημιουργίας της χώρας μας. Η υποδομή αναπτύσσεται από το 2012 με εσωτερική τεχνογνωσία από το Εθνικό Κέντρο Τεκμηρίωσης και Ηλεκτρονικού Περιεχομένου (ΕΚΤ). Σήμερα στο SearchCulture.gr συγκεντρώνονται 856.375 τεκμήρια εικόνας, κειμένου, βίντεο, ήχου και 3D ψηφιοποίησης, όπως φωτογραφικά ιστορικά ντοκουμέντα, αρχαιολογικά μνημεία και λαογραφικά αντικείμενα, χειρόγραφα, ζωγραφικά έργα και έργα θρησκευτικής τέχνης, ιστορικοί χάρτες και άυλη πολιτιστική κληρονομιά, από 91 φορείς από όλη την Ελλάδα, ανάμεσά τους, μουσεία, αρχεία, πινακοθήκες, βιβλιοθήκες, εφορείες αρχαιοτήτων, δήμοι και ερευνητικά κέντρα. Παράλληλα, το ΕΚΤ, μέσω της υποδομής SearchCulture.gr, είναι ο διαπιστευμένος εθνικός συσσωρευτής για την ευρωπαϊκή ψηφιακή πλατφόρμα Europeana, στην οποία διοχετεύει τις συλλογές που συλλέγει, ακολουθώντας τα διεθνή πρότυπα και τις προδιαγραφές και μεριμνώντας για την προβολή της πολιτιστικής μας κληρονομιάς διεθνώς.

Το SearchCulture.gr θεωρείται πρότυπη υποδομή για τα ευρωπαϊκά δεδομένα, χάρη στις συστηματικές διαδικασίες σημασιολογικού εμπλουτισμού που εφαρμόζει στα δεδομένα τα οποία συλλέγει για την αντιμετώπιση της ετερογένειας του μεγάλου όγκου των δεδομένων που συλλέγει. Οι εμπλουτισμοί στο SearchCulture.gr συνολικά βελτιώνουν την ποιότητα και την εγκυρότητα των δεδομένων, επιτρέπουν την αποσαφήνιση εννοιών και προσώπων, υποστηρίζουν τη νοηματική διασύνδεση μεταξύ τεκμηρίων που βρίσκονται σε πολλά διαφορετικά αποθετήρια και αξιοποιούνται στην προηγμένη αναζήτηση και πλοήγηση.

Η παρούσα εισήγηση, εστιάζει στην τελευταία διαδικασία μαζικού αναδρομικού εμπλουτισμού που διενεργήθηκε επί της γεωχωρικής πληροφορίας που αποτυπώνεται στα μεταδεδομένα των τεκμηρίων, καθώς και στον τρόπο με τον οποίο οι σημασιολογικοί εμπλουτισμοί υποστηρίζουν προηγμένες υπηρεσίες τόσο αναζήτησης στα τεκμήρια, όσο και πλοήγησης στο χάρτη.

2. Μεθοδολογία, εργαλεία και ιστορικό εξέλιξης των σημασιολογικών εμπλουτισμών στο SearchCulture.gr.

Η μεθοδολογία που ακολουθείται για τους σημασιολογικούς εμπλουτισμούς συνδυάζει τη χρήση έξυπνων εργαλείων και την επιστημονική επιμέλεια, εφαρμόζεται, δηλαδή, μία ημι-αυτόματη διαδικασία. Η πλατφόρμα που αξιοποιείται είναι το Semantics.gr, μια ολοκληρωμένη υποδομή για την ανάπτυξη, επιμέλεια και δημοσίευση Λεξιλογίων, Θησαυρών όρων και Καταλόγων καθιερωμένων ονομάτων σε οποιοδήποτε σχήμα ως Ανοικτά Διασυνδεδεμένα Δεδομένα, η οποία επίσης έχει αναπτυχθεί από το ΕΚΤ.

Στο Semantics.gr για κάθε διαδικασία εμπλουτισμού δημιουργήθηκαν εξ αρχής ή χρησιμοποιήθηκαν υπάρχοντα λεξιλόγια θησαυρών όρων και καθιερωμένων λεξιλογίων που διατίθενται ως Ανοικτά Διασυνδεδεμένα Δεδομένα (Getty AAT, Unesco Thesaurus, VIAF, κλπ) τα οποία επεκτάθηκαν για να καλύψουν τις ανάγκες των συγκεκριμένων συλλογών. Τα λεξιλόγια είναι ιεραρχικά, όπου είναι εφαρμόσιμο, και όλα είναι δίγλωσσα.

Το Semantics.gr περιλαμβάνει ένα εργαλείο μαζικού σημασιολογικού εμπλουτισμού μεταδεδομένων στο οποίο αντιστοιχίζονται όροι από επιλεγμένα πεδία των μεταδεδομένων του φορέα με

προτεινόμενους από το σύστημα όρους από το λεξιλόγιο-στόχο που έχει δημιουργηθεί από το ΕΚΤ. Το εργαλείο διαθέτει μηχανισμό αυτοβελτίωσης, αποθηκεύει, δηλαδή κάθε νέα αντιστοίχιση ώστε να μπορεί να την αναπαράξει στο μέλλον.

Ξεκινώντας από το 2016 διαδοχικά αναπτύχθηκαν τα λεξιλόγια και έγιναν αναδρομικοί σημασιολογικοί εμπλουτισμοί όσον αφορά τους τύπους περιεχομένου, την ιστορική περίοδο στην οποία εντάσσεται ένα τεκμήριο, τη θεματική κάλυψη, και το φυσικό πρόσωπο που συσχετίζεται με αυτό, είτε ως δημιουργός είτε ως αναφορά. Η πλατφόρμα του Semantics.gr και η λειτουργικότητά της παρουσιάζεται στο Georgiadis et al (2021) και οι ως τώρα διαδικασίες αναδρομικών εμπλουτισμών στα πρόσωπα παρουσιάζονται αναλυτικά στα Georgiadis H. et al (2018, 2020).

3. Στόχοι του σημασιολογικού εμπλουτισμού της γεωχωρικής πληροφορίας

Οι λειτουργικές απαιτήσεις που τέθηκαν για τους γεωχωρικούς εμπλουτισμούς ήταν οι εξής:

- Ανάκτηση όλων των τεκμηρίων που αφορούν μία τοποθεσία από όλες τις συλλογές
- Αποσαφήνιση μεταξύ ομώνυμων τοποθεσιών (π.χ. Τρίπολη στη Λιβύη και Τρίπολη στην Πελοπόννησο), ώστε να μπορεί ένα ερώτημα να επιστρέφει στοχευμένα αποτελέσματα
- Αναζήτηση με εναλλακτική ονομασία (Αϊδίνι=Σελεύκεια επί του Μαιάνδρου=Αντιόχεια)
- Αναζήτηση λαμβάνοντας υπόψη τις Has-a σχέσεις. Π.χ η αναζήτηση “Αττική” θα έπρεπε να επιστρέφει τεκμήρια που έχουν τιμή Αττική αλλά και αυτά που έχουν τοποθεσίες εντός της Αττικής, όπως ο Παρθενώνας και ο Πειραιάς
- Ιεραρχική πλοήγηση στους τόπους, με δυνατότητα εντοπισμού τεκμηρίων σε όσο πιο συγκεκριμένο επίπεδο είναι δυνατό, π.χ. Ελλάδα--> Νήσος Τήνος--> Βώλαξ
- Πλοήγηση στα τεκμήρια με τη χρήση χάρτη

4. Οι διαδικασίες των σημασιολογικών εμπλουτισμών

4.1 Δημιουργία δύο γεωχωρικών λεξιλογίων διασυνδεδεμένων με το Geonames
Για τη δημιουργία του λεξιλογίου των τοποθεσιών χρησιμοποιήθηκε ως βάση το Geonames¹, μία ανοικτή, παγκόσμια βάση δεδομένων τοποθεσιών που εμπλουτίζεται πληθοποριστικά και προσφέρει πολλαπλούς τρόπους αναζήτησης και εξαγωγής των δεδομένων. Διαθέτει πάνω από 11 εκ. τοποθεσίες, με τις αποδόσεις τους σε πολλές γλώσσες, τις συντεταγμένες και τις κατηγορίες τοποθεσιών μαζί με την διοικητική ιεραρχία τους, τύπους οικισμάτων και γεωμορφολογικούς τύπους (π.χ. πόλη, οικισμός, ναός, αρχαιολογικός χώρος, βουνό, ποτάμι, κλπ). Κάθε όρος διαθέτει μοναδικό προσδιοριστή URI.

Από το Geonames δημιουργήθηκαν 2 παράγωγα λεξιλόγια στο Semantics.gr: α. Το “Λεξιλόγιο τοπωνυμίων Geonames” (έκδοση ΕΚΤ)² ξεκίνησε από μία αρχική βάση περίπου 6 χιλιάδων όρων που περιελάμβανε για κάθε χώρα το πρώτο διοικητικό επίπεδο και αρχικά μόνο πόλεις με πληθυσμό πάνω από 100χιλ. παγκοσμίως. Για την Ελλάδα, στο λεξιλόγιο εντάχθηκαν τρία διοικητικά επίπεδα και αρχικά όλοι οι οικισμοί με πληθυσμό άνω των 1.000 κατοίκων. β. Το “Συμπληρωματικό λεξιλόγιο τοπωνυμίων Geonames” (Έκδοση ΕΚΤ)³ για τους τόπους που δεν ενέπιπταν στην παραπάνω ιεραρχία, όπως ιστορικές περιοχές (π.χ. Σοβιετική Ένωση), ευρύτερες περιοχές (π.χ. Βαλκάνια) ή

¹ <https://www.geonames.org/>

² <https://www.semantics.gr/authorities/vocabularies/geonames-places-earth/vocabulary-entries>

³ <https://www.semantics.gr/authorities/vocabularies/geonames-supplementary-places/vocabulary-entries>

γεωμορφολογικά στοιχεία που μπορεί να διασχίζουν περισσότερες από μία χώρες, όπως ποτάμια, κλπ.

Τα δύο λεξιλόγια είναι ιεραρχικά (με βάση την Has-Part σχέση), δίγλωσσα και συμμορφώνονται με την συγκεκριμένη κλάση edm:Place του Europeana Data Model (EDM). Μεταξύ τους συνδέονται με δύο προσαρμοσμένες (custom) ιδιότητες `ekt:isPartOfMatch` και `ekt:hasPartMatch` (κατ'αντιστοιχία των `IS-A skos:broadMatch` και `skos:narrowMatch` από την SKOS οντολογία, αντίστοιχα).

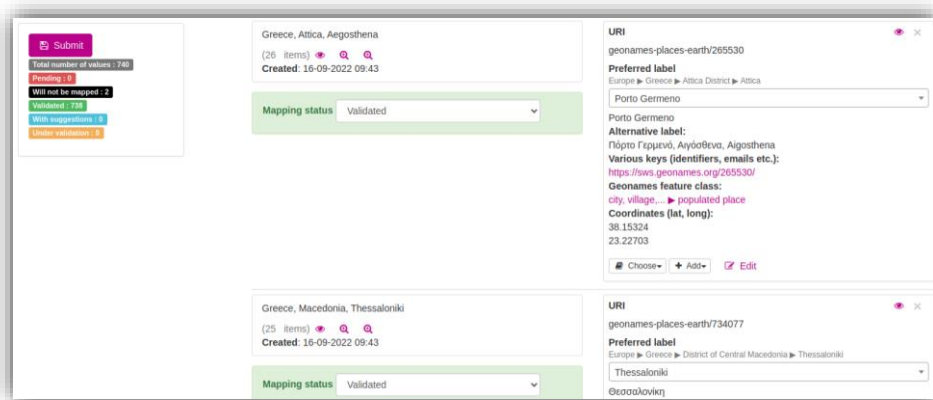
Τα λεξιλόγια σταδιακά εμπλουτίστηκαν με νέους όρους - πάντα διασυνδεδεμένοι με τους αντίστοιχους όρους του Geonames - και έφτασαν στο τέλος των αναδρομικών εμπλουτισμών τους 12.000. Τέλος, υπήρξε σημαντική συνεισφορά ελληνικών όρων στο Geonames, δεδομένου ότι πολλοί νέοι όροι δημιουργήθηκαν κατά τη διαδικασία στη βάση από την ομάδα του SearchCulture.gr και στη συνέχεια, εντάχθηκαν στο λεξιλόγιο του EKT (Papanoti, 2023).

4.2 Επιλογή πεδίων άντλησης γεωχωρικής πληροφορίας

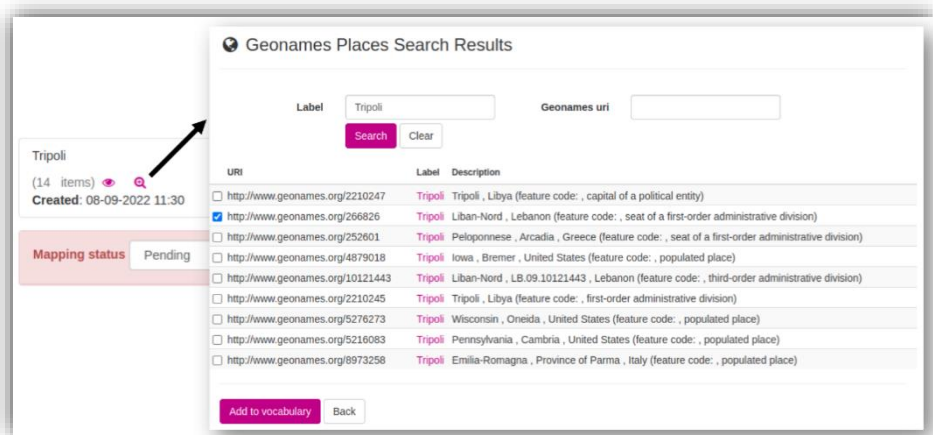
Οι φορείς διάθεσης περιεχομένου στο συσσωρευτή διαθέτουν τις συλλογές μέσα από το πρωτόκολλο OAI-PMH σε ένα από τα μοντέλα διαλειτουργικότητας ESE, EDM, OAI_DC, OAI_QDC ή OAI_DCTERMS που όλα βασίζονται στην οντολογία Dublin Core (EKT, 2024). Εκτός από τα βασικά σχετικά πεδία `dcterms:spatial` και `dc:coverage`, η γεωγραφική πληροφορία μπορεί να έχει ενσωματωθεί σε περιγραφικά πεδία όπως τον τίτλο, την περιγραφή, ή και το θέμα. Για κάθε συλλογή, η επιλογή των πεδίων άντλησης της γεωχωρικής πληροφορίας έγινε κατά περίπτωση. Ένα ελάχιστο ποσοστό των φορέων διέθετε URIs ενώ στην πλειονότητά τους οι τιμές ήταν λεκτικά ενώ, επίσης, σε αρκετές περιπτώσεις είχαμε λεκτικό πιο στοχευμένο απ'ότι η URI αναφορά. Τέλος, πρέπει να σημειωθεί ότι τα πεδία `dcterms:spatial` και `dc:coverage`, εμπεριέχουν εγγενή αμφισημία, δηλαδή, δεν δηλώνεται αν η τιμή αναφέρεται σε τόπο δημιουργίας ή εντοπισμού ενός τεκμηρίου ή τόπου που είναι θέμα σε ένα τεκμήριο, η οποία αμφισημία παραμένει και μετά τους εμπλουτισμούς. Αυτό σημαίνει ότι για τον τελικό χρήστη η σημασιολογία του πεδίου EKT Τόπος ορίζεται ως ο τόπος που σχετίζεται με ένα τεκμήριο.

4.3 Διαδικασία και εργαλεία της αντιστοίχισης

Το Semantics.gr διαθέτει φόρμες αντιστοίχισης στις οποίες επιλέγονται ανά συλλογή και είδος εμπλουτισμού, τα πεδία από τα μεταδεδομένα του φορέα που θα εμπλουτιστούν. Το εργαλείο της αντιστοίχισης προτείνει αυτόματα για κάθε τιμή έναν ή περισσότερους όρους από το λεξιλόγιο-στόχο του EKT. Η αντιστοίχιση βασίζεται στην ομοιότητα λεκτικών (τιμής πεδίου με προτεινόμενη ή εναλλακτική ονομασία όρου του λεξιλογίου) και ο επιμελητής επικυρώνει ή απορρίπτει την αντιστοίχιση (εικόνα 1). Επιπλέον, οι φόρμες αντιστοίχισεων περιλαμβάνουν φίλτρα που κάνουν τις αντιστοιχίσεις πιο στοχευμένες. Διάφορες χρήσιμες επεκτάσεις έγιναν και στη μηχανή ευρετηριασμού Apache Solr, π.χ. ερευνηριάστηκαν επιπλέον όλοι οι γονεϊκοί όροι, έτσι, ώστε π.χ. στην περίπτωση της τιμής "Νάουσα, Ημαθία" το σύστημα να μην προτείνει τη Νάουσα της Πάρου. Ο επιμελητής μπορεί ακόμη να αναζητήσει με χειρωνακτικό τρόπο μία τοποθεσία από το λεξιλόγιο σε περίπτωση που η αυτόματη αντιστοίχιση απέτυχε να επιλέξει τον σωστό όρο (εικόνα 2). Αν ο όρος δεν υπάρχει στο λεξιλόγιο, ο χρήστης μπορεί απευθείας μέσα από τη φόρμα του Semantics.gr να αναζητήσει την τοποθεσία στο Geonames και να την προσθέσει αυτόματα στο λεξιλόγιο ενώ μπορεί εφόσον απαιτείται και να επεξεργαστεί την καρτέλα του όρου. Έτσι, επιτυγχάνεται μία ημι-αυτόματη διαδικασία με επιστημονική επιμέλεια που εξασφαλίζει την υψηλή ποιότητα των αντιστοιχίσεων.



Εικόνα 1. Δύο επικυρωμένες αντιστοιχίσεις σε φόρμα που αντλεί από το dcterms:spatial από επιμέρους συλλογή. Η λειτουργικότητα είναι διαθέσιμη και στα αγγλικά



Εικόνα 2. Αναζήτηση όρου στο GeoNames.org μέσα από το Semantics.gr

Για κάθε όρο των δύο λεξιλογίων αποθηκεύεται στο Semantics το URI, οι προτεινόμενες και οι εναλλακτικές ονομασίες, οι συντεταγμένες, τυχόν βιβλιογραφικές αναφορές, κλπ. Ως αποτέλεσμα, κάθε τεκμήριο εμπλουτίζεται με ένα νέο πεδίο “ΕΚΤ Τόπος” με τη τιμή το URI του όρου από τα δύο λεξιλόγια.

Η διαδικασία των αναδρομικών εμπλουτισμών διήρκεσε από το Σεπτέμβριο του 2022 έως το Φεβρουάριο του 2023, πραγματοποιήθηκε από ομάδα 4 ατόμων και απαιτήθηκε συνολική ανθρωποπροσπάθεια που αντιστοιχεί σε 12 μήνες πλήρους απασχόλησης. Εμπλουτίστηκαν 90 συλλογές που αντιστοιχούσαν σε 585,000 τεκμήρια. Δημιουργήθηκε ένα λεξιλόγιο 12,000 όρων με πολύ μεγάλη κάλυψη ελληνικών τοποθεσιών. Παρήχθησαν 119 φόρμες αντιστοίχισης και 25,000 αντιστοιχίσεις. Η διαδικασία συνεχίζεται για κάθε νέα συλλογή που μπαίνει στο SearchCulture.gr.

4.4 Προκλήσεις ως προς τις αντιστοιχίσεις

Συνηθισμένες προκλήσεις σε όλους τους εμπλουτισμούς περιλαμβάνουν την ύπαρξη γραμματικών λαθών και τη χρήση ξένων χαρακτήρων στις τιμές των πεδίων, οι οποίες προκαλούν θόρυβο στις αντιστοιχίσεις. Ειδικά στους γεωχωρικούς εμπλουτισμούς συνήθεις προκλήσεις ήταν η χρήση κοινών τοπωνυμίων (π.χ. Αγία Άννα, Κάστρο, κλπ) χωρίς λοιπά προσδιοριστικά στοιχεία, η χρήση εναλλακτικών ονομασιών τοποθεσιών, δύσκολων στον εντοπισμό τους, η απόδοση με παράφραση συχνά ξένων ονομάτων, ειδικά όσον αφορά μέρη στη σημερινή Τουρκία. Σε πολλές περιπτώσεις απαιτήθηκε εκτενής έρευνα ταύτισης ενός τόπου με βάση το ψηφιακό αρχείο ή στοιχεία στην περιγραφή. Χρησιμοποιήθηκαν ως πηγές ιδίως η συλλογή “Μετονομασίες των Οικισμών της Ελλάδας” του αποθετηρίου “Πανδέκτης” του Εθνικού Ιδρύματος Ερευνών⁵, οι πλατφόρμες Pleiades

και Pelagios⁶, το VICI.org, ένας αρχαιολογικός άτλαντας της κλασικής αρχαιότητας, αλλά κάποιες φορές και ανεπίσημα blogs μικρασιατών.

5. Υπηρεσίες ανακάλυψης του περιεχομένου

Εκτός από την απλή αναζήτηση, αξιοποιώντας τους σημασιολογικούς εμπλουτισμούς, η πύλη του SearchCulture.gr προσφέρει πολλαπλές δυνατότητες αναζήτησης και πλοήγησης στις εκατοντάδες χιλιάδες των τεκμηρίων που περιλαμβάνονται σε αυτή. Χάρη στους σημασιολογικούς εμπλουτισμούς, τα λεξιλόγια των προσώπων, των θεμάτων, των τύπων των τεκμηρίων, οι ιστορικές περίοδοι, και οι τοποθεσίες χρησιμοποιούνται ως ελεγχόμενα λεξιλόγια στην προηγμένη αναζήτηση και μπορούν να χρησιμοποιηθούν μεμονωμένα ή συνδυαστικά. Μπορεί ο χρήστης π.χ. π.χ. να αναζητήσει “επιστολή” (τύπος τεκμηρίου), “Ελευθέριος Βενιζέλος” (πρόσωπο), “1920-1930” (εύρος χρονολογιών), “Μικρά Ασία” (τόπος) και να ανακτήσει τα σχετικά αποτελέσματα. Με τον ίδιο τρόπο, τα λεξιλόγια χρησιμοποιούνται και ως φίλτρα για τον επιμερισμό των αποτελεσμάτων της αναζήτησης (facets). Τα λεξιλόγια χρησιμοποιούνται επίσης ως κανάλια πλοήγησης στα τεκμήρια μέσω ξεχωριστών σελίδων.

5.1 Αναζήτηση με βάση την τοποθεσία και πλοήγηση στο χάρτη

Το νέο πεδίο ΕΚΤ Τόπος που προστέθηκε στην προηγμένη αναζήτηση αντλεί από την ελεγχόμενη λίστα των όρων των δύο γεωχωρικών λεξιλογίων του ΕΚΤ και περιλαμβάνει λειτουργία αυτόματης συμπλήρωσης. Για κάθε όρο εμφανίζεται η ιεραρχία του, τυχόν εναλλακτικές ονομασίες, η ονομασία στα αγγλικά, κλπ, υποστηρίζοντας την αποσαφήνιση του όρου.

Στα αποτελέσματα μιας αναζήτησης προστέθηκε η δυνατότητα να μπορεί ο χρήστης να επιλέξει να δει τα τεκμήρια ως πλέγμα αποτελεσμάτων ή πάνω σε χάρτη. Στη σελίδα τεκμηρίου προστέθηκε η προβολή της τοποθεσίας του τεκμηρίου σε χάρτη.

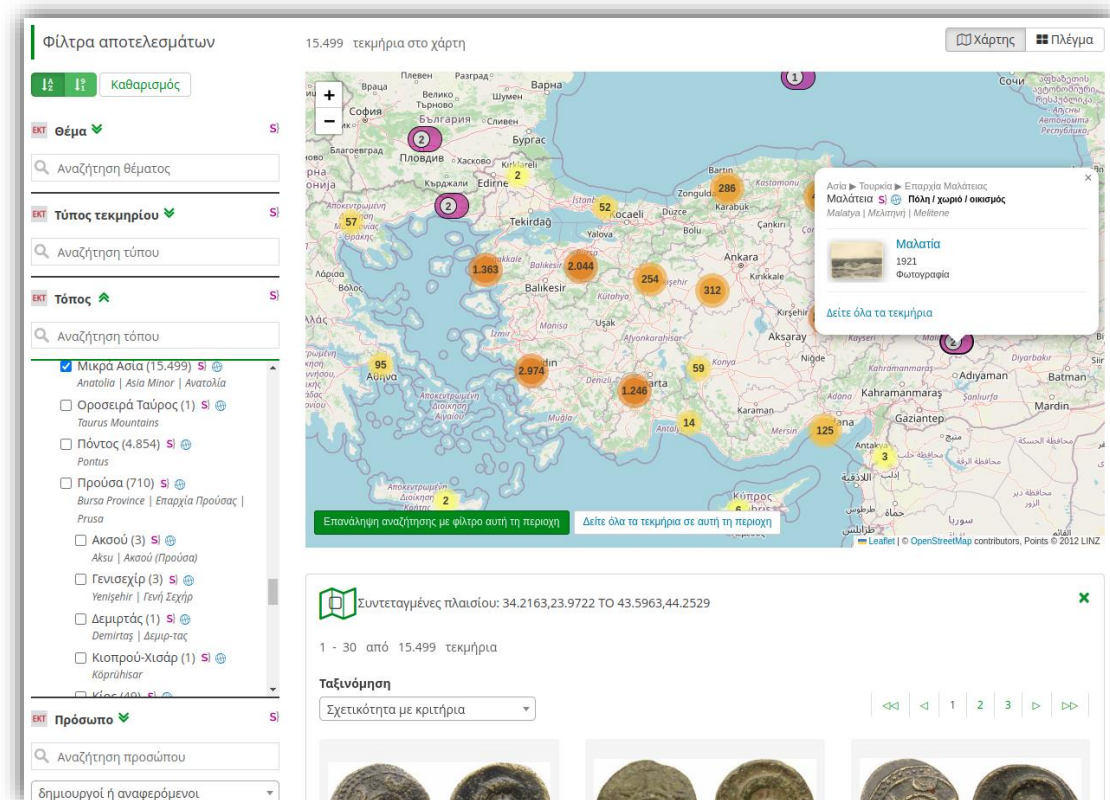
Αναπτύχθηκαν επίσης 2 βασικοί τρόποι πλοήγησης: η πλοήγηση με βάση τα δύο γεωχωρικά λεξιλόγια του ΕΚΤ τα οποία εμφανίζονται ως μία κοινή ιεραρχία οργανωμένη αλφαβητικά και η πλοήγηση σε διαδραστικό χάρτη.

Για την τοποθέτηση των τεκμηρίων στο χάρτη, ευρετηριάστηκαν οι συντεταγμένες των όρων με τους οποίους έχουν εμπλουτιστεί στο ειδικό γεωχωρικό πεδίο “LatLonPointSpatialField” του Solr. Για να υποστηριχθεί η ιεραρχική αναζήτηση και πλοήγηση ευρετηριάζονται για κάθε αντικείμενο και όλοι οι ευρύτεροι όροι. Η υλοποίηση του χάρτη έγινε με τη χρήση του leaflet.js⁴ και του OpenStreetMap⁵.

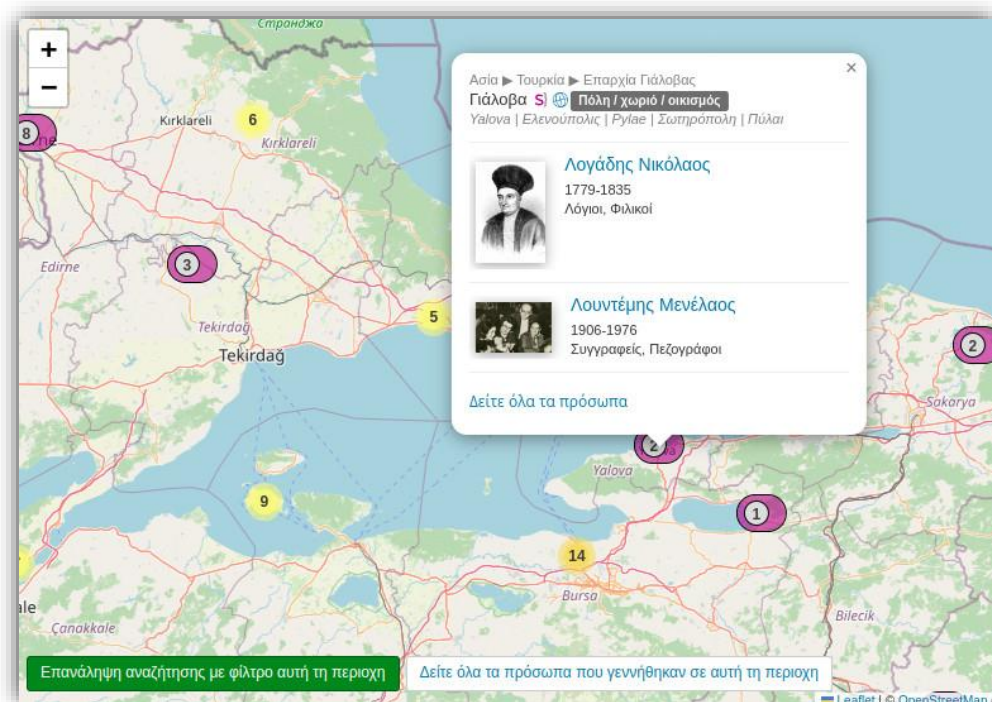
Η εμφάνιση χιλιάδων τεκμηρίων στο χάρτη, ανέγειρε πολλές προκλήσεις, τόσο όσον αφορά την εμπειρία του χρήστη, την ανταπόκριση του συστήματος, την υπολογιστική ισχύ και μνήμη. Αφενός, ο αριθμός των επιμέρους τοποθεσιών που εμφανίζονται σε μία αναζήτηση μπορεί να είναι ιδιαίτερα μεγάλος, και αφετέρου ένας πολύ μεγάλος αριθμός τεκμηρίων μπορεί να σχετίζεται με κάθε τοποθεσία. Για να αντιμετωπιστεί αυτό, τα τεκμήρια εμφανίζονται σε συστάδες (clusters), οι οποίες επεκτείνονται και ανοίγουν βαθμιαία καθώς ο χρήστης κάνει zoom στο χάρτη.

⁴ <https://leafletjs.com/>

⁵ <https://www.openstreetmap.org/#map=7/38.359/23.810>



Εικόνα 3. Αναζήτηση τεκμηρίων στον χάρτη



Εικόνα 4. Αναζήτηση προσώπων στον χάρτη

Οι γεωχωρικοί εμπλουτισμοί ενσωματώθηκαν περαιτέρω στη σελίδα των Προσώπων⁶, όπου έγινε δυνατή η αναζήτηση με βάση τον τόπο γέννησης ή θανάτου. Επιπλέον προστέθηκε η δυνατότητα προβολής και πλοήγησης των προσώπων στον χάρτη, με βάση τον τόπο γέννησης ή θανάτου (Εικόνα 5).

5.2 Θεματικές εκθέσεις και γεωχωρική πληροφορία

Αξιοποιώντας τα πολλαπλά αυτά κριτήρια αναζήτησης, η ομάδα του SearchCulture.gr αναπτύσσει και επιμελείται θεματικές εκθέσεις που αποτελούνται από μικρότερα ή μεγαλύτερα σύνολα τεκμηρίων τα οποία, έχοντας μια νοηματική συνάφεια, αφηγούνται μια ιστορία. Οι θεματικές εκθέσεις αναδεικνύουν το εύρος και την ποικιλία των ιστορικών, καλλιτεχνικών, αρχαιολογικών, κοινωνικών και επιστημονικών θεμάτων που μπορεί ο χρήστης να αναζητήσει στο SearchCulture.gr. Οι σημασιολογικοί εμπλουτισμοί με βάση τη γεωχωρική πληροφορία άνοιξαν νέες διόδους προβολής του πλούτου των πολιτιστικών τεκμηρίων που φιλοξενούνται στο SearchCulture.gr, καθώς η προβολή σε χάρτη είναι ένας από τους πλέον γοητευτικούς τρόπους ανακάλυψης του περιεχομένου.

Πολλές θεματικές εκθέσεις έχουν δημιουργηθεί μετά την ανάπτυξη της νέας λειτουργικότητας, όπως οι 10 θεματικές εκθέσεις -αφιέρωμα στους πολιτισμούς της Μικράς Ασίας, ενώ αξίζει να σημειωθεί και η Έκθεση για τη Μικρασιατική Εκστρατεία⁷ (Εικόνα 5).

H Μικρασιατική Εκστρατεία
Οι επιχειρήσεις του Ελληνικού Στρατού στη Μικρά Ασία και η κατάρρευση του μετώπου

16-05-2022 | Έλενα Λαγούδη | ΕΚΤ

Τα 100 χρόνια από την Επανάσταση του 1821 γιορτάστηκαν στην Ελλάδα και στο στράτευμα της Μικράς Ασίας μέσα σε ένα κλίμα ρευστότητας, πολιτικής και διπλωματικής αστάθειας και απογοήτευσης.

Η Μικρασιατική Εκστρατεία ήταν το τελευταίο χαρτί της Ελλάδας για να λήξει υπέρ της η υπόθεση της Μικράς Ασίας με στρατιωτικά μέσα: η καταδίκη του κεμαλικού στρατού και η πλήρης συντριβή του, μια προσπάθεια που είχε ξεκινήσει από την άνοιξη του 1921.

Στη θεματική έκθεση θα βρείτε πληθώρα φωτογραφιών που απεικονίζουν τις αρχικές, ενθαρρυντικές στιγμές της απόβασης, σκηνές από την προέλαση της Στρατιάς Μικράς Ασίας και από μάχες, καθημερινές σκηνές από τη ζωή στα προμαχώματα και την υποχώρηση προς το Αιγαίο. Τηλεγραφήματα, αναφορές και επιστολές φωτίζουν την πολιτική σκηνή και τις επικοινωνίες ανάμεσα στη Στρατιά και την Ελλάδα, ενώ χάρτες της ευρύτερης περιοχής και σημαντικών πόλεων δίνουν χρήσιμες γεωγραφικές πληροφορίες για την πορεία της Εκστρατείας.

Τα τεκμήρια της έκθεσης εικονογραφούν την προέλαση της Στρατιάς Μικράς Ασίας προς Φιλαδέλφεια - Πάναρμο - Προύσα, την εαρινή επίθεση του Μαρτίου 1921, τη θερινή επίθεση του Ιουνίου 1921, την κρίσιμη σύσκεψη της Κιουτάχειας που κατέληξε στη μοιραία απόφαση της προέλασης στον Σαγγάριο, την πορεία μέσα από την απάνθρωπη Αλμυρά Έρημο και τις φονικές μάχες στον Σαγγάριο μπροστά από τις τουρκικές αμυντικές γραμμές τον Αύγουστο του 1921, την κατάρρευση του Μετώπου στο Αφιόν Καραχισάρ και την αναπάντρετη πορεία προς την ήττα.

Μικρασιατική εκστρατεία: στο μέτωπο, Γαζιάδης, Δημήτριος και Μιχαήλ (Dag Films) Αθήνα, Μητροπόλεως 1, CC BY 4.0
Ελληνικό Λαοτεχνικό και Ιστορικό Αρχείο (ΕΛΙΑ)- Μορφωτικό Ίδρυμα Εθνικής Τραπέζης (ΜΙΕΤ)

Ανακαλύψτε τα 2.059 τεκμήρια της θεματικής έκθεσης

Πληκτρολογήστε λέξη ή φράση
Αναζήτηση

Φίλτρα αποτελεσμάτων
Καθαρισμός

ΕΚΤ Θέμα
Αναζήτηση θέματος

ΕΚΤ Τύπος τεκμηρίου
Αναζήτηση τύπου

ΕΚΤ Τόπος
Αναζήτηση τόπου

2.059 τεκμήρια στο χάρτη

Χάρτης Πλέγμα

Ευρώπη ► Ελλάδα ► Περιφέρεια Βορείου Αιγαίου ► Νάρος Χίου
Νήσος Ψαρά
Psara Island | Nisi Psara | Psara | Ψαρά

6 Ιουνίου 1922 Ο βασιλιάς Κωνσταντίνος Α' συνομιλεί με τον...
1922
Φωτογραφία

Δείτε όλα τα τεκμήρια

Εικόνα 5. Η θεματική έκθεση για την Μικρασιατική Εκστρατεία

⁶ searchculture.gr/aggregator/persons/search?&resultsMode=MAP_PLACE_OF_BIRTH

⁷ https://www.searchculture.gr/aggregator/portal/thematicCollections/mikrasiatiki

6. Σχετικές εργασίες και επίλογος

Οι συσσωρευτές μεγάλου όγκου ετερογενών πολιτιστικών δεδομένων εφαρμόζουν διάφορες στρατηγικές για την ομογενοποίησή τους ώστε να μπορούν να αναπτύξουν υπηρεσίες αναζήτησης και πλοήγησης (Stiller, 2014, EuropeanaTech TF 2015, Smith, 2021). Η στρατηγική του SearchCulture.gr είναι σε συνάφεια με τις Συστάσεις της EuropeanaTech TF για τους πολυγλωσσικούς και σημασιολογικούς εμπλουτισμούς (2014).

Η ίδια η Europeana χρησιμοποιεί αυτόματη αντιστοίχιση μεταξύ των τιμών των πεδίων `dc:terms:spatial` και `dc:coverage` και του GeoNames (Manguinhas, 2016). Εν τούτοις, οι αυτοματοποιημένες διαδικασίες αντιστοίχισης έχουν ως αποτέλεσμα χαμηλή κάλυψη και μεγάλο ποσοστό λαθών (Stiller, 2014) και ως εκ τούτου δεν ενδείκνυνται για την ανάπτυξη εκλεπτισμένων μηχανισμών ανακάλυψης (Europeana TF on Enrichment and Evaluation, 2015). Π.χ. το εργαλείο αντιστοίχισης της Europeana δεν μπορεί να διαχωρίσει ανάμεσα σε διαφορετικούς βαθμούς διοικητικής διαίρεσης με το ίδιο όνομα (π.χ. Νομός Θεσσαλονίκης, πόλη Θεσσαλονίκης) και άρα επιλέγει πάντα τον ίδιο όρο (Stiller, 2014), όπως και δεν μπορεί να ξεχωρίσει μεταξύ ομώνυμων τοπόσημων.

Στην συγκριτική ανάλυση που διηγήρησε το EuropeanaTech Task Force (2014) υποδηλώνεται η αναγκαιότητα ανθρώπινης εμπλοκής στη διαδικασία προκειμένου να επιτυγχάνεται καλή ποιότητα στα αποτελέσματα, το οποίο συμβαδίζει με τη δική μας προσέγγιση. Το SAGE (Kaldeli, 2023) από την άλλη, είναι ένα εργαλείο σημασιολογικού εμπλουτισμού που συνδυάζει μηχανισμούς TN και ανθρώπινη επικύρωση των αυτόματων αντιστοιχίσεων, υπολείπεται όμως σε λειτουργικότητα σε σχέση με το Semantics.gr, όπως τα φίλτρα που μπορούν να χρησιμοποιηθούν στις φόρμες αντιστοίχισης.

Από τους εθνικούς και θεματικούς συσσωρευτές της Europeana, κάποιοι ζητούν τα μεταδεδομένα να διατίθενται εμπλουτισμένα από τους φορείς λίγοι εμπλουτίζουν τα δεδομένα μετά τη συγκομιδή, ενώ οι περισσότεροι ευρετηριάζουν τα δεδομένα χωρίς εμπλουτισμό (Stiller, 2014).

Αναζήτηση με βάση τον τόπο προσφέρεται από τη Γερμανική Ψηφιακή Βιβλιοθήκη⁸. Η CulturalItalia.it προσφέρει φιλτράρισμα των αποτελεσμάτων με βάση τον τόπο και η σουηδική Kringla⁹ προσφέρει αναζήτηση με βάση την επαρχία και πλοήγηση στα τεκμήρια σε χάρτη αλλά μόνο σε ένα μικρό σύνολο τεκμηρίων και μόνο στα σουηδικά.

Ως εκ τούτου, το SearchCulture.gr ξεχωρίζει ως εθνικός διαθεματικός συσσωρευτής μέσω της ολοκληρωμένης μεθοδολογίας σημασιολογικού εμπλουτισμού που υιοθετεί και η οποία προσφέρει μεγάλη κάλυψη σε εμπλουτισμούς ταυτόχρονα με μεγάλη ακρίβεια, γιατί: α. προσαρμόζει τη μεθοδολογία στις ανάγκες κάθε συλλογής, β. υποστηρίζει τους επιμελητές με εργαλεία αυτοματισμού που διαθέτουν μηχανισμούς αυτοβελτίωσης και άλλες υποστηρικτικές λειτουργίες, γ. χρησιμοποιεί ελεγχόμενο λεξιλόγιο-στόχο που αυξάνεται σταδιακά για να καλύψει τις ανάγκες των συγκεκριμένων συλλογών.

⁸ <https://www.deutsche-digitale-bibliothek.de/>

⁹ <https://www.kringla.nu/>

Στο μέλλον προσβλέπουμε να εντάξουμε τη χρονική διάσταση μαζί με τη χωρική αντλώντας έμπνευση από έργα όπως το Φινλανδικό CultureSampo (Hyvonen, 2009) και το πιο πρόσφατο World Historical Gazetteer (Grossner et al, 2023).

Βιβλιογραφία

EKT (2024), Βασικές προδιαγραφές διαλειτουργικότητας για την ένταξη συλλογών στον Εθνικό Συσσωρευτή Ψηφιακών Πολιτιστικών Πόρων, 2η έκδοση, Αθήνα: Εθνικό Κέντρο Τεκμηρίωσης και Ηλεκτρονικού Περιεχομένου

Georgiadis, H. et al. (2022). Enriching the Greek National Cultural Aggregator with Key Figures in Greek History and Culture: Challenges, Methodology, Tools and Outputs. In: Linking Theory and Practice of Digital Libraries. TPDL 2022

Georgiadis, H. et al. (2021). Publishing LOD Vocabularies in Any Schema with Semantics.gr. In: Garoufallou, E., Oualle-Perandonnes, MA., Vlachidis, A. (eds) Metadata and Semantic Research. MTSR 2021.

Georgiadis, H., et al. (2018). Using type and temporal semantic enrichment to boost content discoverability and multilingualism in the Greek cultural aggregator SearchCulture.gr. International Journal of Metadata, Semantics and Ontologies, 13(1), 75–92.

Grossner, K., Grunewald, S. & Mostern, R. Bringing places from the distant past to the present: a report on the World Historical Gazetteer. Int J Digit Libr 24, 159–162 (2023). <https://doi.org/10.1007/s00799-022-00341-2>

Europeana Task Force on Enrichment and Evaluation (2015) Final Report <https://pro.europeana.eu/project/evaluation-and-enrichments>

EuropeanaTech Task Force on a Multilingual and Semantic Enrichment Strategy (2014): final report, <https://pro.europeana.eu/project/multilingual-and-semantic-enrichment-strategy>

Kaldeli, E., (2023) Combining AI tools with human validation to enrich cultural heritage metadata, <https://pro.europeana.eu/post/combining-ai-tools-with-human-validation-to-enrich-cultural-heritage-metadata>

Manguinhas, H. et al. (2016). Exploring Comparative Evaluation of Semantic Enrichment Tools for Cultural Heritage Metadata. In: TPDL 2016.

Manguinhas H. (2016), Europeana Semantic Enrichment Framework, technical documentation, available at <https://pro.europeana.eu/page/europeana-semantic-enrichment>

Papanoti, A., et al. “Greek Culture on the Map: Place-Based Enrichment Scheme at the Greek National Cultural Data Aggregator.” 17th International Conference on Metadata and Semantics Research, 23rd - 27th October 2023, 2023.

Stiller, J., et al. (2014). Automatic Enrichments with Controlled Vocabularies in Europeana: Challenges and Consequences. EuroMed 2014.

Smith, M. (2021): Linked open data and aggregation infrastructure in the cultural heritage sector, A case study of SOCH, a linked data aggregator for Swedish open cultural heritage, in Information and Knowledge Organisation in Digital Humanities, Routledge,

Linked Open Data for Libraries, Archives and Museums, EuropeanaTech Insight, Issue 7
<https://pro.europeana.eu/page/issue-7-lodlam>

Hyvonen, E., et al., CultureSampo (2009) - Finnish Culture on the Semantic Web 2.0: Thematic Perspectives for the End-User.”
<https://www.archimuse.com/mw2009/papers/hyvonen/hyvonen.html>.