



SearchCulture.gr

μία Πρότυπη Πλατφορμα Συσσώρευσης & Αναζήτησης Πολιτιστικού Περιεχομένου: Παρουσίαση – Αρχιτεκτονική



{ Search
Culture.gr }

Ενιαία Αναζήτηση
σε ανοικτό ψηφιακό
περιεχόμενο



Κείμενο: Χάρης Γεωργιάδης, PhD

Σχεδιασμός Έκδοσης: Δήμητρα Πελεκάνου

Επιμέλεια & Οπτικοποίηση Ύλης: Χάρης Γεωργιάδης, Δήμητρα Πελεκάνου

Ομάδα Έργου Aggregator/SearchCulture.gr:

Οργάνωση Σχεδιασμού- Ανάπτυξη Λογισμικού: Χάρης Γεωργιάδης, PhD

Συντονισμός: Δέσποινα Χαρδούβελη, Χάρης Γεωργιάδης

Γραφιστικός Σχεδιασμός: Δήμητρα Πελεκάνου

Συνεισφορά στην διαμόρφωση της αρχικής σελίδας: Ράνια Σταθοπούλου, PhD

Επιμέλεια: Δέσποινα Χαρδούβελη, Χάρης Γεωργιάδης, Παρασκευάς Καμάτσος, Έλενα Λαγούδη, Μαργαρίτης Προέδρου

Σημαιολογικός Εμπλουτισμός: Αγάθη Παπανότη, Χάρης Γεωργιάδης, Μαρία Πάσχου, Αλεξάνδρα Ρουμπάνη

Έλεγχοι συμμόρφωσης περιεχομένου με προδιαγραφές: Ράνια Σταθοπούλου

Ομάδα Συστημάτων και Δικτύων: Μανώλης Μεγαλοοικονόμου, Μιχάλης Σίμου, Γιώργος Ζαχαριάδης

HelpDesk – Επικοινωνία με τους φορείς: Αγάθη Μπελούλη

Σύστημα Harvester:

Οργάνωση Σχεδιασμού- Ανάπτυξη Λογισμικού: Χάρης Γεωργιάδης

Προτεινόμενος τρόπος αναφοράς:

EKT(2016), *SearchCulture.gr: Μία Πρότυπη Πλατφόρμα Συσώρευσης & Αναζήτησης Πολιτιστικού Περιεχομένου, Παρουσίαση – Αρχιτεκτονική*. Αθήνα: Εθνικό Κέντρο Τεκμηρίωσης

Copyright © 2016 Εθνικό Κέντρο Τεκμηρίωσης | ΕΙΕ

δ. Βασιλέως Κωνσταντίνου 48, 11635 Αθήνα | τ.: 210 7273900

| f: 210 7246824 | e: ekt@ekt.gr | www.ekt.gr



Το έργο αυτό διατίθεται με άδεια Creative Commons
Αναφορά-Μη Εμπορική Χρήση-Όχι Παράγωγα Έργα 4.0 Ελλάδα

Προκειμένου να δείτε αντίγραφο της άδειας επισκεφθείτε:

<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.el>

{ Search Culture.gr }

Ενιαία Αναζήτηση σε
ανοικτό ψηφιακό περιεχόμενο



ΠΕΡΙΕΧΟΜΕΝΑ

1	Εισαγωγή	5
1.1	Το ξεκίνημα	7
1.2	Οικοσύστημα / Γενική αρχιτεκτονική.....	7
1.3	Πλεονεκτήματα	14
2	Harvester: Σύστημα Συγκομιδής Ψηφιακού Περιεχομένου	17
2.1	Λειτουργικά Χαρακτηριστικά	18
2.2	Εφαρμογή Ιστού	20
2.3	RESTful Υπηρεσίες	28
2.4	Harvester: αρχιτεκτονική συστήματος.....	33
2.5	Πειραματική Αξιολόγηση Συστήματος Συγκομιδής	35
3	SearchCulture.gr: Σύστημα Συσσωρευτή Περιεχομένου/ Ενιαίας Αναζήτησης	36
3.1	Λειτουργικά Χαρακτηριστικά	37
3.2	Ανοικτή Δικτυακή Πύλη SearchCulture.gr	43
3.2.1	Αρχική Σελίδα	43
3.2.2	Αναζήτηση	45
3.2.3	Καρτέλα τεκμηρίου	49
3.2.4	Φορείς/συλλογές.....	52
3.2.5	Ιεραρχική πλοήγηση περιεχομένου μέσα από τους ΕΚΤ Τύπους	53
3.2.6	Διεπαφή στα Αγγλικά	54
3.3	Διαχειριστικό Περιβάλλον.....	55
3.3.1	Διαχείριση Οργανισμών και Αποθετηρίων/Συλλογών.....	56
3.3.2	Διαδικασίες Εισαγωγής Μεταδεδομένων.....	62
3.3.3	Μαζικός σημασιολογικός εμπλουτισμός από το Semantics.gr.....	67
3.3.4	Διαδικασίες Εισαγωγής Αρχείων Προεπισκόπησης.....	74
3.3.5	Διαχείριση Σημασιολογικών Οντοτήτων.....	76
3.3.6	Αναζήτηση, Προβολή και Επεξεργασία εγγραφών τεκμηρίων.....	79
3.3.7	Παραμετροποίηση EDM.....	82
3.4	SearchCulture.gr: αρχιτεκτονική συστήματος.....	85
3.4.1	Persistent Layer	87
3.4.2	Service Layer	95
3.4.3	To Web layer.....	97
3.5	Πειραματική Αξιολόγηση Συσσωρευτή.....	98

Κατάλογος Σχημάτων	101
Κατάλογος Πινάκων	104
Δημοσιεύσεις	105
Αναφορές	106

1 Εισαγωγή



Το Σύστημα Συσσώρευσης Περιεχομένου / Ενιαίας Αναζήτησης SearchCulture.gr είναι ένας συσσωρευτής περιεχομένου (aggregator) συνοδευόμενος από μία ανοικτή διαδικτυακή πύλη που προσφέρει κεντρική πρόσβαση και ενιαία αναζήτηση σε ψηφιακό πολιτιστικό περιεχόμενο που παράγεται στη χώρα μας από έγκριτους φορείς. Ο συσσωρευτής συλλέγει μεταδεδομένα που εξασφαλίζουν πρόσβαση σε ψηφιακό ή ψηφιοποιημένο περιεχόμενο ποικίλων τύπων από τους πολιτιστικούς φορείς που το παράγουν, τα εμπλουτίζει σημασιολογικά και τα διαθέτει στο ευρύ κοινό μέσω ενός σύγχρονου και φιλικού περιβάλλοντος που διαθέτει προηγμένες λειτουργίες αναζήτησης και εξελιγμένες δυνατότητες πλοήγησης.

Η συλλογή του περιεχομένου πραγματοποιείται από το σύστημα Harvester, ένα αυτόνομο σύστημα συγκομιδής μεταδεδομένων και ψηφιακών αρχείων που ανέπτυξε το ΕΚΤ, το οποίο εξυπηρετεί ήδη διάφορα συστήματα που απαιτούν συλλογή περιεχομένου, όπως το SearchCulture.gr, ο συσσωρευτής CRIS [4] του ΕΚΤ, η υπηρεσία πιστοποίησης προδιαγραφών περιεχομένου – Validator και η

υπηρεσία ασφαλούς διαφύλαξης ψηφιακού περιεχομένου – Preservator. Σε αντίθεση με παρεμφερή εργαλεία όπως το REPOX [3], ο Harvester συλλέγει εκτός από μεταδεδομένα και ψηφιακά αρχεία ενώ διαθέτει και μία υπηρεσία δημιουργίας εικόνων προεπισκόπησης.

Το παρόν έγγραφο αποτελεί μία αναλυτική παρουσίαση της υποδομής του SearchCulture.gr η οποία περιλαμβάνει και το αυτόνομο σύστημα Harvester τις υπηρεσίες του οποίου χρησιμοποιεί ο συσσωρευτής. Η παρουσίαση των δύο συστημάτων περιλαμβάνει περιγραφή των λειτουργικών χαρακτηριστικών τους, αναλυτική περιγραφή των γραφικών διεπαφών χρήστη τους και τεχνική τεκμηρίωση της αρχιτεκτονικής τους.

1.1 Το ξεκίνημα

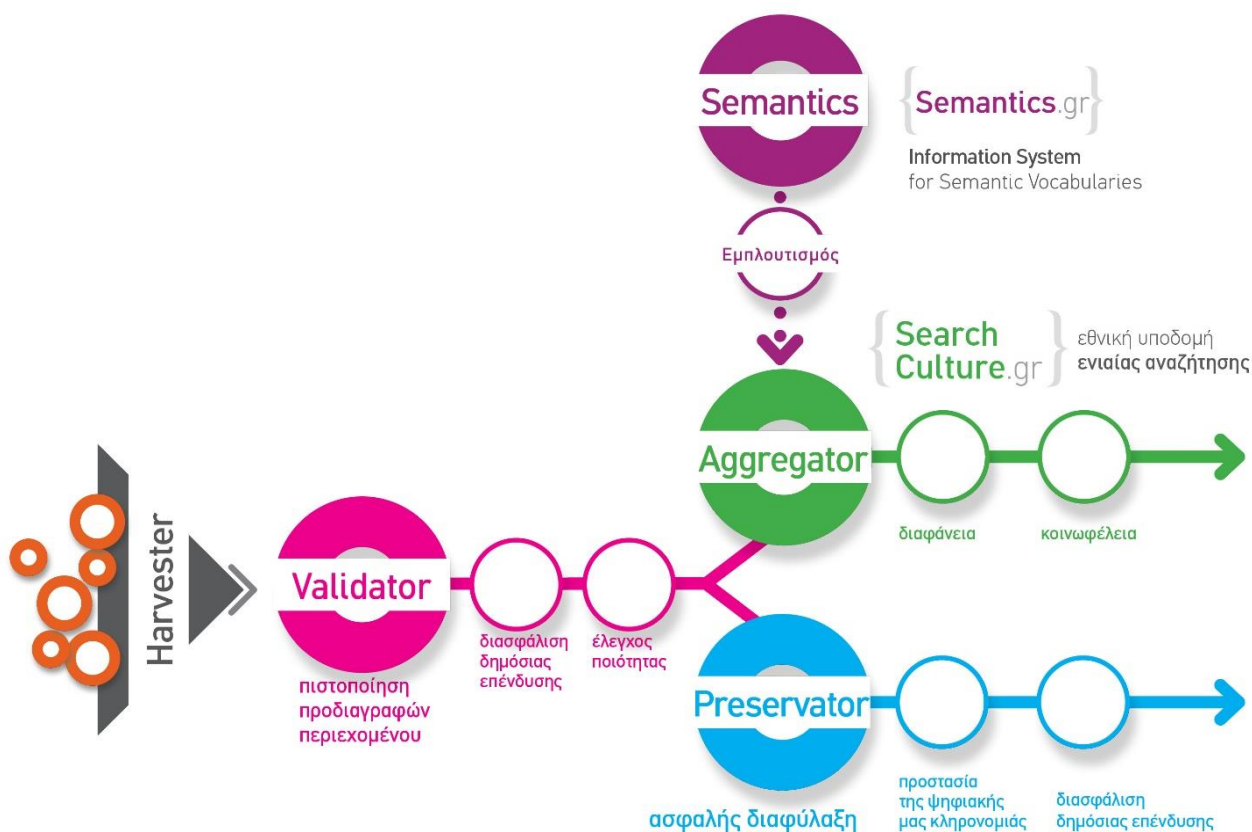
Η ανάπτυξης της υπηρεσίας Ενιαίας Αναζήτησης Ψηφιακού Περιεχομένου SearchCulture.gr έγινε στο πλαίσιο του έργου «Πλατφόρμα Παροχής Υπηρεσιών Κατάθεσης, Διαχείρισης και Διάθεσης Ανοιχτών Δεδομένων και Ψηφιακού Περιεχομένου». Σε πρώτη φάση εντάχθηκε το ψηφιακό υλικό των φορέων των Προσκήσεων 31 και 31.2 του Επιχειρησιακού Προγράμματος «Ψηφιακή Σύγκλιση» της Ειδικής Γραμματείας Ψηφιακού Σχεδιασμού και της Ειδικής Υπηρεσίας Διαχείρισης του Επιχειρησιακού Προγράμματος Ψηφιακή Σύγκλιση του Υπουργείου Ανάπτυξης, Ανταγωνιστικότητας και Ναυτιλίας. Οι φορείς που συμμετείχαν στις προσκλήσεις ψηφιοποίησαν και τεκμηρίωσαν περιεχόμενο και δημιούργησαν συστήματα απόθεσης για την διάχυση του παραγόμενου περιεχομένου. Τα μεταδεδομένα, τα ψηφιακά αρχεία αλλά και τα ίδια τα συστήματα απόθεσης έπρεπε να συμμορφώνονται σε μία σειρά προδιαγραφών. Οι προδιαγραφές αυτές ορίστηκαν από το ΕΚΤ στο έγγραφο «Προδιαγραφές Διαλειτουργικότητας για Ανοιχτό Ψηφιακό Περιεχόμενο» [1] και προέκυψαν από την εμπειρία του ΕΚΤ αλλά και από το διεθνές περιβάλλον και τα πρότυπα στις αντίστοιχες περιοχές. Στα πλαίσια του έργου, το ΕΚΤ είχε αναλάβει ρητά τις υπηρεσίες Ενιαίου Καταλόγου και Ασφαλούς Διαφύλαξης Ψηφιακού Περιεχομένου, την υπηρεσία Πιστοποίησης και Ελέγχου Συμμόρφωσης με τις Προδιαγραφές Διαλειτουργικότητας και την υπηρεσία Ασφαλούς Διαφύλαξης. Στο παρόν έγγραφο παρουσιάζεται η υπηρεσία Ενιαίου Καταλόγου, δηλαδή το Σύστημα Συσσώρευσης Περιεχομένου / Ενιαίας Αναζήτησης SearchCulture.gr καθώς και το σύστημα που ο συσσωρευτής χρησιμοποιεί για την συγκομιδή του περιεχομένου από τις πηγές, το σύστημα Harvester.

Η υπηρεσία ανέβηκε για πρώτη φορά στο διαδίκτυο στις αρχές του 2015. Έκτοτε η υπηρεσία αναβαθμίζεται και επεκτείνεται συνεχώς με νέες λειτουργίες ενώ παράλληλα εμπλουτίζεται με νέες συλλογές που προέρχονται από έγκριτους φορείς της χώρας.

1.2 Οικοσύστημα / Γενική αρχιτεκτονική

Το SearchCulture.gr αποτελεί ουσιαστικά τμήμα μιας επεκτάσιμης και παραμετροποιήσιμης πλατφόρμας η οποία αποτελείται από επιμέρους συστήματα και εργαλεία που ολοκληρώνουν το κύκλο ζωής του περιεχομένου από τη συλλογή και τον έλεγχο, μέχρι τη διάθεση και την ασφαλή διαφύλαξή του. Η πλατφόρμα χρησιμοποιεί το σύστημα συγκομιδής μεταδεδομένων και ψηφιακών αρχείων Harvester, το οποίο συλλέγει μεταδεδομένα και ψηφιακά αρχεία από τους φορείς περιεχομένου και επιπροσθέτως δημιουργεί εικόνες προεπισκόπησης. Μετά την συγκομιδή, το περιεχόμενο ελέγχεται και πιστοποιείται από το Σύστημα Πιστοποίησης Προδιαγραφών Διαλειτουργικότητας και Περιεχομένου Validator και εμπλουτίζεται σημασιολογικά από το Σύστημα Λεξιλογίων και Σημασιολογικού Εμπλουτισμού Semantics.gr. Η διαδικασία εισαγωγής πραγματοποιείται μέσα από το Διαχειριστικό Περιβάλλον του SearchCulture.gr και το περιεχόμενο τελικά δημοσιοποιείται στη Ανοικτή Διαδικτυακή Πύλη Ενιαίας Αναζήτησης και Διάθεσης Περιεχομένου ως Linked Data [9]. Στο τελευταίο βήμα και για τους φορείς που το επιθυμούν, το περιεχόμενο εισάγεται στο Σύστημα Ασφαλούς Διαφύλαξης Περιεχομένου – Preservator. Η πλατφόρμα και τα επιμέρους συστήματα σχεδιάστηκαν και υλοποιήθηκαν εξ ολοκλήρου από το ΕΚΤ.

Στο Σχήμα 1 παρουσιάζεται με παραστατικό τρόπο ο κύκλος ζωής του ψηφιακού περιεχομένου των φορέων μέσα από τις υπηρεσίες και τις υποδομές του ΕΚΤ οι οποίες υποστηρίζουν την συσσώρευση, την πιστοποίηση, την διάχυση, τον εμπλουτισμό και την ασφαλή διαφύλαξη του ψηφιακού περιεχομένου.



Σχήμα 1. Ο κύκλος ζωής του ψηφιακού περιεχομένου στις υποδομές του ΕΚΤ : συλλογή, πιστοποίηση, εμπλουτισμός, συσσώρευση-διάχυση και ασφαλής διαφύλαξη.

Ακολουθεί μία σύντομη περιγραφή των επιμέρους αυτών συστημάτων με έμφαση στην δομή τους και στον τρόπο με τον οποίο αλληλοεπιδρούν μεταξύ τους. Παράλληλα, αναλύονται λειτουργικές αλλά και μη λειτουργικές απαιτήσεις που οδήγησαν σε συγκεκριμένες σχεδιαστικές επιλογές κατά την φάση του σχεδιασμού και της ανάπτυξης της συνολικής πλατφόρμας.



Σύστημα συσσώρευσης περιεχομένου / Υπηρεσία ενιαίας αναζήτησης

Το **SearchCulture.gr** είναι ένα σύστημα συσσώρευσης (**Aggregator**) το οποίο συσσωρεύει μεταδεδομένα και αρχεία προεπισκόπησης και τα διαθέτει μέσα από μία ανοικτή διαδικτυακή πύλη (<http://searchculture.gr>). Η διαδικτυακή πύλη παρέχει προηγμένες δυνατότητες αναζήτησης, πλοήγησης και εξερεύνησης σε ένα ετερογενές ψηφιακό περιεχόμενο και αναδεικνύει το πρωτότυπο περιεχόμενο των φορέων καθώς αυξάνει την ευρεσιμότητά του και διατηρεί συνδέσμους προς τα σημεία σταθερής απόθεσής του. Το

SearchCulture.gr μπορεί να φιλοξενεί σημασιολογικά εμπλουτισμένο περιεχόμενο που διατηρεί συνδέσμους σε λεξιλόγια και θησαυρούς όρων προβάλλοντάς το έτσι και ως Ανοικτά Διασυνδεδεμένα Δεδομένα (Linked Data). Το SearchCulture.gr εμπλουτίστηκε πρόσφατα με νέες λειτουργίες και δυνατότητες, όπως δυνατότητα μαζικού σημασιολογικού εμπλουτισμού και ομογενοποίησης του περιεχομένου που συσσωρεύεται (με την βοήθεια του εργαλείου σημασιολογικού εμπλουτισμού του Συστήματος Σημασιολογικών Λεξιλογίων Semantics.gr), υποστήριξη δίγλωσσου περιβάλλοντος και υποστήριξη πρόσβασης από κινητές συσκευές.

Το SearchCulture.gr διαθέτει ένα φιλικό και λειτουργικό γραφικό διαχειριστικό περιβάλλον μέσω του οποίου ο εξουσιοδοτημένος χρήστης του ΕΚΤ διαχειρίζεται τους φορείς-πηγές, πυροδοτεί εύκολα διαδικασίες εισαγωγής περιεχομένου (μεταδεδομένων και εικόνων προεπισκόπησης), παρακολουθεί την εξέλιξη των διαδικασιών, διαχειρίζεται το συσσωρευμένο περιεχόμενο και πυροδοτεί διαδικασίες σημασιολογικού εμπλουτισμού.

Η υλοποίηση του συστήματος συσσωρευτή SearchCulture.gr τεκμηριώνεται αναλυτικά στο Κεφάλαιο 3.



Σύστημα Ελέγχου Συμμόρφωσης με Προδιαγραφές

Το σύστημα Πιστοποίησης Προδιαγραφών Διαλειτουργικότητας και Περιεχομένου

Validator είναι το πληροφοριακό σύστημα που διατίθεται από το ΕΚΤ και πιστοποιεί τους Ψηφιακούς Πόρους και τις υπηρεσίες του Δικαιούχου για το βαθμό κατά τον οποίο συμμορφώνονται με τις προδιαγραφές που περιγράφονται στο Πλαίσιο Συνεργασίας [1] και αποτελούν εξειδίκευση του συνοδευτικού υλικού των Προσκήσεων 31 και 31.2, ιδίως του Πλαισίου Οδηγιών Προδιαγραφών, Αρχών Διαλειτουργικότητας και Ανοιχτών Δεδομένων.

Το σύστημα αποτελείται από δύο υποσυστήματα, το back-end, το οποίο είναι το σύστημα που πραγματοποιεί αυτούς καθαυτούς τους ελέγχους και δημιουργεί αναλυτικές τεχνικές αναφορές, και το front-end (<http://validator.ekt.gr>), μέσα από το οποίο ο εξουσιοδοτημένος χρήστης πυροδοτεί εύκολα τις διαδικασίες ελέγχου για τα αποθετήρια των εγγεγραμμένων στο σύστημα φορέων, παρακολουθεί τις διαδικασίες και λαμβάνει αναλυτικές αναφορές υψηλού επιπέδου.

Είδη κατά την διάρκεια τις πρώτης φάσης ανάλυσης απαιτήσεων και σχεδιασμού των παραπάνω δύο συστημάτων έγιναν οι εξής παρατηρήσεις:

- I. Τα συστήματα SearchCulture.gr, Ασφαλούς Διαφύλαξης Ψηφιακού Περιεχομένου και το front-end του Validator, προκειμένου να είναι φιλικά προς τον χρήστη και να αυτοματοποιήσουν τις διαδικασίες των υπηρεσιών τους, θα πρέπει να έχουν κοινή πρόσβαση σε ένα πληροφοριακό σύστημα διαχείρισης δεδομένων δικαιούχων/φορέων.
- II. Τα συστήματα SearchCulture.gr, Ασφαλούς Διαφύλαξης Ψηφιακού Περιεχομένου και το back-end του συστήματος Validator θα πρέπει να έχουν πρόσβαση στο ψηφιακό περιεχόμενο των

αποθετηρίων των φορέων μέσω του πρωτοκόλλου OAI-PMH [5]. Κατά την διάρκεια ανάπτυξης και λειτουργίας των συστημάτων αυτών, ο αριθμός των διαδικασιών συγκομιδής περιεχομένου από τα (υπό ανάπτυξη) αποθετήρια των φορέων θα πρέπει να ελαχιστοποιηθεί ώστε τα συστήματα να μπορούν επαναχρησιμοποιούν το περιεχόμενο που συλλέχθηκε και οι υποδομές των αποθετηρίων των φορέων να επιβαρύνονται όσο το δυνατόν λιγότερο.

Το πρώτο συμπέρασμα οδήγησε στο να συμπεριληφθούν τα συστήματα Συσσωρευτή SearchCulture.gr, Ασφαλούς Διαφύλαξης Ψηφιακού Περιεχομένου και το front-end του Validator κάτω από μια κοινή πλατφόρμα, ένα ευρύτερο υπερ-πληροφοριακό σύστημα που να συμπεριλαμβάνει και το ενιαίο σύστημα διαχείρισης δεδομένων φορέων/συλλογών. Η πλατφόρμα ονομάστηκε **Πλατφόρμα Υποστήριξης Υπηρεσιών Συσσώρευσης, Ελέγχου, Ενιαίας Αναζήτησης και Ασφαλούς Διαφύλαξης ψηφιακού περιεχομένου.**

Το δεύτερο συμπέρασμα οδήγησε στην ανάγκη δημιουργίας ενός ξεχωριστού ενιαίου και αυτόνομου συστήματος Συγκομιδής Ψηφιακού Περιεχομένου το οποίο θα αποτελεί έναν ενδιάμεσο κόμβο μεταξύ των συστημάτων SearchCulture.gr, Ασφαλούς Διαφύλαξης Ψηφιακού Περιεχομένου και Validator με τα συστήματα σταθερής απόθεσης των φορέων. Ειδικά για το σύστημα Ασφαλούς Διαφύλαξης Περιεχομένου, διαπιστώθηκε ότι το σύστημα Συγκομιδής Ψηφιακού Περιεχομένου θα μπορούσε να αποτελεί αυτούσιο το back-end σύστημά του αφού υλοποιεί πλήρως τις βασικές low-level λειτουργίες του, δηλαδή, την συγκομιδή μεταδεδομένων και ψηφιακών αρχείων και την αποθήκευσή τους για μετέπειτα χρήση.



Σύστημα Συγκομιδής Ψηφιακού Περιεχομένου

Το σύστημα Συγκομιδής Ψηφιακού Περιεχομένου **Harvester** συλλέγει μεταδεδομένα και ψηφιακά αρχεία, τα αποθηκεύει και τα διαθέτει κατόπιν αίτησης από χρήστη ή από εξωτερικό σύστημα μέσω της γραφική web διεπαφής χρήστη ή των web υπηρεσιών του (RESTful API), αντίστοιχα. Το περιεχόμενο που συλλέγεται αποθηκεύεται αυτούσιο και μπορεί να επαναχρησιμοποιηθεί πολλές φορές από τρίτα συστήματα όπως τα συστήματα SearchCulture.gr, Ασφαλούς Διαφύλαξης Ψηφιακού Περιεχομένου και το back-end του συστήματος Validator. Το σύστημα Harvester υποστηρίζει συλλογή μεταδεδομένων και ψηφιακών αρχείων ενώ ειδικά για τα ψηφιακά αρχεία διαθέτει μία υπηρεσία μετατροπής τους σε εικόνες προεπισκόπησης, λειτουργία που αξιοποιείται από τον συσσωρευτή SearchCulture.gr.

Το Σύστημα Συγκομιδής Ψηφιακού Περιεχομένου, ως υποστηρικτικό σύστημα στο SearchCulture.gr, τεκμηριώνεται αναλυτικά στο Κεφάλαιο 2.



Υπηρεσία ασφαλούς διαφύλαξης

Το σύστημα Ασφαλούς Διαφύλαξης Ψηφιακού Περιεχομένου **Preservator** στόχο έχει την επαύξηση της διαθεσιμότητας και της ακεραιότητας των Ψηφιακών Πολιτιστικών Πόρων των αποθετηρίων των φορέων, σύμφωνα με τις αρχές που θέτουν διεθνή πρότυπα για την ασφαλή παροχή και λειτουργία των Πληροφοριακών Συστημάτων. Αφορά απομακρυσμένα, από το κύριο σημείο διάθεσης, ασφαλή αποθήκευση των Ψηφιακών Πολιτιστικών Πόρων και μηχανισμούς διασφάλισης της διαθεσιμότητας και ακεραιότητας τους σε βάθος χρόνου, σεβόμενη τις απαιτήσεις ελέγχου πρόσβασης.

Το σύστημα Preservator αποτελείται από δύο υποσυστήματα, το back-end, το οποίο είναι ουσιαστικά μία ξεχωριστή εγκατάσταση του Συστήματος Συγκομιδής Harvester και το front-end μέσα από το οποίο ο εξουσιοδοτημένος χρήστης πυροδοτεί εύκολα τις διαδικασίες συλλογής περιεχομένου των αποθετηρίων των εγγεγραμμένων στο σύστημα φορέων και τις διαδικασίες επανάκτησης του περιεχομένου κατόπιν αίτησης του φορέα.

Ένα χρόνο μετά την πρώτη παραγωγική λειτουργία του SearchCulture.gr και καθώς εμπλουτίζονταν συνεχώς με περιεχόμενο από τα αποθετήρια των φορέων που συμμετείχαν στις προσκλήσεις 31 και 31.2, διαπιστώθηκε ότι ήταν επιτακτική η ανάγκη εμπλουτισμού και ομογενοποίησης των τύπων τεκμηρίων του συνόλου του περιεχομένου ως προς ένα συμαπγές, ενιαίο και δίγλωσσο λεξιλόγιο τύπων ψηφιακών τεκμηρίων. Η ανάγκη προέκυψε από το γεγονός ότι οι περισσότεροι φορείς χρησιμοποίησαν δικά τους δεσμευμένα λεξιλόγια για τύπους τεκμηρίων (αντί για καθιερωμένα λεξιλόγια) τα οποία ήταν ετερόκλητα, χαρακτηρίζονταν από διπλότυπα και αμφισημίες, τα περισσότερα δεν ήταν πολυγλωσσικά, είχαν φτωχή τεκμηρίωση και δημιουργήθηκαν με διαφορετικές μεθοδολογίες (για παράδειγμα, άλλα περιείχαν όρους σε πληθυντικό αριθμό ενώ άλλα σε ενικό, κάποια ήταν υπερβολικά λεπτομερή ενώ κάποια άλλα πολύ γενικά, κλπ).

Όλα τα παραπάνω ουσιαστικά μείωναν την ποιότητα και την αισθητική παρουσίαση του περιεχομένου του SearchCulture.gr και συνεπώς και των υπηρεσιών αναζήτησης και εξερεύνησης που αυτό παρέχει, ενώ δεν ευνοούσαν την επικείμενη μετατροπή της διαδικτυακής πύλης σε δίγλωσση. Για το λόγο αυτό αποφασίστηκε το ΕΚΤ να αναλάβει τον εμπλουτισμό/ομογενοποίηση του περιεχομένου που συσσωρεύεται με την βοήθεια του Συστήματος Σημασιολογικών Λεξιλογίων Semantics.gr που αναπτύχθηκε πρόσφατα από το ΕΚΤ με τρόπο μόνιμο, ελεγχόμενο, αποτελεσματικό και όσο το δυνατόν πιο αυτοματοποιημένο. Ειδικά για τον σκοπό αυτό, το Semantics.gr εμπλουτίστηκε με ένα εργαλείο αντιστοίχισης/ομογενοποίησης πεδίων αποθετηρίων σε όρους έγκριτων λεξιλογίων.

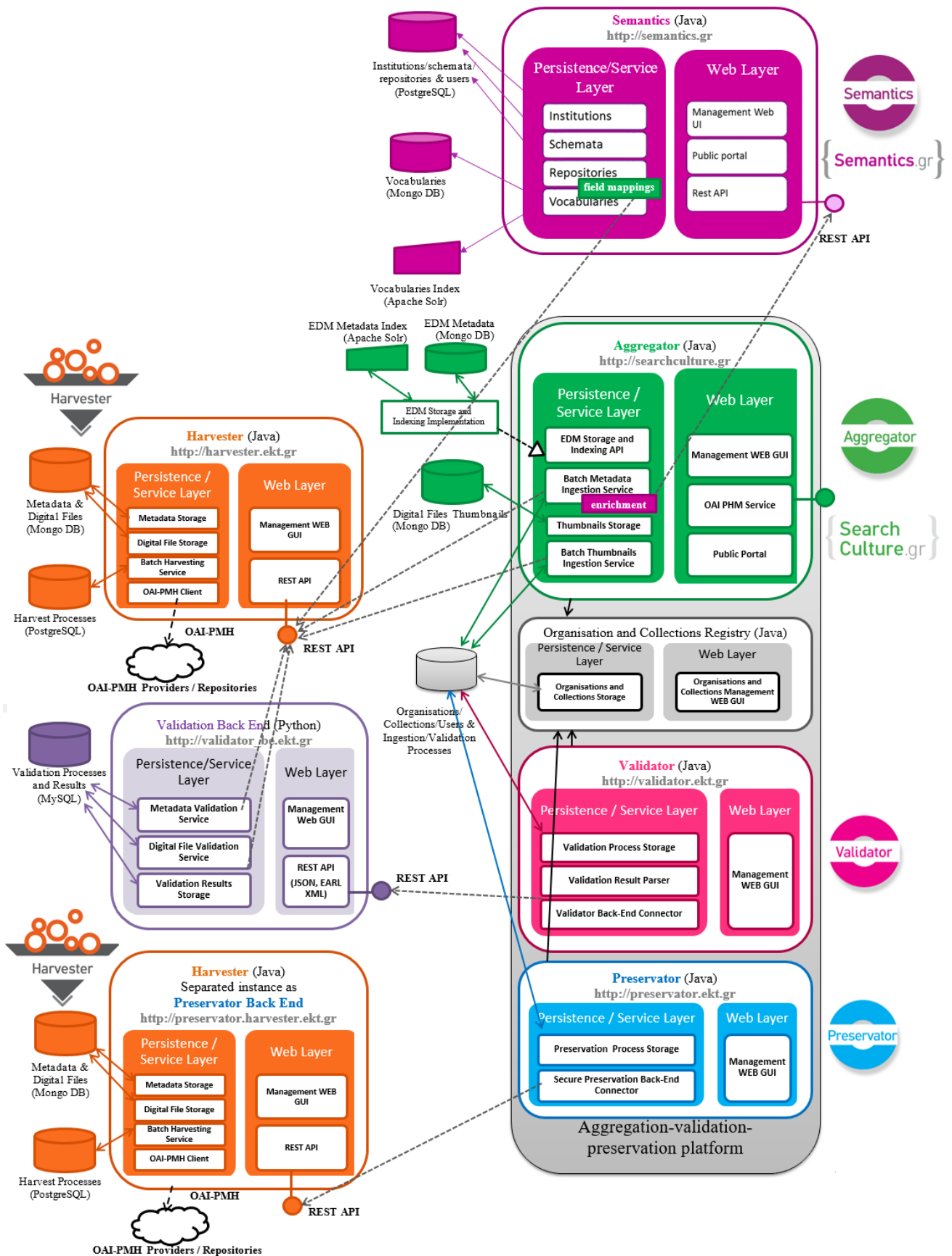


Σύστημα Λεξιλογίων και Εργαλείο Σημασιολογικού εμπλουτισμού

Το **Semantics.gr** είναι ένα Σύστημα Λεξιλογίων και Σημασιολογικού Εμπλουτισμού που ανέπτυξε το ΕΚΤ όπου το ΕΚΤ αλλά και συνεργαζόμενοι εξουσιοδοτημένοι φορείς καταχωρούν, καθιερώνουν και διασυνδέουν λεξιλόγια σημασιολογικών όρων. Παρέχει μία δημόσια διαδικτυακή πύλη (<http://semantics.gr>) προσφέροντας ενιαία αναζήτηση σημασιολογικών όρων σε πολλαπλές πηγές και διαθέτει το περιεχόμενο ως Ανοικτά Διασυνδεδεμένα Δεδομένα (Open Linked Data). Το Semantics.gr διαθέτει ένα φιλικό στην χρήση μηχανισμό παραμετροποίησης για την μοντελοποίηση των σχημάτων των λεξιλογίων που θα δημιουργούνται από τους φορείς.

Παράλληλα, το Semantics.gr διαθέτει ένα πρωτότυπο και ιδιαίτερα φιλικό εργαλείο για μαζικό ημιαυτόματο σημασιολογικό εμπλουτισμό και ομογενοποίηση περιεχομένου που μπορεί να προσφέρεται ως υπηρεσία τόσο σε αποθετήρια πρωτογενούς περιεχομένου όσο και σε συσσωρευτές. Το ΕΚΤ ήδη χρησιμοποιεί το εργαλείο αυτό προκειμένου να εμπλουτίσει το περιεχόμενο που συσσωρεύεται στο SearchCulture.gr με συνδέσμους σε καθιερωμένα λεξιλόγια και θησαυρούς όρων αλλά και για να ομογενοποιήσει το περιεχόμενο αυτό με στόχο την βελτιστοποίηση των υπηρεσιών αναζήτησης και πλοήγησης της διαδικτυακής πύλης του SearchCulture.gr.

Στο Σχήμα 2 παρουσιάζεται ένα διάγραμμα αρχιτεκτονικής που περιλαμβάνει όλα τα συστήματα που παρουσιάστηκαν στην ενότητα αυτή, την βασική τους δομή και τον τρόπο που επικοινωνούν μεταξύ τους. Τα συστήματα Aggregator/SearchCulture.gr, Validator-front-end και Preservator-front-end έχουν ισχυρή εξάρτηση στο σύστημα Harvester με το οποίο επικοινωνούν μέσω των REST API υπηρεσιών του (υπενθυμίζεται ότι το σύστημα Preservator-front-end χρησιμοποιεί ξεχωριστή εγκατάσταση του Harvester). Το σύστημα Semantics.gr και συγκεκριμένα το εργαλείο σημασιολογικού εμπλουτισμού χρειάζεται πρόσβαση στα μεταδεδομένα των αποθετηρίων που εμπλουτίζονται προκειμένου να αντλήσει τις ξεχωριστές τιμές πεδίων μεταδεδομένων. Στην τρέχουσα έκδοση τα μεταδεδομένα διατίθενται από το σύστημα Harvester. Το σύστημα Validator front-end έχει ισχυρή εξάρτηση στο σύστημα Validator-back-end. Και σε αυτήν την περίπτωση η επικοινωνία πραγματοποιείται μέσω REST API. Σημειώνεται ότι τα συστήματα Aggregator/SearchCulture.gr και Validator δεν επικοινωνούν μεταξύ τους. Τα δύο συστήματα μοιράζονται όμως την ίδια σχεσιακή βάση δεδομένων καθώς και εκείνα τα κομμάτια λογισμικού που σχετίζονται με την διαχείριση φορέων και αποθετηρίων/συλλογών.



Σχήμα 2. Συνολική αρχιτεκτονική των συστημάτων του ΕΚΤ για την συλλογή περιεχομένου (Harvester), τον έλεγχο συμμόρφωσης με προδιαγραφές (Validator Back-end / Front-end), την συσσώρευση/ενιαία αναζήτηση (SearchCulture.gr), την ασφαλή διαφύλαξη (Preservator Back-end / Front-end) και τον σημασιολογικό εμπλουτισμό (Semantics.gr).

1.3 Πλεονεκτήματα

Το SearchCulture.gr παρέχει υψηλού επιπέδου υπηρεσίες αναζήτησης, πλοήγησης και παρουσίασης πολιτιστικού περιεχομένου προβάλλοντας και αναδεικνύοντας με τον καλύτερο τρόπο το ψηφιακό περιεχόμενο έγκριτων φορέων πολιτισμού. Βασικοί στόχοι του SearchCulture.gr είναι τόσο η αύξηση της ευρεσιμότητας όσο και η αύξηση του όγκου του ετερόκλητου περιεχομένου που συσσωρεύει. Οι δύο αυτοί στόχοι είναι κοινοί για όλους τους συσσωρευτές περιεχομένου και η ικανοποίηση και των δύο αποτελεί πρόκληση καθώς οι λειτουργικές απαιτήσεις που αυτοί συνεπάγονται δρουν ανταγωνιστικά μεταξύ τους, όπως φαίνεται παραστατικά στο Σχήμα 2.

Το SearchCulture.gr καταφέρνει να ικανοποιήσει και τους δύο αυτούς στόχους σε πολύ ικανοποιητικό βαθμό, χάρις στις υπηρεσίες προστιθέμενης αξίας που παρέχει το ΕΚΤ (δημοσιοποίηση προδιαγραφών διαλειτουργικότητας, έλεγχος/πιστοποίηση, εμπλουτισμός περιεχομένου) αλλά και στα ποιοτικά συστήματα που ανέπτυξε τα οποία καθιστούν εφικτή την υλοποίηση των υπηρεσιών αυτών καθώς και την διαχείριση όλου αυτού του όγκου πληροφορίας (αυτοματοποίηση διαδικασιών ελέγχου, εισαγωγής και εμπλουτισμού, έξυπνος τρόπος μαζικής σημασιολογικής ομογενοποίησης, εξελιγμένο διαχειριστικό περιβάλλον, προηγμένες δυνατότητες αναζήτησης και πλοήγησης περιεχομένου).



Σχήμα 3. Οι δύο βασικοί στόχοι ενός συσσωρευτή και ενδεικτικοί τρόποι ικανοποίησής τους.

Το ΕΚΤ διεξάγοντας αυτοματοποιημένους ελέγχους ποιότητας στο περιεχόμενο που συσσωρεύεται με το σύστημα Validator εξασφαλίζει, σε συνεργασία με τους φορείς, την υψηλή ποιότητα του περιεχομένου που εισάγεται στο SearchCulture.gr. Η ποιότητα της τεκμηρίωσης βελτιώνει την

παρουσίαση, την ευρεσιμότητα και διευκολύνει την αναζήτηση του περιεχομένου και κατ' επέκταση αυξάνει την συνολική ποιότητα των υπηρεσιών που προσφέρει η διαδικτυακή πύλη SearchCulture.gr στους πολίτες και στους φορείς.

Παράλληλα, η ποιότητα υπηρεσιών αυξάνεται περαιτέρω χάρις στον σημασιολογικό εμπλουτισμό και την ομογενοποίηση του περιεχομένου που πραγματοποιεί το EKT με το εργαλείο Semantics.gr. Το σύνολο σχεδόν των τεκμηρίων εμπλουτίστηκε με ένα νέο πεδίο «EKT Τύπος» που περιλαμβάνει αναφορές σε όρους ενός πρότυπου δίγλωσσου ιεραρχικού λεξιλογίου τύπων που δημιουργήθηκε από το EKT. Έτσι το περιεχόμενο κατηγοριοποιήθηκε στους 393 όρους του λεξιλογίου με άμεσο θετικό αντίκτυπο στην ευρεσιμότητα, στην παρουσίαση και στην πολυγλωσσικότητα του περιεχομένου. Συγκεκριμένα:

- Αυξάνεται η ευρεσιμότητα του περιεχομένου αφού ο χρήστης μπορεί να ανακαλύψει περιεχόμενο εξερευνώντας μέσα από ένα ομοιόμορφο και συμπαγές σύνολο τύπων ακόμα και αν δεν γνωρίζει ακριβώς τι ψάχνει.
- Δίνεται η δυνατότητα σε αγγλόφωνο κοινό να ανακαλύψει τεκμήρια βάσει του τύπου τους ακόμα και αν τα μεταδεδομένα τους πρωτοτύπως δεν περιείχαν καθόλου πολυγλωσσική πληροφορία.
- Συλλογές που είναι σημασιολογικά εμπλουτισμένες ομογενοποιούνται με στόχο την ποιοτικότερη πολυγλωσσική ενιαία αναζήτηση και εξερεύνηση.
- Συλλογές που δεν περιλαμβάνουν καθόλου σημασιολογικές αναφορές μπορούν πλέον να εμπλουτιστούν και να αναδειχθούν μέσα από το SearchCulture ως Διασυνδεδεμένα Δεδομένα.

Η διαδικτυακή πύλη SearchCulture.gr προσφέρει προηγμένες δυνατότητες αναζήτησης, ορισμένες πρωτότυπες για συστήματα συσσώρευσης, οι οποίες προσεγγίζουν αυτές που προσφέρονται συνήθως από συστήματα απόθεσης πρωτογενούς περιεχομένου. Οι δυνατότητες περιλαμβάνουν:

- Σύνθετη αναζήτηση
- Δυνατότητα περιορισμού αναζήτησης σε συγκεκριμένα πεδία, όπως EKT τύπος τεκμηρίου, πρωτότυπος τύπος, Europeana τύπος, δημιουργός, θεματική κατηγορία, χρονική και γεωγραφική κάλυψη
- Προτάσεις αυτόματης συμπλήρωσης πεδίων εισαγωγής όρων αναζήτησης (autocomplete)
- Φίλτρα αναζήτησης στα αποτελέσματα βάσει μιας μεγάλης γκάμας βασικών πεδίων που περιλαμβάνει τον ομογενοποιημένο EKT τύπο τεκμηρίου, τον πρωτότυπο τύπο, την θεματική κατηγορία, την άδειας χρήσης, τον Europeana τύπο και τον πάροχο. Έτσι ο χρήστης μπορεί να εντοπίσει πιο εύκολα τα τεκμήρια που αναζητά φιλτράροντας συνδυαστικά τα αποτελέσματα και μάλιστα χρησιμοποιώντας ως φίλτρα, όχι μόνο τεχνικά πεδία ή πεδία που σχετίζονται με την διάθεση του τεκμηρίου (όπως τα δικαιώματα χρήσης και ο πάροχος), αλλά και πεδία που ορίζουν και περιγράφουν σημασιολογικά το ίδιο το τεκμήριο και που βοηθούν καταλυτικά στην ανακάλυψή του (όπως ο ομογενοποιημένος τύπος EKT, ο πρωτότυπος τύπος και η θεματική κατηγορία).

Εκτός από την μηχανή αναζήτησης, η διαδικτυακή πύλη SearchCulture.gr προσφέρει προηγμένες δυνατότητες πλοήγησης μέσα από την αρχική σελίδα οι οποίες περιλαμβάνουν

- Δίγλωσσο διαδραστικό νέφος δημοφιλών ομογενοποιημένων EKT τύπων τεκμηρίων
- Διαδραστική πίτα δημοφιλών EKT τύπων τεκμηρίων










- Σελίδα παρουσίασης των φορέων και των συλλογών/αποθετηρίων τους
- Σελίδα ιεραρχικής περιήγησης στο περιεχόμενο μέσα από τους ΕΚΤ τύπους τεκμηρίων

Παράλληλα, το SearchCulture.gr αποτελεί μία μοντέρνα και εύχρηστη διαδικτυακή πύλη με φιλική και ελκυστική διεπαφή χρήστη, καθώς:

- Υποστηρίζει δίγλωσσο περιβάλλον σε Ελληνικά και Αγγλικά σε στατικό και δυναμικό περιεχόμενο
- Όλα σχεδόν τα τεκμήρια που φιλοξενούνται συνοδεύονται από εικόνες προεπισκόπησης κάνοντας την αναζήτηση πιο ελκυστική αλλά και πιο λειτουργική.
- Οι σελίδες ανταποκρίνονται και προσαρμόζονται ανάλογα με την ανάλυση της οθόνης του χρήστη. Έτσι το SearchCulture.gr είναι προσβάσιμο το ίδιο εύχρηστα και λειτουργικά και μέσα από φορητές συσκευές (κινητά τηλέφωνα, tablets)

Τέλος, πέραν των ποιοτικών υπηρεσιών που προσφέρει η διαδικτυακή πύλη του SearchCulture.gr στους πολίτες και στους φορείς περιεχομένου, αξίζει να τονιστεί ότι το σύστημα διαθέτει ένα πλήρες γραφικό διαχειριστικό περιβάλλον το οποίο επιτρέπει στο προσωπικό του ΕΚΤ να διαχειριστεί εύκολα, μαζικά και αποδοτικά πολύ μεγάλο όγκο δεδομένων προερχόμενα από μεγάλο αριθμό αποθετηρίων/συλλογών. Συγκεκριμένα, το διαχειριστικό περιβάλλον επιτρέπει στον εξουσιοδοτημένο χρήστη να εκτελεί εύκολα και όσο το δυνατόν πιο αυτοματοποιημένα όλες τις απαραίτητες διαδικασίες για την λειτουργία του συσσωρευτή, όπως, να διαχειριστεί πληροφορίες για τις συλλογές/αποθετήρια, να πυροδοτήσει με απόλυτα ελεγχόμενο τρόπο νέες διαδικασίες συλλογής και εισαγωγής μεταδεδομένων ή αρχείων προεπισκόπησης, να επαναχρησιμοποιήσει περιεχόμενο που έχει ήδη συλλεχθεί, να εμπλουτίσει σημασιολογικά συλλογές, να διαχειρίζεται μαζικά το περιεχόμενο ανά συλλογή και, τέλος, να παραμετροποιήσει δυναμικά την διαδικτυακή πύλη.

Στο Σχήμα 4 συνοψίζονται τα βασικότερα λειτουργικά χαρακτηριστικά του SearchCulture.gr.

	Μοντέρνα και ιδιαίτερα αποδοτική μηχανή αναζήτησης
	Υποστηρίζει ιεραρχική πλοήγηση σε τύπους περιεχομένου
	Διάθεση περιεχομένου ως Ανοικτά Διασυνδεδεμένα Δεδομένα (Open Linked Data)
	Υποστηρίζει πλοήγηση στους φορείς και τις συλλογές τους
	Σαφείς άδειες χρήσης ψηφιακού περιεχομένου
	Δίγλωσση πύλη σε στατικό και δυναμικό περιεχόμενο
	Πύλη σχεδιασμένη και για πρόσβαση από κινητά και tablets
	Συνεισφέρει με περιεχόμενο στην Europeana
	Φιλικό διαχειριστικό περιβάλλον για εύκολη εισαγωγή περιεχομένου

Σχήμα 4. Τα κυριότερα λειτουργικά χαρακτηριστικά του SearchCulture.gr

2 Harvester: Σύστημα Συγκομιδής Ψηφιακού Περιεχομένου



Το κεφάλαιο αυτό περιγράφει αναλυτικά το Σύστημα Συγκομιδής Ψηφιακού Περιεχομένου Harvester που αναπτύχθηκε αρχικά στα πλαίσια των προσκλήσεων 31 και 31.2 και το οποίο αποτελεί υποστηρικτικό σύστημα για το Σύστημα Συσσωρευτή SearchCulture.gr, για τον Συσσωρευτή CRIS του ΕΚΤ [4], για το Σύστημα Ελέγχου Συμμόρφωσης με Προδιαγραφές Διαλειτουργικότητας Validator και για το Σύστημα Ασφαλούς Διαφύλαξης Περιεχομένου Preservator. Το σύστημα είναι αυτόνομο και ανεξάρτητο πεδίου (domain-agnostic) και λειτουργεί σε τρεις ξεχωριστές παραγωγικές εγκαταστάσεις: μία που εξυπηρετεί τα συστήματα SearchCulture.gr και Validator, μία που αποτελεί το back-end του Preservator, και μια τρίτη, η οποία εξυπηρετεί τον Συσσωρευτή CRIS. Αρχικά παρουσιάζονται οι λειτουργικές απαιτήσεις που τέθηκαν για το σύστημα και που ικανοποιήθηκαν στο σύνολό τους, ακολουθεί μία πλήρης τεκμηρίωση χρήσης της γραφικής διεπαφής χρήστη και των RESTFul υπηρεσιών του και, τέλος, γίνεται μια τεχνική περιγραφή της αρχιτεκτονικής του συστήματος.

2.1 Λειτουργικά Χαρακτηριστικά

Ο βασικός σκοπός του συστήματος Harvester είναι να συλλέγει και να αποθηκεύει μεταδεδομένα και ψηφιακά αρχεία αποθετηρίων για λογαριασμό κάποιου χρήστη ή ενός τρίτου συστήματος. Επομένως το σύστημα παρέχει τόσο γραφική διεπαφή χρήστη όσο και διεπαφή υπηρεσιών ιστού (RESTful API), για την εκκίνηση διαδικασιών συγκομιδής, την παρακολούθησή τους και την διάθεση των μεταδεδομένων/ψηφιακών αρχείων που συλλέχθηκαν. Για την συλλογή απαραίτητη κρίθηκε η υποστήριξη του πρωτοκόλλου OAI-PMH [5]. Ο Πίνακας 1 παραθέτει αναλυτικά τις λειτουργικές απαιτήσεις τις οποίες το Σύστημα Συγκομιδής Ψηφιακού Περιεχομένου ικανοποιεί στο σύνολό τους.

Πίνακας 1. Οι Λειτουργικές Απαιτήσεις του Συστήματος Συγκομιδής.

Απαίτηση	Περιγραφή
1 Συγκομιδή μεταδεδομένων μέσω OAI-PMH	Δοθέντων παραμέτρων του OAI-PMH πρωτοκόλλου (όπως set, metadataPrefix, from και until) και φυσικά του URL του OAI-PMH εξυπηρετητή, το σύστημα συγκομιδής εκτελεί τις αντίστοιχες OAI-PMH κλίσεις και μεταφορτώνει το σύνολο των μεταδεδομένων.
2 Αποθήκευση των μεταδεδομένων που συλλέγονται σε βάση δεδομένων.	Κάθε εγγραφή μεταδεδομένων αποθηκεύεται σε μία (noSQL) βάση δεδομένων σε δομή η οποία περιλαμβάνει πεδίο για την αποθήκευση της πρωτότυπης XML εγγραφής (XML στοιχείο RECORD του OAI-PMH XML) ως απλό κείμενο. Όλες οι εγγραφές μεταδεδομένων που συλλέχθηκαν από μια διαδικασία ομαδοποιούνται με συγκεκριμένο πεδίο που ονομάζεται location token (προσδιοριστής θέσης). Τα εξωτερικά συστήματα χρησιμοποιούν τον προσδιοριστή θέσης για να ανακτήσουν τα μεταδεδομένα που συλλέχθηκαν στα πλαίσια μιας διαδικασίας.
3 Ροή συγκομιδής μεταδεδομένων.	Υποστηρίζεται ροή συγκομιδής μεταδεδομένων που περιλαμβάνει την ανάκτηση μεταδεδομένων (είτε από OAI-PMH κλίσεις είτε από φάκελο) και την αποθήκευσή τους.
4 Αποθήκευση ιστορικού διαδικασιών συγκομιδής μεταδεδομένων.	Στοιχεία για διαδικασίες συγκομιδής που πραγματοποιήθηκαν φυλάσσονται 1) για λόγους ιστορικότητας 2) για λόγους καταγραφής σφαλμάτων (απαραίτητο τόσο για αποσφαλμάτωση όσο και για δημιουργία αναφοράς προς τον φορέα) 3) για να είναι εφικτό το incremental harvesting που υποστηρίζεται από το πρωτόκολλο OAI-PMH και 4) για να μπορεί κάποιος να αναζητήσει μια παλαιότερη ολοκληρωμένη διαδικασία και εν συνεχεία να αποκτήσει πρόσβαση στα μεταδεδομένα που συλλέχθηκαν.
5 Γραφική διεπαφή χρήστη για την διαχείριση διαδικασιών συγκομιδής μεταδεδομένων.	Το Σύστημα Συγκομιδής διαθέτει εφαρμογή ιστού η οποία επιτρέπει στον εξουσιοδοτημένο χρήστη να εκκινήσει μια διαδικασία συγκομιδής και να παρακολουθήσει την εξέλιξή της. Ο χρήστης μπορεί να αναζητά διαδικασίες που έχουν πραγματοποιηθεί στο παρελθόν και να πλοηγηθεί στα μεταδεδομένα που συλλέχθηκαν.
6 Συγκομιδή ψηφιακών αρχείων από URLs που περιλαμβάνονται σε πεδία μεταδεδομένων που συλλέχθηκαν από το σύστημα Συγκομιδής	Δοθέντος του προσδιοριστή μιας ολοκληρωμένης διαδικασίας συγκομιδής μεταδεδομένων και μιας XPath έκφρασης που προσδιορίζει τα πεδία που περιλαμβάνουν τα URLs των ψηφιακών αρχείων, το σύστημα συγκομιδής μεταφορτώνει και αποθηκεύει το σύνολο των ψηφιακών αρχείων.

7	Συγκομιδή ψηφιακών αρχείων από zip αρχείο που ανεβάζει στο σύστημα εξουσιοδοτημένος χρήστης ή από φάκελο στον εξυπηρετητή	Ο εξουσιοδοτημένος χρήστης μπορεί να ανεβάζει ένα zip αρχείο το οποίο περιλαμβάνει ένα σύνολο ψηφιακών αρχείων και ένα αρχείο κειμένου (csv) με τις αντιστοιχίσεις των ονομάτων των αρχείων με τα ΟΑΙ αναγνωριστικά εγγραφών μεταδεδομένων. Εναλλακτικά, προσωπικό από το Τμήμα Συστημάτων μπορεί να τοποθετήσει τα αρχεία μαζί με το αρχείο αντιστοιχίσεων σε έναν φάκελο στον εξυπηρετητή και το σύστημα συγκομιδής να τα συλλέξει από εκεί.
8	Αποθήκευση των Ψηφιακών Αρχείων που συλλέγονται σε File Storage System.	Κάθε ψηφιακό αρχείο αποθηκεύεται σε ένα File Storage System. Όλα τα ψηφιακά αρχεία που συλλέχτηκαν από μια διαδικασία ομαδοποιούνται με συγκεκριμένο πεδίο που ονομάζεται digital file location token (προσδιοριστής θέσης ψηφιακών αρχείων) η τιμή του οποίου είναι μοναδική για την συγκεκριμένη διαδικασία. Τα εξωτερικά συστήματα χρησιμοποιούν τον προσδιοριστή θέσης ψηφιακών αρχείων για να ανακτήσουν τα ψηφιακά αρχεία που συλλέχθηκαν στα πλαίσια μιας διαδικασίας.
9	Ροή συγκομιδής ψηφιακών αρχείων.	Υποστηρίζεται ροή εισαγωγής που περιλαμβάνει την συγκομιδής ψηφιακών αρχείων και την αποθήκευσή τους.
10	Αποθήκευση ιστορικού διαδικασιών συγκομιδής ψηφιακών αρχείων.	Στοιχεία για διαδικασίες συγκομιδής ψηφιακών αρχείων που πραγματοποιήθηκαν φυλάσσονται 1) για λόγους ιστορικότητας 2) για λόγους καταγραφής σφαλμάτων (απαραίτητο τόσο για αποσφαλμάτωση όσο και για δημιουργία αναφοράς προς τον φορέα) 3) για να είναι εφικτό το incremental harvesting που υποστηρίζεται από το πρωτόκολλο ΟΑΙ-ΡΜΗ και 4) για να μπορεί κάποιος να αναζητήσει μια παλαιότερη ολοκληρωμένη διαδικασία και εν συνεχεία να αποκτήσει πρόσβαση στα αρχεία που συλλέχθηκαν.
11	Γραφική διεπαφή χρήστη για την διαχείριση διαδικασιών συγκομιδής ψηφιακών αρχείων.	Το Σύστημα Συγκομιδής διαθέτει εφαρμογή ιστού η οποία επιτρέπει στον εξουσιοδοτημένο χρήστη να εκκινήσει μια διαδικασία συγκομιδής ψηφιακών αρχείων και να παρακολουθήσει την εξέλιξή της. Ο χρήστης μπορεί να αναζητά διαδικασίες που έχουν πραγματοποιηθεί στο παρελθόν και να αποκτή πρόσβαση στα αρχεία που συλλέχθηκαν.
12	Διεπαφή υπηρεσίας ιστού που επιτρέπουν την πυροδότηση και παρακολούθηση διαδικασιών συγκομιδής μεταδεδομένων / ψηφιακών αρχείων καθώς και την πρόσβαση στις αντίστοιχες εγγραφές μεταδεδομένων / στα αντίστοιχα ψηφιακά αρχεία.	Το Σύστημα Συγκομιδής διαθέτει ένα RESTful API ώστε εξωτερικά συστήματα (όπως το SearchCulture.gr και ο Validator), να πυροδοτήσουν μια νέα διαδικασία συγκομιδής, να ρωτούν για την εξέλιξή της και, εφόσον αυτή έχει ολοκληρωθεί, να ανακτούν τα μεταδεδομένα ή τα ψηφιακά αρχεία που συλλέχθηκαν.
13	Δυνατότητα τερματισμού διαδικασίας συγκομιδής μεταδεδομένων/ψηφιακών αρχείων	Ο εξουσιοδοτημένος χρήστης μπορεί να διακόψει/τερματίσει μια εν εξέλιξη διαδικασία συγκομιδής μέσα από το διαχειριστικό περιβάλλον.
14	Δυνατότητα ορισμού ανώτατου ανεκτού αριθμού αποτυχιών κατά την εκκίνηση διαδικασίας	Για τις διαδικασίες συγκομιδής ψηφιακών αρχείων, ο εξουσιοδοτημένος χρήστης μπορεί να ορίζει ανώτατο επιτρεπτό όριο αποτυχιών. Όταν το όριο ξεπεραστεί η διαδικασία διακόπτεται.

συγκομιδής ψηφιακών αρχείων.

15	Αναλυτική καταγραφή αποτυχιών κατά την διάρκεια διαδικασίας συγκομιδής ψηφιακών αρχείων.	Κατά την διάρκεια συγκομιδής ψηφιακών αρχείων, τυχόν αποτυχίες καταγράφονται μαζί με όλες τις απαραίτητες πληροφορίες που χρειάζονται για αποσφαλμάτωση.
16	Δυνατότητα ορισμού συγκεκριμένων τύπων ψηφιακών αρχείων για συγκομιδή	Στην φόρμα εκκίνησης διαδικασίας συγκομιδής ψηφιακών αρχείων ο χρήστης μπορεί προαιρετικά να επιλέξει συγκεκριμένους τύπους αρχείων (π.χ. μόνο εικόνες jpeg). Όλα τα αρχεία άλλου τύπου θα αγνοούνται.
17	Δυνατότητα επανάληψης προσπάθειας για διαδικασία συγκομιδής ψηφιακών αρχείων που ολοκληρώθηκε με αποτυχίες ή που τερμάτισε ανεπιτυχώς (για τα εναπομείναντα αρχεία)	Ο χρήστης μπορεί να επαναλάβει μια διαδικασία συγκομιδής ψηφιακών αρχείων η οποία είτε ολοκληρώθηκε με κάποιες αποτυχίες, είτε απέτυχε (λόγω περισσότερων του επιτρεπτού ορίου σφαλμάτων). Η επανάληψη θα επιχειρήσει να συλλέξει και να αποθηκεύσει μόνο τα εναπομείναντα αρχεία.
18	Δυνατότητα συνένωσης (join) σε πραγματικό χρόνο των μεταδεδομένων που συλλέχτηκαν από το ίδιο αποθετήριο και διαφέρουν ως προς τον τύπο.	Έστω ότι για το ίδιο αποθετήριο πραγματοποιήθηκαν 2 διαφορετικές διαδικασίες συγκομιδής που αφορούν στις ίδιες εγγραφές μεταδεδομένων αλλά σε διαφορετικό τύπο (π.χ. ESE και DC-DS-XML). Εκτός από την δυνατότητα ανάκτησης των μεταδεδομένων χωριστά για κάθε διαδικασία, το RESTful API παρέχει την δυνατότητα ανάκτησης της συνένωσης (join) των δύο συλλογών μεταδεδομένων. Κάθε συνολική εγγραφή θα περιλαμβάνει και τις δύο επι μέρους εκδοχές της στους δύο διαφορετικούς τύπους μεταδεδομένων.

2.2 Εφαρμογή Ιστού

Στη ενότητα αυτή περιγράφεται η εφαρμογή ιστού του Συστήματος Συγκομιδής. Μέσω της εφαρμογής ο χρήστης μπορεί να πυροδοτήσει μία νέα διαδικασία συγκομιδής, να παρακολουθεί την πρόοδο της διαδικασίας και να αναζητήσει πληροφορίες για διαδικασίες που έχουν εκτελεστεί στο παρελθόν.

Η αρχική σελίδα της εφαρμογής φαίνεται στο Σχήμα 5. Μέσα από την σελίδα αυτή ο χρήστης μπορεί να αναζητήσει διαδικασίες συγκομιδής μεταδεδομένων βάσει του τύπου μεταδεδομένων (metadataPrefix), του προσδιοριστή της διαδικασίας (Harvest Token), του είδους της συγκομιδής (μέσω OAI-PMH ή από φάκελο αρχείων), την OAI-PMH διεύθυνση ή το μονοπάτι του φακέλου των αρχείων και, τέλος, του εύρους ημερομηνιών μέσα στο οποίο ξεκίνησε η διαδικασία. Με το πάτημα του κουμπιού «Αναζήτηση» εμφανίζεται μία λίστα με τις διαδικασίες που πληρούν τα κριτήρια αναζήτησης. Στην πρώτη στήλη του πίνακα εμφανίζεται ο Προσδιοριστής της κάθε διαδικασίας (Harvest Token), ο μοναδικός αναγνωριστικός αριθμός που την ταυτοποιεί.

Σχήμα 5. Η αρχική σελίδα της εφαρμογής ιστού του Συστήματος Συγκομιδής

Επιπλέον, μέσα από την ίδια σελίδα, ο χρήστης μπορεί να εκκινήσει μία νέα διαδικασία συγκομιδής πατώντας το κουμπί «Εκκίνηση Νέας Διαδικασίας Συγκομιδής». Με το πάτημα του κουμπιού αυτού εμφανίζεται μια φόρμα στην οποία ο χρήστης καλείται να συμπληρώσει τις παραμέτρους της συγκομιδής, όπως φαίνεται στο Σχήμα 6. Συγκεκριμένα, ο χρήστης καλείται να συμπληρώσει το URL της OAI-PMH υπηρεσίας, τον τύπο μεταδεδομένων (παραμέτρος `metadataPrefix` του OAI-PMH πρωτοκόλλου), το σύνολο μεταδεδομένων (παραμέτρος `set` του OAI-PMH πρωτοκόλλου) και το εύρος ημερομηνιών τελευταίας ενημέρωσης (παραμέτροι `from` και `until` του πρωτοκόλλου OAI-PMH).

Με το πάτημα του κουμπιού «Εκκίνηση» πραγματοποιείται έλεγχος της ορθότητας των παραμέτρων (π.χ. αν δόθηκαν τιμές σε υποχρεωτικά πεδία ή αν η URL διεύθυνση είναι έγκυρη). Σε περίπτωση σφαλμάτων ως προς την εγκυρότητα των παραμέτρων, ο χρήστης ενημερώνεται και καλείται να τα διορθώσει (Σχήμα 7). Διαφορετικά η διαδικασία ξεκινά, η φόρμα κλείνει και στην λίστα των διαδικασιών εμφανίζεται η νέα διαδικασία με κατάσταση «Σε εξέλιξη». Το πεδίο «Στατιστικά Συγκομιδής» για την συγκεκριμένη διαδικασία εμφανίζει μία μπάρα προόδου που ανανεώνεται σχεδόν σε πραγματικό χρόνο (Σχήμα 8).

Εκκίνηση νέας διαδικασίας συγκομιδής

Τύπος Μεταδεδ.

ΟΑΙ συνολο

ΟΑΙ ημερομηνίες Από Έως

Τύπος συγκομιδής

ΟΑΙ URL

Καθυστέρηση μεταξύ κλίσεων (σε δευτερόλεπτα)

January 2015

Su	Mo	Tu	We	Th	Fr	Sa
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

Σχήμα 6. Φόρμα συμπλήρωσης παραμέτρων για την εκκίνηση νέας διαδικασίας συγκομιδής

Εκκίνηση νέας διαδικασίας συγκομιδής

Τύπος Μεταδεδ.

*Το πεδίο Τύπος μεταδεδομένων είναι υποχρεωτικό.

ΟΑΙ συνολο

ΟΑΙ ημερομηνίες Από Έως

Τύπος συγκομιδής

ΟΑΙ URL

*Το πεδίο ΟΑΙ URL είναι υποχρεωτικό για συγκομιδή μέσω ΟΑΙ-PMH.

Καθυστέρηση μεταξύ κλίσεων (σε δευτερόλεπτα)

Σχήμα 7. Μηνύματα σφαλμάτων για μη έγκυρες παραμέτρους εκκίνησης νέας διαδικασίας συγκομιδής.

Επιλογή μιας στατιστικής συγκομιδής

« 1 2 3 4 5 » ↻

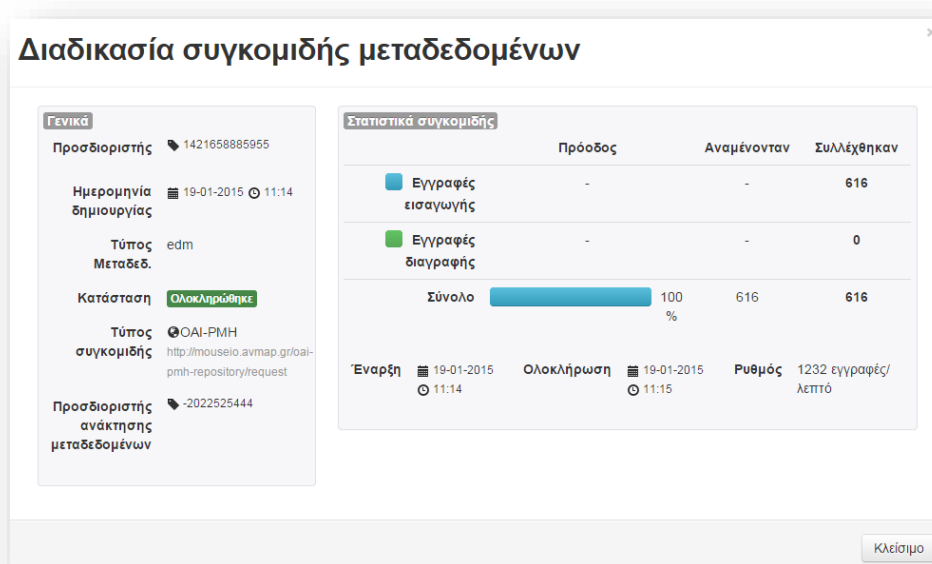
Προσδιοριστής	Ημερομηνία Δημιουργίας	Τύπος Μεταδεδομένων	ΟΑΙ Από-Εως	ΟΑΙ Συνολο	Τύπος Συγκομιδής	Κατάσταση	Στατιστικά Συγκομιδής	Ενέργειες
1392295356245	13-02-2014 02:42	edm	01-01-2014 - 15-01-2014	esegen	ΟΑΙ-PMH http://hispana.mcu.es/i18n/oai/oai.cmd	Σε Εξέλεση	500 εισαγωγής 0 διαγραφής	Ενέργειες

Σχήμα 8 Παρακολούθηση της προόδου μιας εν εξελίξει διαδικασίας σε πραγματικό χρόνο

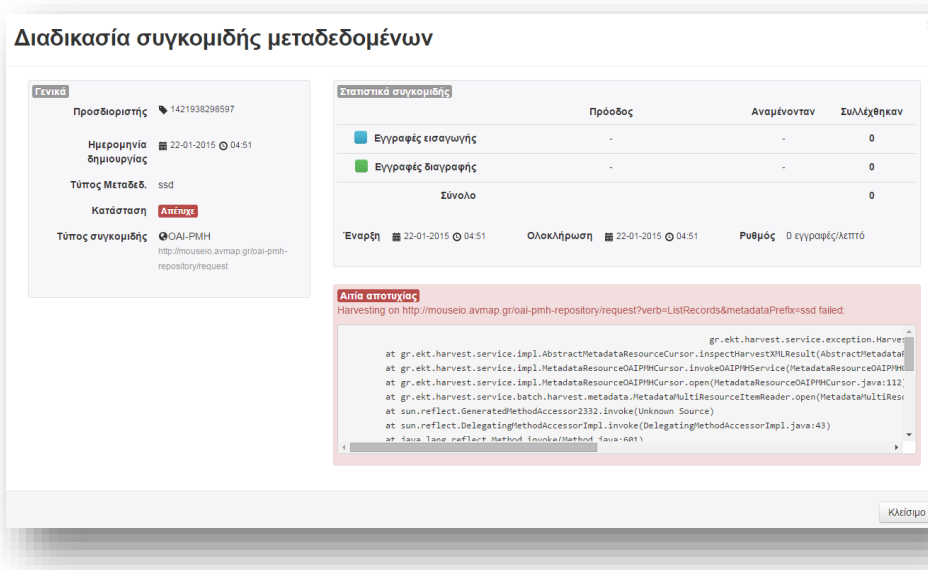
Επιλέγοντας «Ενέργειες»→«Προβολή» για μια συγκεκριμένη διαδικασία, ο χρήστης μπορεί να δει όλες τις σχετικές λεπτομέρειες και στατιστικά στοιχεία, όπως αριθμός εγγραφών που συλλέχθηκαν, ημερομηνίες και ώρες έναρξης και ολοκλήρωσης της διαδικασίας, ταχύτητα συλλογής (αριθμός εγγραφών ανά λεπτό) κα. (Σχήμα 9).

Σε περίπτωση που για οποιονδήποτε λόγο μία διαδικασία αποτύχει, αμέσως θα εμφανιστεί με κατάσταση «Απέτυχε» συνοδευόμενη από το μήνυμα σφάλματος. Εάν ο χρήστης πατήσει το εικονίδιο δίπλα ακριβώς από το μήνυμα σφάλματος ή εάν επιλέξει να δει τις λεπτομέρειες της συγκεκριμένης διαδικασίας («Ενέργειες»→«Προβολή»), θα δει λεπτομερές μήνυμα σφάλματος καθώς και αναλυτικό Exception Stack Trace (Σχήμα 10).

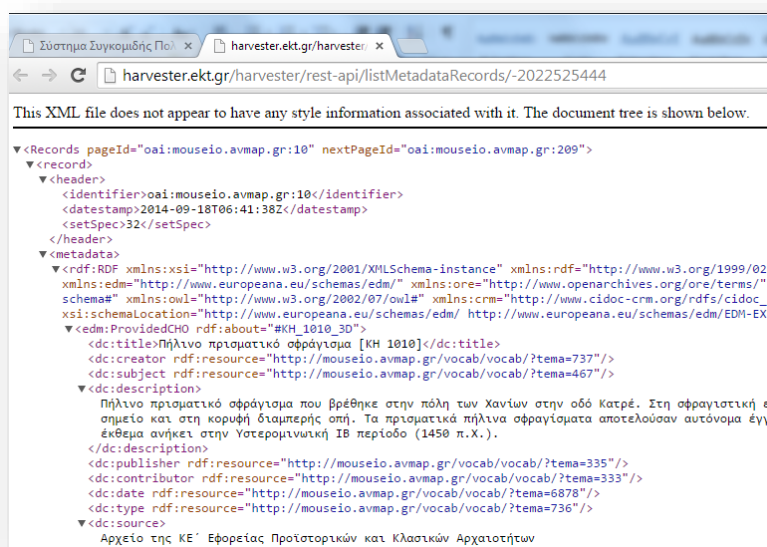
Ειδικά για τις ολοκληρωμένες διαδικασίες, ο χρήστης μπορεί να πληροφορηθεί για τον προσδιοριστή θέσης μεταδεδομένων (Location Token) με τον οποίο μπορεί να αναζητήσει και να ανακτήσει τα μεταδεδομένα είτε απευθείας χτυπώντας την Mongo βάση του Συστήματος Συγκομιδής, είτε καλώντας την μέθοδο /rest-api/listMetadataRecords του RESTful API (βλέπε Ενότητα 2.3).



Σχήμα 9. Προβολή λεπτομερών στατιστικών διαδικασίας συγκομιδής



Σχήμα 10. Εμφάνιση πληροφορίας αποσφαλμάτωσης για διαδικασία που απέτυχε



Σχήμα 11. Προβολή μεταδεδομένων που συλλέχθηκαν στα πλαίσια μιας διαδικασίας συγκομιδής

Επιλέγοντας «Διαδικασίες Συγκομιδής Ψηφιακών Αρχείων» από το βασικό μενού, ο χρήστης μεταφέρεται στην σελίδα διαχείρισης διαδικασιών συγκομιδής ψηφιακών αρχείων (Σχήμα 12).

Μέσα από την σελίδα αυτή ο χρήστης μπορεί να αναζητήσει διαδικασίες συγκομιδής αρχείων βάσει του προσδιοριστή της διαδικασίας (Digital File Harvest Token), του είδους της συγκομιδής (μέσω URL πεδίου μεταδεδομένων ή φακέλου στον εξυπηρετητή), του προσδιοριστή διαδικασίας συγκομιδής μεταδεδομένων (για συγκομιδή ψηφιακών αρχείων από πεδίο μεταδεδομένων) και, τέλος, του εύρους ημερομηνιών μέσα στο οποίο ξεκίνησε η διαδικασία. Με το πάτημα του κουμπιού «Αναζήτηση» εμφανίζεται μία λίστα με τις διαδικασίες που πληρούν τα κριτήρια αναζήτησης. Στην πρώτη στήλη του

πίνακα εμφανίζεται ο Προσδιοριστής της κάθε διαδικασίας (Digital Harvest Token), ο μοναδικός αναγνωριστικός αριθμός που την ταυτοποιεί.

Επιπλέον, μέσα από την αρχική σελίδα, ο χρήστης μπορεί να πυροδοτήσει μία νέα διαδικασία συγκομιδής ψηφιακών αρχείων πατώντας το κουμπί «Εκκίνηση Νέας Διαδικασίας Συγκομιδής Ψηφιακών Αρχείων». Με το πάτημα του κουμπιού αυτού εμφανίζεται μια φόρμα στην οποία ο χρήστης καλείται να συμπληρώσει τις παραμέτρους της συγκομιδής. Ο χρήστης επιλέγει αν θέλει να υποβάλει τα αρχεία με ανέβασμα zip αρχείου (Σχήμα 13) ή να ορίσει έναν φάκελο στον εξυπηρετητή όπου έχουν τοποθετηθεί από τους διαχειριστές των συστημάτων ή μέσω συγκομιδής μεταδεδομένων, εισάγοντας τον αντίστοιχο προσδιοριστή της διαδικασίας συγκομιδής μεταδεδομένων και ορίζοντας με XPath τα πεδία που περιέχουν τα αντίστοιχα URLs (Σχήμα 14). Ο χρήστης μπορεί προαιρετικά να επιλέξει συγκεκριμένους τύπους αρχείων (π.χ. μόνο εικόνες jpeg). Όλα τα αρχεία άλλου τύπου θα αγνοούνται. Επιπλέον ο χρήστης μπορεί να προσδιορίσει όριο αποδεκτών σφαλμάτων, δηλαδή πόσα αρχεία μπορούν να αποτύχουν (είτε στην μεταφόρτωση είτε στην εισαγωγή) χωρίς να τερματίσει ολόκληρη η διαδικασία σε κατάσταση «Απέτυχε». Με το πάτημα του κουμπιού «Εκκίνηση» πραγματοποιείται έλεγχος της ορθότητας των παραμέτρων (π.χ. αν δόθηκαν τιμές σε υποχρεωτικά πεδία ή αν η URL διεύθυνση είναι έγκυρη). Σε περίπτωση σφαλμάτων ως προς την εγκυρότητα των παραμέτρων, ο χρήστης ενημερώνεται και καλείται να τα διορθώσει. Διαφορετικά η διαδικασία ξεκινά, η φόρμα κλείνει και στην λίστα των διαδικασιών εμφανίζεται η νέα διαδικασία με κατάσταση «Σε εξέλιξη». Το πεδίο «Στατιστικά Συγκομιδής» για την συγκεκριμένη διαδικασία εμφανίζει μία μπάρα προόδου που ανανεώνεται σχεδόν σε πραγματικό χρόνο.

Προσδιοριστής	Ημερομηνία Δημιουργίας	Τύπος συγκομιδής αρχείου	Κατάσταση	Στατιστικά συγκομιδής	Ενέργειες
1421658536842	19-01-2015 11:08	Από URL πεδίου μεταδεδομένων 1421658533831 /!ese:object	Ολοκληρώθηκε	616 αρχεία	Ενέργειες
1421400392489	16-01-2015 11:26	Από URL πεδίου μεταδεδομένων 1421400389453 /!europa:object	Ολοκληρώθηκε	1881 αρχεία	Ενέργειες
1421339263716	15-01-2015 06:27	Από URL πεδίου μεταδεδομένων 1421315671376 /!ese:isShownBy	Ολοκληρώθηκε	616 αρχεία	Ενέργειες
1421244811389	14-01-2015 04:13	Από URL πεδίου μεταδεδομένων 1421244742127	Ολοκληρώθηκε με αποτυχία	616 αρχεία	Ενέργειες

Σχήμα 12. Σελίδα διαχείρισης διαδικασιών συγκομιδής Ψηφιακών Αρχείων

Επιλέγοντας «Ενέργειες» → «Προβολή» για μια συγκεκριμένη διαδικασία, ο χρήστης μπορεί να δει όλες τις σχετικές λεπτομέρειες και στατιστικά στοιχεία, όπως αριθμός αρχείων που συλλέχτηκαν,

ημερομηνίες και ώρες έναρξης και ολοκλήρωσης της διαδικασίας, ταχύτητα συλλογής (αριθμός αρχείων ανά λεπτό) κα. (Σχήμα 15).

The dialog box is titled "Εκκίνηση νέας διαδικασίας συγκομιδής αρχείων". It contains the following fields and options:

- Τύπος συγκομιδής αρχείων:** A dropdown menu set to "Από ZIP αρχείο".
- ZIP αρχείο:** A "Choose File" button and the text "No file chosen".
- Τύποι αρχείων που θα συλλεχθούν:** A list of file types with checkboxes. "Όλοι οι τύποι" is checked. Other options include application/pdf, application/zip, audio/basic, audio/L24, audio/mp4, audio/mpeg, image/gif, image/jpeg, image/pjpeg, image/png, image/svg+xml, image/tiff, image/vnd.djvu, video/avi, video/mpeg, video/mp4, video/ogg, video/quicktime, video/webm, video/x-matroska, video/x-ms-wmv, and video/x-flv.
- Όριο αποδεκτών σφαλμάτων εισαγωγής:** A text input field containing the number "10".
- Buttons:** "Εκκίνηση" (Start) and "Ακύρωση" (Cancel).

Σχήμα 13. Φόρμα εκκίνησης νέας διαδικασίας συγκομιδής Ψηφιακών Αρχείων από ZIP αρχείο.

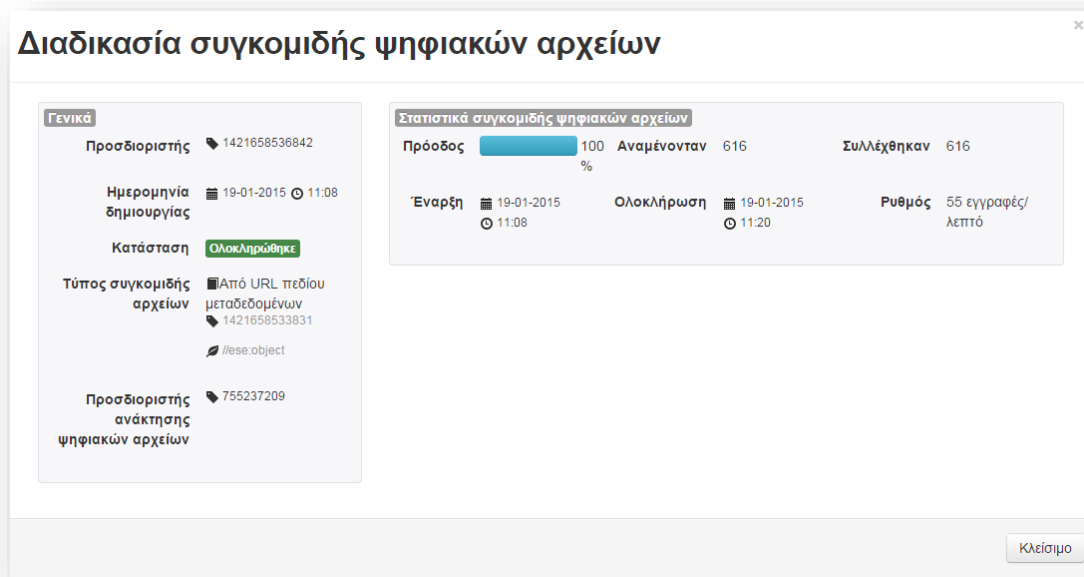
The dialog box is titled "Εκκίνηση νέας διαδικασίας συγκομιδής αρχείων". It contains the following fields and options:

- Τύπος συγκομιδής αρχείων:** A dropdown menu set to "Από URL πεδίου μεταδεδομένων".
- Προσδ. συγκομιδής μεταδεδομένων:** A text input field containing "1421658533831".
- Πεδίο σε μεταδεδομένα (XPath π.χ. //edm:object/@rdf:resource):** A text input field containing "//europeana:isShownBy".
- Τύποι αρχείων που θα συλλεχθούν:** A list of file types with checkboxes. "Όλοι οι τύποι" is checked. Other options include application/pdf, application/zip, audio/basic, audio/L24, audio/mp4, audio/mpeg, image/gif, image/jpeg, image/pjpeg, image/png, image/svg+xml, image/tiff, image/vnd.djvu, video/avi, video/mpeg, video/mp4, video/ogg, video/quicktime, video/webm, video/x-matroska, video/x-ms-wmv, and video/x-flv.
- Όριο αποδεκτών σφαλμάτων εισαγωγής:** A text input field containing the number "10".
- Buttons:** "Εκκίνηση" (Start) and "Ακύρωση" (Cancel).

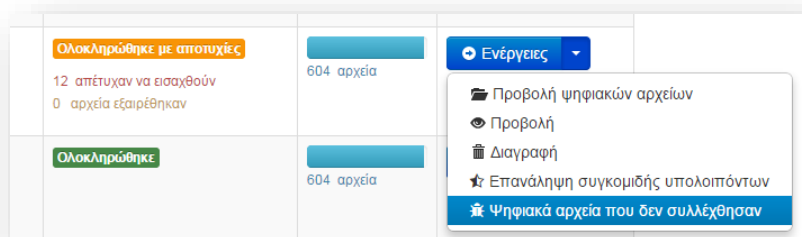
Σχήμα 14. Φόρμα εκκίνησης νέας διαδικασίας συγκομιδής Ψηφιακών Αρχείων από ανάγνωση URL από πεδία μεταδεδομένων

Στην περίπτωση που η διαδικασία ολοκληρώθηκε σε κατάσταση «Ολοκληρώθηκε με αποτυχίες» ή τερματίστηκε ανεπιτυχώς (αν ο αριθμός των αποτυχιών ξεπέρασε το όριο που έθεσε ο χρήστης)

(κατάσταση «Επέτυχε»), στο κουμπί «Ενέργειες» εμφανίζονται δύο νέες επιλογές: «Ψηφιακά αρχεία που δεν συλλέχθηκαν» και «Επανάληψη συγκομιδής υπολειπόντων» (Σχήμα 16).



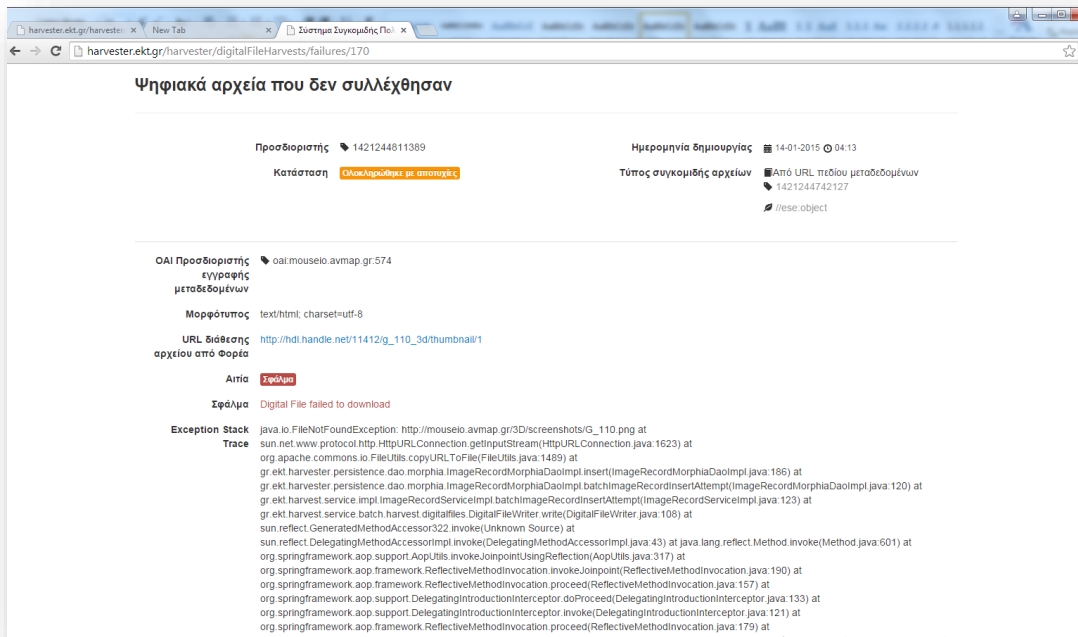
Σχήμα 15. Λεπτομέρειες και στατιστικά μιας διαδικασίας συγκομιδής Ψηφιακών Αρχείων



Σχήμα 16. Ενέργειες για διαδικασίες που είτε απέτυχαν είτε ολοκληρώθηκαν με αποτυχίες

Με την επιλογή «Ψηφιακά αρχεία που δεν συλλέχθηκαν», ο εξουσιοδοτημένος χρήστης μπορεί να δει αναλυτικά ποια ψηφιακά αρχεία απέτυχαν, για ποιό λόγο και σε ποια φάση παρουσιάστηκε το σφάλμα (π.χ. κατά την μεταφόρτωση ή την αποθήκευση), όπως φαίνεται στο Σχήμα 17. Με την επιλογή «Επανάληψη συγκομιδής υπολοίπων» θα ξεκινήσει η προσπάθεια συλλογής όλων των ψηφιακών αρχείων που δεν συλλέχθηκαν.

Τέλος, ειδικά για τις ολοκληρωμένες διαδικασίες, ο χρήστης μπορεί να πληροφορηθεί για τον προσδιοριστή θέσης ψηφιακών αρχείων (digital file location token) με τον οποίο μπορεί είτε να αναζητήσει τα μεταδεδομένα απευθείας χτυπώντας την Mongo βάση του Συστήματος Συγκομιδής, είτε καλώντας την μέθοδο /rest-api/listDigitalFileRecords του RESTful API (βλέπε Ενότητα 2.3).



Σχήμα 17. Προβολή σφαλμάτων συλλογής ψηφιακών αρχείων

2.3 RESTful Υπηρεσίες

Το REST API του Συστήματος Συγκομιδής παρέχει μία σειρά από μεθόδους που επιτρέπουν σε ένα εξωτερικό σύστημα (συσσωρευτής SearchCulture.gr, συσσωρευτής CRIS, Validator, Preservator) να χρησιμοποιήσει τις υπηρεσίες του συσσωρευτή. Ακολουθούν οι προδιαγραφές του API.

POST /rest-api/startHarvestProcess

Περιγραφή Ξεκινά μία νέα διαδικασία συγκομιδής μεταδεδομένων

Response Content Type: Application/json

Επιστρέφει: HarvestStartedResponse {
errorsMap (List[Entry«String,List«String»]], optional),
token (long, optional),
state = ['SUCCESS' or 'FAILED' or 'VALIDATION_ERRORS']
}

Παράμετροι: HarvestSchedule {
harvestType = ['OAI_PMH_HARVESTING' or 'FOLDER'],
metadataPrefix (string),
oaiDateFrom (Date),
oaiDateUntil (Date),
oaiURL (string, optional),
sourceFolder (string, optional),
dataSet (string, optional)
}

POST /rest-api/validateHarvestRequest

Περιγραφή Ελέγχει αν οι παράμετροι για έναρξη διαδικασίας συγκομιδής μεταδεδομένων είναι έγκυρες

Response Content Type: Application/json

Επιστρέφει: HarvestValidationResponse {
errorsMap (List[Entry«String,List«String»»], optional),
state = ['VALID' or 'VALIDATION_ERRORS']
}

Παράμετροι: HarvestSchedule {
harvestType = ['OAI_PMH_HARVESTING' or 'FOLDER'],
metadataPrefix (string),
oaiDateFrom (Date),
oaiDateUntil (Date),
oaiURL (string, optional),
sourceFolder (string, optional),
dataSet (string, optional)
}

GET /rest-api/getStatusOfHarvest/{harvestToken}

Περιγραφή Επιστρέφει πληροφορίες σχετικά με την κατάσταση/πρόοδο μιας διαδικασίας συγκομιδής μεταδεδομένων

Response Content Type: Application/json

Επιστρέφει: HarvestProgressResponse {
token (long, optional),
dateCreated (timestamp, optional),
status = ['IN_PROCESS' or 'COMPLETED' or 'FAILED'],
metadataPrefix (string, optional),
dataSet (string, optional),
oaiDateFrom (Date, optional),
oaiDateUntil (Date, optional),
harvestType = ['OAI_PMH_HARVESTING' or 'FOLDER'],
oaiURL (string, optional),
sourceFolder (string, optional),
harvestingStart (Date, optional),
harvestingEnd (Date, optional),
completeListExpectedSize (int, optional),
numberOfInsertRecordsHarvested (int, optional),
numberOfDeleteRecordsHarvested (int, optional),
errorMessage (string, optional),
exceptionStackTrace (string, optional),
metadataLocationToken (long, optional)
}

Παράμετροι: harvestToken (path)

GET /rest-api/getStatusOfMultipleHarvests/{harvestTokens}

Περιγραφή Επιστρέφει πληροφορίες σχετικά με την κατάσταση/πρόοδο πολλαπλών διαδικασιών συγκομιδής μεταδεδομένων

Response Content Type: Application/json

Επιστρέφει: Array Of HarvestProgressResponse {

```

token (long, optional),
dateCreated (timestamp, optional),
status = ['IN_PROCESS' or 'COMPLETED' or 'FAILED'],
metadataPrefix (string, optional),
dataSet (string, optional),
oaiDateFrom (Date, optional),
oaiDateUntil (Date, optional),
harvestType = ['OAI_PMH_HARVESTING' or 'FOLDER'],
oaiURL (string, optional),
sourceFolder (string, optional),
harvestingStart (Date, optional),
harvestingEnd (Date, optional),
completenessExpectedSize (int, optional),
numberOfInsertRecordsHarvested (int, optional),
numberOfDeleteRecordsHarvested (int, optional),
errorMessage (string, optional),
exceptionStackTrace (string, optional),
metadataLocationToken (long, optional)
}

```

Παράμετροι: harvestTokens (path) : Λίστα από harvest tokens χωριζόμενα από ',' (π.χ. '3323, 23332, 3355')

GET

/rest-api/listMetadataRecords/{metadataLocationToken}

Περιγραφή Επιστρέφει εγγραφές μεταδεδομένων που συλλέχθηκαν από μία ολοκληρωμένη διαδικασία συγκομιδής με το συγκεκριμένο location token, ανά σελίδα.

Response Content Type: Application/xml

Επιστρέφει: MetadataRecordsPage {
pageId (string, optional),
nextPageId (string, optional),
list (List[MetadataRecord], optional),
harvestId (long, optional)
}

Παράμετροι: metadataLocationToken : το location token

page.pagelId : προσδιοριστής σελίδας αποτελεσμάτων

page.size : μέγεθος σελίδας

POST

/rest-api/startDigitalFileHarvestProcess

Περιγραφή Ξεκινά μία νέα διαδικασία συγκομιδής ψηφιακών αρχείων

Response Content Type: Application/json

Επιστρέφει: HarvestStartedResponse {
errorsMap (List[Entry«String,List«String»»], optional),
token (long, optional),
state = ['SUCCESS' or 'FAILED' or 'VALIDATION_ERRORS']
}

Παράμετροι: DigitalFileHarvestSchedule {

```
zipFile (string, optional),
sourceFolder (string, optional),
digitalFileHarvestType (DigitalFileHarvestType, optional)
    = ['METADATA_FIELD' or 'ZIP' or 'FOLDER' or 'FTP'],
metadataToken (long, optional),
metadataField (string, optional)
}
```

POST /rest-api/validateDigitalFileHarvestRequest

Περιγραφή Ελέγχει αν οι παράμετροι για έναρξη διαδικασίας συγκομιδής ψηφιακών αρχείων είναι έγκυρες

Response Content Type: Application/json

Επιστρέφει: HarvestValidationResponse {
errorsMap (List[Entry«String,List«String»»], optional),
state = ['VALID' or 'VALIDATION_ERRORS']
}

Παράμετροι: DigitalFileHarvestSchedule {
zipFile (string, optional),
sourceFolder (string, optional),
digitalFileHarvestType (DigitalFileHarvestType, optional)
 = ['METADATA_FIELD' or 'ZIP' or 'FOLDER' or 'FTP'],
metadataToken (long, optional),
metadataField (string, optional)
}

GET /rest-api/getStatusOfDigitalFileHarvest/{harvestToken}

Περιγραφή Επιστρέφει πληροφορίες σχετικά με την κατάσταση/πρόοδο μιας διαδικασίας συγκομιδής ψηφιακών αρχείων

Response Content Type: Application/json

Επιστρέφει: DigitalFileHarvestProgressResponse {
errorMessage (string, optional),
status (HarvestStatus, optional) = ['IN_PROCESS' or
'COMPLETED' or 'FAILED'],
numberOfDigitalFilesHarvested (int, optional),
exceptionStackTrace (string, optional),
sourceFolder (string, optional),
digitalFileHarvestingEnd (Date, optional),
lastUpdateDate (Date, optional),
digitalFileLocationToken (long, optional),
token (long, optional),
zipFile (string, optional),
digitalFileHarvestingStart (Date, optional),
completelistExpectedSize (int, optional),
dateCreated (timestamp, optional),
digitalFileHarvestType (DigitalFileHarvestType, optional)
 = ['METADATA_FIELD' or 'ZIP' or 'FOLDER' or
 'FTP'],
metadataToken (long, optional)
}

Παράμετροι: digitalFileHarvestToken (path)

GET**/rest-api/getStatusOfMultipleDigitalFileHarvests/{digitalFileHarvestTokens}**

Περιγραφή Επιστρέφει πληροφορίες σχετικά με την κατάσταση/πρόοδο πολλαπλών διαδικασιών συγκομιδής ψηφιακών αρχείων

Response Content Type: Application/json

Επιστρέφει: Array Of DigitalFileHarvestProgressResponse {
 errorMessage (string, optional),
 status (HarvestStatus, optional) = ['IN_PROCESS' or 'COMPLETED' or 'FAILED'],
 numberOfDigitalFilesHarvested (int, optional),
 exceptionStackTrace (string, optional),
 sourceFolder (string, optional),
 digitalFileHarvestingEnd (Date, optional),
 lastUpdateDate (Date, optional),
 digitalFileLocationToken (long, optional),
 token (long, optional),
 zipFile (string, optional),
 digitalFileHarvestingStart (Date, optional),
 completeListExpectedSize (int, optional),
 dateCreated (timestamp, optional),
 digitalFileHarvestType (DigitalFileHarvestType, optional) = ['METADATA_FIELD' or 'ZIP' or 'FOLDER' or 'FTP'],
 metadataToken (long, optional)
}

Παράμετροι: digitalFileHarvestTokens (path) : Λίστα από digital file harvest tokens χωριζόμενα από ',' (π.χ. '3323, 23332, 3355')

GET**/rest-api/listDigitalFileRecords/{digitalFileLocationToken}**

Περιγραφή Επιστρέφει εγγραφές ψηφιακών αρχείων (η κάθε εγγραφή περιλαμβάνει το URL του αρχείου που εξυπηρετείται από το σύστημα συγκομιδής) που συλλέχθηκαν από μία ολοκληρωμένη διαδικασία συγκομιδής με το συγκεκριμένο digital file location token, ανά σελίδα.

Response Content Type: Application/xml

Επιστρέφει: DigitalFilesPage {
 pageId (string, optional),
 empty (boolean, optional),
 nextPageId (string, optional),
 locationToken (long, optional),
 list (List[DigitalFileRecordDTO], optional)
}

Παράμετροι: digitalFileLocationToken : το digital file location token
page.pageId : προσδιοριστής σελίδας αποτελεσμάτων
page.size : μέγεθος σελίδας

2.4 Harvester: αρχιτεκτονική συστήματος

Το Σύστημα Συγκομιδής Ψηφιακού Περιεχομένου Harvester είναι μία Java εφαρμογή ιστού πακεταρισμένη ως war αρχείο και μπορεί να εγκατασταθεί σε οποιονδήποτε Servlet container, όπως ο Apache Tomcat. Ταυτόχρονα είναι ένα Maven project αποτελούμενο από τρία modules, το harvester-persistence, το harvester-service και το harvester-web. Τα modules αυτά αντιστοιχούν στα στρώματα persistence (ή αλλιώς infrastructure layer), service (ή αλλιώς business layer) και presentation (εν προκειμένου web layer), αντίστοιχα, της πολυ-διαστρωματωμένης αρχιτεκτονικής (multi-layered architecture). Τα modules harvester-persistence και harvester-service πακετάρονται ως jar αρχεία, με το δεύτερο να έχει εξάρτηση στο πρώτο. Το module harvester-web πακετάρεται ως war και έχει εξάρτηση στο harvester-service module και κατ' επέκταση και στο harvester-persistence.

Το persistence layer παρέχει τις βασικές CRUD μεθόδους (Create-Read-Update-Delete) καθώς και μεθόδους αναζήτησης για τις οντότητες πεδίου που αποθηκεύονται μόνιμα σε βάσεις. Για κάθε βασική οντότητα πεδίου ορίζεται από ένα DAO αντικείμενο, σύμφωνα με το σχεδιαστικό πρότυπο DAO [16]. Το Σύστημα Συγκομιδής χρησιμοποιεί 3 βάσεις, μία NoSQL βάση (Mongo DB) για την αποθήκευση των εγγραφών μεταδεδομένων (οντότητα πεδίου MetadataRecord) που συλλέγονται, μία noSQL βάση (MongoDB GridFS [22]) για την αποθήκευση των ψηφιακών αρχείων που συλλέγονται (οντότητα πεδίου DigitalFile) και μία σχεσιακή βάση (PostgreSQL) για την αποθήκευση πληροφορίας σχετικά με τις διαδικασίες συγκομιδής και την κατάστασή τους (οντότητες πεδίου Harvest και DigitalFileHarvest για διαδικασίες συγκομιδής μεταδεδομένων και ψηφιακών αρχείων, αντίστοιχα).

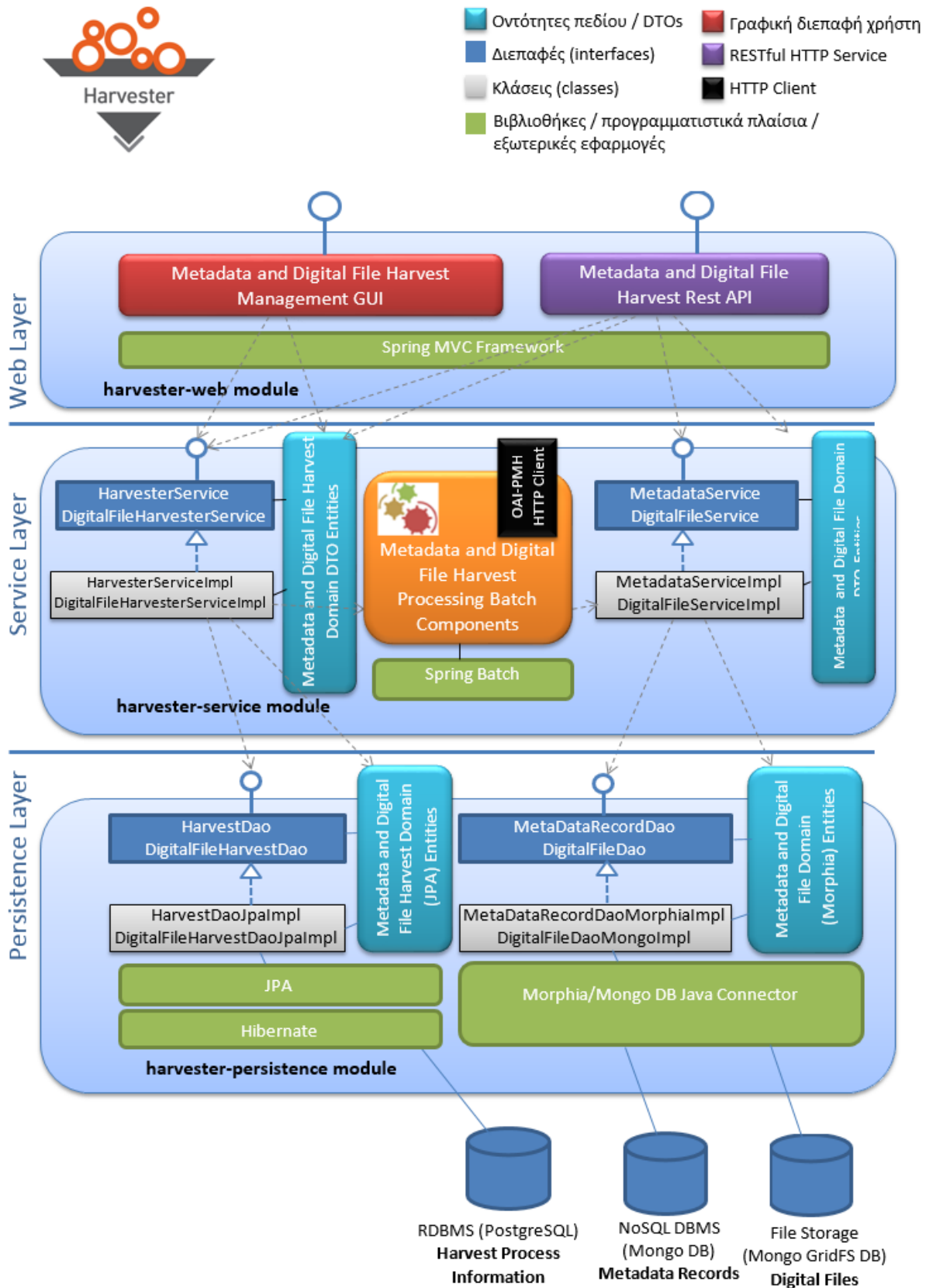
Το service layer παρέχει μία σειρά από υπηρεσίες (services) στο επόμενο layer (harvester-web module) οι οποίες υλοποιούν όλες τις business διαδικασίες του συστήματος: εκκίνηση νέας διαδικασίας συγκομιδής μεταδεδομένων, εκκίνηση νέας διαδικασίας συγκομιδής ψηφιακών αρχείων, παρακολούθηση και καταγραφή της προόδου της κάθε διαδικασίας συγκομιδής κα. Για να επιτευχθεί η αποδέσμευση του επόμενου layer από το persistent layer, άρα και από τις οντότητες πεδίου που αυτό ορίζει, δημιουργήθηκαν Data Access Services για κάθε αντικείμενο DAO και υιοθετήθηκε αυστηρά το DTO προγραμματιστικό υπόδειγμα [23]: δημιουργήθηκαν δηλαδή νέες οντότητες πεδίου στο Service Layer, τα DTOs. Τα services χρησιμοποιούν ως είσοδο ή/και επιστρέφουν αποκλειστικά DTO αντικείμενα.

Για την υλοποίηση της ροής συγκομιδής μεταδεδομένων και ψηφιακών αρχείων χρησιμοποιήθηκε το Spring Batch framework [15] και υλοποιήθηκαν μια σειρά από διεπαφές που ορίζονται στο framework.

Τέλος το web layer υλοποιεί την γραφική διεπαφή ιστού και τις RESTfull υπηρεσίες που επιτρέπουν σε ένα χρήστη ή σε ένα εξωτερικό σύστημα, αντίστοιχα, να πυροδοτούν διαδικασίες συγκομιδής, να ενημερώνονται για την εξέλιξη τους και να ανακτούν τα μεταδεδομένα ή τα ψηφιακά αρχεία που συλλέχθηκαν από ολοκληρωμένες διαδικασίες.

Και τα τρία module χρησιμοποιούν το Spring framework [13] για Dependency Injection και για την διαχείριση των διεπαφών (interfaces) και των κλάσεων που τις υλοποιούν. Για τον συντονισμό και ενορχήστρωση των κλάσεων που αναπτύχθηκαν και που υλοποιούν τις βασικές ροές συγκομιδής χρησιμοποιήθηκε το Spring Batch Framework [15]. Για την ταυτοποίηση και αυθεντικοποίηση χρήστη χρησιμοποιήθηκε το προγραμματιστικό πλαίσιο Spring Security [35]. Για την ανάπτυξη της γραφικής εφαρμογής ιστού και των RESTful υπηρεσιών χρησιμοποιήθηκε το Spring MVC framework [14]. Οι ιστοσελίδες της γραφικής εφαρμογής γράφτηκαν σε JSP (με χρήση JSTL 2.0 [33] και JSP custom tags 2.0

[33]) σε συνδυασμό με διάφορες front-end τεχνολογίες, με κυριότερες το προγραμματιστικό πλαίσιο JQuery [29], την εργαλειοθήκη Bootstrap [30] και την εργαλειοθήκη για εικονίδια Font awesome [34].



Σχήμα 18. Συνολική Αρχιτεκτονική του Συστήματος Συγκομιδής Ψηφιακού Περιεχομένου

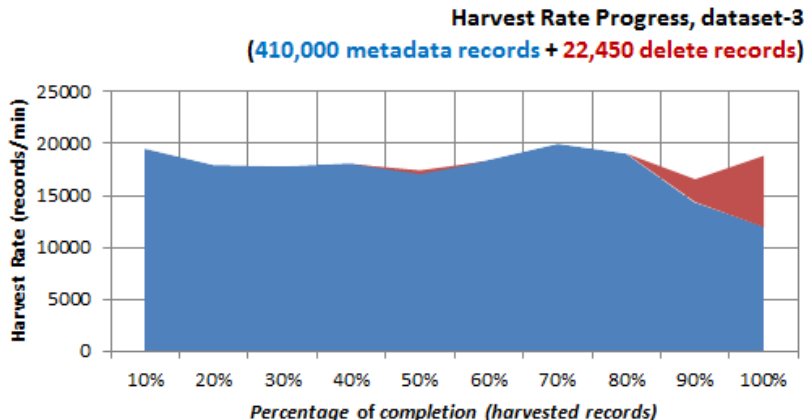
2.5 Πειραματική Αξιολόγηση Συστήματος Συγκομιδής

Στην ενότητα αυτή παρουσιάζεται η πειραματική αξιολόγηση της απόδοσης του συστήματος Συγκομιδής. Χρησιμοποιήθηκαν τρία διαφορετικού μεγέθους σύνολα μεταδεδομένων σε μορφή OAI-PMH XML απόκρισης, τα χαρακτηρίστηκα των οποίων συγκεντρώνει ο Πίνακας II.

Το διάγραμμα στο Σχήμα 19 παρουσιάζει την εξέλιξη του ρυθμού συλλογής και αποθήκευσης των εγγραφών μεταδεδομένων κατά την διάρκεια μια ολοκληρωμένης διαδικασίας συγκομιδής των μεταδεδομένων του συνόλου dataset-3. Παρατηρούμε ότι ο ρυθμός παραμένει σχεδόν σταθερός γεγονός που αποδεικνύει ότι το σύστημα συγκομιδής κλιμακώνει ιδιαίτερα ικανοποιητικά. Το κόκκινο χρώμα στο διάγραμμα αντιστοιχεί σε εγγραφές διαγραφής (οι οποίες στο συγκεκριμένο δείγμα έτυχε να συγκεντρώνονται στο τέλος των εγγραφών). Οι εγγραφές διαγραφής έχουν πολύ μικρό όγκο με αποτέλεσμα να αυξάνουν την ταχύτητα της συλλογής και αποθήκευσης.

Πίνακας II. Χαρακτηριστικά δεδομένων που χρησιμοποιήθηκαν στα πειράματα

	Total records	"deleted" records	records with metadata	size in MB
dataset-1	82,004	4	82,000	138 MB
dataset-2	205,920	920	205,000	344 MB
dataset-3	432,435	22,435	410,000	679 MB



Σχήμα 19. Η εξέλιξη της ροής εισαγωγής κατά την διάρκεια μιας διαδικασίας συγκομιδής μεταδεδομένων ως προς το ποσοστό εξέλιξης αυτής

3 SearchCulture.gr:

Σύστημα Συσσωρευτή Περιεχομένου/ Ενιαίας Αναζήτησης



Το κεφάλαιο αυτό περιγράφει αναλυτικά το Σύστημα Συσσωρευτή Περιεχομένου/Ενιαίας Αναζήτησης SearchCulture.gr. Αρχικά παρατίθενται όλα τα λειτουργικά χαρακτηριστικά του συστήματος, ακολουθεί μία πλήρης τεκμηρίωση χρήσης τόσο της ανοικτής διαδικτυακής πύλης όσο και του διαχειριστικού περιβάλλοντος και, τέλος, γίνεται μια διεξοδική τεχνική περιγραφή της αρχιτεκτονικής του συστήματος.

3.1 Λειτουργικά Χαρακτηριστικά

Το Σύστημα Συσσωρευτή Περιεχομένου/ Ενιαίας Αναζήτησης SearchCulture.gr αποτελεί ένα κεντρικό σημείο πρόσβασης σε ψηφιακό περιεχόμενο, διασυνδέοντας τα μεταδεδομένα του ψηφιακού υλικού του κάθε φορέα που συμμετέχει με το ψηφιακό υλικό καθαυτό στο σημείο σταθερής απόθεσής του καθώς και με εξωτερικές πηγές σημασιολογικής τεκμηρίωσης (θησαυρούς όρων, δεσμευμένα λεξιλόγια).

Ο συσσωρευτής διαθέτει μια διαχειριστική web εφαρμογή η οποία επιτρέπει στον εξουσιοδοτημένο χρήστη να καταχωρήσει φορείς περιεχομένου και συλλογές και να πραγματοποιήσει διαδικασίες εισαγωγής μεταδεδομένων και ψηφιακών αρχείων (ως εικόνες προεπισκόπησης) από τα σημεία σταθερής απόθεσής τους. Ο εξουσιοδοτημένος χρήστης μπορεί να εμπλουτίσει τις εισηγμένες εγγραφές σημασιολογικά ως προς τον τύπο τους χρησιμοποιώντας το εργαλείο σημασιολογικού εμπλουτισμού του Semantics.gr. Με την διαδικασία αυτή, το σύνολο σχεδόν των τεκμηρίων εμπλουτίστηκε με ένα νέο πεδίο «ΕΚΤ Τύπος» που περιλαμβάνει αναφορές σε όρους ενός πρότυπου δίγλωσσου ιεραρχικού λεξιλογίου τύπων που δημιουργήθηκε από το ΕΚΤ.

Το SearchCulture.gr αποθηκεύει τα μεταδεδομένα που συσσωρεύει σε μορφή συμβατή με τον τύπο μεταδεδομένων EDM [7]. Το EDM είναι το μοντέλο που προτείνεται από την Europeana για την αναπαράσταση μεταδεδομένων πολιτιστικής κληρονομιάς και αποτελεί μετεξέλιξη του ESE [6]. Το EDM είναι ένα RDF μοντέλο [8] και ενσωματώνει μηχανισμούς που επιτρέπουν στα μεταδεδομένα να διατηρούν σημασιολογικές αναφορές σε λεξιλόγια και θησαυρούς όρων καθιστώντας εφικτή την διάθεσή τους ως Διασυνδεδεμένα Δεδομένα. Επομένως, το SearchCulture.gr μπορεί να φιλοξενήσει εγγενώς σημασιολογικά εμπλουτισμένο περιεχόμενο. Εσωτερικά η Europeana χρησιμοποιεί μια παραλλαγή του δημόσιου EDM τύπου, γνωστή ως EDM-internal [10], η οποία είναι προσαρμοσμένη για τις ανάγκες ενός συσσωρευτή. Το SearchCulture.gr, ως συσσωρευτής, αποθηκεύει τα μεταδεδομένα στο εσωτερικό EDM σχήμα.

Το SearchCulture.gr υποστηρίζει εισαγωγή μεταδεδομένων σε τρεις εναλλακτικούς τύπους. Αποθετήρια τα οποία δεν περιέχουν σημασιολογικά εμπλουτισμένο περιεχόμενο μπορούν να διαθέσουν τα μεταδεδομένα τους σε τύπο ESE, το οποίο ουσιαστικά βασίζεται στο εμπλουτισμένο Dublin Core (Metadata Terms) [37] μοντέλο επεκτείνοντάς το με επιπρόσθετα πεδία που αφορούν στα ψηφιακά αρχεία. Τα αποθετήρια που διαθέτουν σημασιολογικά εμπλουτισμένο περιεχόμενο μπορούν να διαθέσουν τα μεταδεδομένα τους σε EDM (στην δημόσια έκδοσή του). Εναλλακτικά, μπορούν να διαθέσουν τα μεταδεδομένα τους συνδυαστικά σε δύο τύπους, σε ESE και σε DC-DS-XML [11]. Ο τύπος DC-DS-XML είναι μία XML αναπαράσταση του Dublin Core (Metadata Terms) με δυνατότητα σημασιολογικών αναφορών σε όρους λεξιλογίων και θησαυρών. Στην περίπτωση αυτή, οι πληροφορίες για τα ψηφιακά αρχεία αντλούνται από τα μεταδεδομένα σε τύπο ESE. Η βασική ροή εισαγωγής μεταδεδομένων του SearchCulture.gr μετασχηματίζει τα πρωτότυπα μεταδεδομένα από EDM, ή από ESE ή από ESE σε συνδυασμό με DC-DS-XML στον τύπο EDM-internal προκειμένου να τα αποθηκεύσει στην βάση δεδομένων που χρησιμοποιεί.

Η δημόσια υπηρεσία του συστήματος είναι η διαδικτυακή πύλη Ενιαίας Αναζήτησης Ψηφιακού Περιεχομένου όπου ο χρήστης μπορεί να αναζητήσει τεκμήρια και για κάθε τεκμήριο να δει αναλυτικές πληροφορίες που περιλαμβάνουν μεταξύ άλλων σημασιολογικές παραπομπές, σύνδεσμο στο ψηφιακό αρχείο και στην σελίδα παρουσίασης στο σημείο σταθερής απόθεσής του καθώς και το αποτέλεσμα του σημασιολογικού εμπλουτισμού του ΕΚΤ. Τα αποτελέσματα της αναζήτησης συνοδεύονται από εικόνες προεπισκόπησης ενώ παρέχεται η δυνατότητα περαιτέρω φιλτραρίσματος των αποτελεσμάτων βάσει

συγκεκριμένων πεδίων (facet queries [21]). Η δικτυακή πύλη της Ενιαίας Αναζήτησης Ψηφιακού Περιεχομένου παρέχει το περιεχόμενο που έχει συσσωρευτεί και ως Ανοιχτά Διασυνδεδεμένα Δεδομένα. Επιπλέον, η διαδικτυακή πύλη διαθέτει εναλλακτικούς τρόπους ανακάλυψης περιεχομένου, όπως σελίδα παρουσίασης φορέων και συλλογών και ιεραρχικής πλοήγησης στους ΕΚΤ τύπους τεκμηρίων, σελίδες οι οποίες αξιοποιούν στο έπακρον τον σημασιολογικό εμπλουτισμό.

Ο Πίνακας III περιλαμβάνει όλα τα λειτουργικά χαρακτηριστικά της δημόσιας διαδικτυακής πύλης SearchCulture.gr ενώ ο Πίνακας IV περιλαμβάνει τα λειτουργικά χαρακτηριστικά του συστήματος συσσώρευσης και του διαχειριστικού του περιβάλλοντος.

Πίνακας III. Οι Λειτουργικές Απαιτήσεις για την δημόσια διαδικτυακή πύλη του SearchCulture.gr

Απαίτηση	Περιγραφή
1 Μηχανή αναζήτησης	Το SearchCulture.gr παρέχει μία δικτυακή πύλη η οποία προσφέρει στους χρήστες δυνατότητες αναζήτησης και πλοήγησης του συσσωρευμένου υλικού, μέσα από ένα φιλικό και εύχρηστο περιβάλλον. Η αναζήτηση μπορεί να εξειδικευθεί ανά φορέα και ανά σημασιολογική κατηγορία (π.χ. τύπος υλικού, θέμα, χρονολογικός προσδιορισμός) αξιοποιώντας φίλτρα αναζήτησης. Τα πεδία αναζήτησης υποστηρίζουν μηχανισμό αυτόματης συμπλήρωσης (autocomplete) για την εμφάνιση προτεινόμενων τιμών από δυναμική λίστα. Τα αποτελέσματα παρουσιάζονται σελιδοποιημένα, ενώ για κάθε τεκμήριο προβάλλεται εικόνα προεπισκόπησης, συνοδευόμενη από βασικές πληροφορίες όπως τίτλος, πάροχος και δικαιώματα διάθεσης.
2 Δυνατότητα φιλτραρίσματος αποτελεσμάτων βάσει συγκεκριμένων πεδίων – Facet Queries	Ο χρήστης μπορεί να φιλτράρει τα αποτελέσματα βάσει τιμών σε συγκεκριμένα πεδία (προτυποποιημένου τύπου τεκμηρίου ΕΚΤ, άδειας χρήσης, φορέα κλπ) οι οποίες συνοδεύονται με τον πληθάρισμά τους. Ο χρήστης μπορεί να επιλέξει αν επιθυμεί τα φίλτρα να εμφανίζονται με σειρά δημοφιλίας (πληθαρίστου) ή αλφαβητικά.
3 Καρτέλα τεκμηρίου	Η καρτέλα τεκμηρίου περιλαμβάνει αναλυτικές πληροφορίες για το τεκμήριο με δομημένο τρόπο, εικόνα προεπισκόπησης καθώς και συνδέσμους προς ψηφιακά αρχεία που παρουσιάζουν το τεκμήριο στην δικτυακή πύλη του παρόχου. Τυχόν σημασιολογικές αναφορές εμφανίζονται επίσης ως σύνδεσμοι παραπέμποντας σε σελίδα τεκμηρίωσης της αντίστοιχης σημασιολογικής οντότητας. Στην καρτέλα τεκμηρίου μεταφέρεται κανείς είτε από τα αποτελέσματα αναζήτησης είτε εισάγοντας στον φυλλομετρητή το URI του τεκμηρίου, είτε από εξωτερικούς συνδέσμους από Linked Data πηγές που περιέχουν αναφορές στο συγκεκριμένο τεκμήριο. Τα πεδία που εμφανίζονται και η σειρά εμφάνισής τους ορίζονται παραμετρικά από τον εξουσιοδοτημένο χρήστη μέσα από το διαχειριστικό περιβάλλον.
4 Δημόσια Διαδικτυακή Πύλη - Διάθεση ως Linked Data [9]	Για κάθε τεκμήριο αντιστοιχίζεται ένα μοναδικό URI το οποίο εφόσον αιτηθεί με HTML header Accept=html/text (π.χ. μέσω ενός κοινού φυλλομετρητή) επιστρέφει την ιστοσελίδα παρουσίασης του (καρτέλα τεκμηρίου), ενώ όταν αιτείται με HTML header Accept=application/rdf+xml επιστρέφει την EDM rdf/xml αναπαράσταση του τεκμηρίου.

5	Δημόσια Διαδικτυακή Πύλη – Προσαρμοστικότητα (responsiveness) ώστε η διαδικτυακή πύλη να είναι κατάλληλη και για πρόσβαση από φορητές συσκευές	Οι ιστοσελίδες του SearchCulture.gr ανταποκρίνονται και προσαρμόζονται ανάλογα με την ανάλυση της οθόνης του χρήστη. Έτσι το SearchCulture.gr μπορεί να χρησιμοποιηθεί και από φορητές συσκευές (κινητά τηλέφωνα, tablets).
6	Δημόσια Διαδικτυακή Πύλη - Δίγλωσσο περιβάλλον (Ελληνικά-Αγγλικά) σε στατικό αλλά και δυναμικό περιεχόμενο.	Η διαδικτυακή πύλη του SearchCulture.gr είναι δίγλωσση, υποστηρίζοντας Ελληνικά και Αγγλικά τόσο στο στατικό όσο και στο δυναμικό περιεχόμενο (ονομασίες φορέων, φίλτρα αναζήτησης, ομογενοποιημένοι τύποι τεκμηρίων κλπ.)
7	Αξιοποίηση και διακριτή προβολή του σημασιολογικού εμπλουτισμού που πραγματοποιεί το ΕΚΤ	Ο σημασιολογικός εμπλουτισμός που πραγματοποιείται στο SearchCulture.gr με όρους λεξιλογίων του Semantics.gr και σύμφωνα με τους κανόνες αντιστοίχισης που ορίστηκαν στο εργαλείο εμπλουτισμού/ομογενοποίησης του Semantics.gr ανά φορέα, αξιοποιούνται πλήρως αναβαθμίζοντας την προβολή του περιεχομένου και τις δυνατότητες πολυγλωσσικής αναζήτησης και πλοήγησης του SearchCulture.gr. Ο χρήστης μπορεί να πλοηγηθεί αλλά και να φιλτράρει τα αποτελέσματα βάσει των ομογενοποιημένων τύπων τεκμηρίων, ενώ το αποτέλεσμα του σημασιολογικού εμπλουτισμού εμφανίζεται με διακριτό τρόπο στην καρτέλα τεκμηρίου ξεχωριστά από την πρωτότυπη τεκμηρίωση.
8	Σελίδα φορέων/συλλογών	Το SearchCulture.gr διαθέτει μία ξεχωριστή σελίδα παρουσίασης των φορέων και των συλλογών τους που φιλοξενούνται στην διαδικτυακή πύλη. Ο χρήστης μπορεί να προηγηθεί στα τεκμήρια της κάθε συλλογής όπως παρουσιάζονται στο SearchCulture.gr
9	Σελίδα ΕΚΤ Τύπων	Το SearchCulture.gr διαθέτει μία ξεχωριστή σελίδα περιήγησης στο περιεχόμενο μέσω όλων των ΕΚΤ Τύπων στους οποίους ταξινομήθηκαν τα τεκμήρια μέσω του σημασιολογικού εμπλουτισμού

Πίνακας IV. Οι Λειτουργικές Απαιτήσεις για το Σύστημα Συσώρευσης και το Διαχειριστικό Περιβάλλον του SearchCulture.gr

	Απαίτηση	Περιγραφή
1	Συγκομιδή μεταδεδομένων από πηγές	Ενοποίηση με το Σύστημα Συγκομιδής Περιεχομένου Harvester μέσω της RESTful διεπαφής του.
2	Διαχειριστικό περιβάλλον - Σύστημα διαχείρισης Φορέων - Αποθετηρίων.	Για κάθε αποθετήριο/συλλογή αποθηκεύονται μόνιμα οι βασικές πληροφορίες που απαιτούνται για την συγκομιδή (π.χ. OAI-PMH URL, τύποι μεταδεδομένων που υποστηρίζονται, OAI-PMH σύνολα) αλλά και τον έλεγχο συμμόρφωσης με προδιαγραφές διαλειτουργικότητας. Υπάρχει γραφική διεπαφή χρήστη για το σύστημα διαχείρισης Φορέων – Αποθετηρίων μέσα στο διαχειριστικό περιβάλλον του συσσωρευτή.
3	Υποσύστημα εισαγωγής /ευρετηρίασης και αναζήτησης EDM	Το υποσύστημα αυτό αποθηκεύει μεταδεδομένα που βρίσκονται σε μορφή EDM—internal [11]. Εκτός από την αποθήκευση υποστηρίζει ευρετηρίαση των εγγραφών και παρέχει βασικές υπηρεσίες αναζήτησης.

μεταδεδομένων
(EDM Storage System)

-
- 4** **Ροή εισαγωγής μεταδεδομένων ως EDM-internal από μεταδεδομένα που συλλέχθηκαν σε EDM [7], ESE[6] ή συνδυασμό ESE και DC-DS-XML [11] .**
- Υποστηρίζεται ροή εισαγωγής μεταδεδομένων ενός αποθετηρίου (συλλογής) που περιλαμβάνει την ανάκτηση μεταδεδομένων του αποθετηρίου του φορέα από το Σύστημα Harvester, τον μετασχηματισμό τους σε EDM-internal μορφή, τον σημασιολογικό εμπλουτισμό τους σύμφωνα με κανόνες αντιστοίχισης που ορίστηκαν στο Semantics.gr ειδικά για το αποθετήριο και την αποθήκευσή τους στο EDM Storage System. Η βασική ροή υποστηρίζει εγγενώς μεταδεδομένα σε αρχική μορφή EDM-external [7] ή ESE[6] ή ESE σε συνδυασμό με DC-DS-XML [11]. Υποστηρίζεται ο μετασχηματισμός μεταδεδομένων από αυτές τις μορφές σε EDM-internal προκειμένου να αποθηκευτούν στο EDM-Storage System.
-
- 5** **Διαχειριστικό περιβάλλον Διαχείρισης διαδικασιών συγκομιδής/ εισαγωγής μεταδεδομένων.**
- Στοιχεία για διαδικασίες εισαγωγής που πραγματοποιήθηκαν φυλάσσονται 1) για λόγους ιστορικότητας 2) για λόγους καταγραφής σφαλμάτων (απαραίτητο τόσο για αποσφαλμάτωση όσο και για δημιουργία αναφοράς προς τον φορέα) και 3) για να είναι εφικτό το incremental harvesting που υποστηρίζεται από το πρωτόκολλο OAI-PMH. Συγκεκριμένα η παράμετρος 'OAI from' της συγκομιδής της επόμενης διαδικασίας εισαγωγής υπολογίζεται αυτόματα βάσει της ημερομηνίας διεξαγωγής της τελευταίας επιτυχούς αντίστοιχης διαδικασίας.
- Το διαχειριστικό περιβάλλον του SearchCulture.gr διαθέτει γραφική διεπαφή χρήστη για το σύστημα πυροδότησης και διαχείρισης διαδικασιών συγκομιδής - εισαγωγής μεταδεδομένων. Ο εξουσιοδοτημένος χρήστης μπορεί να επιλέξει να πραγματοποιηθεί μια νέα συγκομιδή από το Σύστημα Harvester ή να χρησιμοποιήσει ως πηγή μεταδεδομένα που έχει ήδη συλλέξει και αποθηκεύσει στο παρελθόν το Σύστημα Harvester.
-
- 6** **Βασική ροή εισαγωγής ψηφιακών αρχείων ως εικόνες προεπισκόπησης.**
- Προκειμένου η μηχανή αναζήτησης του SearchCulture.gr να είναι φιλική στο χρήστη τα αποτελέσματα συνοδεύονται από εικόνες προεπισκόπησης. Το σύστημα συσσωρευτή δίνει την δυνατότητα μαζικής δημιουργίας και αποθήκευσης αρχείων προεπισκόπησης για τις εγγραφές μεταδεδομένων που έχουν εισαχθεί. Παρέχονται τρεις εναλλακτικοί τρόποι εισαγωγής εικόνων. Τα αρχεία συλλέγονται και διατίθενται από το Σύστημα Συσσωρευτή Harvester το οποίο είναι υπεύθυνο για την διάθεσή τους ως εικόνες προεπισκόπησης.
-
- 7** **Διαχειριστικό περιβάλλον - Διαχείριση διαδικασιών συγκομιδής / εισαγωγής αρχείων προεπισκόπησης.**
- Στοιχεία για διαδικασίες εισαγωγής αρχείων προεπισκόπησης που πραγματοποιήθηκαν φυλάσσονται για λόγους ιστορικότητας και για λόγους καταγραφής σφαλμάτων (απαραίτητο τόσο για αποσφαλμάτωση όσο και για δημιουργία αναφοράς προς τον φορέα). Το διαχειριστικό περιβάλλον του SearchCulture.gr διαθέτει γραφική διεπαφή χρήστη για το σύστημα πυροδότησης και διαχείρισης διαδικασιών συγκομιδής - εισαγωγής αρχείων προεπισκόπησης.
-
- 8** **HTTP διάθεση των αρχείων προεπισκόπησης που αποθηκεύτηκαν στον Συσσωρευτή.**
- Ένα αρχείο προεπισκόπησης που αντιστοιχεί σε μια αποθηκευμένη EDM εγγραφή, διατίθεται μέσω URL το οποίο περιέχει το αναγνωριστικό της αντίστοιχης EDM εγγραφής. Έχει οριστεί μια σταθερή ένα-προς-ένα αντιστοίχιση από το αναγνωριστικό της EDM εγγραφής του τεκμηρίου στο URL της εικόνας προεπισκόπησης του.
-

9	Διαχειριστικό περιβάλλον - Φόρμα δημιουργίας/τροποποίησης EDM εγγραφών	Παρέχεται η δυνατότητα δημιουργίας ή τροποποίησης μιας EDM εγγραφής τεκμηρίου από τον εξουσιοδοτημένο χρήστη μέσω δυναμικής φόρμας που παρέχεται από το διαχειριστικό περιβάλλον. Η φόρμα με φιλικό προς τον χρήστη και δομημένο τρόπο μοντελοποιεί πλήρως έναν οποιονδήποτε EDM γράφο.
10	Διαχειριστικό περιβάλλον - Σελίδες για αυτόνομη διαχείριση EDM σημασιολογικών οντοτήτων edm:Agent, skos:Concept, edm:Place και edm:Timespan	Στο διαχειριστικό περιβάλλον, υπάρχουν ξεχωριστές σελίδες για κάθε EDM σημασιολογική κλάση (edm:Agent, skos:Concept, edm:Place και edm:Timespan [7]) που δίνουν την δυνατότητα στον εξουσιοδοτημένο χρήστη να αναζητά, να επεξεργάζεται, να δημιουργεί και να διαγράφει σημασιολογικές οντότητες της αντίστοιχης κλάσης.
11	Διαχειριστικό περιβάλλον - Ανέβασμα ψηφιακών αρχείων μέσα από την φόρμα δημιουργίας/τροποποίησης EDM εγγραφών.	Ο εξουσιοδοτημένος χρήστης μπορεί να ανεβάσει ψηφιακά αρχεία για μία EDM εγγραφή. Στα αρχεία αντιστοιχείται μοναδικό ενεργό URL το οποίο περιλαμβάνει το αναγνωριστικό της εγγραφής του τεκμηρίου και θα εξυπηρετείται μόνιμα από το σύστημα.
12	Διαχειριστικό περιβάλλον - Δυνατότητα χειροκίνητου σημασιολογικού εμπλουτισμού μιας EDM εγγραφής μέσω της φόρμας δημιουργίας/τροποποίησης EDM εγγραφών	Μέσα από την φόρμα δημιουργίας/τροποποίησης EDM εγγραφών του διαχειριστικού περιβάλλοντος, ο εξουσιοδοτημένος χρήστης μπορεί, για εκείνα τα πεδία που - σύμφωνα με το EDM μοντέλο – υποστηρίζουν σημασιολογικές αναφορές, να επιλέγει μεταξύ συμπλήρωσης περιγραφικής τιμής ή συσχέτισης με σημασιολογική οντότητα. Στην δεύτερη περίπτωση, μπορεί εύκολα να αναζητήσει και να επιλέξει μία ήδη αποθηκευμένη σημασιολογική οντότητα ή να δημιουργήσει επί τόπου μία νέα ή να επιλέξει από κάποιο λεξιλόγιο που φιλοξενείται στο Σύστημα Λεξιλογίων του EKT, Semantics.gr.
13	Διαχειριστικό περιβάλλον - Διαχείριση περιεχομένου ανά αποθετήριο (συλλογή)	Ο εξουσιοδοτημένος χρήστης μπορεί να διαχειρίζεται το περιεχόμενο του συσσωρευτή ανά αποθετήριο (συλλογή) φορέα. Υπάρχουν λειτουργίες διαγραφής του συνόλου των μεταδεδομένων, επανάληψης ευρετηρίασης και διαγραφής του συνόλου των αρχείων προεπισκόπησης.
14	Διαχειριστικό περιβάλλον - Ορισμός δεσμευμένων λεξιλογίων ανά Φορέα.	Ο εξουσιοδοτημένος χρήστης μπορεί στην καρτέλα κάθε φορέα να ορίζει ποια δεσμευμένα λεξιλόγια χρησιμοποιεί ο φορέας ανά σημασιολογική κατηγορία και ποια πεδία αφορά. Η πληροφορία αυτή χρησιμοποιείται για των έλεγχο των αντίστοιχων προδιαγραφών από το Σύστημα Validator.
15	Διαχειριστικό περιβάλλον - Προεπισκόπηση των τεκμηρίων που εισήχθησαν στα πλαίσια μιας διαδικασίας εισαγωγής	Ο εξουσιοδοτημένος χρήστης μπορεί να βλέπει το σύνολο των τεκμηρίων που εισήχθησαν σε μια διαδικασία με φιλικό τρόπο (σελιδοποιημένα και συνοδευόμενα από τυχόν εικόνες προεπισκόπησης και με δυνατότητα αναζήτησης).
16	Διαχειριστικό περιβάλλον - Παραμετροποίηση της διαδικτυακής πύλης.	Ο εξουσιοδοτημένος χρήστης μπορεί να αλλάζει τα πεδία και την σειρά εμφάνισής τους τόσο στην σελίδα παρουσίασης τεκμηρίων (καρτέλα τεκμηρίου) όσο και στην σελίδα αποτελεσμάτων της μηχανής αναζήτησης. Μπορεί επίσης να παραμετροποιεί την φόρμα επεξεργασίας τεκμηρίου: να ορίζει ποια EDM πεδία θα υποστηρίζουν σημασιολογικές αναφορές, για ποιά θα χρησιμοποιείται πλαίσιο κειμένου (text area) αντί απλού πεδίου και, τέλος, ποιών οι τιμές θα υπόκεινται σε απόπειρα μετάφρασης.

<p>17 Διαχειριστικό περιβάλλον - Διαδικασίες εισαγωγής (μεταδεδομένων και εικόνων προεπισκόπησης): Καταγραφή αποτυχιών σε πίνακα και προβολή από το UI, δυνατότητα τερματισμού μιας εξ εξελίξει διαδικασίας</p>	<p>Ο εξουσιοδοτημένος χρήστης μπορεί αν θέλει να δώσει ένα κατώφλι (threshold) του αριθμού αποδεκτών αποτυχιών πριν αποτύχει και τερματιστεί ολόκληρη η διαδικασία εισαγωγής. Αν υπάρχουν αποτυχίες, αυτές αποθηκεύονται. Ο χρήστης μπορεί να δει αναλυτική αναφορά μέσα από το διαχειριστικό περιβάλλον. Επίσης ο χρήστης μπορεί να σταματήσει σε οποιαδήποτε στιγμή μια διαδικασία που βρίσκεται σε εξέλιξη.</p>
<p>18 Διαχειριστικό περιβάλλον - Διασύνδεση με το Σύστημα Λεξιλογίων Semantics.gr</p>	<p>Μέσα από το διαχειριστικό περιβάλλον του SearchCulture.gr, στην καρτέλα διαχείρισης περιεχομένου αποθετηρίου υπάρχει σύνδεσμος με το πάτημα του οποίου το αποθετήριο εγγράφεται αυτόματα στο Σύστημα Λεξιλογίων Semantics.gr.</p>
<p>19 Δυνατότητα μαζικού αυτόματου σημασιολογικού εμπλουτισμού/ομογενοποίησης ανά αποθετήριο (συλλογή) για συγκεκριμένα πεδία ως προς συγκεκριμένα σημασιολογικά λεξιλόγια, βάσει κανόνων αντιστοίχισης.</p>	<p>Για τα αποθετήρια φορέων που υπόκεινται σε σημασιολογικό εμπλουτισμό (το οποίο ορίζεται μαζί με σχετικές παραμέτρους στην καρτέλα του αποθετηρίου), ο εξουσιοδοτημένος χρήστης μπορεί να πυροδοτήσει την διαδικασία αυτόματου σημασιολογικού εμπλουτισμού. Η διαδικασία συνδέεται με το σύστημα Semantics.gr αντλώντας όλη την απαραίτητη πληροφορία (κανόνες αντιστοίχισης καθώς και όροι λεξιλογίων ως XML αναπαραστάσεις συμβατές με τις σημασιολογικές κλάσεις του EDM) και στην συνέχεια κάθε εγγραφή εμπλουτίζεται και επαναποθηκεύεται.</p>
<p>20 Υπηρεσία διάθεσης μεταδεδομένων για συγκομιδή από τρίτα συστήματα μέσω του πρωτοκόλλου OAI-PMH</p>	<p>Τα μεταδεδομένα που συσσωρεύονται στο SearchCulture.gr διατίθενται για συγκομιδή μέσω του πρωτοκόλλου OAI-PMH [5] από τρίτα συστήματα και συσσωρευτές, όπως η Europeana.</p>

3.2 Ανοικτή Δικτυακή Πύλη SearchCulture.gr

Η διαδικτυακή πύλη του SearchCulture.gr είναι ελεύθερα προσβάσιμη μέσα από την διεύθυνση <http://searchculture.gr>. Ένα μενού επιλογών επιτρέπει στον χρήστη να μεταβεί στην σελίδα αναζήτησης και στην σελίδα παρουσίασης των φορέων και των συλλογών, να επιλέξει γλώσσα και να επιστρέψει στην αρχική σελίδα. Στο κάτω μέρος εμφανίζονται πληροφορίες για το έργο, τους όρους χρήσης, την πνευματική ιδιοκτησία καθώς και στοιχεία επικοινωνίας.

3.2.1 Αρχική Σελίδα

Η αρχική σελίδα, η οποία παρουσιάζεται στο Σχήμα 20, περιλαμβάνει μεταξύ άλλων ένα μενού επιλογών, μία φόρμα αναζήτησης, ένα διαδραστικό νέφος δημοφιλών ομογενοποιημένων από το ΕΚΤ τύπων τεκμηρίων (που προέρχονται ένα ενιαίο λεξιλόγιο τύπων τεκμηρίων που δημιούργησε το ΕΚΤ στο Semantics.gr), δυναμικά διαγράμματα που απεικονίζουν ποιοτικά και ποσοτικά χαρακτηριστικά του περιεχομένου του συσσωρευτή και, τέλος, μια περιληπτική παρουσίαση ενός τεκμηρίου που επιλέγεται από το σύστημα τυχαία και ανανεώνεται σε ημερήσια βάση. Οι σελίδες του SearchCulture.gr ανταποκρίνονται και προσαρμόζονται ανάλογα με την ανάλυση της οθόνης του χρήστη. Έτσι το SearchCulture.gr μπορεί να χρησιμοποιηθεί και μέσω φορητών συσκευών (κινητά τηλέφωνα, tablets). Στο Σχήμα 21 παρουσιάζεται η αρχική σελίδα όπως φαίνεται μέσα από την οθόνη κινητού τηλεφώνου.

Search Culture.gr

Ενιαία Αναζήτηση σε ανοικτό ψηφιακό περιεχόμενο

ΕΚΤ ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΠΑΡΑΜΟΝΕΣ ΚΑΙ ΕΠΙΣΤΑΣΕΩΝ ΔΟΚΙΜΑΤΩΝ Ε.Σ.Τ.Κ.Ε. SaaS Services aggregator

Αρχική σελίδα Αναζήτηση Φορείς και συλλογές ΕΚΤ τύποι τεκμηρίων Το έργο Ελληνικά | English

Πληκτρολογήστε λέξη ή φράση Σε όλα τα πεδία Αναζήτηση Καθαρισμός

Επιλέξτε ΕΚΤ τύπους Επιλέξτε άδεια χρήσης ψηφιακού περιεχομ... Αναζήτηση σε όλες τις συλλογές

Εξερευνήστε 159178 τεκμήρια ποικίλων τύπων από 42 φορείς περιεχομένου...

Τμήμα βιβλίου Μηνύημα Αρχαιολογικός χώρος Επιγραφή Νόμισμα Θραύσμα Σελίδα βιβλίου Κεραμικό αντικείμενο Χαρακτικό έργο Εργαλεία και εξοπλισμός Ζωγραφική Αγγείο Διακόσμηση εγγράφου Κάσμημα Χειρόγραφο Σκεύος Έγγραφο Σχέδιο/Σκίτσο Πρωτότυπο Ανάγλυφο Ενδυμασία Φωτογραφία Αλληλογραφία Αρχαιολογικό αντικείμενο Βιβλίο Ειδώλιο Εκτύπωση Άρθρο Οικιακός εφπίλαμος Νομική υπόθεση και σημείωμα Φωτογραφική διαφάνεια Φωτοσχηματισμός Έργο θρησκευτικής τέχνης Αρχιτεκτονικό μέλος Προσωπικά αντικείμενα (αξσουάρ) Περιοδική έκδοση Αντίγραφο

Η ταξινόμηση του περιεχομένου σε ενιαίο λεξιλόγιο τύπων πραγματοποιήθηκε από το ΕΚΤ

54 συλλογές / αποθετήρια

Δείτε όλους τους φορείς και τις συλλογές τους εδώ

Αποθετήριο Υπηρεσίας Συντήρησης Μνημείων Ακρόπολης

Αποθετήριο Ε.Δ.Κ. - Ένωση Δημοκρατικού Κέντρου

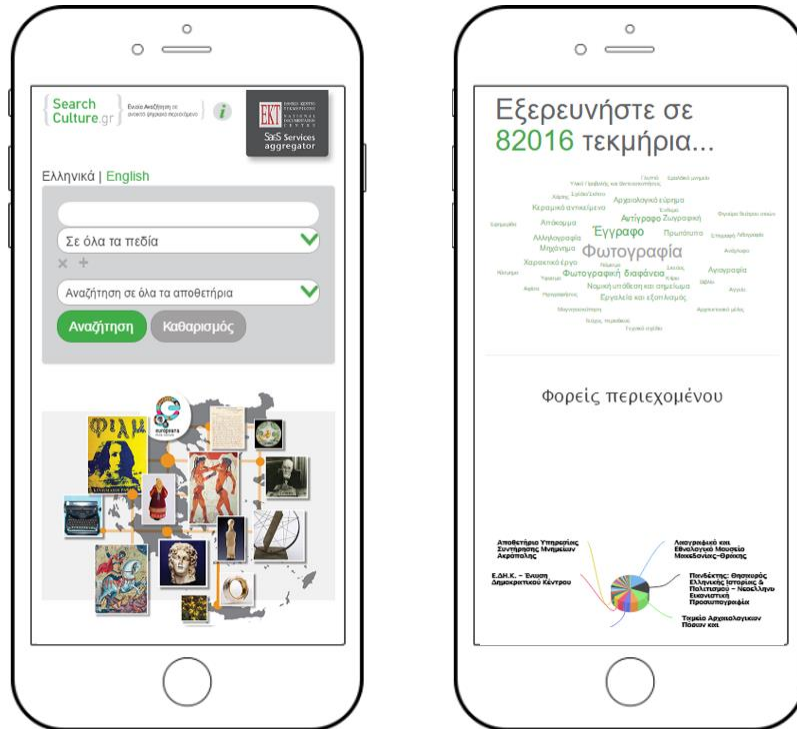
Αποθετήριο Αμερικανικής Σχολής Κλασικών Σπουδών

Αποθετήριο Ιστορικής & Εθνολογικής Εταιρείας της Ελλάδος (ΕΙΜ)

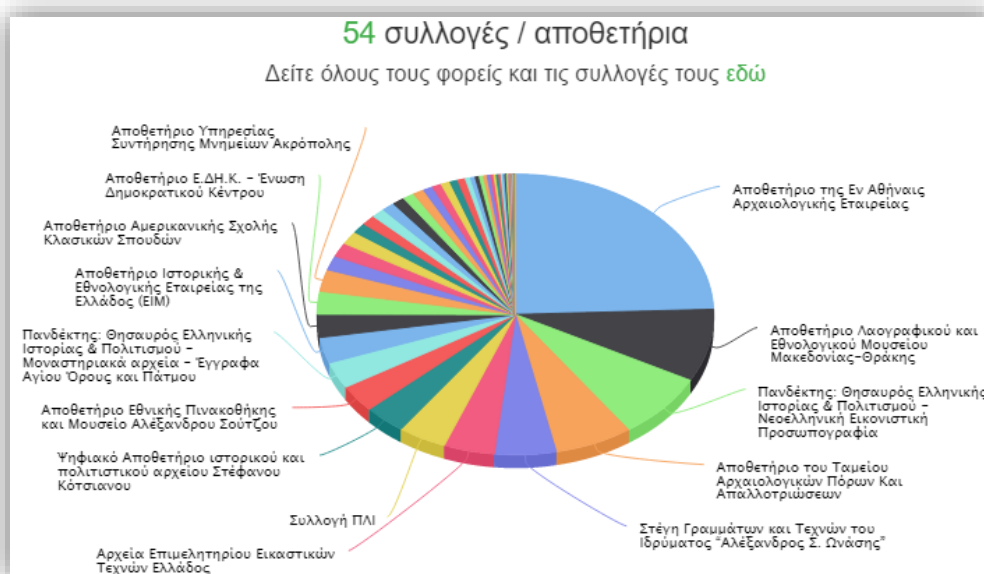
Αποθετήριο της Εν Αθηνών Αρχαιολογικής Εταιρείας

Σχήμα 20. Η διαδικτυακή πύλη του SearchCulture.gr - Αρχική σελίδα

Τα διαγράμματα (πίτες) εμφανίζονται εναλλασσόμενα (carousel widget), και περιλαμβάνουν μία πίτα με τις μεγαλύτερες, σε αριθμό τεκμηρίων, συλλογές (Σχήμα 22), ένα ενημερωτικό πλαίσιο για τον σημασιολογικό εμπλουτισμό που διεξάγει το ΕΚΤ (Σχήμα 23) και ένα ενημερωτικό πλαίσιο για τους ελέγχους ποιότητας/πιστοποίηση που διεξάγει το ΕΚΤ. (Σχήμα 24). Το νέφος ομογενοποιημένων τύπων, και η πίτα με τις συλλογές είναι διαδραστικά στοιχεία, επιτρέποντας στον χρήστη να πλοηγηθεί στα τεκμήρια συγκεκριμένου ΕΚΤ τύπου, συγκεκριμένης συλλογής και συγκεκριμένου Euroreana τύπου, αντίστοιχα, και αποτελούν έτσι εναλλακτικούς τρόπους εξερεύνησης περιεχομένου πέρα από τον κύριο τρόπο που προσφέρεται από την φόρμα αναζήτησης.



Σχήμα 21. Η αρχική σελίδα του SearchCulture.gr όπως φαίνεται από κινητά τηλέφωνα.



Σχήμα 22. Διαδραστική πίτα με τις μεγαλύτερες συλλογές

Περισσότερα από 155973 τεκμήρια έχουν εμπλουτιστεί και ομογενοποιηθεί από το EKT

Το EKT εμπλουτίζει και ομογενοποιεί το περιεχόμενο που συσσωρεύει βάσει ενός πρότυπου διγλωσσίου ιεραρχικού **λεξιλογίου τύπων** που δημιούργησε.

Τα τεκμήρια εμπλουτίστηκαν με ένα νέο πεδίο «EKT Τύπος» που περιλαμβάνει τις αναφορές στους όρους του λεξιλογίου. Έτσι, το περιεχόμενο του SearchCulture.gr ταξινομήθηκε σε ένα συμπαγές και ισορροπημένο σύνολο τύπων αποκτώντας σημασιολογική συνοχή με θετικό αντίκτυπο στην ευρεσιμότητα, στην παρουσίαση και στην πολυγλωσσικότητά του.

Semantics.gr



Σχήμα 23. Ενημερωτικό πλαίσιο για τον σημασιολογικό εμπλουτισμό που διεξάγει το EKT

159178 τεκμήρια έχουν ελεγχθεί και πιστοποιηθεί από το EKT

Το EKT όρισε μια σειρά από προδιαγραφές διαλειτουργικότητας και περιεχομένου βάσει των οποίων ελέγχει και πιστοποιεί σε καθημερινή βάση την ποιότητα των μεταδεδομένων και των ψηφιακών αρχείων που συσσωρεύει καθώς και τις υπηρεσίες διαχείρισης περιεχομένου των φορέων.

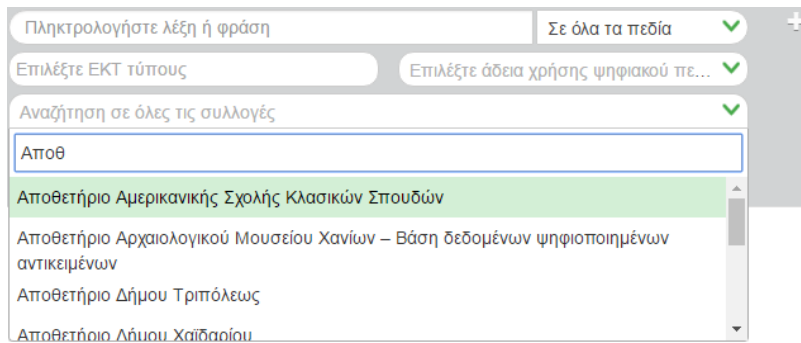
Οι προδιαγραφές αφορούν σε πρότυπα μεταδεδομένων, ψηφιοποίηση υλικού και πρωτόκολλα διαλειτουργικότητας και είναι αποτέλεσμα μελέτης η οποία βασίστηκε στην τεχνογνωσία του EKT, σε καλές πρακτικές από το διεθνές περιβάλλον και σε καθιερωμένα πρότυπα.



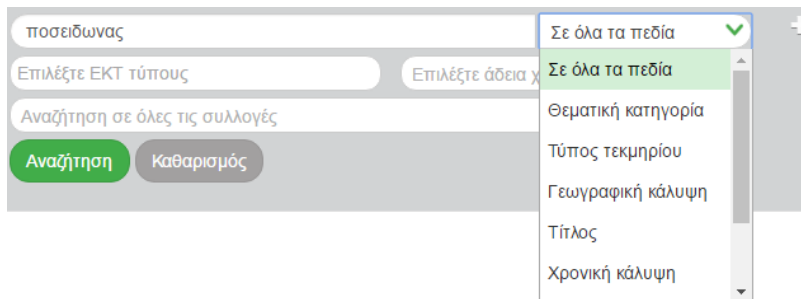
Σχήμα 24. Ενημερωτικό πλαίσιο για τους ελέγχους ποιότητας/πιστοποίηση που διεξάγει το EKT.

3.2.2 Αναζήτηση

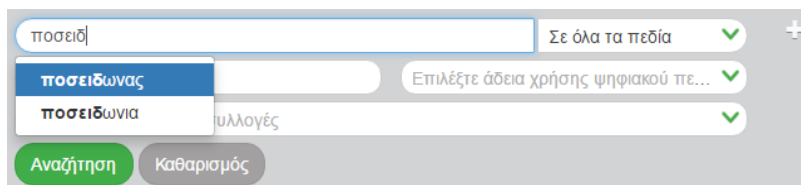
Η φόρμα αναζήτησης εμφανίζεται τόσο στην αρχική σελίδα όσο και στην σελίδα αναζήτησης (μενού επιλογών → «Αναζήτηση»). Η φόρμα επιτρέπει στον χρήστη να περιορίσει την αναζήτηση του σε συγκεκριμένη μόνο συλλογή/αποθετήριο (Σχήμα 25). Εάν ο χρήστης επιλέξει αποθετήριο δεν χρειάζεται να πληκτρολογήσει όρο αναζήτησης και μπορεί να δει έτσι το σύνολο των τεκμηρίων του συγκεκριμένου αποθετηρίου. Ο χρήστης μπορεί, αν θέλει, να συγκεκριμενοποιήσει την αναζήτηση σε συγκεκριμένα πεδία μεταδεδομένων (Σχήμα 26-α), όπως τύπος τεκμηρίου EKT, Europeana τύπος, πρωτότυπος τύπος τεκμηρίου, δημιουργός/συντονιστής, θεματική κατηγορία, χρονική και γεωγραφική κάλυψη. Το πεδίο εισαγωγής όρου διαθέτει βοηθητικό μηχανισμό αυτόματης συμπλήρωσης τιμών: καθώς ο χρήστης πληκτρολογεί εμφανίζεται και ανανεώνεται λίστα με πιθανές τιμές (Σχήμα 26-β). Το ίδιο ισχύει και για την λίστα επιλογής αποθετηρίου/συλλογής. Ο χρήστης μπορεί να πραγματοποιήσει σύνθετη αναζήτηση εισάγοντας περισσότερους από έναν όρους, όπως φαίνεται στο Σχήμα 27. Πατώντας το κουμπί «+» προστίθεται ένα επιπλέον πεδίο για εισαγωγή όρου ενώ πατώντας το κουμπί «-» αφαιρείται το αντίστοιχο πεδίο.



Σχήμα 25. Προαιρετική επιλογή αποθετηρίου

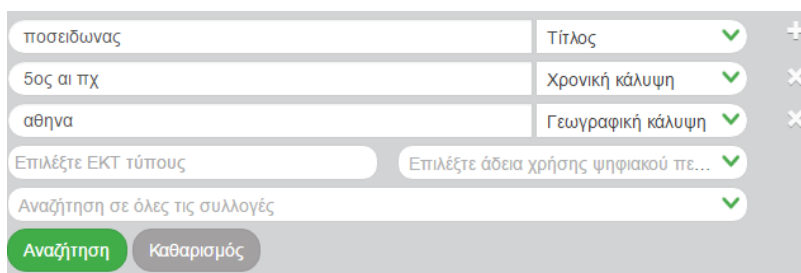


(α)



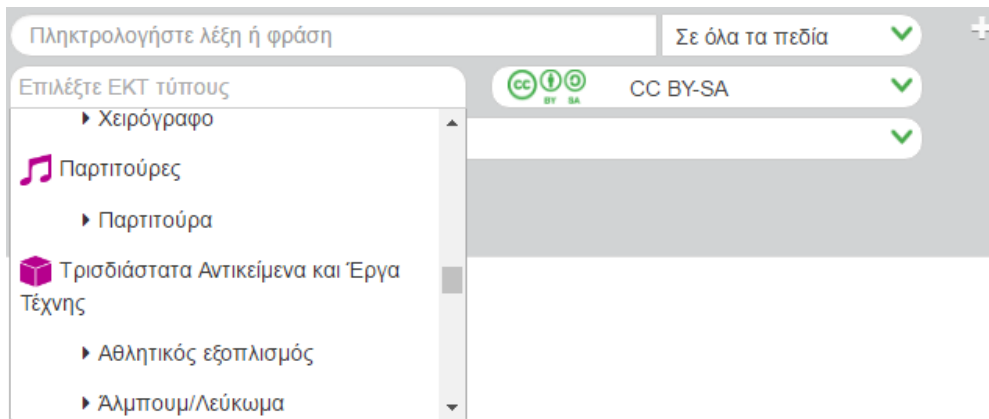
(β)

Σχήμα 26. Συγκεκριμενοποίηση αναζήτησης (α) και αυτόματη συμπλήρωση όρου αναζήτησης (β)

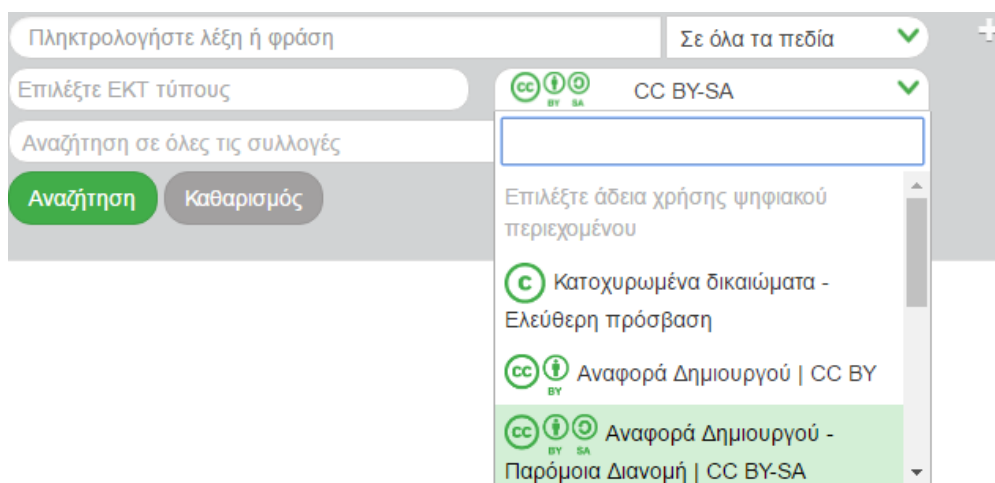


Σχήμα 27. Σύνθετη αναζήτηση

Η φόρμα αναζήτησης δίνει την δυνατότητα στον χρήστη να επιλέξει έναν ή περισσότερους ΕΚΤ Τύπους τεκμηρίων μέσα από μενού επιλογής το οποίο παρουσιάζει τους ΕΚΤ τύπους ιεραρχικά και διαθέτει βοηθητικό μηχανισμό αναζήτησης (Σχήμα 28) . Επιπλέον ο χρήστης μπορεί να επιλέξει άδεια χρήσης ψηφιακού περιεχομένου μέσα από λίστα επιλογής, όπως φαίνεται στο Σχήμα 29.

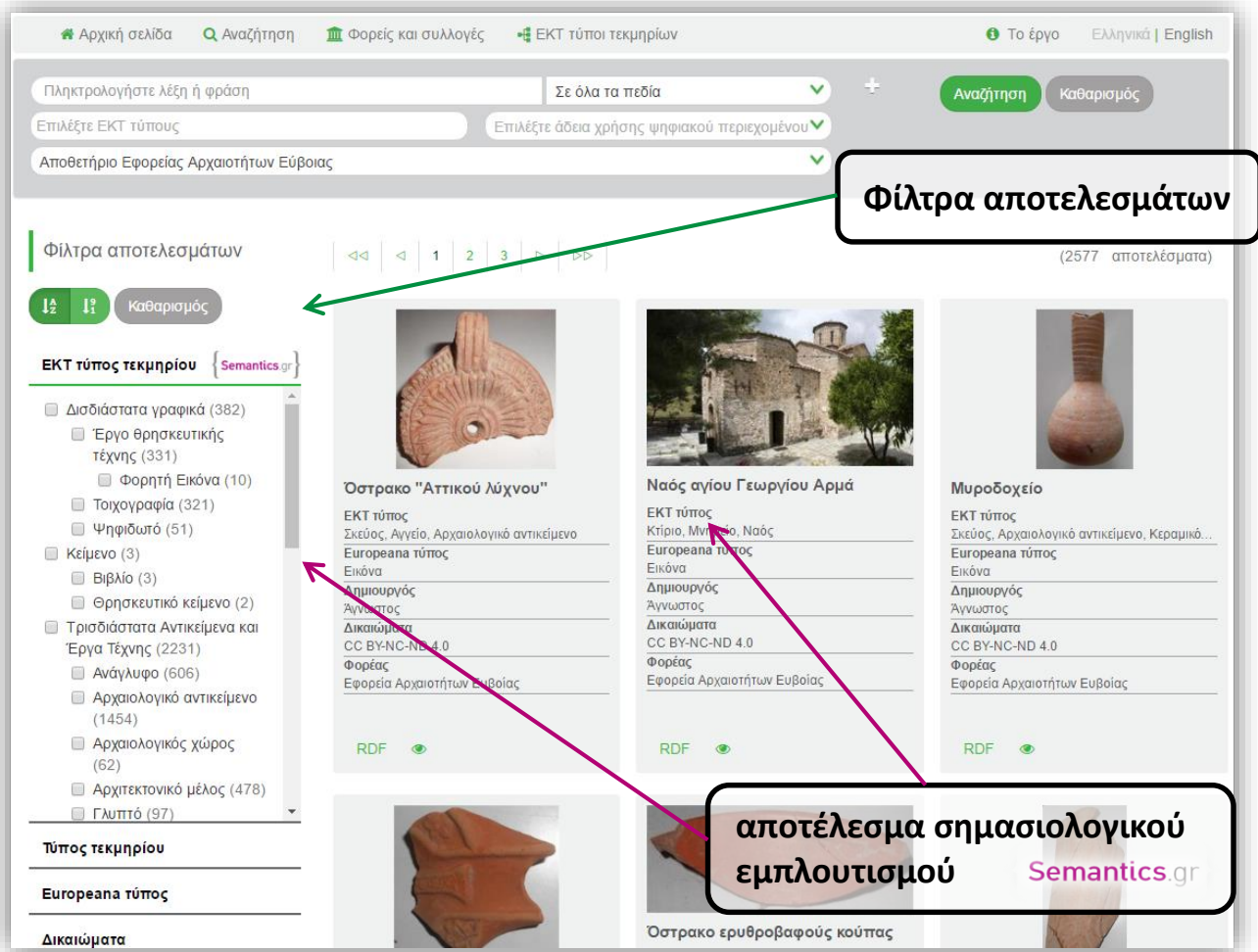


Σχήμα 28. Προαιρετική επιλογή ΕΚΤ Τύπου

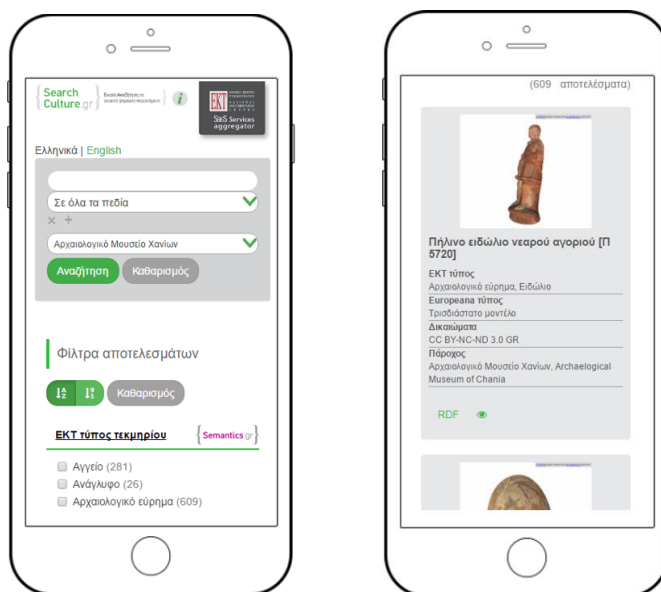


Σχήμα 29. Προαιρετική επιλογή άδειας χρήσης ψηφιακού περιεχομένου

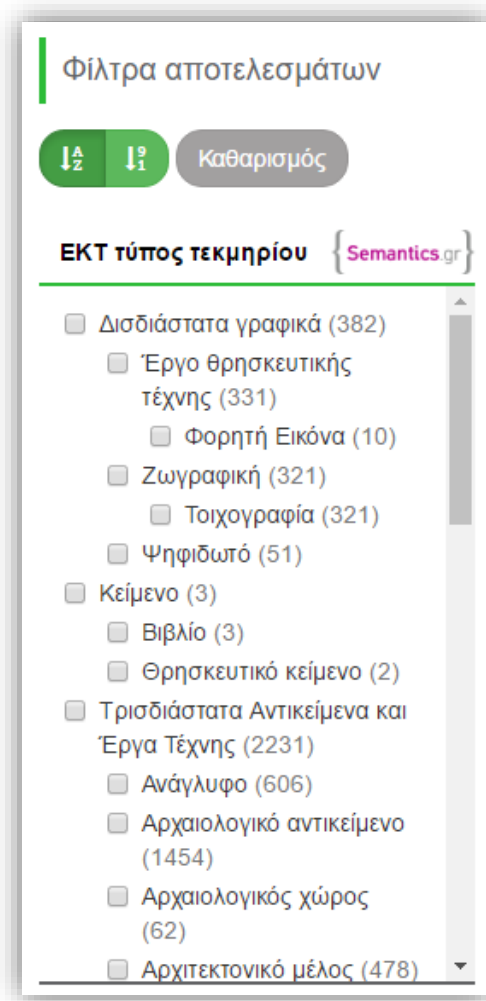
Τα αποτελέσματα της αναζήτησης εμφανίζονται σελιδοποιημένα ανά 12 τεκμήρια. Για κάθε τεκμήριο εμφανίζεται μία εικόνα προεπισκόπησης καθώς και οι τιμές επιλεγμένων πεδίων, όπως φαίνεται στο Σχήμα 30 από οθόνη υπολογιστή και στο Σχήμα 31 από οθόνη κινητής συσκευής.





Σχήμα 30. Αποτελέσματα αναζήτησης






Σχήμα 31. Αποτελέσματα αναζήτησης όπως φαίνονται από οθόνη κινητού τηλεφώνου



Σχήμα 32. Φιλτράρισμα αποτελεσμάτων (Facet Queries)

Η αναζήτηση τεκμηρίων διαθέτει δυνατότητα φιλτραρίσματος των αποτελεσμάτων με μηχανισμό *facet queries* [21] σε συγκεκριμένα πεδία συμπεριλαμβανομένων του ομογενοποιημένου τύπου τεκμηρίου του ΕΚΤ, του πρωτότυπου τύπου τεκμηρίου της συλλογής, του Europeana τύπου, της θεματικής κατηγορίας, του παρόχου (συλλογής) και των δικαιωμάτων διάθεσης ψηφιακού περιεχομένου. Οι τιμές του φίλτρου «ΕΚΤ τύπος τεκμηρίου» εμφανίζονται ιεραρχικά. Το φιλτράρισμα γίνεται μέσω ενός βοηθητικού πλαισίου που συνοδεύει τα αποτελέσματα μιας αναζήτησης στο αριστερό τμήμα της σελίδας (Σχήμα 32). Στα φίλτρα αναζήτησης υπάρχουν κουμπιά με τα οποία ο χρήστης μπορεί να επιλέξει αν οι τιμές των φίλτρων θα εμφανίζονται σε αλφαβητική σειρά («») ή σε φθίνουσα σειρά δημοφιλίας («»).

3.2.3 Καρτέλα τεκμηρίου

Για κάθε αποτέλεσμα ο χρήστης μπορεί να επιλέξει να δει λεπτομέρειες στη Καρτέλα Τεκμηρίου σε ξεχωριστή σελίδα, πατώντας τον τίτλο (Σχήμα 33 από οθόνη υπολογιστή, Σχήμα 34 από οθόνη κινητής συσκευής), ή σε ξεχωριστό παράθυρο, πατώντας το κουμπί «» (Σχήμα 35). Η καρτέλα τεκμηρίου σε ξεχωριστό παράθυρο δίνει στον χρήστη την δυνατότητα εύκολης πλοήγησης («», «») στα αποτελέσματα μιας σελίδας αποτελεσμάτων αναζήτησης.

The image shows a digital record card for an archaeological object. The card includes a photograph of a bronze votive offering depicting two young athletes. The record is structured as follows:

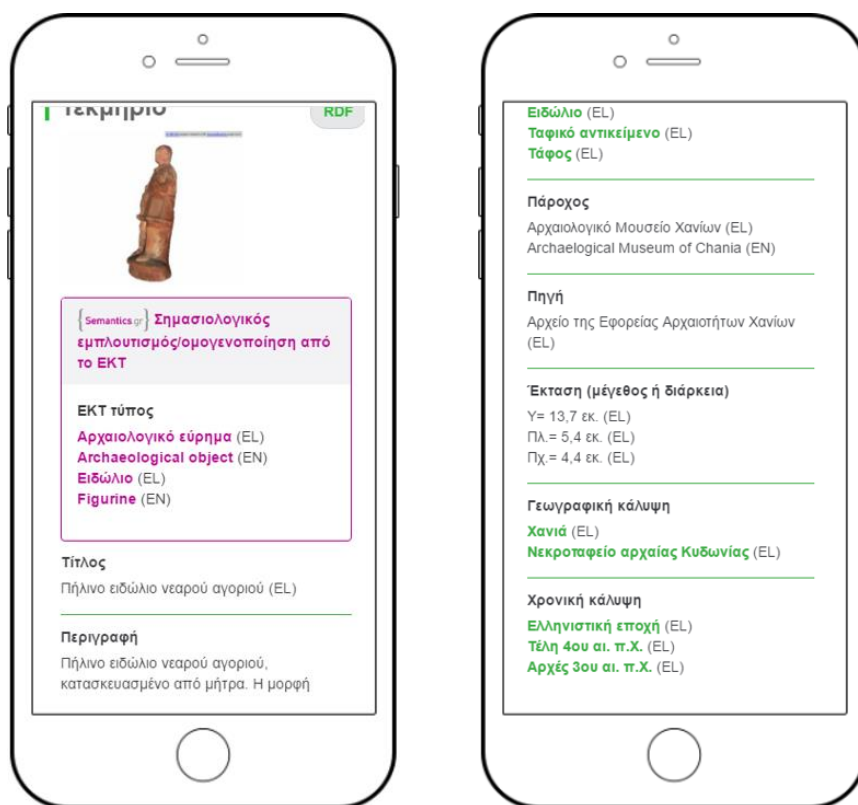
- Header:** 'Τεκμήριο' (Record) and 'RDF' icon with Creative Commons license icons (CC BY NC SA).
- Left Column:**
 - Download icon: 'κατεβάστε το αρχείο' (download the file) with a link to the repository.
 - Star icon: 'χρησιμοποιήστε' (use) with a link to the repository and CC BY-NC-SA 4.0 license.
 - Building icon: 'Το τεκμήριο παρέχεται από τον Φορέα: Εφορεία Αρχαιοτήτων Φωκίδος' (The record is provided by the Provider: Ephorate of Antiquities of Phocis).
 - Link icon: 'δείτε την καρτέλα τεκμηρίου' (view the record card) with a link to the repository.
- Right Column (Main Record):**
 - ΕΚΤ τύπος (EKT type):** Αρχαιολογικό αντικείμενο (EL), Archaeological object (EN), Γλυπτό (EL), Sculpture (EN). Callout: 'Διάθεση ως RDF Linked Data' (Available as RDF Linked Data).
 - Τίτλος (Title):** Ζεύγος αθλητών (EL), A pair of athletes (EN), Une paire d'athlètes (FR). Callout: 'αποτέλεσμα σημασιολογικού εμπλουτισμού' (result of semantic enrichment).
 - Δημιουργός (Creator):** Άγνωστος δημιουργός (EL), Unknown creator (EN), créateur inconnu (FR).
 - Περιγραφή (Description):** Χάλκινο ανάθημα που εικονίζει δύο νεαρούς αθλητές. Ο αριστερός μάλλον ήταν αθλητής στο άλμα εις μήκος, αφού κρατά αλτήρα. Έργο πιθανόν τοπικού εργαστηρίου, που βρέθηκε μαζί με άλλα στον αποθέτη της Ιεράς Οδού. (EL) Bronze votive offering which depicts two young athletes. The left one was probably an athlete in the long jump, since he holds a jumping weight. It was probably made by a local workshop, which was found together with other votive offerings of the Sacred Way. (EN) Bronze ex-voto figurine représentant deux athlètes. Celui de gauche était probablement un sauteur en longueur, car il détient un poids de saut. L'œuvre a été découverte avec d'autres travaux similaires dans le dépôt de la Voie Sacrée. (FR).
 - Τύπος (Type):** Ανάθημα (EL), Votive object (EN), votive object (FR).
 - Θέμα (Topic):** Metallurgy (EN), métallurgie (FR).
 - Πάροχος (Provider):** Εφορεία Αρχαιοτήτων Φωκίδος (EL).

Σχήμα 33. Καρτέλα τεκμηρίου

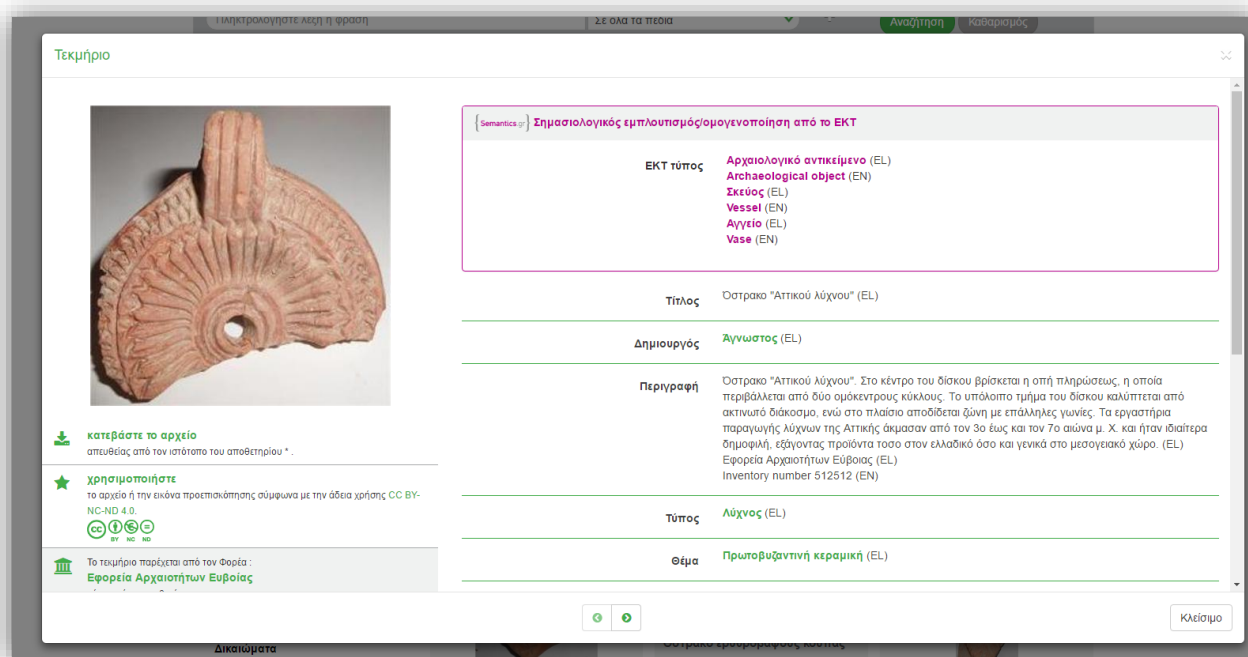
Η καρτέλα τεκμηρίου περιλαμβάνει μια εικόνα προεπισκόπησης και ένα σύνολο από γνωρίσματα και τιμές (το σύνολο των γνωρισμάτων που παρουσιάζονται καθώς και η σειρά με την οποία εμφανίζονται ορίζονται από τον διαχειριστή και αποτελούν μέρος των παραμέτρων του συστήματος που περιγράφονται στην Ενότητα 3.3.7). Οι ΕΚΤ Τύποι του τεκμηρίου, το αποτέλεσμα του σημασιολογικού εμπλουτισμού του ΕΚΤ, εμφανίζονται ξεχωριστά από τα υπόλοιπα μεταδεδομένα, σε διακριτικό πλαίσιο. Τιμές που περιλαμβάνουν αναφορά σε σημασιολογικές οντότητες (edm:Agent, skos:Concept, edm:Place και edm:Timespan), συμπεριλαμβανομένων των ΕΚΤ τύπων, εμφανίζονται ως σύνδεσμοι οι οποίοι οδηγούν στις αντίστοιχες σελίδες τεκμηρίωσης όρων των αντίστοιχων θησαυρών/λεξιλογίων.

Κάτω από την εικόνα προεπισκόπησης εμφανίζονται σύνδεσμοι προς τον ψηφιακό αρχείο (edm:isShownBy), προς την καρτέλα τεκμηρίου στο αποθετήριο του φορέα (edm:isShownAt), το όνομα του φορέα και της συλλογής/αποθετηρίου, καθώς και η άδεια χρήσης του ψηφιακού αρχείου.

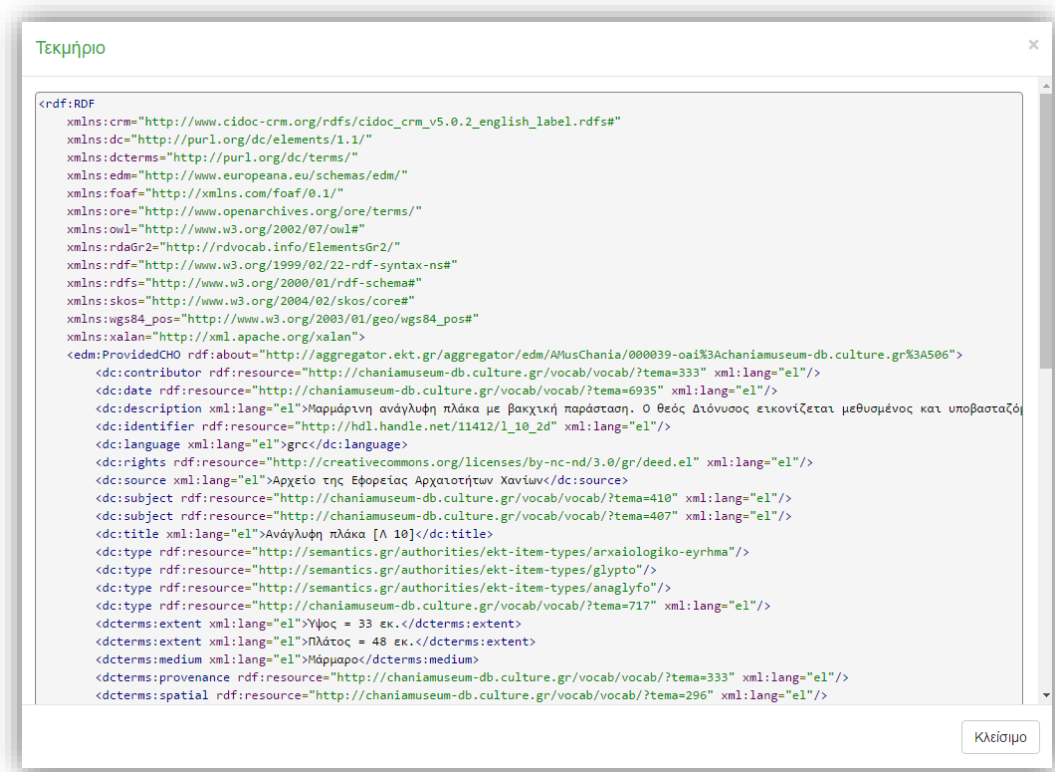
Τέλος, ο χρήστης μπορεί να δει την EDM RDF/XML αναπαράστασή του τεκμηρίου σε ξεχωριστό παράθυρο (Σχήμα 36) πατώντας το κουμπί «RDF».



Σχήμα 34. Καρτέλα τεκμηρίου όπως φαίνεται από οθόνη κινητού τηλεφώνου



Σχήμα 35. Καρτέλα τεκμηρίου σε παράθυρο – πλοήγηση σε προηγούμενα / επόμενα τεκμήρια



Σχήμα 36. RDF απεικόνιση τεκμηρίου - Διάθεση ως Linked Data

3.2.4 Φορείς/συλλογές

Το SearchCulture.gr διαθέτει μία ξεχωριστή σελίδα παρουσίασης των φορέων και των συλλογών τους που φιλοξενούνται στην διαδικτυακή πύλη (Σχήμα 37). Η σελίδα είναι προσβάσιμη από το μενού επιλογών επιλέγοντας «Φορείς/συλλογές».

Η σελίδα παρουσιάζει την λίστα με τους συνεργαζόμενους φορείς περιεχομένου (σελιδοποιημένη ανά δέκα φορείς). Ο χρήστης μπορεί να αναζητήσει και να επιλέξει συγκεκριμένο φορέα μέσα από δυναμική λίστα αναζήτησης. Κάτω από κάθε φορέα παρατίθενται οι συλλογές του. Οι πληροφορίες για τους φορείς και τις συλλογές περιλαμβάνουν τίτλο, διευθύνσεις στην διαδικτυακή σελίδα τόσο του φορέα όσο και του αποθετηρίου/συλλογής, περιγραφές και λογότυπα.

Ο χρήστης μπορεί να προηγηθεί στα τεκμήρια της κάθε συλλογής όπως παρουσιάζονται στο SearchCulture.gr, είτε επιλέγοντας τον σύνδεσμο «δείτε την συλλογή μέσα από το SC» ή επιλέγοντας συγκεκριμένο τύπο τεκμηρίου μέσα από όλους τους τύπους που διαθέτει η συγκεκριμένη συλλογή. Οι τύποι που εμφανίζονται είναι οι «ΕΚΤ τύποι τεκμηρίων», δηλαδή προέρχονται όρους του δίγλωσσου λεξιλογίου τύπων του ΕΚΤ στους οποίους κατηγοριοποιήθηκε το περιεχόμενο μέσα από τον εμπλουτισμό/ομογενοποίηση που διεξήγαγε το ΕΚΤ.

Αρχική σελίδα Αναζήτηση Φορείς και συλλογές ΕΚΤ τύποι τεκμηρίων Το έργο Ελληνικά | English

Φορείς 42 φορείς διαθέτουν συνολικά 56 συλλογές / αποθετήρια

Αναζήτηση φορέα

Επιτροπή Ποντιακών Μελετών (Ε.Π.Μ.)
<http://www.epm.gr/>
 Η Επιτροπή Ποντιακών Μελετών είναι σωματείο επιστημονικό, μη κερδοσκοπικού χαρακτήρα, που ιδρύθηκε το 1927 στην Αθήνα από ομάδα Ποντίων διανοουμένων και πρωτοβουλία του γιατρού Θεοφύλακτου και την ηγεσία του τελευταίου Μητροπολίτη της Τραπεζούντας Χρυσανθού (Φιλίππιδης) και μετέπειτα Αρχιεπισκόπου Αθηνών και Πάσης Ελλάδος (Χρυσόστομος Β΄).
αποθετήριο

Ψηφιακό Μουσείο του Ποντιακού Ελληνισμού
<http://www2.epm.gr/web/guest/9>
 Το «Ψηφιακό Μουσείο του Ποντιακού Ελληνισμού» συνδυάζει την έκθεση των κειμηλίων που υπάρχουν στη διάθεση της Ε.Π.Μ. μαζί με εκπαιδευτικό υλικό και αναπτύσσει μία σειρά από εφαρμογές .

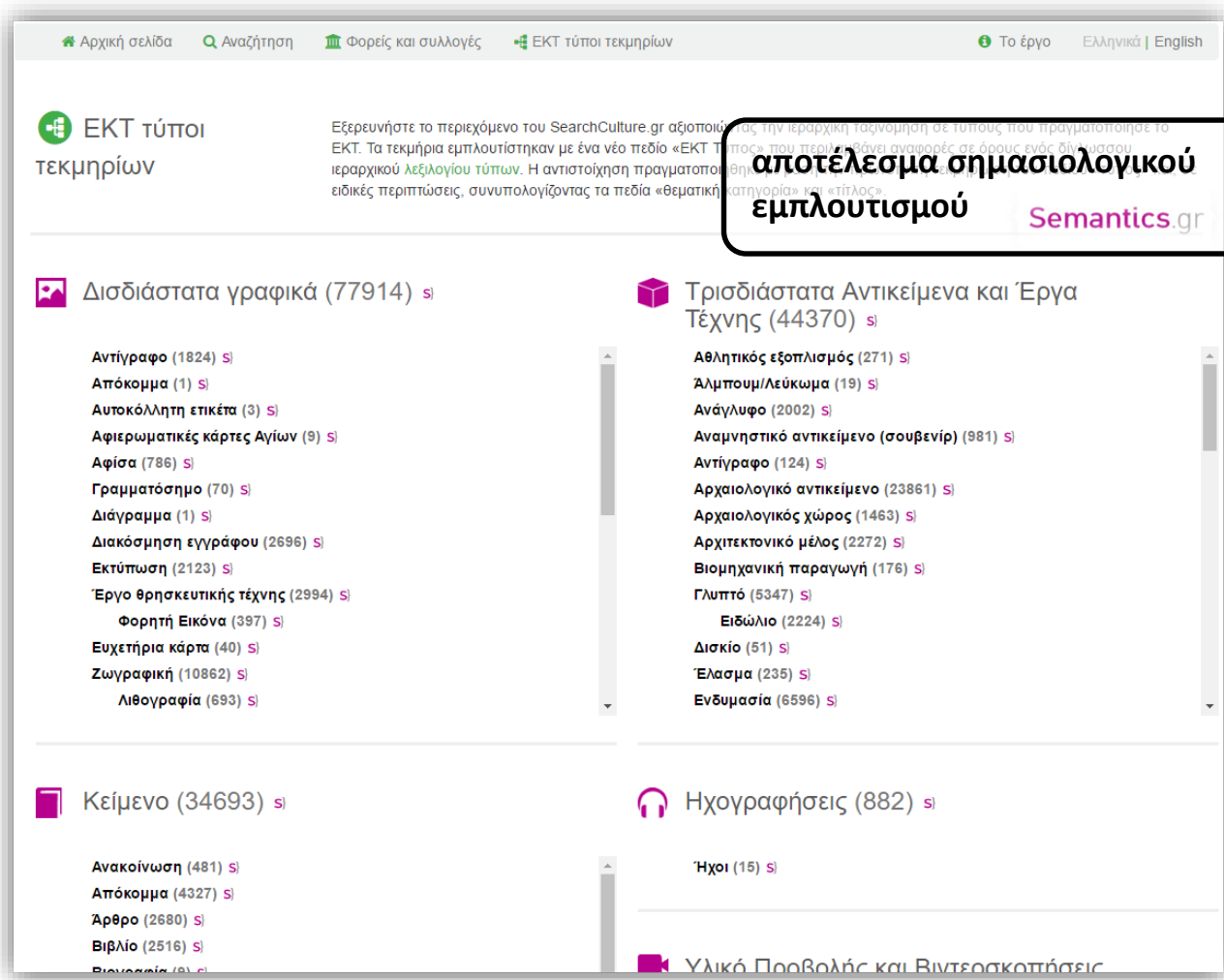
Φωτογραφία (256), Δισδιάστατα γραφικά (256), Κείμενο (80), Τρισδιάστατα Αντικείμενα και Έργα Τέχνης (69), Έγγραφο (39), Ενδυμασία (30), Βιβλίο (29), Οικεικός εξοπλισμός (27), Εργόχειρο (Κεντητική, Πλεκτική) (20), Προσωπικά αντικείμενα (αξεσουάρ) (14), Περιοδικό επιστημονικό (11), Ύφασμα (8), Σκεύος (7), Μετάλλιο (5), Κτίριο (3), Αλληλογραφία (3), Φυλλάδιο (1), Κόσμημα (1), Χάρτης (1), Άλμπουμ/Λεύκωμα (1), Επίπλωση (1), Έργο θρησκευτικής τέχνης (1), Σφραγίδα (1), Χαρτογραφικό Υλικό (1)

Ευρωπαϊκό Πολιτιστικό Κέντρο Δελφών (Ε.Π.Κε.Δ.)
<http://eccd.gr/>
 Το Ευρωπαϊκό Πολιτιστικό Κέντρο Δελφών (Ε.Π.Κε.Δ.) ιδρύθηκε το 1977 ως Νομικό Πρόσωπο Ιδιωτικού Δικαίου, υπό την εποπτεία του Υπουργείου Πολιτισμού και υπό την αιγίδα του Συμβουλίου της Ευρώπης με στόχο την δημιουργία στους Δελφούς ενός κέντρου Ευρωπαϊκού πολιτισμού. Οι πολιτιστικές δραστηριότητες και εκπαιδευτικές

Σχήμα 37. Παρουσίαση των φορέων και των συλλογών τους

3.2.5 Ιεραρχική πλοήγηση περιεχομένου μέσα από τους ΕΚΤ Τύπους

Το SearchCulture.gr διαθέτει μία ξεχωριστή σελίδα περιήγησης στο περιεχόμενο μέσω όλων των ΕΚΤ Τύπων στους οποίους ταξινομήθηκαν τα τεκμήρια (Σχήμα 38). Η σελίδα είναι προσβάσιμη από το μενού επιλογών επιλέγοντας «ΕΚΤ Τύποι τεκμηρίων». Οι ΕΚΤ Τύποι εμφανίζονται ιεραρχικά ομαδοποιημένοι σε 7 κατηγορίες, συγκεκριμένα «Δισδιάστατα γραφικά», «Τρισδιάστατα Γραφικά», «Κείμενο», «Ηχογραφήσεις», «Υλικό Προβολής και Βιντεοσκοπήσεις», «Παρτιτούρες» και «Χαρτογραφικό Υλικό». Κάθε τύπος είναι ένας σύνδεσμος που οδηγεί στη σελίδα αποτελεσμάτων με τα αντίστοιχα τεκμήρια. Ο κάθε τύπος συνοδεύεται από τον αριθμό τεκμηρίων και από έναν σύνδεσμο «S» ο οποίος οδηγεί στην σελίδα του αντίστοιχου όρου του Λεξιλογίου Τύπων του ΕΚΤ στο semantics.gr.



Σχήμα 38. Περιήγηση – EKT τύποι

3.2.6 Διεπαφή στα Αγγλικά

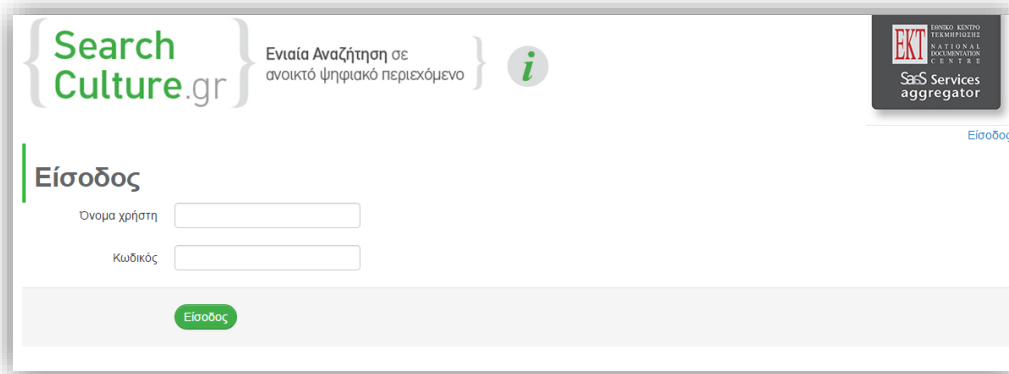
Η διαδικτυακή πύλη SearchCulture.gr υποστηρίζει πολυγλωσσικότητα τόσο σε στατικό όσο και σε δυναμικό περιεχόμενο. Στην τρέχουσα έκδοση υποστηρίζονται Ελληνικά και Αγγλικά.

Όλα τα στατικά μηνύματα/ετικέτες της εφαρμογής έχουν μεταφραστεί και στις δύο γλώσσες. Η πολυγλωσσικότητα όμως υποστηρίζεται και στο δυναμικό περιεχόμενο. Αν ο χρήστης επιλέξει άλλη γλώσσα (στο πάνω δεξί μέρος της εφαρμογής), τότε εκτός από τις στατικές ετικέτες και τα μηνύματα της εφαρμογής, θα «μεταφραστούν» και δυναμικές πληροφορίες, όπως οι ονομασίες των φορέων και των συλλογών, το νέφος δημοφιλών τύπων του EKT, οι πίτες της αρχικής σελίδας καθώς και όλες οι τιμές στα φίλτρα αναζήτησης.

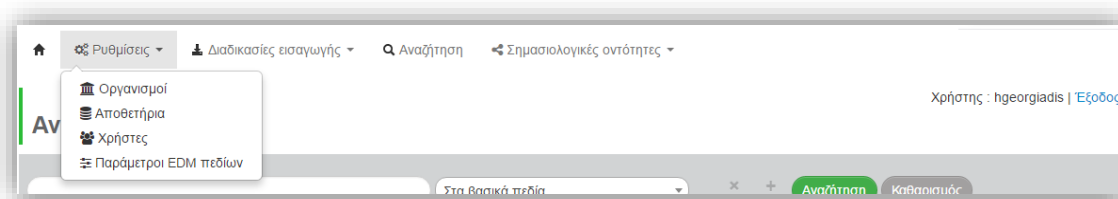
Σχήμα 39. Πολυγλωσσικότητα: Η αρχική σελίδα του SearchCulture.gr στα Αγγλικά.

3.3 Διαχειριστικό Περιβάλλον

Στη ενότητα αυτή περιγράφεται η διαχειριστική εφαρμογή ιστού του Συστήματος Συσσωρευτή SearchCulture.gr. Μέσω της εφαρμογής ο εξουσιοδοτημένος χρήστης μπορεί να διαχειριστεί πληροφορίες για τις συλλογές/αποθετήρια, να πυροδοτήσει νέες διαδικασίες εισαγωγής μεταδεδομένων ή αρχείων προεπισκόπησης, να εμπλουτίσει σημασιολογικά μια συλλογή, να παρακολουθεί την πρόοδο των διαδικασιών, να αναζητήσει πληροφορίες για διαδικασίες που έχουν εκτελεστεί στο παρελθόν, να αναζητήσει και να τροποποιήσει εγγραφές τεκμηρίων και σημασιολογικών οντοτήτων που έχουν εισαχθεί, να διαχειριστεί συνολικά το περιεχόμενο ανά αποθετήριο και, τέλος, να παραμετροποιήσει τόσο την διαδικτυακή πύλη όσο και το διαχειριστικό περιβάλλον.



Σχήμα 40. Είσοδος στο διαχειριστικό περιβάλλον



Σχήμα 41. Βασικό μενού επιλογών

Με την είσοδο στο διαχειριστικό περιβάλλον εμφανίζεται φόρμα εισαγωγής ονόματος χρήστη και συνθηματικού, καθώς μόνο εξουσιοδοτημένοι χρήστες από το προσωπικό του ΕΚΤ έχουν πρόσβαση στο διαχειριστικό περιβάλλον. Με την επιτυχή είσοδο στο σύστημα, εμφανίζεται η αρχική σελίδα με το βασικό μενού επιλογών (Σχήμα 41).

3.3.1 Διαχείριση Οργανισμών και Αποθετηρίων/Συλλογών

Επιλέγοντας «Οργανισμοί» από το βασικό μενού, ο χρήστης μεταφέρεται στην σελίδα διαχείρισης οργανισμών (Σχήμα 42). Στην σελίδα αυτή εμφανίζεται μια λίστα με όλους τους φορείς που έχουν καταχωρηθεί στο σύστημα. Ο χρήστης μπορεί να φιλτράρει την λίστα συμπληρώνοντας το όνομα ή την σύντομη ονομασία στην φόρμα αναζήτησης και επιλέγοντας στην συνέχεια «Αναζήτηση». Μπορεί επίσης να δημιουργήσει (πατώντας το κουμπί «Προσθήκη»), να επεξεργαστεί και να διαγράψει εγγραφές οργανισμών (επιλέγοντας «Ενέργειες» → «Επεξεργασία» ή «Ενέργειες» → «Διαγραφή», αντίστοιχα).

Επιλέγοντας «Αποθετήρια» από το βασικό μενού, ο χρήστης μεταφέρεται στην σελίδα διαχείρισης αποθετηρίων/συλλογών (Σχήμα 42). Κάθε αποθετήριο/συλλογή ανήκει σε έναν οργανισμό (φορέα). Στην σελίδα αποθετηρίων εμφανίζεται μια λίστα με όλα τα αποθετήρια που έχουν καταχωρηθεί στο σύστημα. Ο χρήστης μπορεί να φιλτράρει την λίστα συμπληρώνοντας το όνομα ή την σύντομη ονομασία στην φόρμα αναζήτησης και επιλέγοντας στην συνέχεια «Αναζήτηση». Από την στήλη «Ενέργειες», ο χρήστης μπορεί να επιλέξει «Ενέργειες» → «Ενημέρωση» για να τροποποιήσει τις πληροφορίες για συγκεκριμένο αποθετήριο μέσω της φόρμας που φαίνεται στο Σχήμα 43. Η ίδια φόρμα εμφανίζεται και όταν επιλέξουμε «Προσθήκη» από την αρχική σελίδα των αποθετηρίων για την καταχώρηση ενός νέου αποθετηρίου.

Χρήστης : hgeorgiadis | Έξοδος

Λίστα αποθετηρίων

Όνομα Σύντομο όνομα Ολοκληρώθηκε ο έλεγχος Προς δημοσίευση στην Europeana Αναζήτηση

[+ Προσθήκη](#)

< 1 2 3 4 5 6 > (53 αποτελέσματα)

Σύντομο όνομα	Όνομα	OAI-PMH URL	Γενικά	Τύποι μεταδεδομένων	Ενέργειες
ASCSA	Αμερικανική Σχολή Κλασικών Σπουδών	http://dlibrary.ascsa.edu.gr/oai/provider-wrapper/oaip/provider/asksa/	Εμφανίζεται στον Ενιαίο Κατάλογο ✓	oai_dc cidoc ese dc-ds-xml	Ενέργειες
Platania	Αποθετήριο Δήμου Πλατανιά	http://www.libraryoc.gr/oai/request	Εμφανίζεται στον Ενιαίο Κατάλογο ✓ Ολοκληρώθηκε ο έλεγχος ✓ Προς δημοσίευση στην Europeana ✓	edm oai_dc ese	Ενέργειες
ysma	Αποθετήριο Υπηρεσίας Συντήρησης Μνημείων Ακρόπολης	http://repository-ysma.ekt.gr/ysma-oai/request	Εμφανίζεται στον Ενιαίο Κατάλογο ✓	oai_dc ese	Ενέργειες
AMusChania	Αρχαιολογικό Μουσείο Χανίων	http://chaniamuseum-db.culture.gr/oai-pmh-repository/request	Εμφανίζεται στον Ενιαίο Κατάλογο ✓ Ολοκληρώθηκε ο έλεγχος ✓ Προς δημοσίευση στην Europeana ✓	edm oai_dc cidoc ese	Ενέργειες
Mandra	Δήμος Μάνδρας – Ειδυλλίας	http://services.mandra.evolution-isa.gr/mandra/oai	Εμφανίζεται στον Ενιαίο Κατάλογο ✓ Ολοκληρώθηκε ο έλεγχος ✓ Προς δημοσίευση στην	edm oai_dc ese	Ενέργειες

Σχήμα 42. Η σελίδα με τους παρόχους

Μέσα από την φόρμα δημιουργίας/ενημέρωσης αποθετηρίου, ο χρήστης συμπληρώνει όλες τις παραμέτρους του αποθετηρίου που απαιτούνται τόσο για την συγκομιδή (που πραγματοποιείται μέσω του συστήματος Harvester) και την εισαγωγή περιεχομένου στο SearchCulture.gr όσο και για να είναι εφικτός ο αυτοματοποιημένος έλεγχος που διεξάγει το Σύστημα Validator.

Στο Σχήμα 43 φαίνεται η φόρμα καταχώρησης/επεξεργασίας αποθετηρίου και συγκεκριμένα η καρτέλα γενικών στοιχείων η οποία περιλαμβάνει βασικά πεδία, όπως όνομα, σύντομο όνομα, URLs των OAI-PMH και SRU/SRW υπηρεσιών και τρόπο υποβολής ψηφιοποιημένου υλικού. Ο εξουσιοδοτημένος χρήστης επίσης μπορεί να χαρακτηρίσει ένα αποθετήριο ως έργο που πραγματοποιήθηκε στα πλαίσια των προσκλήσεων 31 και 31.2, να ορίσει αν έχει ολοκληρώσει τον έλεγχο συμμόρφωσης με τις προδιαγραφές διαλειτουργικότητας (που πραγματοποιεί το σύστημα Validator) και αν τα μεταδεδομένα της συλλογής θα διατίθενται μέσω της OAI-PMH υπηρεσίας στην Europeana ή όχι.

Χρήστης : hgeorgiadis | Έξοδος

Αποθετήριο Αρχαιολογικό Μουσείο Χανίων

Γενικά <> Τύποι μεταδεδομένων [OAI σύνολα](#) [Εισαγωγή στον ενιαίο κατάλογο](#) [Αναφορές σε δεσμευμένα λεξιλόγια/θησαυρούς](#)

Οργανισμός: ✓

Όνομα:

Όνομα (Αγγλικά):

Σύντομο όνομα:

URL αποθετηρίου:

OAI-PMH URL:

SRU/SRW URL:

URL διόθεσης μεταδεδομένων ως RDF (Συνδεδεμένα Δεδομένα):

Τρόπος υποβολής ψηφιακών αρχείων για επικύρωση: ✓

Προέλευση αποθετηρίου: ✓

Ολοκληρώθηκε ο έλεγχος:

Προς δημοσίευση στην Europeana:

Σχήμα 43. Φόρμα αποθετηρίου φορέα – Κάρτα γενικών στοιχείων.

Χρήστης : hgeorgiadis | Έξοδος

Αποθετήριο Αρχαιολογικό Μουσείο Χανίων

Γενικά <> Τύποι μεταδεδομένων [OAI σύνολα](#) [Εισαγωγή στον ενιαίο κατάλογο](#) [Αναφορές σε δεσμευμένα λεξιλόγια/θησαυρούς](#)

Τύποι μεταδεδομένων

edm ✓ OAI-PMH Αρχεία

Πρόθεμα που χρησιμοποιεί το αποθετήριο στην OAI-PMH υπηρεσία του:

Πηγή για ψηφιακά αρχεία
 Πηγή για αρχεία προεπισκόπησης
 Πηγή αρχείων για ασφαλή διαφύλαξη

oai_dc ✓ OAI-PMH Αρχεία

Πρόθεμα που χρησιμοποιεί το αποθετήριο στην OAI-PMH υπηρεσία του:

Πηγή για ψηφιακά αρχεία
 Πηγή για αρχεία προεπισκόπησης
 Πηγή αρχείων για ασφαλή διαφύλαξη

ese ✓ OAI-PMH Αρχεία

Πρόθεμα που χρησιμοποιεί το αποθετήριο στην OAI-PMH υπηρεσία του:

Πηγή για ψηφιακά αρχεία
 Πηγή για αρχεία προεπισκόπησης
 Πηγή αρχείων για ασφαλή διαφύλαξη

Πεδίο με URL (XPath) //ese:isShownBy
 Πεδίο με URL (XPath) //ese:object
 Πεδίο με URL (XPath) //ese:isShownBy

Σχήμα 44. Φόρμα αποθετηρίου φορέα – Κάρτα καταχώρησης τύπων μεταδεδομένων

Στο Σχήμα 44 φαίνεται η καρτέλα καταχώρησης υποστηριζόμενων τύπων μεταδεδομένων. Ο χρήστης προσθέτει τους τύπους μεταδεδομένων που υποστηρίζονται από την OAI-PMH υπηρεσία του αποθετηρίου. Για κάθε τύπο, ο χρήστης επιλέγει αν θα χρησιμοποιείται για την ανάκτηση ψηφιακών αρχείων (προκειμένου αυτά να ελεγχθούν από το Σύστημα Validator ή να αποθηκευτούν από το σύστημα Preservator) και τον τύπο μεταδεδομένων που θα χρησιμοποιείται για τη ανάκτηση εικόνων από τις οποίες το SearchCulture.gr θα δημιουργεί και θα αποθηκεύει εικόνες προεπισκόπησης. Η επιλογή τύπου μεταδεδομένων είτε ως πηγή για ψηφιακά αρχεία είτε ως πηγή για εικόνες προεπισκόπησης συνοδεύεται από τον ορισμό του συγκεκριμένου πεδίου με την χρήση XPath εκφράσεων (π.χ. “//ese:isShownBy” για τύπο ESE ή “//edm:isShownBy/@rdf:resource” για τύπο EDM) .

Σχήμα 45. Φόρμα αποθετηρίου φορέα – OAI Σύνολα ανά dc:type και ανά τύπο ψηφιοποίησης

Στην καρτέλα «OAI σύνολα» ο χρήστης ορίζει τα OAI σύνολα ανά τιμή του πεδίου dc:type, τα OAI σύνολα ανά κατηγορία ψηφιοποιημένου υλικού και, τέλος, το OAI σύνολο των τεκμηρίων που περιλαμβάνουν ενσωματωμένο κείμενο. Οι πληροφορίες αυτές είναι απαραίτητες για τον έλεγχο των αντίστοιχων προδιαγραφών από το Σύστημα Validator.

Στην καρτέλα «Εισαγωγή στον ενιαίο κατάλογο» ο χρήστης ορίζει ποιο τύπο (ή ποιος συνδυασμός τύπων) μεταδεδομένων θα χρησιμοποιηθεί ως πηγή από τον Συσσωρευτή SearchCulture.gr για την εισαγωγή των μεταδεδομένων του αποθετηρίου (Σχήμα 46). Οι τύποι που υποστηρίζονται μέχρι στιγμής είναι οι EDM, ESE [6] και συνδυασμός DC-DS-XML [11] με ESE .

Χρήστης : hgeorgiadis | Έξοδος

Αποθετήριο Αρχαιολογικό Μουσείο Χανίων

Γενικά </> Τύποι μεταδεδομένων OAI σύνολα Εισαγωγή στον ενιαίο κατάλογο Αναφορές σε δεσμευμένα λεξιλόγια/θησαυρούς

Πηγή για ενιαίο κατάλογο edm

Να εμφανίζεται στον Ενιαίο Κατάλογο

Κοινό πρόθεμα OAI προσδιοριστιών oai:museio.avmap.gr

Εμπλουτισμός από **Semantics.gr**

Υποστηρίζεται

Πεδία dc:type

Αποθήκευση Επιστροφή

Σχήμα 46. Φόρμα αποθετηρίου φορέα – Στοιχεία για εισαγωγή μεταδεδομένων

Χρήστης : hgeorgiadis | Έξοδος

Αποθετήριο Αρχαιολογικό Μουσείο Χανίων

Γενικά </> Τύποι μεταδεδομένων OAI σύνολα Εισαγωγή στον ενιαίο κατάλογο Αναφορές σε δεσμευμένα λεξιλόγια/θησαυρούς

Επικύρωση αναφορών σε δεσμευμένα λεξιλόγια Σε μεταδεδομένα σε τύπο EDM

Αναφορές σε λεξιλόγια/θησαυρό για ανθρώπους

Υποστηρίζεται Αξιοποίηση

Πεδία dc:creator | dc:contributor

URL λεξιλογίου/θησαυρού http://chaniamuseum-db.culture.gr/vocab/vocab/?tema=

Δημόσιο/ανοιχτό Ιδιωτικό

Συσχέτιση με δεσμ. λεξιλόγιο ενιαίου καταλόγου Αρχαιολογικό Μουσείο Χανίων - Λεξιλόγιο για ανθρώπους

Αναφορές σε λεξιλόγια/θησαυρό για τύπους

Υποστηρίζεται Αξιοποίηση

Πεδία dc:type

URL λεξιλογίου/θησαυρού http://chaniamuseum-db.culture.gr/vocab/vocab/?tema=

Δημόσιο/ανοιχτό Ιδιωτικό

Συσχέτιση με δεσμ. λεξιλόγιο ενιαίου καταλόγου Αρχαιολογικό Μουσείο Χανίων - Λεξιλόγιο για τύπους

Αναφορές σε λεξιλόγια/θησαυρό για θεματικές ενότητες

Υποστηρίζεται Αξιοποίηση

Πεδία dc:subject

URL λεξιλογίου/θησαυρού http://chaniamuseum-db.culture.gr/vocab/vocab/?tema=

Δημόσιο/ανοιχτό Ιδιωτικό

Συσχέτιση με δεσμ. λεξιλόγιο ενιαίου καταλόγου Αρχαιολογικό Μουσείο Χανίων - Λεξιλόγιο για θεματικές ενότητες

Αναφορές σε λεξιλόγια/θησαυρό για οργανισμούς

Υποστηρίζεται Αξιοποίηση

Αναφορές σε λεξιλόγια/θησαυρό για τόπους

Υποστηρίζεται Αξιοποίηση

Αναφορές σε λεξιλόγια/θησαυρό για χρονικές περιόδους

Υποστηρίζεται Αξιοποίηση

Πεδία dc:date | dcterms:temporal

URL λεξιλογίου/θησαυρού http://chaniamuseum-db.culture.gr/vocab/vocab/?tema=

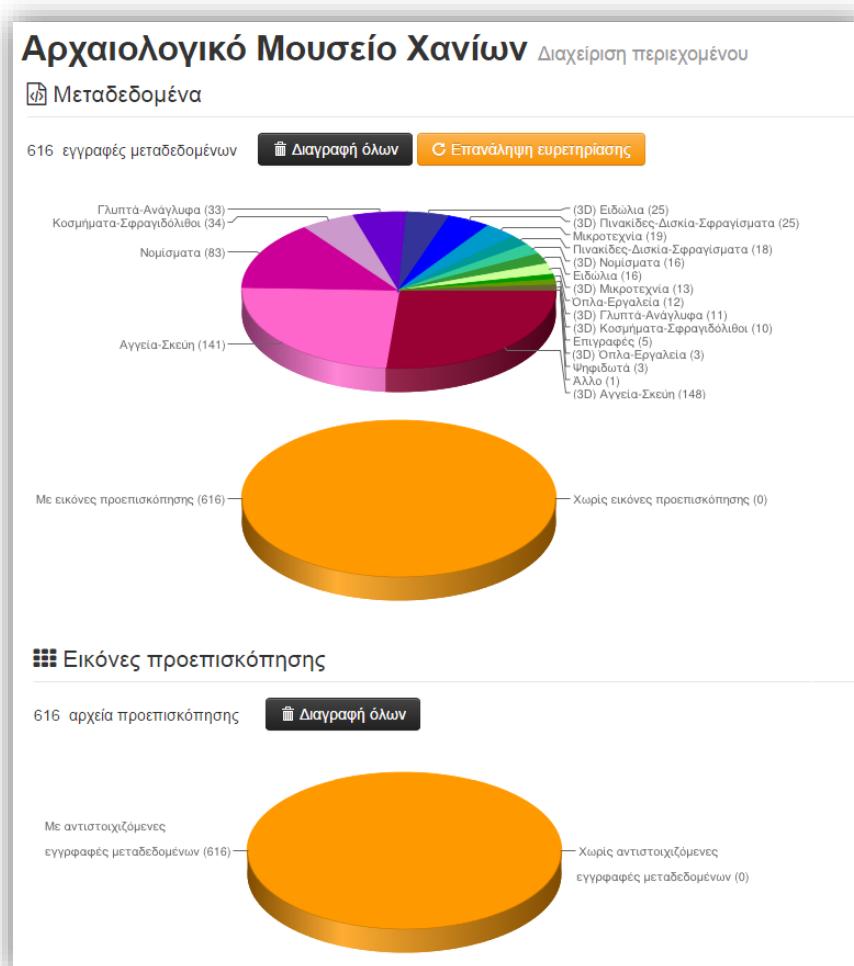
Δημόσιο/ανοιχτό Ιδιωτικό

Συσχέτιση με δεσμ. λεξιλόγιο ενιαίου καταλόγου Αρχαιολογικό Μουσείο Χανίων - Λεξιλόγιο για χρονικές περιόδους

Αποθήκευση Επιστροφή

Σχήμα 47. Φόρμα αποθετηρίου φορέα – Κάρτα δεσμευμένων λεξιλογίων/θησαυρών.

Στο Σχήμα 47 φαίνεται η καρτέλα καθορισμού των δεσμευμένων λεξιλογίων/θησαυρών που χρησιμοποιεί ο φορέας για κάθε μία από τις 6 κατηγορίες που ορίζονται στις προδιαγραφές διαλειτουργικότητας. Για κάθε υποστηριζόμενη κατηγορία, ο χρήστης ορίζει ποια πεδία θα περιλαμβάνουν αναφορές σε όρους του λεξιλογίου/θησαυρού, το URL του λεξιλογίου/θησαυρού και το αν πρόκειται για καθιερωμένο λεξιλόγιο/θησαυρό ή αν αναπτύχθηκε από τον ίδιο τον φορέα. Οι πληροφορίες σχετικά με τα δεσμευμένα λεξιλόγια χρησιμοποιούνται από το σύστημα Validator για τον έλεγχο των αντίστοιχων προδιαγραφών.

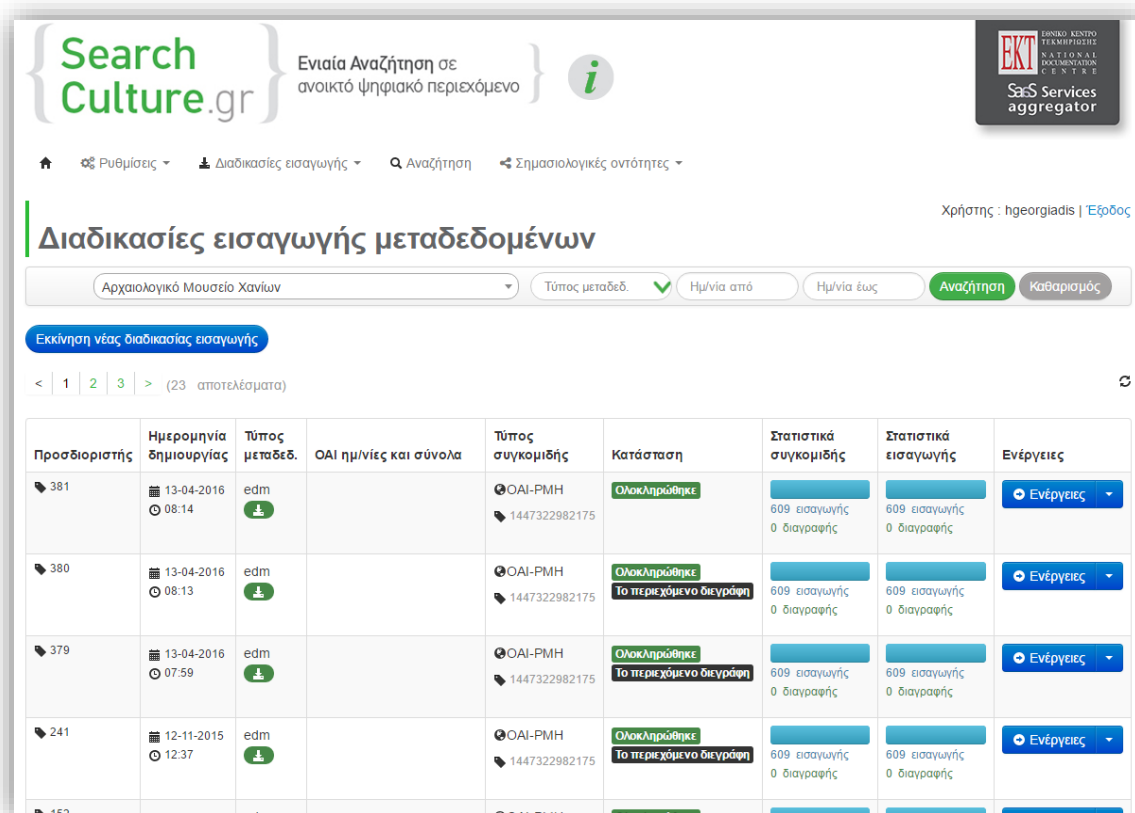


Σχήμα 48. Διαχείριση περιεχομένου αποθετηρίου

Από την σελίδα των αποθετηρίων, ο χρήστης μπορεί να επιλέξει «Ενέργειες»→«Διαχείριση περιεχομένου» για συγκεκριμένο αποθετήριο προκειμένου να διαχειριστεί μαζί το περιεχόμενό του (Σχήμα 48). Μέσα από την σελίδα που εμφανίζεται ο χρήστης έχει μια εποπτική εικόνα των μεταδεδομένων και των αρχείων προεπισκόπησης που έχουν εισαχθεί από το συγκεκριμένο αποθετήριο. Για τα μεταδεδομένα μπορεί είτε να διαγράψει εντελώς το περιεχόμενο από τον ενιαίο κατάλογο είτε απλώς να επαναλάβει την ευρετηρίασή του. Η δεύτερη επιλογή είναι ιδιαίτερα χρήσιμη στην περίπτωση που τροποποιηθεί η παραμετροποίηση του συστήματος ευρετηρίασης (Apache Solr). Τέλος, ο χρήστης μπορεί να διαγράψει και τις εικόνες προεπισκόπησης που εισήχθησαν για το συγκεκριμένο αποθετήριο.

3.3.2 Διαδικασίες Εισαγωγής Μεταδεδομένων

Επιλέγοντας «Διαδικασίες Εισαγωγής» → «Διαδικασίες Εισαγωγής Μεταδεδομένων» από το βασικό μενού, ο χρήστης μεταφέρεται στην σελίδα διαχείρισης διαδικασιών εισαγωγής μεταδεδομένων (Σχήμα 49). Μέσα από την σελίδα αυτή ο χρήστης μπορεί να αναζητήσει διαδικασίες εισαγωγής βάσει του αποθετηρίου, του τύπου μεταδεδομένων και του εύρους ημερομηνιών μέσα στο οποίο ξεκίνησε η διαδικασία. Με το πάτημα του κουμπιού «Αναζήτηση» εμφανίζεται μία λίστα με τις διαδικασίες που πληρούν τα κριτήρια αναζήτησης. Για κάθε διαδικασία φαίνονται βασικά στοιχεία, όπως οι ΟΑΙ-ΡΜΗ παράμετροι με τις οποίες πραγματοποιήθηκε η συγκομιδή, η κατάσταση της διαδικασίας («Συγκομιδή σε Εξέλιξη», «Η Συγκομιδή Ολοκληρώθηκε», «Η Συγκομιδή Απέτυχε», «Εισαγωγή σε Εξέλιξη», «Ολοκληρώθηκε», «Ολοκληρώθηκε με Αποτυχίες»), καθώς και συνοπτικά στατιστικά τόσο για την φάση της Συγκομιδής (τα στατιστικά αυτά προέρχονται από το Σύστημα Συγκομιδής) όσο και για την δεύτερη φάση, αυτήν της εισαγωγής στο Σύστημα Συσσωρευτή.



Προσδιοριστής	Ημερομηνία δημιουργίας	Τύπος μεταδεδ.	ΟΑΙ ημ/νίες και σύνολα	Τύπος συγκομιδής	Κατάσταση	Στατιστικά συγκομιδής	Στατιστικά εισαγωγής	Ενέργειες
381	13-04-2016 08:14	edm		ΟΑΙ-ΡΜΗ 1447322982175	Ολοκληρώθηκε	609 εισαγωγής 0 διαγραφής	609 εισαγωγής 0 διαγραφής	Ενέργειες
380	13-04-2016 08:13	edm		ΟΑΙ-ΡΜΗ 1447322982175	Ολοκληρώθηκε Το περιεχόμενο διεγράφη	609 εισαγωγής 0 διαγραφής	609 εισαγωγής 0 διαγραφής	Ενέργειες
379	13-04-2016 07:59	edm		ΟΑΙ-ΡΜΗ 1447322982175	Ολοκληρώθηκε Το περιεχόμενο διεγράφη	609 εισαγωγής 0 διαγραφής	609 εισαγωγής 0 διαγραφής	Ενέργειες
241	12-11-2015 12:37	edm		ΟΑΙ-ΡΜΗ 1447322982175	Ολοκληρώθηκε Το περιεχόμενο διεγράφη	609 εισαγωγής 0 διαγραφής	609 εισαγωγής 0 διαγραφής	Ενέργειες

Σχήμα 49. Διαδικασίες Εισαγωγής Μεταδεδομένων

Ο χρήστης μπορεί να εκκινήσει μια νέα διαδικασία εισαγωγής πατώντας το κουμπί «Εκκίνηση Νέας Διαδικασίας Εισαγωγής». Με το πάτημα του κουμπιού αυτού εμφανίζεται μία φόρμα στην οποία ο χρήστης καλείται να συμπληρώσει τις παραμέτρους οι οποίες θα μεταβιβαστούν στο Σύστημα Συγκομιδής προκειμένου αυτό να συλλέξει τα μεταδεδομένα και στην συνέχεια το Σύστημα Συσσωρευτή να τα «χωνέψει». Ο χρήστης μπορεί επίσης να επιλέξει το ανώτατο επιτρεπτό όριο μεμονωμένων αποτυχιών προτού τερματίσει ολόκληρη η διαδικασία σε κατάσταση «Αποτυχία». Η φόρμα με ενδεικτικές τιμές στα πεδία φαίνεται στο Σχήμα 50.

Ο χρήστης αρχικά επιλέγει το αποθετήριο από το οποίο θα εισαχθούν τα μεταδεδομένων (επομένως εμμέσως επιλέγεται το URL της OAI-PMH υπηρεσίας και ο τύπος μεταδεδομένων «για εισαγωγή» όπως ορίστηκε στην καρτέλα του Αποθετηρίου).

The screenshot shows a web interface titled "Εκκίνηση νέας διαδικασίας εισαγωγής". It contains several input fields and controls:

- Apothetirio: ZoggoroulosF
- Typos metasdō: dc-ds-xml.ese
- Όλα τα OAI σύνολα:
- OAI-PMH παράμετροι: Two rows, each with a dropdown menu (set-zwgrafiko_erg, set-glypto), "OAI από" and "OAI έως" date pickers, and a "Hδη συλλεχθέντα" checkbox.
- Οριο αποδεκτών ασφαμάτων εισαγωγής: 10
- Buttons: Εκκίνηση (blue), Ακύρωση (grey)

Σχήμα 50. Φόρμα εκκίνησης νέας διαδικασίας εισαγωγής για πολλαπλά OAI σύνολα

Στην συνέχεια, και εφόσον το κουτί ελέγχου «Όλα τα OAI Σύνολα» δεν είναι επιλεγμένο, ο χρήστης επιλέγει τα OAI-PMH σύνολα (set-specs) που επιθυμεί να συμπεριλάβει στην συγκομιδή. Στην πράξη για κάθε σύνολο που προστίθεται θα πραγματοποιηθεί ξεχωριστή διαδικασία συγκομιδής στο Σύστημα Συγκομιδής Harvester και, αντιστοίχως, ξεχωριστή διαδικασία εισαγωγής στο Σύστημα Συσσωρευτή. Επομένως, για κάθε σύνολο που προστίθεται, ο χρήστης θα πρέπει να συμπληρώσει το εύρος ημερομηνιών τελευταίας ενημέρωσης, δηλαδή τις παραμέτρους 'from' («OAI Από») και 'until' («OAI Έως») του πρωτοκόλλου OAI-PMH. Προκειμένου να εξασφαλιστεί ότι τα μεταδεδομένα είναι όλα ενημερωμένα αξιοποιώντας τον μηχανισμό για incremental harvesting που παρέχει το πρωτόκολλο OAI-PMH, η ημερομηνία «OAI Από» συμπληρώνεται αυτόματα από το σύστημα με την μέγιστη τιμή. Ο χρήστης δεν μπορεί να θέσει πιο πρόσφατη ημερομηνία. Το όριο αυτό προκύπτει από την «OAI Έως» ημερομηνία με την οποία εκτελέστηκε η τελευταία ολοκληρωμένη διαδικασία εισαγωγής που αφορά το ίδιο αποθετήριο, τύπο μεταδεδομένων και OAI σύνολο.

Εάν ο χρήστης επιλέξει «Όλα τα OAI Σύνολα», τότε θα τρέξει μία μόνο διαδικασία συγκομιδής στο Σύστημα Συγκομιδής Harvester (και αντιστοίχως μία διαδικασία εισαγωγής στο Σύστημα Συσσωρευτή) χωρίς να προσδιορίζεται συγκεκριμένο OAI-PMH σύνολο. Η φόρμα προσαρμόζεται ώστε να μην επιτρέπεται η προσθήκη συνόλων (Σχήμα 51).

Τέλος, δίνεται η δυνατότητα στον χρήστη να επαναχρησιμοποιήσει μεταδεδομένα τα οποία έχουν συλλεχθεί στο παρελθόν από το Σύστημα Συγκομιδής Harvester. Η δυνατότητα αυτή είναι πολύ σημαντική, ιδιαίτερα όταν ο χρήστης επιθυμεί να επαναλάβει μια διαδικασία εισαγωγής η οποία απέτυχε στην φάση της εισαγωγής (αφού προφανώς διορθώθηκε το πρόβλημα) ή όταν επιθυμεί να επανεισάγει μία συλλογή προκειμένου να εμπλουτιστεί σημασιολογικά (βλ. Ενότητα 3.3.3). Επίσης,

μπορεί τα μεταδεδομένα να έχουν συλλεχθεί προηγουμένως κατ' εντολή άλλου συστήματος (π.χ. του Συστήματος Validator) και επομένως η επανάληψη της συγκομιδής τους να είναι περιττή. Για την επαναχρησιμοποίηση μεταδεδομένων που έχουν συλλεχθεί στο παρελθόν, ο χρήστης θα πρέπει πρώτα να επισκεφθεί το Σύστημα Συγκομιδής Harvester, να εντοπίσει την διαδικασία και να σημειώσει τον προσδιοριστή της (Harvest Token). Στην συνέχεια επιστρέφοντας στην φόρμα εκκίνησης νέας διαδικασίας εισαγωγής (είτε έχει επιλέξει 'Όλα τα ΟΑΙ Σύνολα' είτε όχι), θα πρέπει να επιλέξει το κουτί ελέγχου 'Ήδη συλλεχθέντα' και να συμπληρώσει τον προσδιοριστή της διαδικασίας συγκομιδής στο πεδίο που εμφανίζεται, όπως φαίνεται στο Σχήμα 52.

Su	Mo	Tu	We	Th	Fr	Sa
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24

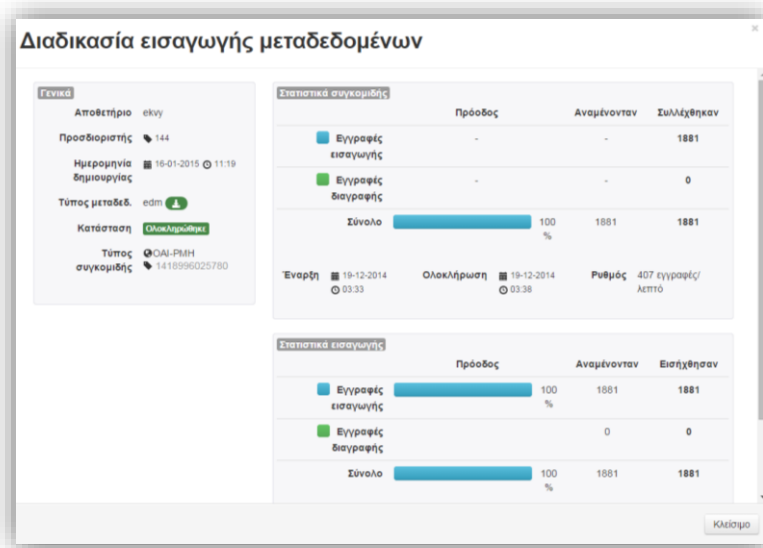
Σχήμα 51. Νέα διαδικασία εισαγωγής χωρίς καθορισμό συγκεκριμένου ΟΑΙ συνόλου

Su	Mo	Tu	We	Th	Fr	Sa

Σχήμα 52. Εισαγωγή μεταδεδομένων που έχουν ήδη συλλεχθεί στο παρελθόν από το Σύστημα Συγκομιδής.

Με το πάτημα του κουμπιού «Εκκίνηση» πραγματοποιείται έλεγχος της ορθότητας των παραμέτρων (π.χ. αν δόθηκαν τιμές σε υποχρεωτικά πεδία). Σε περίπτωση σφαλμάτων ως προς την εγκυρότητα των παραμέτρων, ο χρήστης ενημερώνεται και καλείται να τα διορθώσει (Σχήμα 53). Διαφορετικά η διαδικασία ξεκινά, η φόρμα κλείνει και στην λίστα των διαδικασιών εμφανίζεται η νέα διαδικασία με κατάσταση «Συγκομιδή σε εξέλιξη». Τα πεδία «Στατιστικά Συγκομιδής» και «Στατιστικά Εισαγωγής» για την συγκεκριμένη διαδικασία εμφανίζουν από μία μπάρα προόδου που ανανεώνεται σχεδόν σε πραγματικό χρόνο.

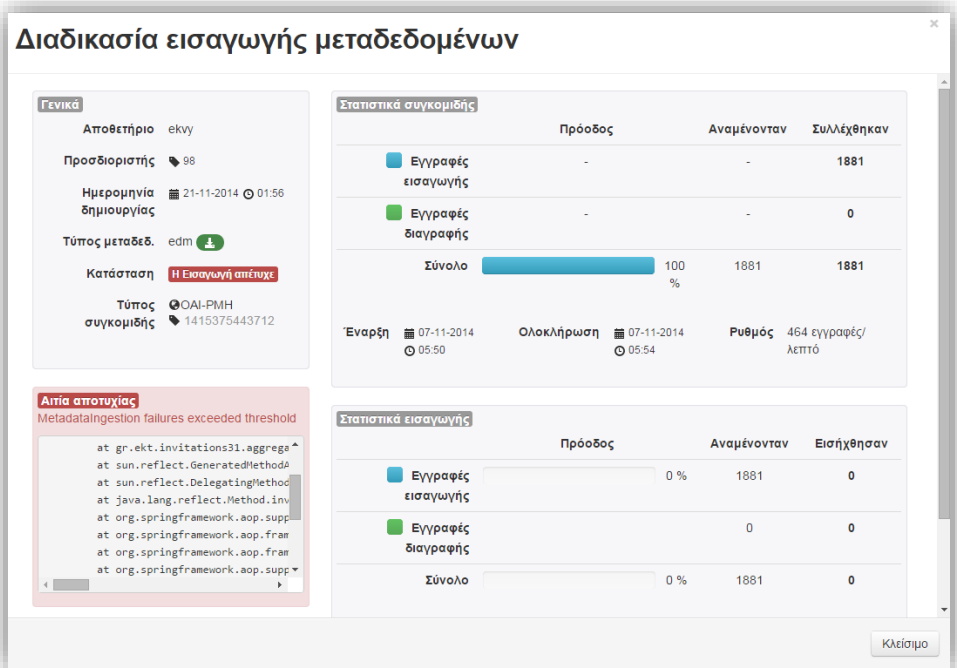
Σχήμα 53. Έλεγχος ορθότητας παραμέτρων εκκίνησης εισαγωγής



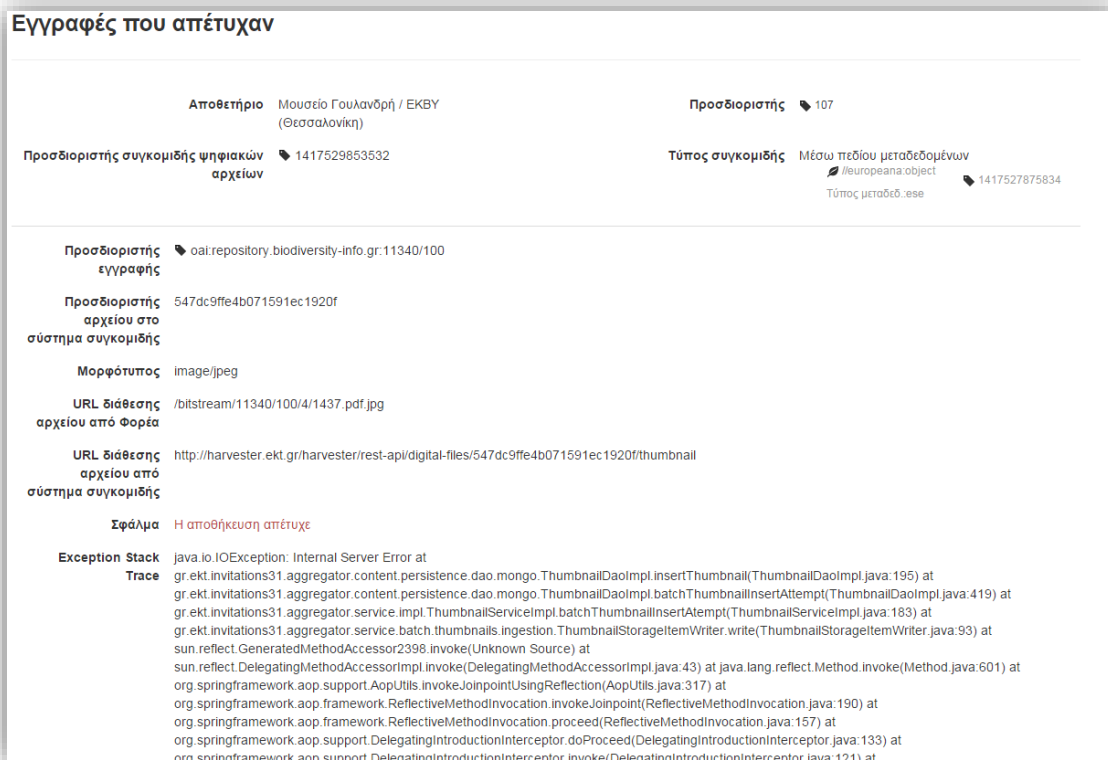
Σχήμα 54. Προβολή λεπτομερών στατιστικών διαδικασίας εισαγωγής

Σε περίπτωση που για οποιονδήποτε λόγο μία διαδικασία αποτύχει, εμφανίζεται με κατάσταση «Απέτυχε» συνοδευόμενη από το μήνυμα σφάλματος. Εάν ο χρήστης πατήσει το εικονίδιο δίπλα ακριβώς από το μήνυμα σφάλματος ή εάν επιλέξει να δει τις λεπτομέρειες της συγκεκριμένης διαδικασίας («Ενέργειες»→«Προβολή»), θα δει λεπτομερές μήνυμα καθώς και αναλυτικό Exception Stack Trace, όπως φαίνεται στο Σχήμα 55.

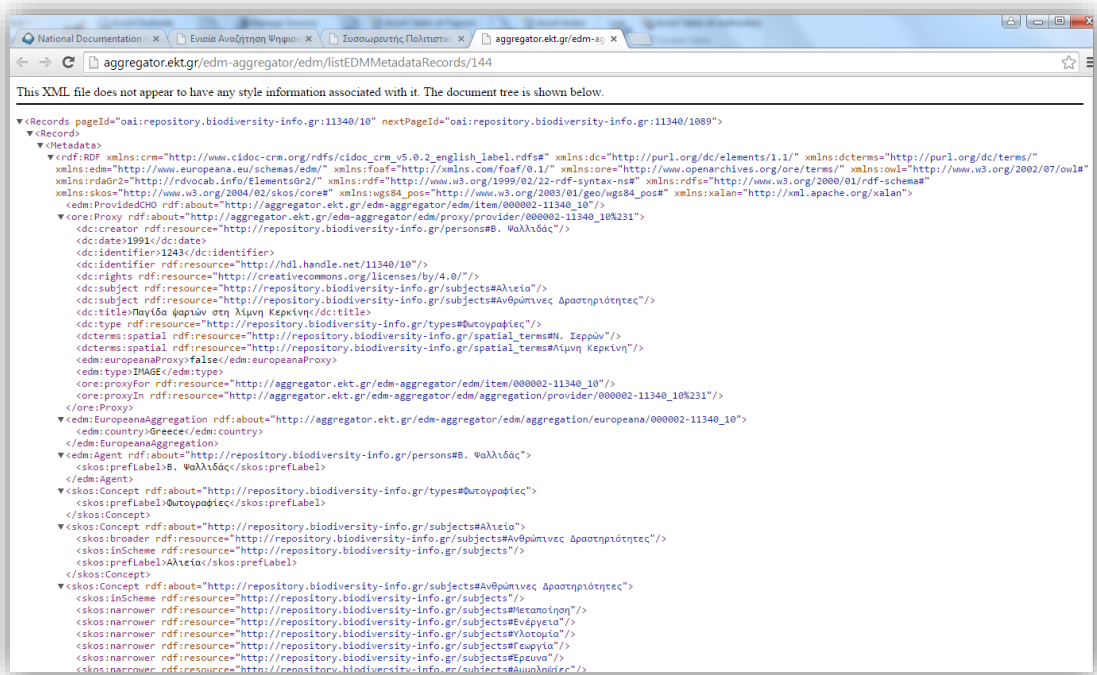
Για περισσότερες πληροφορίες για μία διαδικασία εισαγωγής, ο χρήστης μπορεί να επιλέξει «Ενέργειες»→«Προβολή», οπότε και θα εμφανιστεί ένα παράθυρο (Σχήμα 54) με πληροφορίες όπως, στατιστικά συγκομιδής (αριθμός εγγραφών εισαγωγής/διαγραφής που αναμένονταν και που συλλέχθηκαν, ρυθμός συγκομιδής) και στατιστικά εισαγωγής (αριθμός εγγραφών εισαγωγής/διαγραφής που εισήχθησαν/διαγράφηκαν, ρυθμός εισαγωγής).



Σχήμα 55. Εμφάνιση μηνυμάτων για αποσφαλμάτωση για διαδικασίες που απέτυχαν



Σχήμα 56. Αναφορά εγγραφών μεταδεδομένων που απέτυχαν.



Σχήμα 57. Προβολή EDM εγγραφών που εισήχθησαν στα πλαίσια μιας διαδικασίας εισαγωγής

Τέλος, ειδικά για τις ολοκληρωμένες διαδικασίες, ο χρήστης μπορεί να δει το σύνολο των EDM εγγραφών που εισήχθησαν («Ενέργειες»→«Προβολή EDM μεταδεδομένων») ή να πλοηγηθεί σε όλα τα τεκμήρια με δυνατότητες αναζήτησης και προεπισκόπησης («Ενέργειες»→«Προεπισκόπηση εγγραφών»).

3.3.3 Μαζικός σημασιολογικός εμπλουτισμός από το Semantics.gr

Το SearchCulture.gr παρέχει την δυνατότητα ελεγχόμενου σημασιολογικού εμπλουτισμού πεδίων μεταδεδομένων των αποθετηρίων που φιλοξενεί με αναφορές σε λεξιλόγια όρων που δημιουργήθηκαν και διατίθενται στο Σύστημα Σημασιολογικών Λεξιλογίων του EKT Semantic.gr. Ο εμπλουτισμός γίνεται ανά αποθετήριο για συγκεκριμένα πεδία και προϋποθέτει τον ορισμό κανόνων αντιστοίχισης σε ειδικό εργαλείο που παρέχει το Semantics.gr. Μόλις ολοκληρωθούν οι αντιστοιχισμοί απαιτείται η επανεισαγωγή των μεταδεδομένων του αποθετηρίου στο SearchCulture.gr προκειμένου να πραγματοποιηθεί μαζικά ο εμπλουτισμός βάσει των κανόνων αντιστοίχισης που ορίστηκαν.

Για τον ορισμό κανόνων αντιστοίχισης, θα πρέπει ο εξουσιοδοτημένος χρήστης να εγγράψει το αποθετήριο στο Semantic.gr με την ίδια σύντομη ονομασία που έχει το αποθετήριο στο SearchCulture.gr. Η εγγραφή και η διαχείριση αποθετηρίων γίνεται μέσα από την σελίδα αποθετηρίων του Semantics.gr (Μενού διαχειριστή →«Αποθετήρια»). Από εκεί ο χρήστης μπορεί να επιλέξει «Προσθήκη» για να εγγράψει το αποθετήριο μέσα από ειδική φόρμα. Για κάθε αποθετήριο, ο χρήστης θα πρέπει να ορίσει συγκεκριμένα πεδία μεταδεδομένων του αποθετηρίου που προορίζονται για εμπλουτισμό και να επιλέξει συγκεκριμένα λεξιλόγια για τον εμπλουτισμό τους. Στο Σχήμα 58 παρουσιάζεται η φόρμα εγγραφής/επεξεργασίας αποθετηρίου.

Εναλλακτικά, μέσα από το διαχειριστικό περιβάλλον του SearchCulture.gr, στην καρτέλα διαχείρισης περιεχομένου αποθετηρίου υπάρχει σύνδεσμος με το πάτημα του οποίου το αποθετήριο εγγράφεται αυτόματα στο Σύστημα Λεξιλογίων Semantics.gr.

Αποθετήριο

Σύντομη ονομασία: AMusChania

Φορέας:

- Σύντομη ονομασία: KE_eforeia
- Όνομασία: ΚΕ Έφορεία Προϊστορικών και Κλασικών Αρχαιοτήτων

Όνομασία:

- Αρχαιολογικό Μουσείο Χανίων
- Ελληνικά

Περιγραφή:

- Επιλέξτε Γλώσσα

Ιστότοπος αποθετηρίου: <http://chaniamuseum-db.culture.gr/oai-pmh-repository/request>

URL προβολής όλων των τεκμηρίων: <http://searchculture.gr/aggregator/portal/search?keywords%5B0%5D.value=&keywords%5B0%5D.field=&providerShortName=AMusChania&page.page=1>

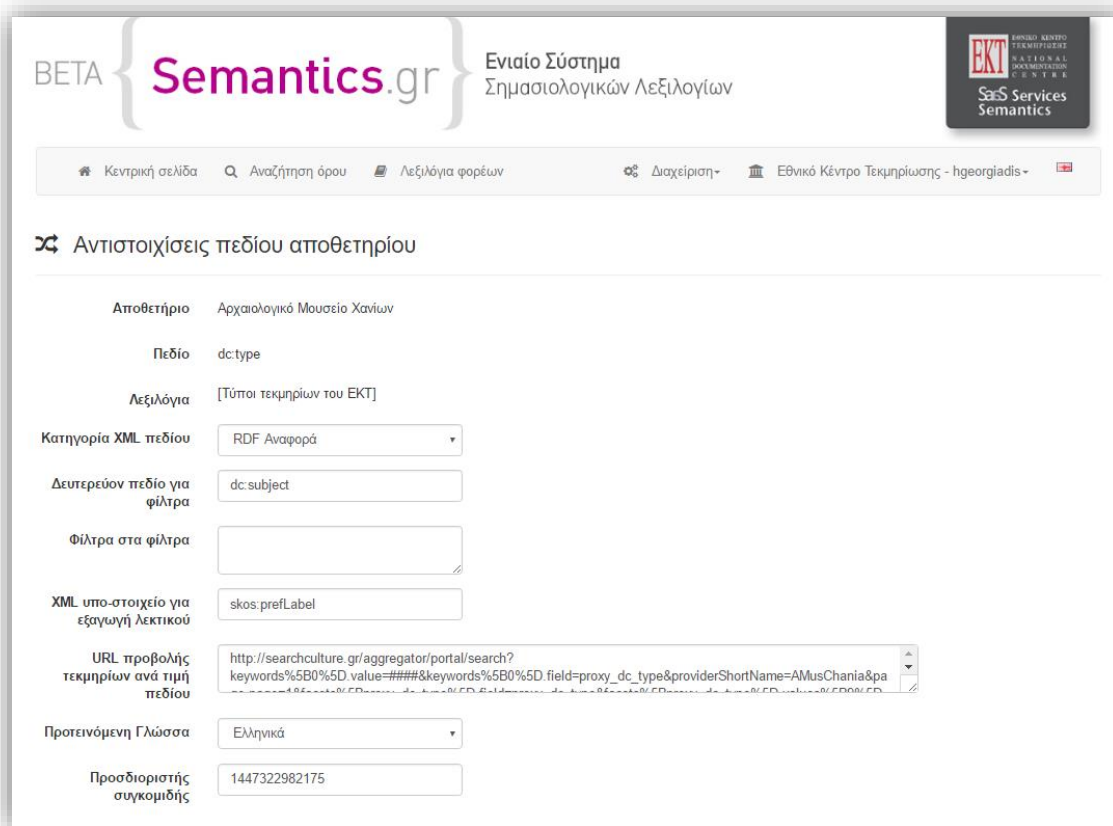
Πεδία μεταδεδομένων αποθετηρίου:

- Πεδίο: dc:type
- Σημασιολογικές κατηγορίες: Έννοιες
- Λεξιλόγια: Τύποι τεκμηρίων του ΕΚΤ
- Δεσμευμένο
- Πεδίο: dc:subject
- Σημασιολογικές κατηγορίες

Σχήμα 58. Semantics.gr: φόρμα εγγραφής αποθετηρίου

Στην συνέχεια ο εξουσιοδοτημένος χρήστης θα πρέπει να ορίσει αντιστοιχίσεις χωριστά για καθένα από τα πεδία που ορίστηκαν στην φόρμα αποθετηρίου. Για να το κάνει αυτό θα πρέπει από το βασικό μενού να επιλέξει «Διαχείριση» → «Αντιστοιχίσεις Πεδίων Αποθετηρίων». Στην συνέχεια θα πρέπει να επιλέξει «Προσθήκη» για να δημιουργήσει ένα σύνολο αντιστοιχίσεων για συγκεκριμένο πεδίο μέσα από ειδική φόρμα η οποία παρουσιάζεται στο Σχήμα 59. Ο χρήστης θα πρέπει να ορίσει το αποθετήριο, το πεδίο μεταδεδομένων προς εμπλουτισμό και ορισμένες οδηγίες που χρειάζεται το σύστημα για την εξαγωγή των ξεχωριστών τιμών του πεδίου μεταδεδομένων από το περιεχόμενο του αποθετηρίου (μέσω του συστήματος Harvester). Συγκεκριμένα, θα πρέπει να οριστεί η κατηγορία του πεδίου, αν δηλαδή είναι απλό λεκτικό (όπως το dc:type στο σχήμα μεταδεδομένων ESE) ή υποστηρίζει RDF αναφορές (όπως το dc:type στο RDF-based σχήμα μεταδεδομένων EDM). Στην δεύτερη περίπτωση, απαιτείται να δοθεί το XML υπο-στοιχείο του RDF πόρου που περιέχει το λεκτικό (π.χ. skos:prefLabel). Ο χρήστης μπορεί να

ορίσει και δευτερεύον πεδίο οι τιμές του οποίου επίσης θα συλλεχθούν και θα μπορούν να χρησιμοποιηθούν ως φίλτρο για πιο λεπτομερή αντιστοίχιση.



BETA { Semantics.gr } Ενιαίο Σύστημα Σημασιολογικών Λεξιλογίων

Κεντρική σελίδα Αναζήτηση όρου Λεξιλόγια φορέων Διαχείριση Εθνικό Κέντρο Τεκμηρίωσης - hgeorgiadis

Αντιστοιχίσεις πεδίου αποθετηρίου

Αποθετήριο Αρχαιολογικό Μουσείο Χανίων

Πεδίο dc:type

Λεξιλόγια [Τύποι τεκμηρίων του ΕΚΤ]

Κατηγορία XML πεδίου RDF Αναφορά

Δευτερεύον πεδίο για φίλτρα dc:subject

Φίλτρα στα φίλτρα


XML υπο-στοιχείο για εξαγωγή λεκτικού skos:prefLabel

URL προβολής τεκμηρίων ανά τιμή πεδίου http://searchculture.gr/aggregator/portal/search?keywords%5B0%5D.value#####&keywords%5B0%5D.field=proxy_dc_type&providerShortName=AMusChania&pa...

Προτεινόμενη Γλώσσα Ελληνικά

Προσδιοριστής συγκομιδής 1447322982175

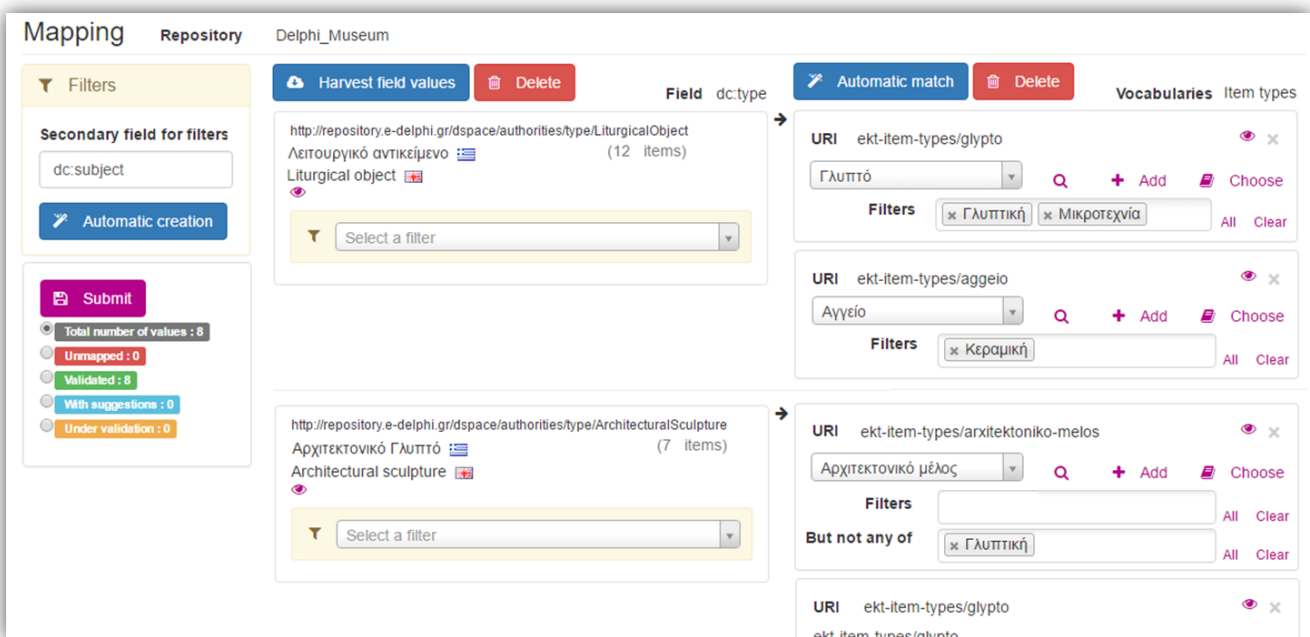
Σχήμα 59. Semantics.gr: φόρμα ορισμού αντιστοιχίσεων πεδίου αποθετηρίου – εισαγωγή βασικών στοιχείων

Ο χρήστης προαιρετικά καλείται να συμπληρώσει προτεινόμενη γλώσσα για πολυγλωσσικά πεδία ώστε οι αντιστοιχίσεις να γίνουν βάσει του λεκτικού σε αυτή την γλώσσα καθώς και ένα μοτίβο (pattern) ενός URL το οποίο θα αποτελεί οδηγό για την εμφάνιση όλων των τεκμηρίων με συγκεκριμένη τιμή στο αντιστοιχιζόμενο πεδίο, δυνατότητα που θα δίνεται με τον ειδικό σύνδεσμο «  » σε κάθε μία τιμή χωριστά. Ο σύνδεσμος αυτός είναι ιδιαίτερα χρήσιμος καθώς ο χρήστης που ορίζει τις αντιστοιχίσεις μπορεί εύκολα να πλοηγηθεί στις εγγραφές τεκμηρίων για κάθε μία ξεχωριστή τιμή του πρωτεύοντος πεδίου μεταδεδομένων. Σημειώνεται ότι για αποθετήρια των οποίων το περιεχόμενο έχει ήδη συλλεχθεί από το SearchCulture.gr, το URL μπορεί να είναι ένα κατάλληλο μοτίβο αναζήτησης στο SearchCulture.gr. Για παράδειγμα, για ένα αποθετήριο με σύντομη ονομασία “AMusChania” και για το πεδίο dc:type, το URL θα μπορούσε να έχει την μορφή:

http://searchculture.gr/aggregator/portal/search?keywords%5B0%5D.value#####&keywords%5B0%5D.field=proxy_dc_type&providerShortName=AMusChania

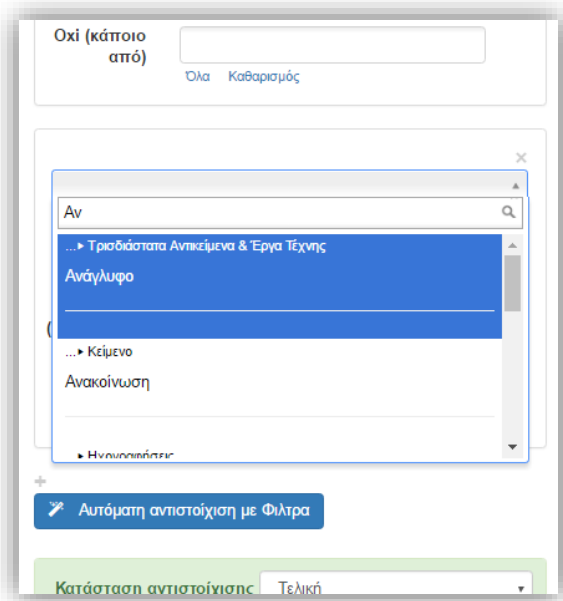
όπου “#####” το σημείο το οποίο θα αντικαθίσταται από το σύστημα δυναμικά με την εκάστοτε τιμή. Τέλος, ο χρήστης θα πρέπει να συμπληρώσει τον προσδιοριστή της διαδικασίας συγκομιδής των μεταδεδομένων του αποθετηρίου από το Σύστημα Συγκομιδής Harvester (ο χρήστης θα πρέπει πρώτα να επισκεφθεί το Σύστημα Συγκομιδής Harvester, να εντοπίσει την διαδικασία και να σημειώσει τον προσδιοριστή της).

Στην συνέχεια ο χρήστης θα πρέπει να πατήσει το κουμπί «Συγκομιδή τιμών» προκειμένου το σύστημα Semantics.gr να επικοινωνήσει με το σύστημα Harvester (μέσω των REST υπηρεσιών του δεύτερου) να αντλήσει τα μεταδεδομένα και να δημιουργήσει μία λίστα με διαφορετικές τιμές για το προς εμπλουτισμό πεδίο. Στην περίπτωση που έχει οριστεί δευτερεύον πεδίο, κάθε τιμή του πρωτεύοντος θα συνδεθεί με ένα σύνολο τιμών του δευτερεύοντος πεδίου τις οποίες ο χρήστης μπορεί να χρησιμοποιήσει στη συνέχεια ως φίλτρα. Ο συνδυασμός των τιμών των δύο πεδίων μπορεί να αυξήσει εντυπωσιακά τον αριθμό των αντιστοιχίσεων και κατ' επέκταση να οδηγήσει σε ποιοτικότερο και λεπτομερέστερο εμπλουτισμό. Μόλις ολοκληρωθεί η συγκομιδή, όλες οι ξεχωριστές τιμές του πρωτεύοντος πεδίου εμφανίζονται στο αριστερό τμήμα της οθόνης, όπως φαίνεται στο Σχήμα 60. Κάθε τιμή συνοδεύεται από τον σύνδεσμο («👁») που επιστρέφει σε ξεχωριστή καρτέλα τα τεκμήρια που περιλαμβάνουν την συγκεκριμένη τιμή στο πεδίο μεταδεδομένων που ορίστηκε ως πρωτεύον.



Σχήμα 60. Semantics.gr: η φόρμα αντιστοιχίσεων πεδίου αποθετηρίου αμέσως μετά την συγκομιδή τιμών

Η φόρμα αντιστοίχισης παρέχει την δυνατότητα αυτόματης αντιστοίχισης των τιμών που συλλέχθηκαν με όρους του λεξιλογίου που ορίστηκε με το κουμπί «Αυτόματη αντιστοίχιση». Η αυτόματη αντιστοίχιση βασίζεται στην ομοιότητα αλφαριθμητικών τιμών (string matching) των λεκτικών τιμών του πεδίου μεταδεδομένων με τις προτεινόμενες ονομασίες των όρων του λεξιλογίου. Στην συνέχεια ο εξουσιοδοτημένος χρήστης καλείται να αντιστοιχίσει τιμές που δεν ήταν εφικτό να αντιστοιχηθούν αυτόματα ή για τις οποίες η αυτόματη αντιστοίχιση δεν έδωσε σωστό αποτέλεσμα (Σχήμα 61) . Μόλις ολοκληρωθεί μία αντιστοίχιση, ο χρήστης μπορεί να ορίσει την κατάστασή της ως «Τελική». Εάν ο χρήστης εκτιμήσει ότι θα πρέπει να δημιουργηθεί νέος όρος στο λεξιλόγιο προκειμένου να καλύψει κάποια τιμή πεδίου, μπορεί να ορίσει την κατάσταση της αντιστοίχισης ως «Προτάσεις για δημιουργία όρων» και να συμπληρώσει τα λεκτικά των όρων που προτείνει να εισαχθούν στο λεξιλόγιο μαζί με κάποιο σχόλιο.



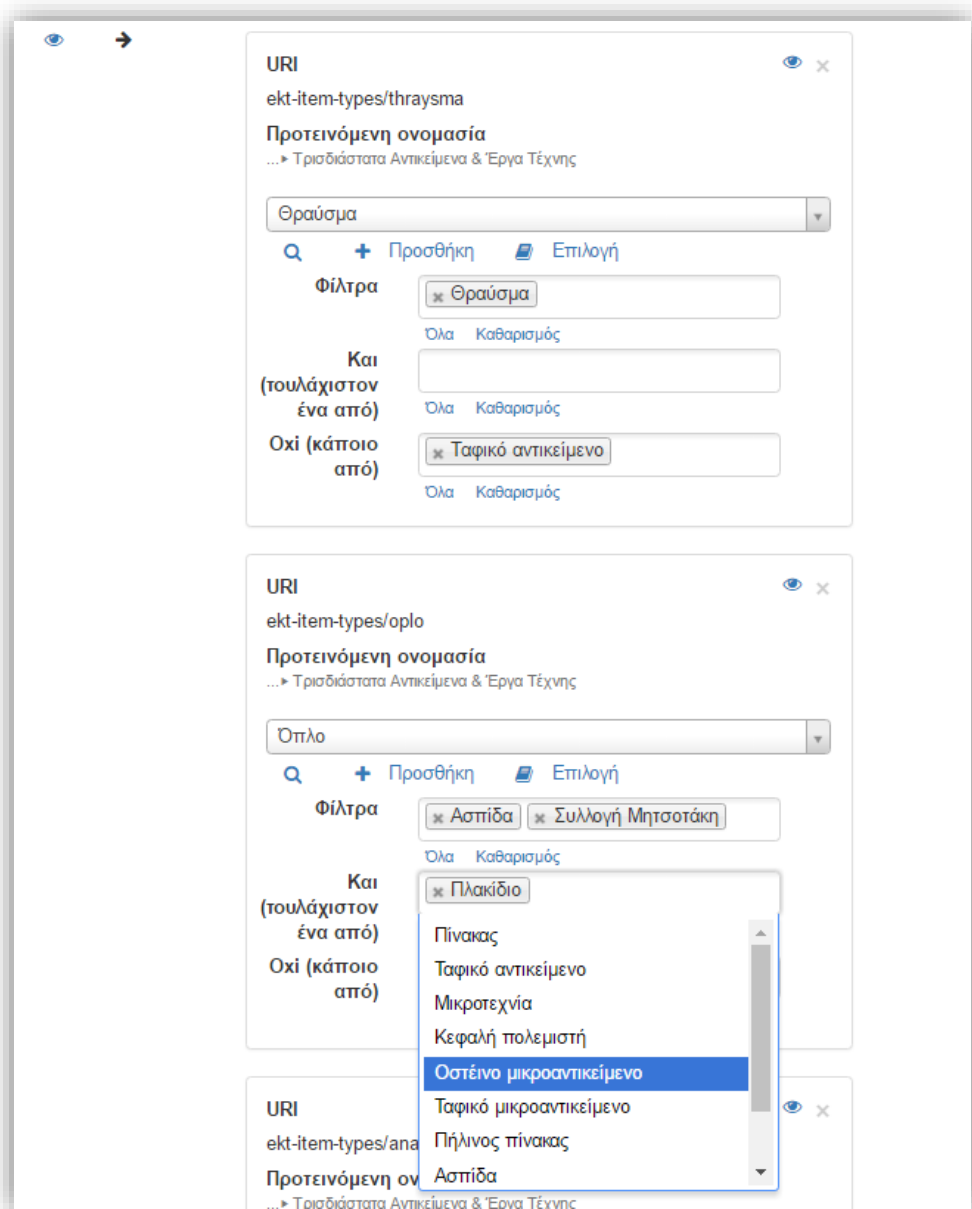
Σχήμα 61. Semantics.gr: χειροκίνητη αντιστοίχιση με όρο λεξιλογίου

Ο χρήστης μπορεί να αντιστοιχίσει μία τιμή με πολλούς όρους λεξιλογίου πατώντας το κουμπί «+». Αν μάλιστα έχει επιλέξει δευτερεύον πεδίο, ο χρήστης μπορεί να ορίσει διαφορετικούς όρους λεξιλογίων ανάλογα με το αν η εγγραφή έχει ή δεν έχει κάποιες τιμές του δευτερευόντως πεδίου. Συγκεκριμένα, ο χρήστης μπορεί να δημιουργήσει σύνθετους κανόνες (με λογικό «Η», λογικό «ΚΑΙ» και λογικό «ΟΧΙ») πάνω στις τιμές των φίλτρων για να δρομολογήσει με ακρίβεια τον εμπλουτισμό του πρωτεύοντος πεδίου, όπως φαίνεται στο Σχήμα 62.

Αφού ολοκληρωθούν οι αντιστοιχίσεις για ένα αποθετήριο, τότε ο εξουσιοδοτημένος χρήστης του διαχειριστικού περιβάλλοντος του SearchCulture.gr επανεισάγει το περιεχόμενο του αποθετηρίου (χωρίς να απαιτηθεί η επανάληψη της συγκομιδής αφού μπορούν να χρησιμοποιηθούν τα μεταδεδομένα της τελευταίας συγκομιδής από το σύστημα Harvester), σύμφωνα με τις οδηγίες της παραγράφου 3.3.2. Κατά την διάρκεια της εισαγωγής, ο συσσωρευτής SearchCulture.gr για κάθε εγγραφή συμβουλευεται τους κανόνες αντιστοίχισης που ορίστηκαν στο Semantics.gr και εμπλουτίζει την εγγραφή με νέες σημασιολογικές αναφορές. Τονίζεται ότι οι νέες τιμές-αναφορές αποθηκεύονται σε ξεχωριστή δομή (edm:EuropeanaProxy) από την δομή που συγκεντρώνει τα πρωτότυπα μεταδεδομένα του αποθετηρίου (edm:ProvidedCHO).

Το ΕΚΤ δημιούργησε ένα πρότυπο ιεραρχικό και δίγλωσσο (ελληνικά-αγγλικά) Λεξιλόγιο Τύπων Ψηφιακού Περιεχομένου αποτελούμενο από 392 όρους το οποίο διατίθεται ανοικτά μέσα στο το Semantics.gr στην διεύθυνση <http://semantics.gr/authorities/vocabularies/ekt-item-types/vocabulary-entries/tree>. Στην συνέχεια δημιουργήθηκαν εγγραφές αποθετηρίων στο Semantics.gr που αντιστοιχούν στα αποθετήρια/συλλογές που καταχωρήθηκαν στο Σύστημα Διαχείρισης Φορέων/Αποθετηρίων του SearchCulture.gr (το οποίο υπενθυμίζεται ότι χρησιμοποιείται και από τα συστήματα Validator και Preservator). Για καθένα από αυτά τα αποθετήρια και συγκεκριμένα για το πεδίο dc:type ορίστηκαν κανόνες αντιστοίχισης προς όρους του Λεξιλογίου Τύπων του ΕΚΤ. Πολλά αποθετήρια χαρακτηρίζονται από φτωχή τεκμηρίωση του πεδίου dc:type και για τον λόγο αυτό για τα συγκεκριμένα αποθετήρια χρησιμοποιήθηκε ως «φίλτρο» αντιστοίχισης το πεδίο dc:subject (δευτερεύον πεδίο). Το αποτέλεσμα πολλές φορές ήταν εντυπωσιακό, όπως φαίνεται από το πραγματικό παράδειγμα που

παρουσιάζεται στο Σχήμα 63. Ο εμπλουτισμός σε πολλές περιπτώσεις είχε τον αντίθετο στόχο: κάποια αποθετήρια παρείχαν πολύ λεπτομερή και εξειδικευμένη τεκμηρίωση του πεδίου dc:type και έπρεπε να γενικευτούν. Σε αυτές τις περιπτώσεις οι αντιστοιχίσεις έγιναν σε γενικότερους όρους του Λεξιλογίου Τύπων του ΕΚΤ, ώστε το συνολικό αποτέλεσμα να είναι ομοιογενές και ο χρήστης του SearchCulture.gr να προηγείται μέσα από ένα ελεγχόμενο σύνολο τύπων.




Σχήμα 62. Semantics.gr: προσθήκη φίλτρων

Πραγματικό παράδειγμα εμπλουτισμού τύπων τεκμηρίου συγκεκριμένου πολιτιστικού φορέα

Τύπος τεκμηρίου

- Slides (8643)
- Βίντεο (779)
- Ήχοι (867)
- Κείμενα (101)
- Μουσικά αντικείμενα (1823)
- Φωτογραφίες (1110)

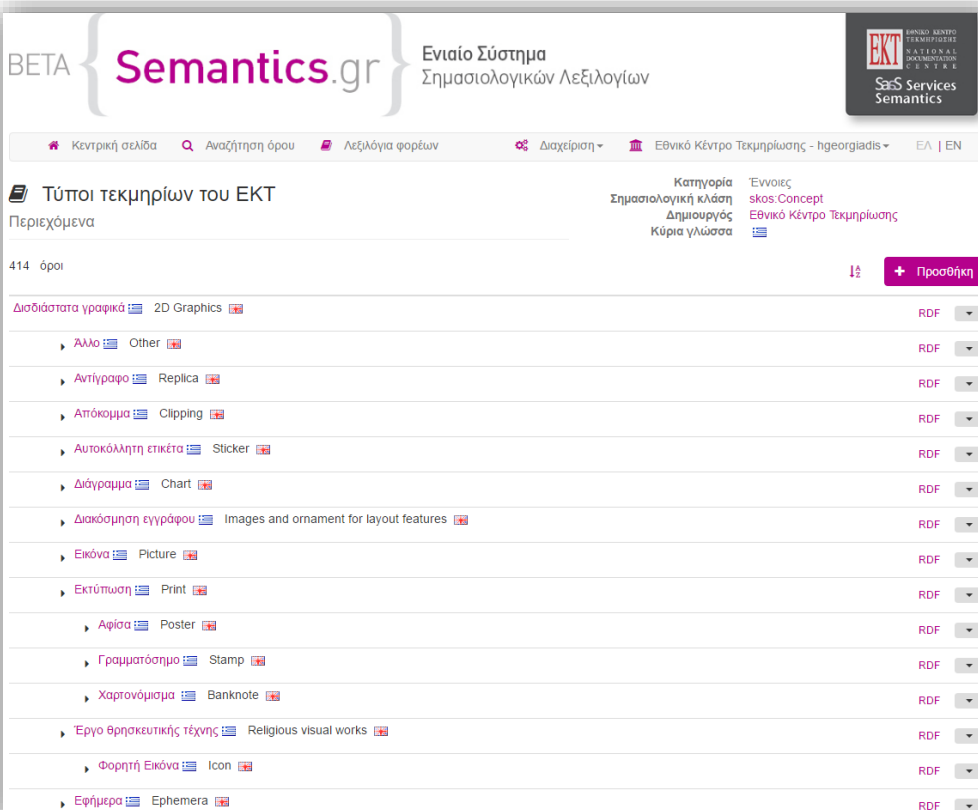


Εμπλουτισμός τύπων βάσει πρωτότυπων τύπων και θεματικών κατηγοριών συνδυαστικά

ΕΚΤ τύπος τεκμηρίου	EKT type
Δισδιάστατα γραφικά (10154)	2D Graphics (10154)
Φωτογραφική διαφάνεια (8643)	Slide (8643)
Τρισδιάστατα Αντικείμενα και Έργα Τέχνης (4107)	3D Artefacts and Realia (4107)
Φωτογραφία (1458)	Photo (1458)
Κεραμικό αντικείμενο (1294)	Ceramic ware (1294)
Ένδυμα (1104)	Clothes (1104)
Σκεύος (1073)	Vessel (1073)
Ήχογραφήσεις (867)	Sound Recordings (867)
Ύφασμα (804)	Textile (804)
Μαγνητοσκόπηση (779)	Projected and Video Material (779)
Υλικό Προβολής και Βιντεοσκοπήσεις (779)	Videorecording (779)
Κείμενο (629)	Text (629)
Φιγούρα θεάτρου σκιών (522)	Shadow puppet (522)
Έγγραφο (401)	Document (401)
Ανάγλυφο (346)	Relief (346)
Ζωγραφική (325)	Painting (325)
Ταχυδρομική κάρτα (202)	Postcard (202)
Αλληλογραφία (141)	Correspondence (141)
Χειρόγραφο (126)	Manuscript (126)
Αγιογραφία (77)	Icon (77)
Εργαλεία και εξοπλισμός (71)	Tools and equipment (71)
Μακέτα (47)	Scale model (47)
Κόσμημα (44)	Jewellery (44)
Γραμματόσημο (43)	Stamp (43)

Το εργαλείο σημασιολογικού εμπλουτισμού του **Semantics.gr** έχει την δυνατότητα να εμπλουτίσει ένα πεδίο με περιορισμένες τιμές (π.χ. τύπος τεκμηρίου) λαμβάνοντας υπόψη πολλαπλά πεδία (π.χ. τύπος τεκμηρίου και θεματική κατηγορία συνδυαστικά).

Σχήμα 63. Πραγματικό παράδειγμα εμπλουτισμού αποθετηρίου του SearchCulture.gr. Το αποθετήριο πρωτοτύπως είχε μόνο 6 διαφορετικούς τύπους χωρίς πολυγλωσσική πληροφορία. Μετά τον εμπλουτισμό, τα τεκμήρια κατανεμήθηκαν σε 55 δίγλωσσους τύπους.



BETA Semantics.gr Ενιαίο Σύστημα Σημασιολογικών Λεξιλογίων

Κεντρική σελίδα Αναζήτηση όρου Λεξιλόγια φορέων Διαχείριση Εθνικό Κέντρο Τεκμηρίωσης - hgeorgiadis EA | EN

Τύποι τεκμηρίων του ΕΚΤ

Κατηγορία: Σημασιολογική κλάση Δημιουργός: Κύρια γλώσσα

Έννοιες: skos:Concept Εθνικό Κέντρο Τεκμηρίωσης

414 όροι

Δισδιάστατα γραφικά	2D Graphics	RDF
Άλλο	Other	RDF
Αντίγραφο	Replica	RDF
Απόκομμα	Clipping	RDF
Αυτοκόλλητη ετικέτα	Sticker	RDF
Διάγραμμα	Chart	RDF
Διακόσμηση εγγράφου	Images and ornament for layout features	RDF
Εικόνα	Picture	RDF
Εκτύπωση	Print	RDF
Αφίσα	Poster	RDF
Γραμματόσημο	Stamp	RDF
Χαρτονόμισμα	Banknote	RDF
Έργο θρησκευτικής τέχνης	Religious visual works	RDF
Φορητή Εικόνα	Icon	RDF
Εφήμερα	Ephemera	RDF

Σχήμα 64. Το Λεξιλόγιο Τύπων Ψηφιακού Περιεχομένου του ΕΚΤ μέσα από την διαδικτυακή πύλη του Semantcs.gr

Πρέπει να τονιστεί ότι η πρωτότυπη τεκμηρίωση των φορέων δεν μεταβάλλεται από την διαδικασία εμπλουτισμού γεγονός πολύ σημαντικό ιδιαίτερα για τις περιπτώσεις φορέων που έκαναν πολύ αναλυτική και λεπτομερή τεκμηρίωση (λεπτομέρεια όμως που βοηθά μεν στην αναζήτηση αλλά δεν βοηθά στην ομοιόμορφη πλοήγηση ενός χρήστη στο περιεχόμενο ενός συσσωρευτή). Οι εμπλουτισμένοι τύποι διαχωρίζονται σαφώς και με διακριτό τρόπο στην διαδικτυακή πύλη του SearchCulture.gr. Εξάλλου, οι πρωτότυποι τύποι είναι αναζητήσιμοι (ευρετηριάζονται όπως όλα τα πεδία) και εξακολουθούν να εμφανίζονται τόσο στην καρτέλα τεκμηρίου όσο και ως φίλτρο αναζήτησης.

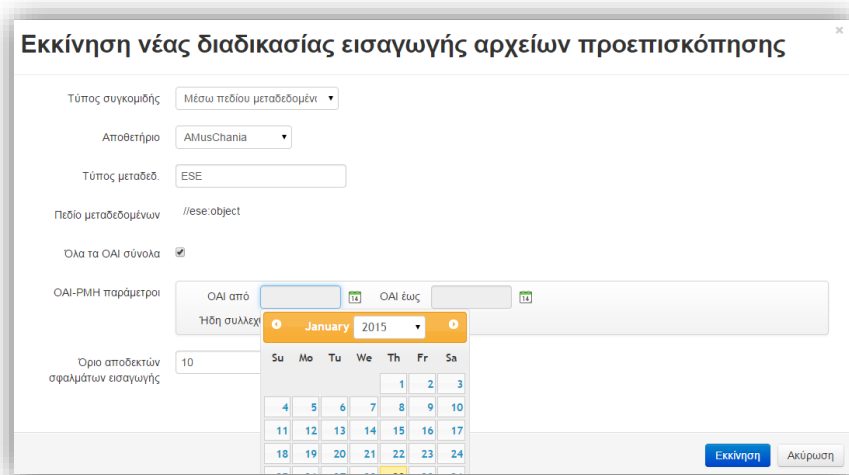
3.3.4 Διαδικασίες Εισαγωγής Αρχείων Προεπισκόπησης

Επιλέγοντας «Διαδικασίες Εισαγωγής Αρχείων Προεπισκόπησης» από το βασικό μενού, ο χρήστης μεταφέρεται στην σελίδα διαχείρισης διαδικασιών εισαγωγής εικόνων προεπισκόπησης (Σχήμα 65).

Μέσα από την σελίδα αυτή ο χρήστης μπορεί να αναζητήσει διαδικασίες εισαγωγής αρχείων προεπισκόπησης βάσει του αποθετηρίου/συλλογής, του προσδιοριστή της διαδικασίας και του εύρους ημερομηνιών μέσα στο οποίο ξεκίνησε η διαδικασία. Με το πάτημα του κουμπιού «Αναζήτηση» εμφανίζεται μία λίστα με τις διαδικασίες που πληρούν τα κριτήρια αναζήτησης. Για κάθε διαδικασία φαίνονται βασικά στοιχεία, όπως πληροφορίες σχετικά με την συγκομιδή των πρωτότυπων ψηφιακών αρχείων, την κατάσταση της διαδικασίας («Συγκομιδή σε Εξέλιξη», «Η Συγκομιδή Ολοκληρώθηκε», «Η Συγκομιδή Απέτυχε», «Εισαγωγή σε Εξέλιξη», «Ολοκληρώθηκε», «Ολοκληρώθηκε με Αποτυχίες»), καθώς και συνοπτικά στατιστικά τόσο για την φάση της Συγκομιδής (τα στατιστικά αυτά προέρχονται και ενημερώνονται on-line από το Σύστημα Harvester) όσο και για την δεύτερη φάση, αυτήν της δημιουργίας των αρχείων προεπισκόπησης και της εισαγωγής τους στο SearchCulture.gr.

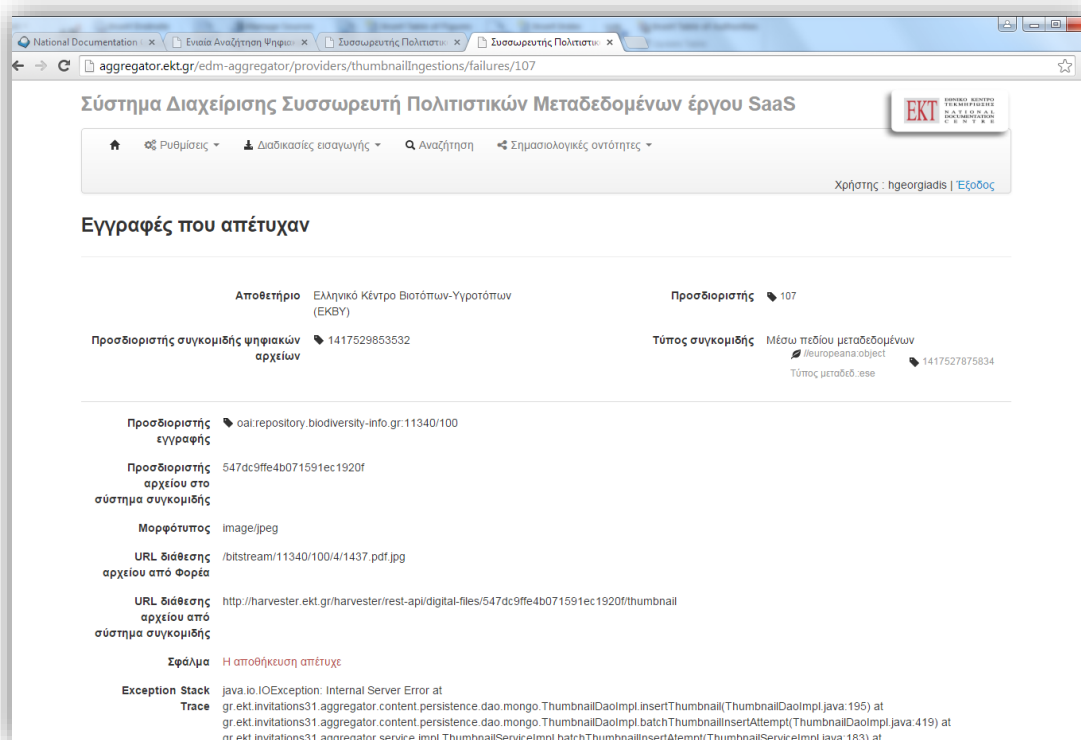
Προσδιοριστής	Ημερομηνία δημιουργίας	Προσδιοριστής συγκομιδής ψηφιακών αρχείων	Τύπος συγκομιδής	Κατάσταση	Στατιστικά συγκομιδής	Στατιστικά εισαγωγής	Ενέργειες
140	19-01-2015 11:38	1421658536842	Πεδίο μεταδεδομένων # esse:object 1421658533831 Τύπος μεταδιδ. esse OAI από-έως 01-01-2005 - 21-01-2015	Ολοκληρώθηκε	616 ψηφιακά αρχεία	616 αρχεία προεπισκόπησης	Ενέργειες
141	15-01-2015 02:24	1421055432930	Πεδίο μεταδεδομένων # esse:object 1420809723940 Τύπος μεταδιδ. esse	Ολοκληρώθηκε Το περιεχόμενο διεγρύφη	604 ψηφιακά αρχεία	604 αρχεία προεπισκόπησης	Ενέργειες

Σχήμα 65. Σελίδα διαχείρισης διαδικασιών εισαγωγής Αρχείων Προεπισκόπησης



Σχήμα 66. Φόρμα εκκίνησης νέας διαδικασίας εισαγωγής Αρχείων Προεπισκόπησης

Σε περίπτωση που για οποιονδήποτε λόγο μία διαδικασία αποτύχει, εμφανίζεται με κατάσταση «Απέτυχε» συνοδευόμενη από το μήνυμα σφάλματος. Εάν ο χρήστης πατήσει το εικονίδιο δίπλα ακριβώς από το μήνυμα σφάλματος ή εάν επιλέξει να δει τις λεπτομέρειες της συγκεκριμένης διαδικασίας (Ενέργειες→Προβολή), θα δει λεπτομερές μήνυμα καθώς και αναλυτικό Exception Stack Trace.



Σχήμα 67. Αρχεία που απέτυχαν να εισαχθούν ως εικόνες προεπισκόπησης

Μέσα από την σελίδα διαχείρισης διαδικασιών εισαγωγής αρχείων προεπισκόπησης, ο χρήστης μπορεί να εκκινήσει μια νέα διαδικασία εισαγωγής πατώντας το κουμπί «Εκκίνηση Νέας Διαδικασίας Εισαγωγής

Αρχείων Προεπισκόπησης». Με το πάτημα του κουμπιού αυτού εμφανίζεται μία φόρμα στην οποία ο χρήστης καλείται να συμπληρώσει τις παραμέτρους οι οποίες θα μεταβιβαστούν στο Σύστημα Συγκομιδής Harvester προκειμένου αυτό να συλλέξει τα ψηφιακά αρχεία και στην συνέχεια το Σύστημα Συσσωρευτή SearchCulture.gr να τα ζητήσει ως εικόνες προεπισκόπησης και να τα «χωνέψει». Η φόρμα με ενδεικτικές τιμές στα πεδία φαίνεται στο Σχήμα 66. Ο χρήστης αρχικά συμπληρώνει τον πάροχο και στην συνέχεια επιλέγει την πηγή των ψηφιακών αρχείων: 1) ανέβασμα zip αρχείο με εικόνες 2) από πεδίο μεταδεδομένων τα οποία είτε θα συλλεχθούν επί τόπου ή έχουν ήδη συλλεχθεί από το Σύστημα Συγκομιδής (στην δεύτερη περίπτωση ο χρήστης θα πρέπει να ορίσει τον προσδιοριστή της αντίστοιχης διαδικασίας συγκομιδής μεταδεδομένων (όπως εμφανίζεται στο σύστημα Harvester) 3) από ψηφιακά αρχεία που έχουν ήδη συλλεχθεί από το Σύστημα Συγκομιδής, οπότε ο χρήστης θα πρέπει να ορίσει τον προσδιοριστή της αντίστοιχης διαδικασίας συγκομιδής ψηφιακών αρχείων (όπως εμφανίζεται στο σύστημα Harvester).

Με το πάτημα του κουμπιού «Εκκίνηση» πραγματοποιείται έλεγχος της ορθότητας των παραμέτρων (π.χ. αν δόθηκαν τιμές σε υποχρεωτικά πεδία). Σε περίπτωση σφαλμάτων ως προς την εγκυρότητα των παραμέτρων, ο χρήστης ενημερώνεται και καλείται να τα διορθώσει. Διαφορετικά η διαδικασία ξεκινά, η φόρμα κλείνει και στην λίστα των διαδικασιών εμφανίζεται η νέα διαδικασία με κατάσταση «Συγκομιδή σε εξέλιξη». Το πεδία «Στατιστικά Συγκομιδής» και «Στατιστικά Εισαγωγής» για την συγκεκριμένη διαδικασία εμφανίζουν από μία μπάρα προόδου που ανανεώνεται σε τακτά χρονικά διαστήματα.

Στην περίπτωση που μια διαδικασία ολοκληρώθηκε σε κατάσταση «Ολοκληρώθηκε με αποτυχίες» ή τερματίστηκε ανεπιτυχώς (σε περίπτωση που ο αριθμός των αποτυχιών ξεπέρασε το όριο που έθεσε ο χρήστης), στο κουμπί «Ενέργειες» εμφανίζονται δύο νέες επιλογές: «Αρχεία προεπισκόπησης που δεν εισήχθησαν» και «Επανάληψη εισαγωγής υπολειπόντων».

Με την επιλογή «Αρχεία προεπισκόπησης που δεν εισήχθησαν», ο εξουσιοδοτημένος χρήστης μπορεί να δει αναλυτικά ποια ψηφιακά αρχεία απέτυχαν να μετατραπούν σε/να εισαχθούν ως εικόνες προεπισκόπησης, για ποιο λόγο απέτυχαν και σε ποια φάση (π.χ. κατά την μεταφόρτωση ή την αποθήκευση), όπως φαίνεται στο Σχήμα 67. Με την επιλογή «Επανάληψη εισαγωγής υπολειπόντων» θα ξεκινήσει η προσπάθεια εισαγωγής όλων των εικόνων προεπισκόπησης που δεν εισήχθησαν.

3.3.5 Διαχείριση Σηματολογικών Οντοτήτων

Ο εξουσιοδοτημένος χρήστης έχει την δυνατότητα να διαχειριστεί ξεχωριστά τις σηματολογικές οντότητες των EDM σηματολογικών κλάσεων edm:Agent, skos:Concept, edm:Place και edm:Timespan οι οποίες αποθηκεύονται στο Σύστημα Αποθήκευσης και Ευρετηρίασης EDM μεταδεδομένων (EDM Storage and Indexing System, με backend το Corelib της Europeana). Συγκεκριμένα, για κάθε σηματολογική κλάση δίνεται η δυνατότητα αναζήτησης, δημιουργίας, επεξεργασίας και διαγραφής των αντίστοιχων οντοτήτων. Τονίζεται ότι η κύρια πηγή δημιουργίας σηματολογικών οντοτήτων είναι οι διαδικασίες συγκομιδής/εισαγωγής σηματολογικά εμπλουτισμένων μεταδεδομένων αποθετηρίων (EDM ή DC-DS-XML).

Επιλέγοντας από το βασικό μενού «Άνθρωποι/Οργανισμοί» μεταβαίνουμε στην σελίδα αναζήτησης των οντοτήτων που ανήκουν στην σηματολογική κλάση edm:Agent, η οποία παρουσιάζεται στο Σχήμα 68. Ομοίως, επιλέγοντας 'Τόποι', 'Έννοιες' και 'Χρονικές Περιόδους' από το βασικό μενού μεταβαίνουμε στις σελίδες διαχείρισης edm:Place, skos:Concept και edm:Timespan οντοτήτων, αντίστοιχα (Σχήμα 69).

Στις σελίδες αναζήτησης σημασιολογικών οντοτήτων, ο χρήστης μπορεί να αναζητήσει μια σημασιολογική οντότητα είτε συμπληρώνοντας το URL της (γνώρισμα `rdf:about`), είτε συμπληρώνοντας μία λέξη κλειδί. Τα αποτελέσματα εμφανίζονται σελιδοποιημένα. Ο χρήστης μπορεί για κάποιο αποτέλεσμα να δει την EDM RDF/XML όψη της αντίστοιχης οντότητας, επιλέγοντας «Ενέργειες→Προβολή» (Σχήμα 70), ή να επεξεργαστεί την οντότητα, επιλέγοντας «Ενέργειες→Επεξεργασία», οπότε μεταβαίνει στην φόρμα επεξεργασίας.

URL (rdf:about)	Προτεινόμενες ονομασίες	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:3305#ent6	[Parroquia de Santa María de Roo (Noia, A Coruña)]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:3322#ent3	[Mosteiro de Santa María de Sobrado dos Monxes (A Coruña)]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:2820#ent13	[Monasterio de Santa María de Meira (Lugo)]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:3420#ent6	[Alfonso X, Rei de Castela, 1221-1284]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:4834#ent2	[Asociación de la Obra de la Santa Infancia]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:4405#ent3	[Teresa de Jesús , Santa]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:6397#ent3	[Salomé , Santa]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:3335#ent11	[María, Santa Virxen]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:6212#ent3	[Librada , santa]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:3999#ent3	[Vicente de Santa María , O C.D.]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:3265#ent4	[Colegio de Santa Catalina (Úbeda, Jaén)]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:3416#ent6	[Alfonso X, Rei de Castela, 1221-1284]	Ενέργειες
http://hispana.mcu.es/lod/oi/galiciana-bibliotecadegaliccia.xunta.es:3331#ent10	[Monasterio de Santa María de Oseira (Orense)]	Ενέργειες

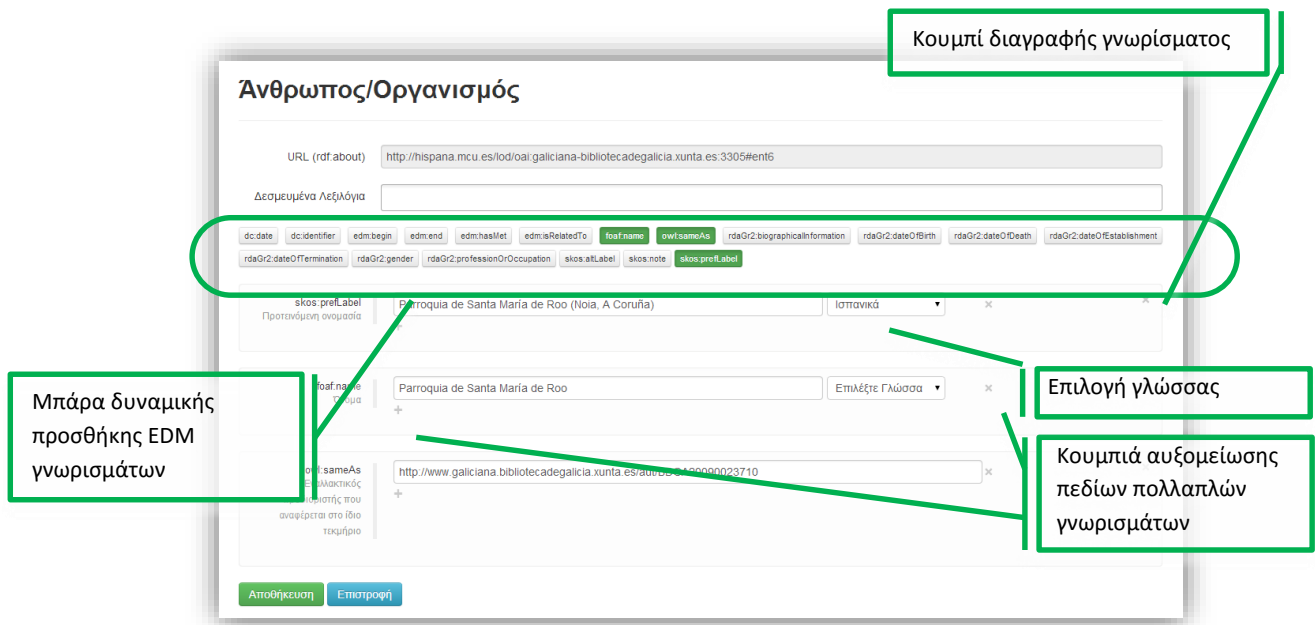
Σχήμα 68. Σελίδα σημασιολογικών οντοτήτων της κατηγορίας 'Ανθρωποι/Οργανισμοί' (edm:Agent)

URL (rdf:about)	Προτεινόμενες ονομασίες	Ενέργειες
http://semium.org/time/greek_classical	[Κλασική Περίοδος] [Classical period]	Ενέργειες

Σχήμα 69. Αντίστοιχες σελίδες σημασιολογικών οντοτήτων για `skos:Concept`, `edm:Place` και `edm:Timespan`



Σχήμα 70. Προβολή μιας σημασιολογικής οντότητας ως RDF



Σχήμα 71. Φόρμα εισαγωγής/επεξεργασίας σημασιολογικής οντότητας της κατηγορίας "Άνθρωποι/Οργανισμοί" (edm:Agent)

Στο Σχήμα 71 παρουσιάζεται η φόρμα επεξεργασίας μιας σημασιολογικής οντότητας της κλάσης «Άνθρωποι/Οργανισμοί» (edm:Agent). Ο χρήστης μπορεί να επιλέξει οποιοδήποτε πεδίο της κλάσης edm:Agent επιλέγοντας το αντίστοιχο κουμπί από την μπάρα δυναμικής προσθήκης γνωρισμάτων. Αμέσως εμφανίζεται το γνώρισμα στο οποίο ο χρήστης καλείται να συμπληρώσει τιμή. Μπορεί να αφαιρέσει το γνώρισμα επιλέγοντας το κουμπί «x», πάνω δεξιά. Αν το γνώρισμα είναι πολλαπλό εμφανίζονται αντίστοιχα κουμπιά («+», «x») για την προσθαφαίρεση πεδίων. Για ορισμένες κατηγορίες γνωρισμάτων (όπως ορίζει το EDM μοντέλο), ο χρήστης μπορεί να συνοδέψει μια τιμή με την γλώσσα. Αντίστοιχες είναι και οι φόρμες που εμφανίζονται για τις υπόλοιπες σημασιολογικές κλάσεις, με μόνη διαφορά ότι στην μπάρα προσθήκης γνωρισμάτων εμφανίζονται τα γνωρίσματα που ορίζονται στην αντίστοιχη κλάση από το EDM. Στο Σχήμα 72, φαίνεται η αντίστοιχη φόρμα για οντότητα της κλάσης «Τόπος» (edm:Place). Ειδικά για αυτήν την σημασιολογική κλάση, αν τα πεδία «Γεωγραφικό Μήκος»

(wgs84_pos:long) και «Γεωγραφικό Πλάτος» (wgs84_pos:lat) είναι συμπληρωμένα, εμφανίζεται αυτόματα πάνω δεξιά μια δυναμική εικόνα από το Google Maps με το στίγμα της τοποθεσίας στον χάρτη.

Σχήμα 72. Φόρμα εισαγωγής/επεξεργασίας σημασιολογικής οντότητας της κατηγορίας 'Τόποι' (edm:Place)

3.3.6 Αναζήτηση, Προβολή και Επεξεργασία εγγραφών τεκμηρίων.

Επιλέγοντας «Αναζήτηση» από το βασικό μενού, ο χρήστης μεταφέρεται στην σελίδα αναζήτησης τεκμηρίων που έχουν εισαχθεί στο SearchCulture.gr. Η αναζήτηση έχει ακριβώς την ίδια λειτουργικότητα με αυτήν της διαδικτυακής πύλης του (περιγράφηκε στην Ενότητα 3.2) .

Ο εξουσιοδοτημένος χρήστης του διαχειριστικού περιβάλλοντος του συσσωρευτή μπορεί να επεξεργαστεί πλήρως μια αποθηκευμένη EDM εγγραφή (που εισήχθη στα πλαίσια μιας διαδικασίας εισαγωγής μεταδεδομένων) και να την εμπλουτίσει σημασιολογικά. Αυτό το επιτυγχάνει μέσω μιας ευέλικτης και δυναμικής φόρμας φτιαγμένης ειδικά για να μοντελοποιεί έναν περίπλοκο EDM γράφο με φιλικό και δομημένο όμως τρόπο.

Στο Σχήμα 73 φαίνεται η σελίδα επεξεργασίας για μια συγκεκριμένη αποθηκευμένη εγγραφή τεκμηρίου. Η σελίδα έχει 3 καρτέλες που αντιστοιχούν στις βασικές EDM κλάσεις edm:ProvidedCHO, ore:Aggregation και (λίστα από), edm:webResource με τίτλους, «Τεκμήριο», «Συλλογή Ψηφιακών Πόρων» και «Ψηφιακοί Πόροι». Η σελίδα έχει επίσης 4 καρτέλες με λίστες σημασιολογικών αναφορών σε οντότητες των 4 σημασιολογικών κλάσεων edm:Agent, edm:Place, edm:Timespan και skos:Concept, με τίτλους «Άνθρωποι/Οργανισμοί», «Τόποι», «Χρονικές Περίοδοι» και «Έννοιες», αντίστοιχα.


Σε κάθε ενότητα που αντιστοιχεί σε EDM κλάση, ο χρήστης μπορεί εύκολα να προσθαφαιρέσει EDM γνωρίσματα της αντίστοιχης κλάσης χρησιμοποιώντας την μπάρα δυναμικής προσθήκης γνωρισμάτων. Για τα γνωρίσματα που έχουν ήδη προστεθεί (ή είχαν από πριν τιμές), τα αντίστοιχα κουμπιά εμφανίζονται με πράσινο χρώμα και είναι ανενεργά.

Τεκμήριο

Προσδιοριστής εγγραφής: 000039-693

Αποθετήριο: Αρχαιολογικό Μουσείο Χανίων

Ημ/νία εισαγωγής: 16-01-2015 11:21 Προσδιοριστής διαδικασίας εισαγωγής: 145



Τρέξιμο

dc.contributor
 dc.coverage
 dc.creator
 dc.date
 dc.description
 dc.format
 dc.identifier
 dc.language
 dc.publisher
 dc.relation
 dc.rights
 dc.source
 dc.subject
 dc.title
 dc.type
 dc.items.alternative
 dc.items.conformTo
 dc.items.created
 dc.items.extent
 dc.items.hasFormat
 dc.items.hasPart
 dc.items.hasVersion
 dc.items.isFormatOf
 dc.items.isPartOf
 dc.items.isReferenceBy
 dc.items.isVersionOf
 dc.items.isBased
 dc.items.medium
 dc.items.provenance
 dc.items.references
 dc.items.replaces
 dc.items.requires
 dc.items.spatial
 dc.items.tableOfContents
 dc.items.temporal
 edm.currentLocation
 edm.hasMet
 edm.hasType
 edm.incorporates
 edm.isDerivativeOf
 edm.isIntrinsicSequence
 edm.isRelatedTo
 edm.isRepresentationOf
 edm.isSimilarTo
 edm.isSuccessorOf
 edm.realizes
 edm.rights
 edm.type
 edm.userTag
 edm.wasPresentAt
 edm.year
 owl:sameAs

dc:type Τύπος

dc:spatial Γεωγραφική κάλυψη

dc:subject Θέμα

dc:creator Δημιουργός

dc:source Πηγή Επιλέξτε Γλώσσα

dc:contributor Συνεισφέρτής

dc:description Περιγραφή Επιλέξτε Γλώσσα

dc:rights Δικαιώματα Επιλέξτε Γλώσσα

dc:publisher Εκδότης

dc:temporal Χρονική κάλυψη

dc:date Ημερομηνία

dc:identifier Προσδιοριστής Επιλέξτε Γλώσσα

dc:language Γλώσσα Επιλέξτε Γλώσσα

dc:title Τίτλος Επιλέξτε Γλώσσα

edm:type Ευρωπαϊκά τύπος

Μπάρα δυναμικής προσθήκης EDM γνωρισμάτων

Κουμπιά εναλλαγής τιμής από λεκτική περιγραφή σε σημασιολογική αναφορά και αντίστροφα

Κουμπί διαγραφής γνωρισματος

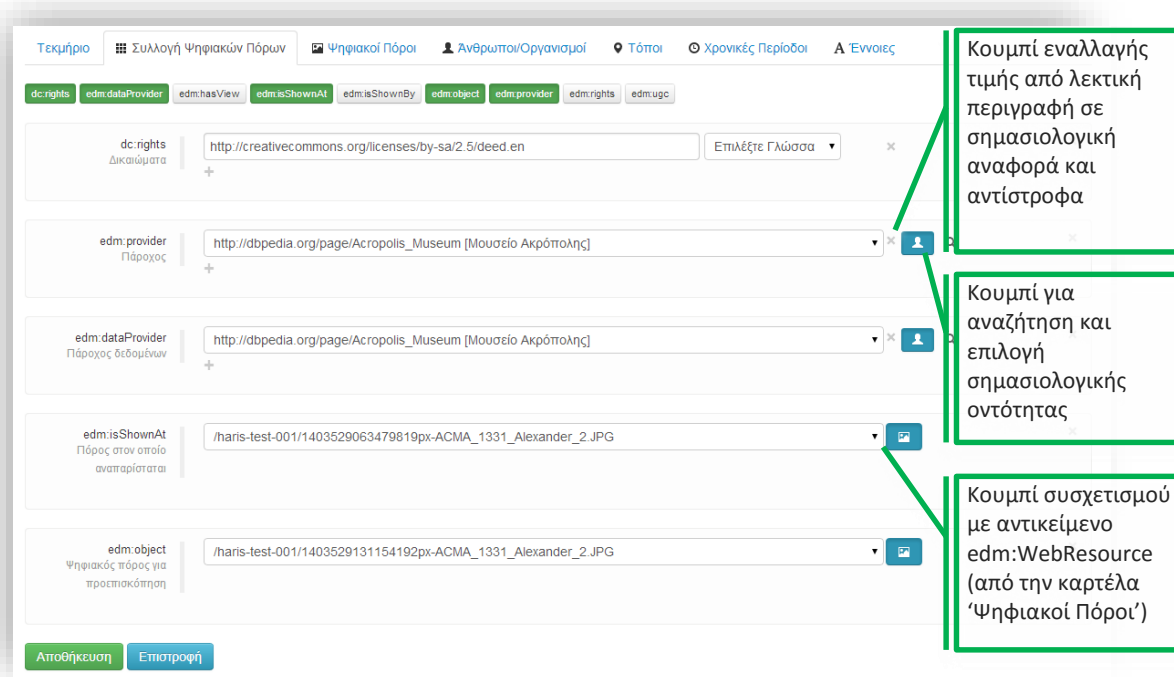
Επιλογή γλώσσας

Κουμπιά αυξομείωσης πεδίων πολλαπλών γνωρισμάτων

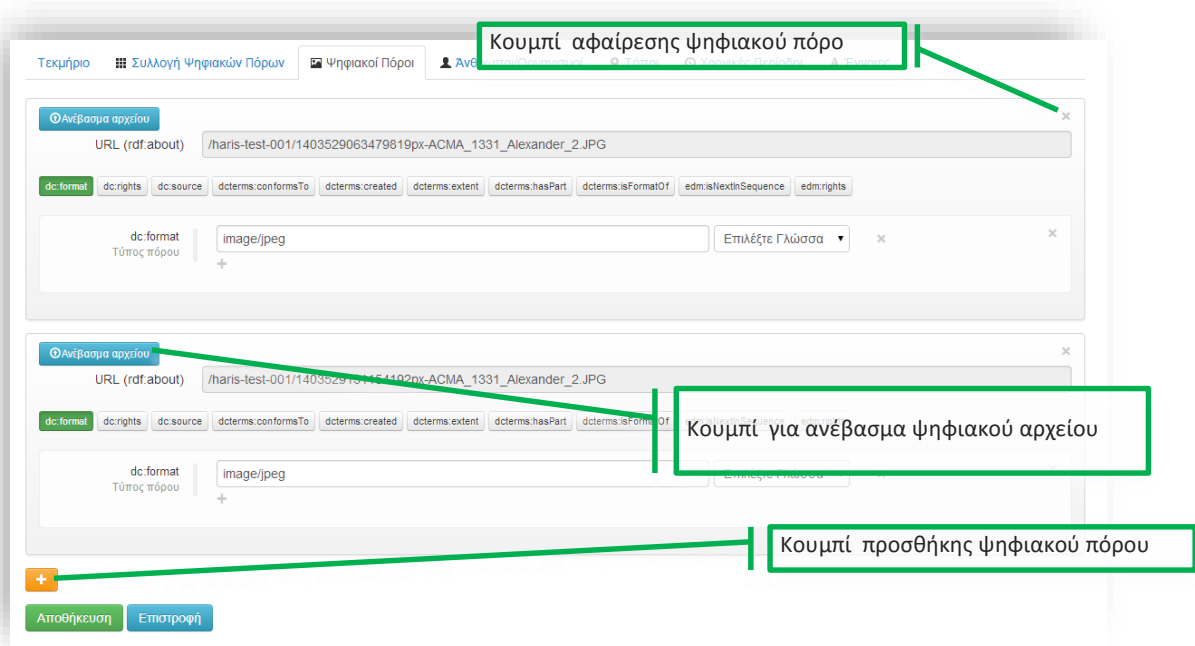
Σχήμα 73. Φόρμα επεξεργασίας EDM τεκμηρίου - Κартέλα 'Τεκμήριο' (edm:ProvidedCHO)

Για να αφαιρέσει κανείς ένα γνώρισμα αρκεί να πατήσει το κουμπί «x» πάνω δεξιά του γνωρίσματος (το πεδίο εξαφανίζεται και το αντίστοιχο κουμπί στην μπάρα γνωρισμάτων γίνεται ξανά γκρι και ενεργό). Όπως ορίζεται στο EDM μοντέλο, κάποια γνωρίσματα είναι πολλαπλά. Αυτά συνοδεύονται από τα χαρακτηριστικά «+» και «x» κουμπιά που επιτρέπουν την αυξομείωση της πολλαπλότητά τους. Επίσης κάποια πεδία συνοδεύονται από επιλογή γλώσσας (τα πεδία αυτά και πάλι ορίζονται από το EDM μοντέλο).

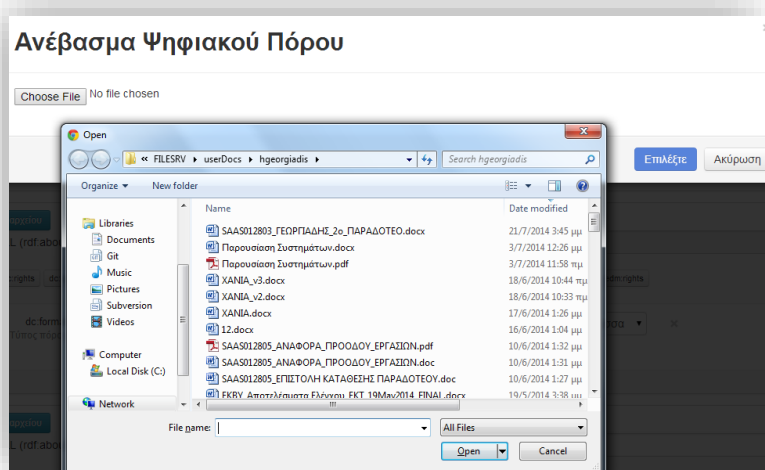
Στο Σχήμα 74 βλέπουμε την καρτέλα που αντιστοιχεί στην EDM κλάση ore: Aggregation. Για τα πεδία edm:object, edm:isShownAt, edm:isShownBy και edm:hasView, τα οποία παίρνουν τιμές URL διευθύνσεις ψηφιακών πόρων, ο χρήστης μπορεί είτε να συμπληρώσει απλά το URL είτε, εφόσον επιθυμεί να περάσει ποιο λεπτομερείς πληροφορίες για τον ψηφιακό πόρο ως edm:WebResource αντικείμενο, να επιλέξει ένα ψηφιακό πόρο που έχει καταχωρήσει στην καρτέλα «Ψηφιακοί Πόροι», η οποία φαίνεται στο Σχήμα 75. Στην καρτέλα αυτή, ο χρήστης πατώντας το κουμπί Προσθήκη («+») που βρίσκεται στο κάτω μέρος μπορεί να καταχωρήσει ένα αντικείμενο της κλάσης edm:WebResource. Η ενότητα της οντότητας που εμφανίζεται έχει ίδια λειτουργικότητα με κάθε ενότητα αντικείμενου EDM κλάσης (μπάρα δυναμικής προσθήκης γνωρισμάτων κλπ.). Για έναν νέο ψηφιακό πόρο ο χρήστης έχει την δυνατότητα είτε να συμπληρώσει στο πεδίο edm:about το URL του πόρου είτε να ανεβάσει ένα ψηφιακό αρχείο από τον υπολογιστή του (Σχήμα 76). Στην δεύτερη περίπτωση, μετά την αποθήκευση της εγγραφής, θα αντιστοιχηθεί μοναδική URL διεύθυνση για τον ψηφιακό πόρο η οποία θα εξυπηρετείται μόνιμα από το σύστημα. Σημειώνεται ότι αυτή η δυνατότητα προορίζεται μόνο για την μελλοντική λειτουργία του συστήματος ως Αποθετήριο.



Σχήμα 74. Φόρμα επεξεργασίας EDM τεκμηρίου - Καρτέλα 'Συλλογή Ψηφιακών Πόρων' (ore:Aggregation)



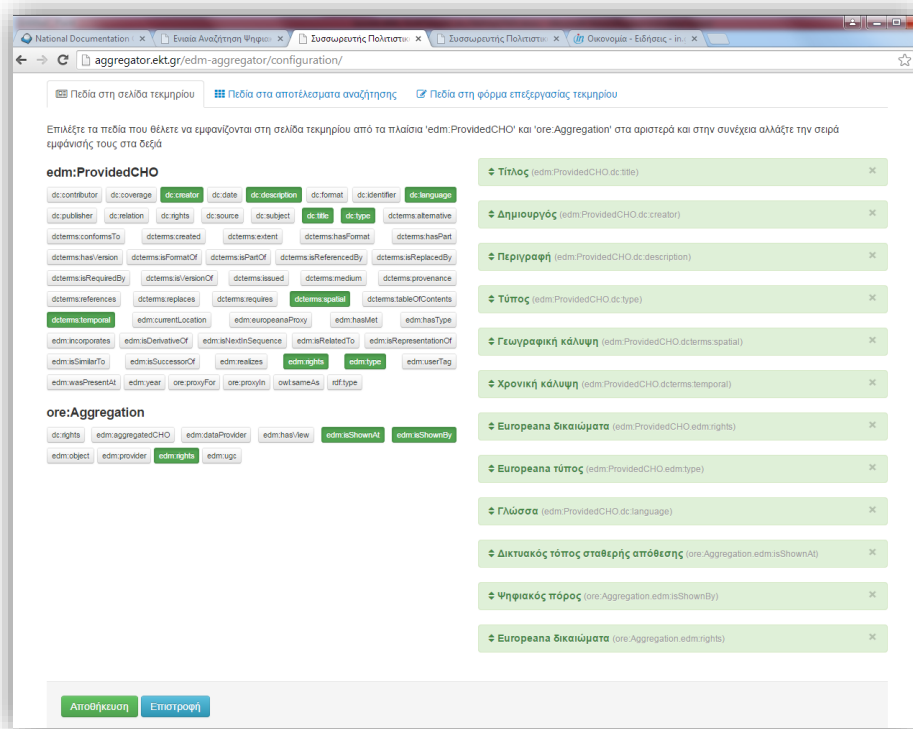
Σχήμα 75. Φόρμα επεξεργασίας EDM τεκμηρίου - Καρτέλα 'Ψηφιακοί Πόροι' (emd:WebResource)



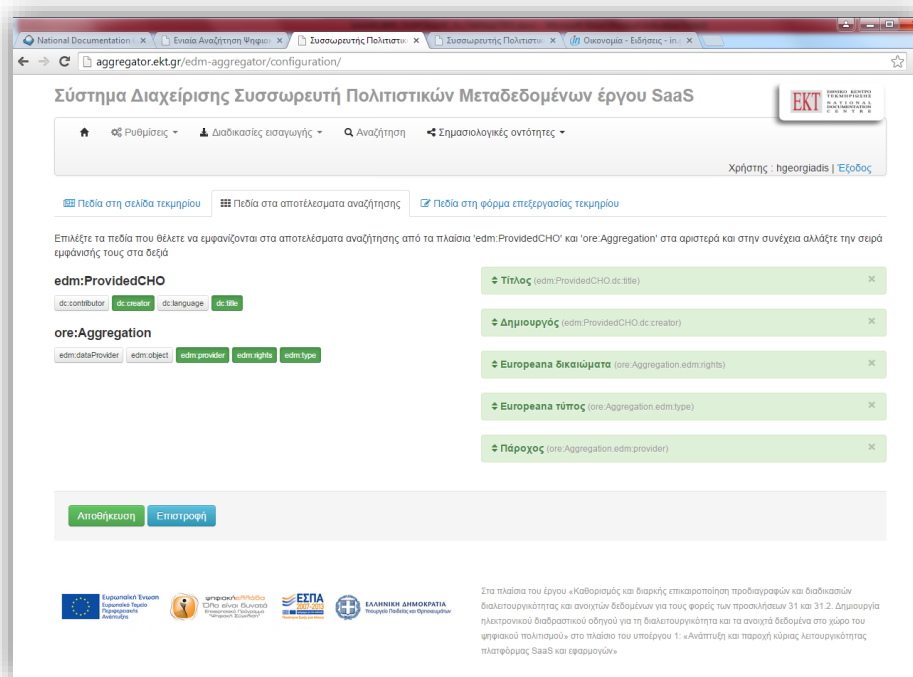
Σχήμα 76. Ανέβασμα αρχείου κατά την δημιουργία νέου Ψηφιακού Πόρου (edm:WebResource)

3.3.7 Παραμετροποίηση EDM

Ο εξουσιοδοτημένος χρήστης μπορεί να παραμετροποιήσει την διαδικτυακή πύλη καθώς και την φόρμα επεξεργασίας εγγραφών του διαχειριστικού περιβάλλοντος, επιλέγοντας από το βασικό μενού «Ρυθμίσεις»→«Παραμετροποίηση EDM πεδίων». Στην σελίδα που θα εμφανιστεί, μέσα από την καρτέλα «Πεδία στην καρτέλα τεκμηρίου» (Σχήμα 77) θα μπορεί να αλλάζει τα πεδία και την σειρά εμφάνισής τους στην σελίδα παρουσίασης τεκμηρίων (βλ. Ενότητα 3.2).



Σχήμα 77. Επιλογή πεδίων και σειράς εμφάνισής τους στη σελίδα παρουσίασης τεκμηρίου



Σχήμα 78. Επιλογή πεδίων και σειράς εμφάνισής τους στα αποτελέσματα αναζήτησης

Στην καρτέλα «Πεδία στα αποτελέσματα αναζήτησης» (Σχήμα 78) θα μπορεί να αλλάζει τα πεδία και την σειρά εμφάνισής στην σελίδα αποτελεσμάτων της μηχανής αναζήτησης (βλ. Ενότητα 3.2).

Τέλος, μέσα από την καρτέλα «Πεδία στην φόρμα επεξεργασίας τεκμηρίου» Θα μπορεί επίσης να παραμετροποιεί την φόρμα επεξεργασίας τεκμηρίου (βλ. ενότητα 3.3.6): να ορίζει ποια EDM πεδία θα υποστηρίζουν σημασιολογικές αναφορές, για ποιά θα χρησιμοποιείται πλαίσιο κειμένου (text area) αντί απλού πεδίου και ποιών οι τιμές θα υπόκεινται σε απόπειρα μετάφρασης.

The screenshot shows a web browser window with the URL `agggregator.ekt.gr/edm-agggregator/configuration/`. The page is titled "Πεδία στην φόρμα επεξεργασίας τεκμηρίου" and contains several configuration sections:

- Πεδία που υποστηρίζουν αναφορές σε έννοιες:** `dc:subject`, `dc:type`
- Πεδία που υποστηρίζουν αναφορές σε ανθρώπους/οργανισμούς:** `dc:contributor`, `dc:creator`, `dc:publisher`
- Πεδία που υποστηρίζουν αναφορές σε τοποθεσίες:** `dc:coverage`, `dcterms:spatial`, `edm:currentLocation`
- Πεδία που υποστηρίζουν αναφορές σε χρονικές περιόδους:** `dc:coverage`, `dc:date`, `dcterms:created`, `dcterms:temporal`
- Πεδία που υποστηρίζουν αναφορές σε ψηφιακούς πόρους:** `edm:hasView`, `edm:isShownAt`, `edm:isShownBy`, `edm:object`
- Πεδία κειμένου που θα εμφανίζονται ως πλαίσια κειμένου (κατάλληλο για εισαγωγή κειμένου πολλαπλών γραμμών):** `dc:description`, `dc:title`
- Πεδία που δεν θα εμφανίζονται στην φόρμα:** `edm:aggregatedCHO`, `edm:europaanaProxy`, `ore:proxyFor`, `ore:proxyIn`, `rdf:type`
- Πεδία των οποίων οι τιμές θα μεταφράζονται:** `edm:rights`, `dc:language`, `edm:type`

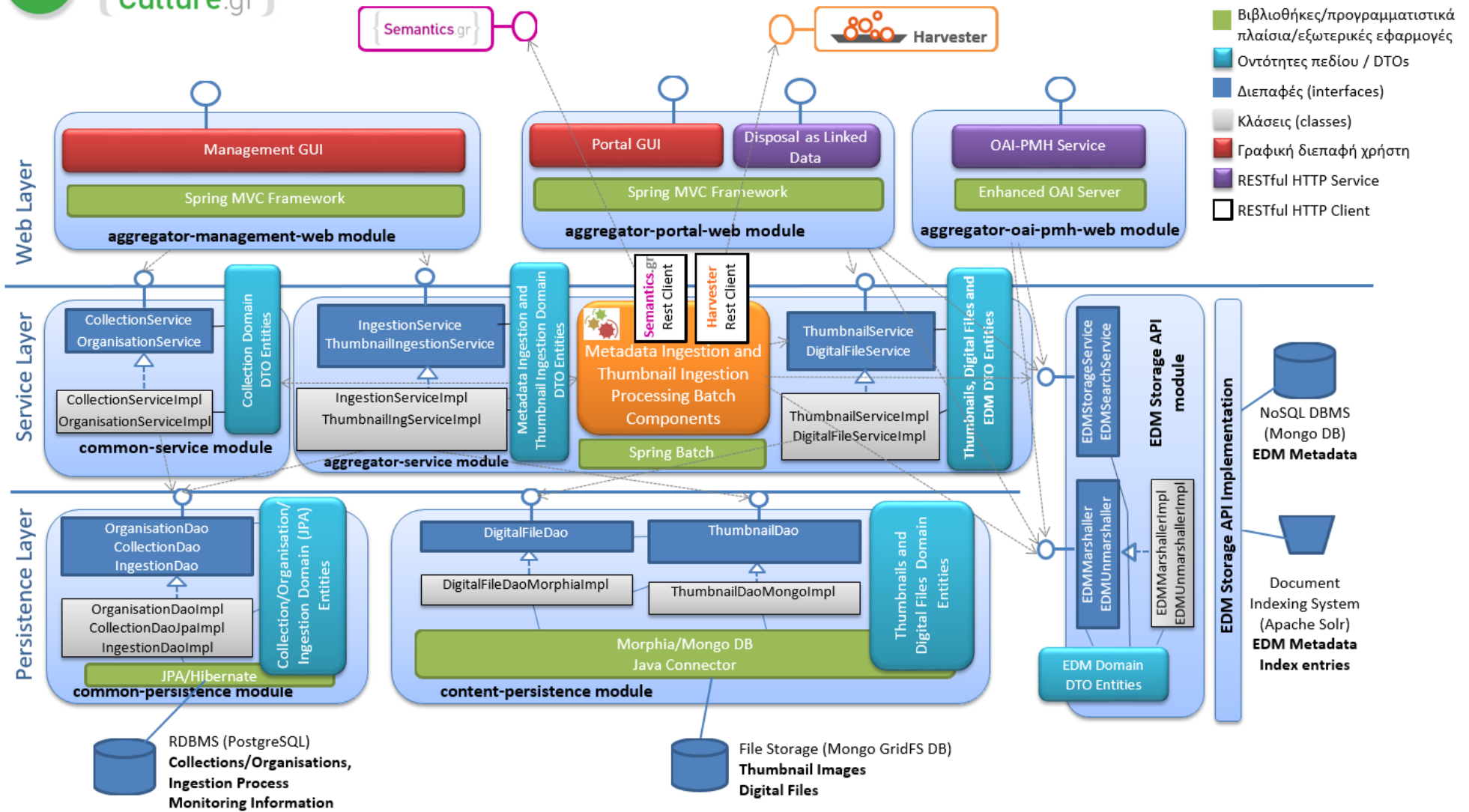
At the bottom of the page, there are two buttons: "Αποθήκευση" (Save) and "Επιστροφή" (Back).

Σχήμα 79. Παραμετροποίησης φόρμας επεξεργασίας εγγραφής τεκμηρίου

3.4 SearchCulture.gr: αρχιτεκτονική συστήματος

Το Σύστημα Συσσωρευτή /Ενιαίας Αναζήτησης SearchCulture.gr είναι μία Java εφαρμογή ιστού πακεταρισμένη ως τρία ξεχωριστά war αρχεία (διαδικτυακή πύλη, διαχειριστικό περιβάλλον, ΟΑΙ-ΡΜΗ υπηρεσία) και μπορεί να εγκατασταθεί σε οποιονδήποτε Servlet container, όπως ο Apache Tomcat. Ταυτόχρονα είναι ένα Maven project αποτελούμενο από τα εξής 9 modules: common-persistence, content-persistence, edm-storage-api, edm-storage-implementation, common-service, aggregator-service, aggregator-management-web, aggregator-portal-web και aggregator-oai-pmh-web. Το Σχήμα 80 παρουσιάζει την επικαιροποιημένη αρχιτεκτονική.

Η σύστημα SearchCulture.gr σχεδιάστηκε σύμφωνα με τις αρχές της πολυ-διαστρωματωμένης αρχιτεκτονικής (multi-layered architecture) και χωρίζεται στα στρώματα persistence (ή αλλιώς infrastructure layer), service (ή αλλιώς business layer) και presentation (εν προκειμένου web layer). Τα modules common-persistence, content-persistence, edm-storage-api ανήκουν στο persistence layer, τα modules common-service και aggregator-service ανήκουν στο service layer ενώ τα aggregator-portal-web, aggregator-management-web και aggregator-oai-pmh-web ανήκουν στο web (presentation) layer. Το module edm-storage-implementation υλοποιεί το edm-storage-api ως edm storage back-end. Και τα δύο αυτά modules αποτελούν επίσης μέρος του persistence layer. Όλα τα modules χρησιμοποιούν το Spring framework [13] για Dependency Injection και για την διαχείριση των διεπαφών (interfaces) και των κλάσεων που τις υλοποιούν. Για τον συντονισμό και ενορχήστρωση των κλάσεων που αναπτύχθηκαν και που υλοποιούν τις ροές συγκομιδής/εισαγωγής μεταδεδομένων και εικόνων προεπισκόπησης χρησιμοποιήθηκε το Spring Batch Framework [15]. Για την ανάπτυξη της εφαρμογής ιστού χρησιμοποιήθηκε το Spring MVC framework [14] σε συνδυασμό με front-end τεχνολογίες όπως JQuery και Bootstrap. Για την ανάπτυξη της ΟΑΙ-ΡΜΗ υπηρεσίας ιστού χρησιμοποιήθηκε η βιβλιοθήκη Enhanced OAI Server [23] που ανέπτυξε το ΕΚΤ και που έχει χρησιμοποιηθεί ήδη σε πληθώρα αποθετηρίων που δημιούργησε το ΕΚΤ για την διάθεση του περιεχομένου τους για συγκομιδή.



Σχήμα 80. Αρχιτεκτονική του Συστήματος Συσσωρευτή/Ενιαίας Αναζήτησης SearchCulture.

3.4.1 Persistent Layer

Το persistence layer παρέχει τις βασικές CRUD μεθόδους καθώς και μεθόδους αναζήτησης για τις οντότητες πεδίου που αποθηκεύονται μόνιμα σε βάσεις δεδομένων. Το SearchCulture.gr χρησιμοποιεί άμεσα τρεις βάσεις δεδομένων – οι οποίες απεικονίζονται σχηματικά στο Σχήμα 81– και ένα σύστημα ευρετηρίασης:

1. Μία σχεσιακή βάση (PostgreSQL) για την αποθήκευση πληροφορίας σχετικά με τους φορείς/παρόχους και τις συλλογές τους, τις διαδικασίες εισαγωγής μεταδεδομένων και εικόνων προεπισκόπησης, την καταγραφή σφαλμάτων κατά την εισαγωγή και την παραμετροποίηση του συστήματος (οντότητες πεδίου Organisation, Collection, MetadataIngestion / ThumbnailIngestion, MetadataIngestionFailure / ThumbnailIngestionFailure και και Configuration, αντίστοιχα).
2. Στα πλαίσια του EDM Storage Implementation που υλοποιεί το EDM Storage API, μία noSQL Mongo DB βάση δεδομένων για την αποθήκευση EDM μεταδεδομένων.
3. Στα πλαίσια του EDM Storage Implementation που υλοποιεί το EDM Storage API, ένα σύστημα ευρετηρίασης Apache Solr, για την ευρετηρίαση των EDM μεταδεδομένων.
4. Ένα File Storage (Mongo DB GridFS) για την αποθήκευση εικόνων προεπισκόπησης και ψηφιακών αρχείων.

Στο persistent layer, για όλες τις βάσεις δεδομένων που χρησιμοποιούνται, ορίζονται οι οντότητες πεδίου ως Java POJOs οι οποίες αντιστοιχίζονται με τις αντίστοιχες δομές (PostgreSQL πίνακες ή Mongo DB collections) μέσω ειδικών προγραμματιστικών προτύπων/βιβλιοθηκών (JPA ή Morphia [19]). Για κάθε βασική οντότητα πεδίου ορίζεται από ένα αντικείμενο με CRUD μεθόδους και μεθόδους αναζήτησης, σύμφωνα με το σχεδιαστικό πρότυπο DAO [16].

Ο Πίνακας V περιλαμβάνει τις βασικές οντότητες πεδίου που αποθηκεύονται στις βάσεις δεδομένων και τα αντίστοιχα DAOs, Services και DTOs.

Σχεσιακή Βάση

Οι πίνακες ORGANISATION, COLLECTION, COLLECTION_OAI_SET_DC_TYPE και PROVIDR_OAI_SET_IMAGE_TYPE είναι μέρος του υποσυστήματος Διαχείρισης Δεδομένων Φορέων και Συλλογών (Organisations and Collections Registry) το οποίο είναι κοινό κομμάτι των συστημάτων SearchCulture.gr, του Front-End του Validator και του Front-End του Preservator. Οι πίνακες CONFIGURATION, INGESTION και INGESTION_FAILURE χρησιμοποιούνται αποκλειστικά από το SearchCulture.gr.

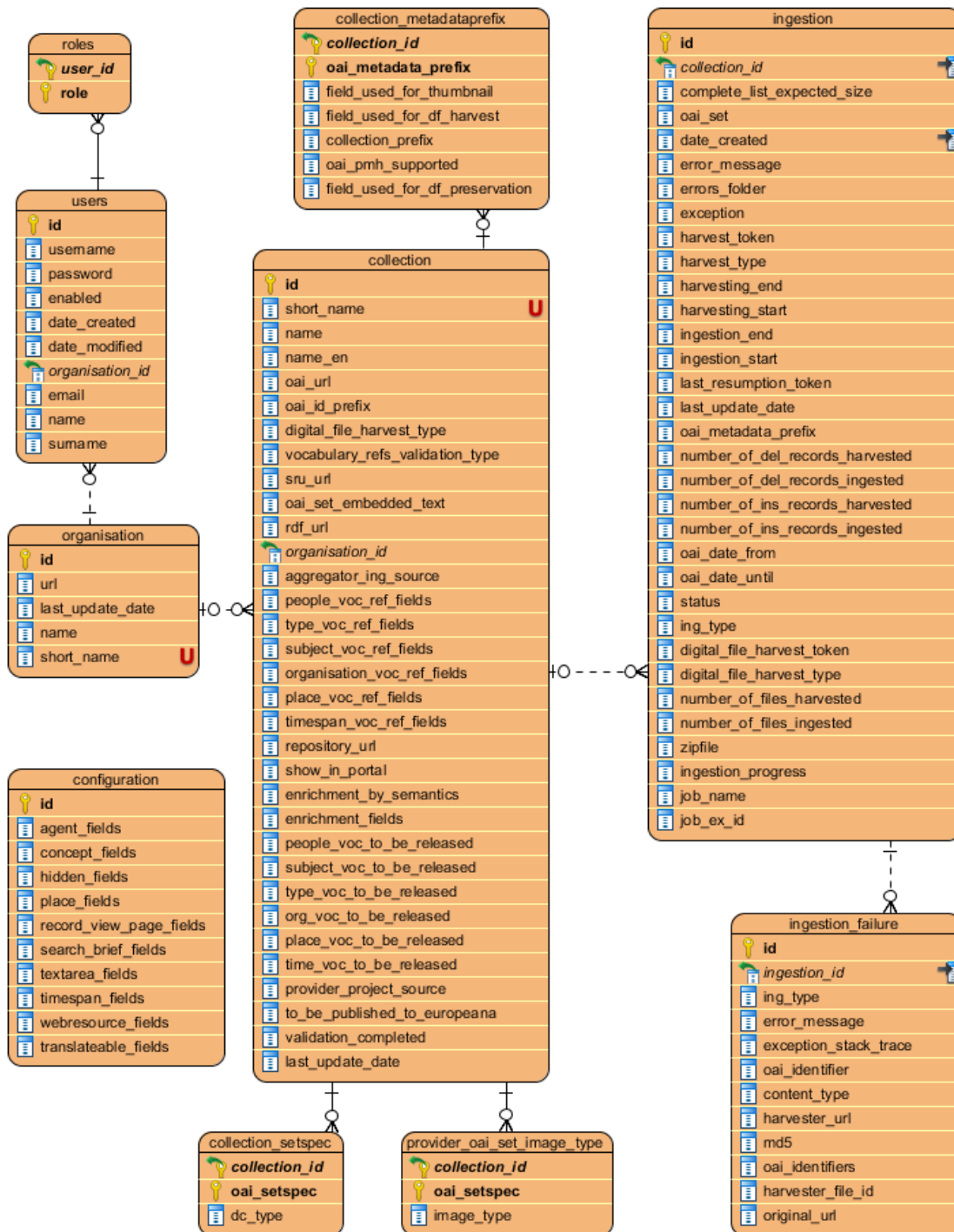
Η οντότητα ORGANISATION αποθηκεύει τους εγγεγραμμένους φορείς με κυριότερα πεδία, το όνομα (NAME) και μία σύντομη ονομασία (SHORT_NAME).

Ο πίνακας COLLECTION αποθηκεύει πληροφορίες για τις συλλογές/αποθετήρια των φορέων και περιλαμβάνει:

- κάποια βασικά πεδία όπως το όνομα (NAME) και η σύντομη ονομασία (SHORT_NAME)
- πεδία απαραίτητα για την συγκομιδή μεταδεδομένων, όπως η διεύθυνση της OAI-PMH υπηρεσίας (OAI_URL)

- μία ένα-προς-πολλά συσχέτιση με τον πίνακα `COLLECTION_METADATAPREFIX`, ο οποίος αποθηκεύει τους τύπους μεταδεδομένων που υποστηρίζει η κάθε συλλογή (metadata formats του πρωτοκόλλου OAI-PMH) μαζί με πληροφορίες για την συγκομιδή ψηφιακών αρχείων. Συγκεκριμένα, τα πεδία `FIELD_USED_FOR_THUMBNAILS`, `FIELD_USED_FOR_DF_HARVEST` και `FIELD_USED_FOR_DF_PRESERVATION` ορίζουν ως XPath εκφράσεις τα πεδία μεταδεδομένων του αντίστοιχου τύπου (π.χ. `//ese:isShownBy` για τύπο μεταδεδομένων `ese`) που περιέχουν URL ψηφιακών αρχείων για δημιουργία εικόνων προεπισκόπησης, για την ανάκτηση ψηφιακών αρχείων προς έλεγχο και για την ανάκτηση ψηφιακών αρχείων για ασφαλή διαφύλαξη, αντίστοιχα.
- πεδία και συσχετίσεις που είναι απαραίτητα για τον έλεγχο συμμόρφωσης με τις προδιαγραφές διαλειτουργικότητας. Μερικά ενδεικτικά πεδία είναι η διεύθυνση της SRU υπηρεσίας αναζήτησης (`SRU`), η διεύθυνση που επιστρέφει το σύνολο των μεταδεδομένων ως Linked Data (`RDF_URL`), ο τρόπος συλλογής ψηφιακών αρχείων (`DIGITAL_FILE_HARVEST_TYPE`), ο τρόπος ελέγχου αναφορών σε δεσμευμένα λεξιλόγια (`VOCABULARY_REFS_VALIDATION_TYPE`) και, τέλος, για κάθε κατηγορία λεξιλογίου, σε ποια πεδία εφαρμόζεται το συγκεκριμένο λεξιλόγιο (`PEOPLE_VOC_REF_FIELDS`, `TYPE_VOC_REF_FIELDS`, `SUBJECT_VOC_REF_FIELDS`, `ORGANISATION_VOC_REF_FIELDS`, `PLACE_VOC_REF_FIELDS`, `TIMESPAN_VOC_REF_FIELDS`). Η ένα-προς-πολλά συσχέτιση με τον πίνακα `PROVIDER_OAI_SET_IMAGE_TYPE` αποθηκεύει ανά συλλογή αντιστοιχίσεις OAI συνόλων (`setSpecs`) με κατηγορίες ψηφιοποιημένων εικόνων. Τέλος, η ένα-προς-πολλά συσχέτιση με τον πίνακα `COLLECTION_SETSPEC` αποθηκεύει ανά συλλογή αντιστοιχίσεις OAI συνόλων (`setSpecs`) με τύπους τεκμηρίων (τιμές στο πεδίο μεταδεδομένων `dc:type`).
- πεδία απαραίτητα για την εισαγωγή μεταδεδομένων στον συσσωρευτή, όπως ο τύπος μεταδεδομένων (π.χ. `ESE`, `EDM`) ή ο συνδυασμός τύπων μεταδεδομένων (π.χ. `DC-DS-XML` και `ESE`) που χρησιμοποιείται ως πηγή για την εισαγωγή (`AGGREGATOR_ING_SOURCE`).
- πεδία με πληροφορίες απαραίτητες για τον σημασιολογικό εμπλουτισμό και την σωστή διασύνδεση με το σύστημα `Semantics.gr`, όπως αν η συλλογή εμπλουτίζεται σημασιολογικά (`ENRICHMENT_BY_SEMANTICS`) και, εάν ναι, ποια πεδία υπόκεινται σε εμπλουτισμό και επομένως υπάρχουν για αυτά κανόνες αντιστοίχισης σε όρους λεξιλογίων στο εργαλείο εμπλουτισμού του `Semantics.gr` (`ENRICHMENT_FIELDS`)
- πεδία με διάφορες διαχειριστικές πληροφορίες, όπως αν η συλλογή προέρχεται από τις προσκλήσεις 31 και 31.2 (`PROVIDER_PROJECT_SOURCE`), αν έχει ολοκληρώσει τον έλεγχο (`VALIDATION_COMPLETED`), αν εμφανίζεται στην διαδικτυακή πύλη του `SearchCulture.gr` (`SHOW_IN_PORTAL`) και αν θα διατίθεται μέσα από την υπηρεσία OAI-PMH για συγκομιδή από την `Europeana` (`TO_BE_PUBLISHED_TO EUROPEANA`).

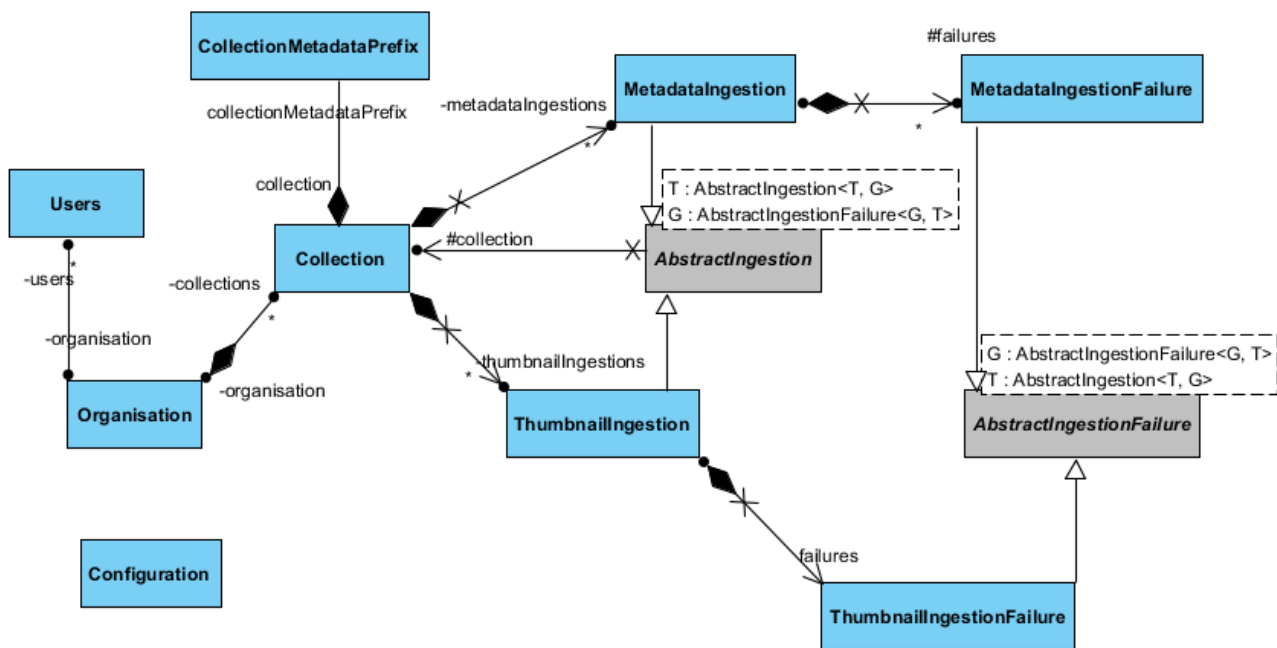
Η πίνακας `INGESTION` αποθηκεύει διαδικασίες εισαγωγής μεταδεδομένων ή εικόνων προεπισκόπησης και περιλαμβάνει στοιχεία για την κατάσταση της διαδικασίας, διάφορα στατιστικά και μια αναφορά στην διαδικασία συγκομιδής μέσω του προσδιοριστή της η οποία εκτελείται από το Σύστημα Συγκομιδής `Harvester`. Το πεδίο `TYPE` με επιτρεπόμενες τιμές «M» και «T» δηλώνει αν μια εγγραφή αφορά διαδικασία εισαγωγής μεταδεδομένων ή εικόνων προεπισκόπησης, αντίστοιχα.



Σχήμα 81. Διάγραμμα οντοτήτων συσχετίσεων της σχεσιακής βάσης του SearchCulture.gr.

Τέλος, ο πίνακας CONFIGURATION θα περιέχει πάντα μόνο μια εγγραφή με πεδία που επιτρέπουν μια σειρά από παραμετροποιήσεις, όπως ο καθορισμός των πεδίων και της σειράς εμφάνισής τους τόσο στην σελίδα παρουσίασης τεκμηρίων (RECORD_VIEW_PAGE_FIELDS) όσο και στην σελίδα αποτελεσμάτων της μηχανής αναζήτησης (SEARCH_BRIEF_FIELDS), ποια EDM πεδία θα υποστηρίζουν σημασιολογικές αναφορές (AGENT_FIELDS, CONCEPT_FIELDS, PLACE_FIELDS, TIMESPAN_FIELDS), για ποιά πεδία θα χρησιμοποιείται πλαίσιο κειμένου αντί απλού πεδίου (TEXTAREA_FIELDS) και ποιών πεδίων οι τιμές θα υπόκεινται σε απόπειρα μετάφρασης (TRANSLATEABLE_FIELDS).

Το διάγραμμα κλάσεων των οντοτήτων πεδίου της σχεσιακής βάσης παρουσιάζεται στο Σχήμα 82.



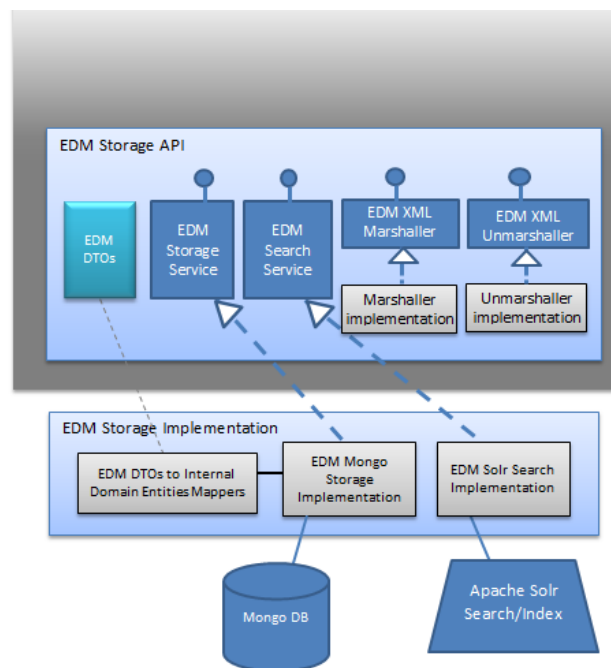
Σχήμα 82: UML διάγραμμα κλάσεων οντοτήτων πεδίου Σχεσιακής Βάσης

EDM Storage Implementation: Mongo DB βάση και Apache Solr

Κατά την φάση της ανάλυσης και σχεδιασμού του Συστήματος Συσσωρευτή SearchCulture.gr και με δεδομένο ότι ο συσσωρευτής θα έπρεπε να αποθηκεύει τα μεταδεδομένα σε μορφή συμβατή με το EDM μοντέλο [7], αποφασίστηκε η χρησιμοποίηση ενός τμήματος του υποσυστήματος Corelib του συστήματος συσσωρευτή της Europeana ως υποσύστημα για την αποθήκευση και ευρετηρίαση/αναζήτηση EDM μεταδεδομένων. Για να διασφαλιστεί η όσο το δυνατόν μικρότερη εξάρτηση του Συστήματος Συσσωρευτή SearchCulture.gr στο Corelib δημιουργήθηκε ένα ενδιάμεσο API που ονομάστηκε EDM Storage API και της αντίστοιχης υλοποίησης του η οποία σχεδιάστηκε να λειτουργεί αποκλειστικά πάνω από το Corelib. Ωστόσο, κατά την διάρκεια της ανάπτυξης και λόγω επικείμενων επεκτάσεων, αποφασίστηκε ότι ήταν πλέον προτιμότερη η πλήρης αποδέσμευση του Συστήματος SearchCulture.gr από το Corelib της Europeana και η αντικατάστασή του από ένα νέο συμπαγές και ευέλικτο υποσύστημα που θα εξακολουθούσε να βασίζεται σε μία Mongo DB βάση – για την αποθήκευση – και μια Apache Solr εγκατάσταση – για την ευρετηρίαση – των EDM μεταδεδομένων, προσαρμοσμένο όμως στις συγκεκριμένες λειτουργικές ανάγκες του συσσωρευτή. Η μετάβαση πραγματοποιήθηκε ιδιαίτερα εύκολα επιβεβαιώνοντας την ορθότητα της σχεδιαστικής επιλογής της δημιουργίας του EDM Storage API.

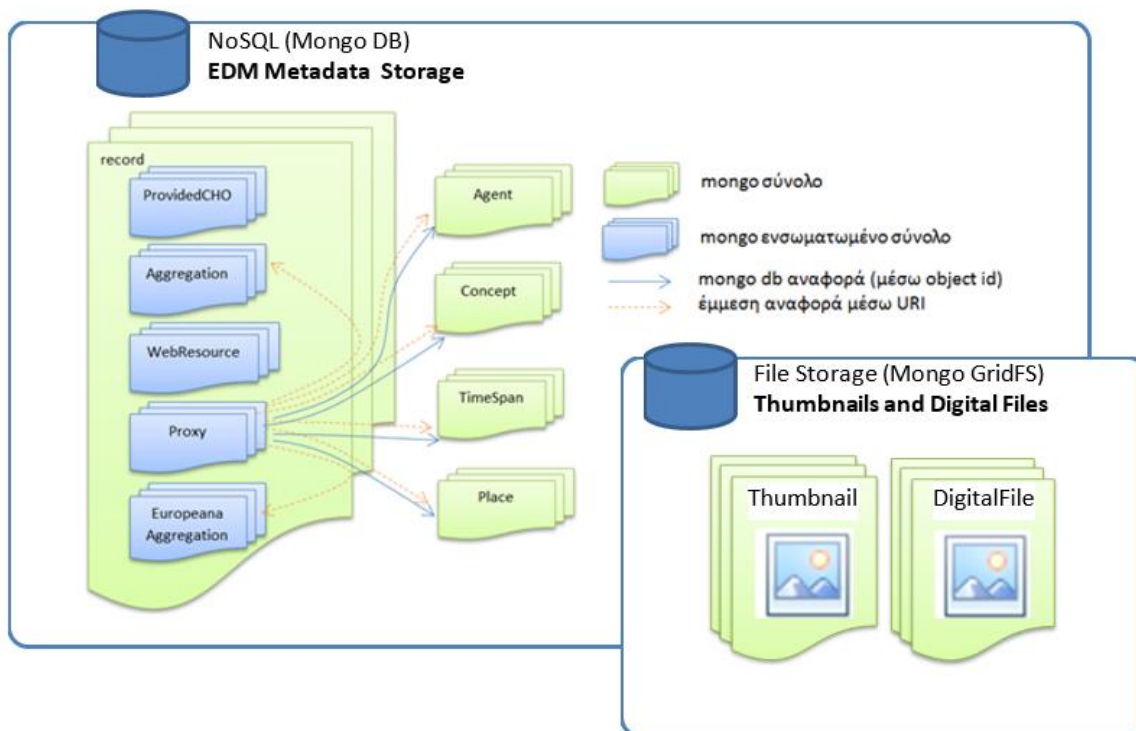
Το νέο component που υλοποιήθηκε ονομάστηκε EDM Storage Implementation (Σχήμα 83) . Πρόκειται για ένα πολύ ευέλικτο υποσύστημα που υλοποιεί μόνο τις μεθόδους του API. Το φυσικό μοντέλο δεδομένων στην Mongo DB βάση δεδομένων περιγράφεται από το Σχήμα 84. Το κεντρικό σύνολο (mongo collection) έχει όνομα «record» και αποθηκεύει έγγραφα τα οποία εμπεριέχουν ως ενσωματωμένα έγγραφα τις δομές edm:ProvidedCHO, ore:Aggregation, ore:Proxy, edm:EuropeanaProxy και edm:WebResource. Σημειώνεται ότι οι τιμές που παράγονται από τον σημασιολογικό εμπλουτισμό που

πραγματοποιεί το EKT αποθηκεύονται στα αντίστοιχα πεδία αλλά σε διαφορετική δομή (edm:EuropeanaProxy αντί της edm:ProvidedCHO) προκειμένου να διαχωρίζονται από τα πρωτότυπα μεταδεδομένα του κάθε φορέα. Η ανάκτηση ενός εγγράφου record εμπεριέχει όλη σχεδόν την EDM εγγραφή. Εξάιρεση αποτελούν οι οντότητες edm:Agent, skos:Concept, edm:TimeSpan και edm:Place οι οποίες εξακολουθούν να αποθηκεύονται σε ξεχωριστά σύνολα καθώς μια τέτοια οντότητα πιθανώς να αναφέρεται από πολλαπλά records. Στο Σχήμα 85 παρουσιάζεται το UML διάγραμμα κλάσεων των οντοτήτων πεδίου στις οποίες αντιστοιχίζονται οι Mongo δομές.

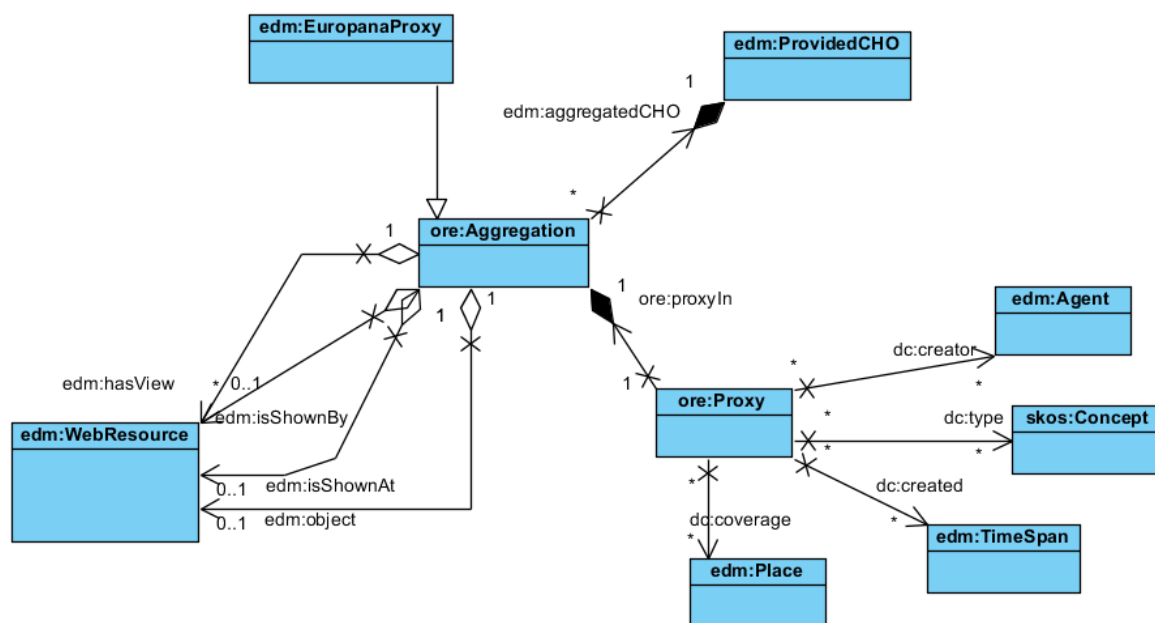


Σχήμα 83. Η τελική λύση: το EDM Storage Implementation

Κατά την εισαγωγή της στον συσσωρευτή SearchCulture.gr, μία EDM εγγραφή ευρετηριάζεται από ένα ειδικά παραμετροποιημένο σύστημα Apache Solr. Όλα τα πεδία του EDM ευρετηριάζονται για keyword-based αναζήτηση, για ένα μικρό υποσύνολο πεδίων οι τιμές αποθηκεύονται ώστε να επιστρέφονται ως μέρος της πληροφορίας που επιστρέφεται ως αποτέλεσμα σε μία αναζήτηση (dc:type εμπλουτισμού από το edm:EuropeanaProxy, dc:rights, dc:creator, edm:rights, edm:type) και, τέλος, για ένα άλλο μικρό υποσύνολο θα υποστηρίζεται μηχανισμός facet queries [21], για το περαιτέρω φιλτράρισμα των αποτελεσμάτων αναζήτησης (dc:type, dc:type εμπλουτισμού από το edm:EuropeanaProxy, dc:subject, edm:rights, dc:language, edm:provider).



Σχήμα 84. Αποθήκευση EDM μεταδεδομένων σε collections της Mongo και εικόνων προεπισκόπησης και ψηφιακών αρχείων σε GridGS collections της Mongo



Σχήμα 85. Αφαιρετικό UML διάγραμμα κλάσεων οντοτήτων πεδίου Mongo Βάσης (Αποθήκευση EDM μεταδεδομένων)

File Storage (Mongo DB GridFS)

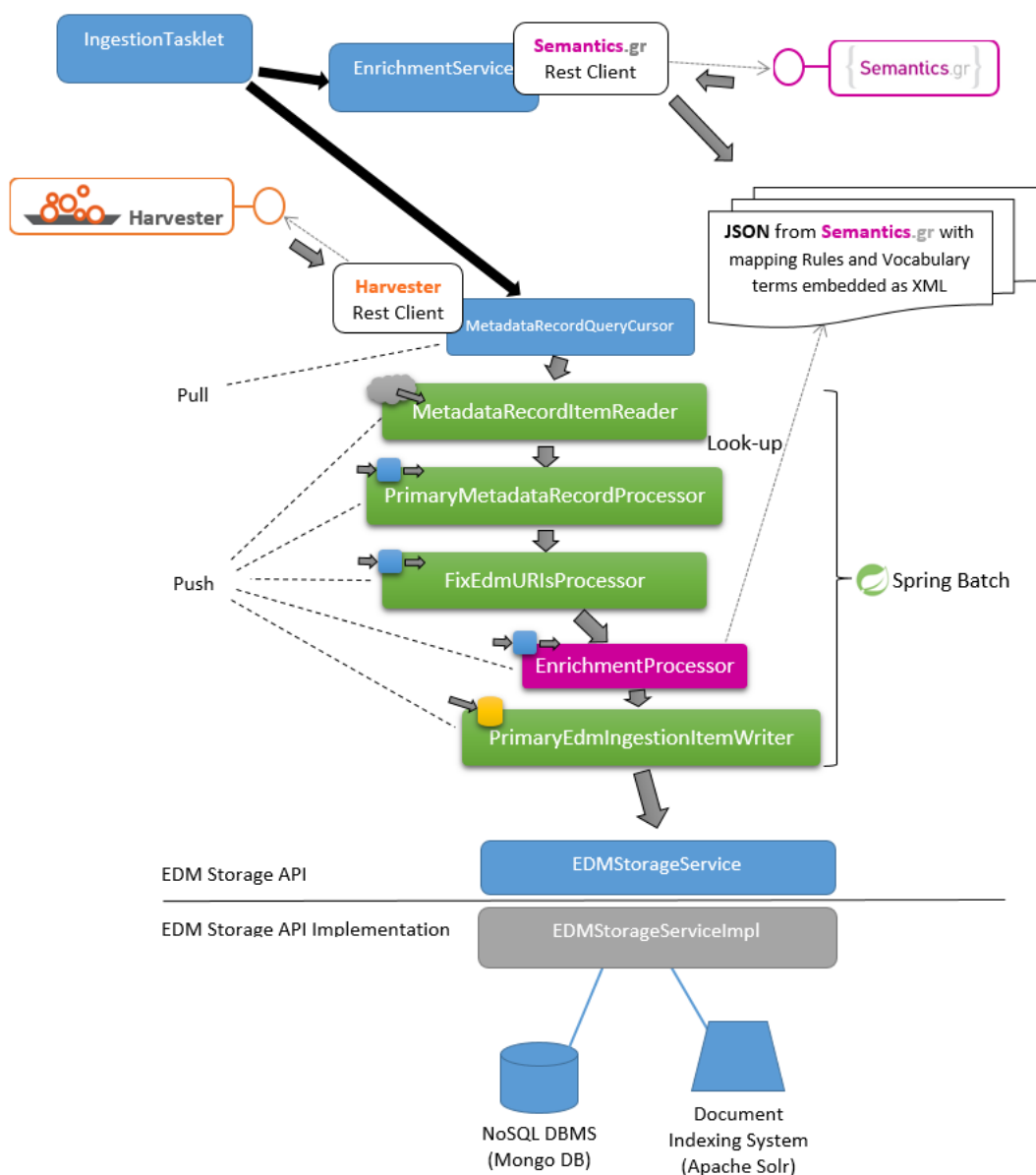
Το File Storage σύστημα δημιουργήθηκε για την αποθήκευση των εικόνων προεπισκόπησης αλλά και τον ψηφιακών αρχείων των τεκμηρίων που δημιουργούνται από την φόρμα (δηλαδή των εγγραφών που δεν προέρχονται από συσσώρευση περιεχομένου). Η τεχνολογία που επιλέχθηκε είναι η GridFS [22] της Mongo DB που επιτρέπει την αποδοτική αποθήκευση και ανάκτηση αρχείων με ενσωματωμένα και ευρετηριαζόμενα μεταδεδομένα. Οι GridFS συλλογές έχουν όνομα Thumbnails και DigitalFiles (Σχήμα 84) και οι αντίστοιχες οντότητες πεδίου, Thumbnail και DigitalFile.

Πίνακας V. Οι κυριότερες οντότητες πεδίου και τα αντίστοιχα αντικείμενα διαχείρισης στο Persistent και στο Service Layer του SearchCulture.gr

Persistent Layer				Service Layer	
Βάση Δεδομένων	Οντότητα Πεδίου	DAO	Πίνακας/ Collection	DTO	Service
PostgreSQL (aggregatorDB)	Organisation	OrganisationDao	ORGANISATION	OrganisationDTO	OrganisationService
	Collection	CollectionDao	COLLECTION COLLECTION_OAI_SET_DC_TYPE COLLECTION_OAI_SET_IMAGE_TYPE	CollectionDTO	CollectionService
	CollectionMetadataPrefix		COLLECTION_METADATAPREFIX	CollectionMetadataPrefixDTO	
	MetadataIngestion	MetadataIngestionDAO	INGESTION	MetadataIngestionDTO	MetadataIngestionService
	ThumbnailIngestion	ThumbnailIngestionDAO		ThumbnailIngestionDTO	ThumbnailIngestionService
	MetadataIngestionFailure	MetadataIngestionFailureDAO	INGESTION_FAILURE	MetadataIngestionFailureDTO	MetadataIngestionFailureService
	ThumbnailIngestionFailure	ThumbnailIngestionFailureDAO		ThumbnailIngestionFailureDTO	ThumbnailIngestionFailureService
	Configuration	ConfigurationDAO	CONFIGURATION	ConfigurationDTO	ConfigurationService
Mongo DB (metadataDB)	EDMRecord	EDMStorageDAO	records	EDMRecordDTO	EDMStorageService (EDM Storage API)
	ProviderCHO			ProviderCHODTO	
	Aggregation			AggregationDTO	
	WebResource			WebResourceDTO	
	Agent		agents	AgentDTO	
	TimeSpan		timeSpans	TimeSpanDTO	
	Place		places	PlaceDTO	
	Concept		concepts	ConceptDTO	
Mongo DB GridFS (digitalFile SDB)	Thumbnail	ThumbnailDAO	thumbnails	ThumbnailRecord	ThumbnailService
	DigitalFile	EdmDigitalFileDAO	digitalFiles	EdmDigitalFileRecord	EdmDigitalFileService

3.4.2 Service Layer

Το service layer περιλαμβάνει modules που παρέχουν μία σειρά από υπηρεσίες (services) στα modules του επόμενου layer (aggregator-portal-web, aggregator-management-web και aggregator-oai-pmh-web modules) οι οποίες υλοποιούν όλες τις business διαδικασίες του συσσωρευτή: εκκίνηση νέας διαδικασίας εισαγωγής, REST κλήσεις στο σύστημα Συγκομιδής Harvester, REST κλήσεις στο σύστημα λεξιλογίων και σημασιολογικού εμπλουτισμού Semantics.gr, παρακολούθηση και καταγραφή της προόδου της κάθε διαδικασίας, διάθεση αποθηκευμένων εγγραφών μεταδεδομένων, αναζήτηση περιεχομένου κα. Για να επιτευχθεί η αποδέσμευση του επόμενου layer από το persistent layer, άρα και από τις οντότητες πεδίου που αυτό ορίζει, δημιουργήθηκαν Data Access Services για κάθε αντικείμενο DAO και υιοθετήθηκε αυστηρά το DTO προγραμματιστικό υπόδειγμα [23], δημιουργήθηκαν δηλαδή νέες οντότητες πεδίου στο Service Layer, τα DTOs. Τα services χρησιμοποιούν ως είσοδο ή/και επιστρέφουν αποκλειστικά DTO αντικείμενα.



Σχήμα 86. Spring Batch ροή εισαγωγής Μεταδεδομένων στο EDM Storage

Οι βασικές υπηρεσίες που ορίζονται στο service layer είναι η OrganisationService και CollectionService, στο common-service module, και οι IngestionService, ThumbnailIngestionService, ThumbnailService ConfigurationService, και HarvestingService, στο aggregator-service module. Η τελευταία υλοποιεί την επικοινωνία με το Σύστημα Συγκομιδής Harvester.

Για την υλοποίηση της ροής εισαγωγής μεταδεδομένων και εικόνων προεπισκόπησης χρησιμοποιήθηκε το Spring Batch framework. Υλοποιήθηκαν μια σειρά από διεπαφές του framework και συγκεκριμένα Tasklets (IngestionTasklet, ThumbnailIngestionTasklet), ItemReaders (MetadataRecordItemReader, ThumbnailRecordItemReader), Processors (PrimaryMetadataIngestionProcessor, FixEdmURIs, EnrichmentProcessor) και ItemWriters (PrimaryEdmIngestionItemWriter, ThumbnailIngestionItemWriter).

Στο πρώτο βήμα της ροής εισαγωγής μεταδεδομένων μιας συλλογής και εφόσον η συγκεκριμένη συλλογή/αποθετήριο υπόκειται σε σημασιολογικό εμπλουτισμό, πραγματοποιείται επικοινωνία με το σύστημα Semantic.gr μέσω του REST-API του. Σκοπός είναι να μεταφορτωθούν σε μορφή ενός JSON εγγράφου όλοι οι κανόνες αντιστοίχισης που έχουν οριστεί για τα προς εμπλουτισμό πεδία του συγκεκριμένου αποθετηρίου/συλλογής καθώς και όλοι οι σχετικοί όροι λεξιλογίων (οι οποίοι περιλαμβάνονται στις αντιστοιχίσεις). Οι όροι λεξιλογίων είναι ενσωματωμένοι σε μορφή XML συμβατή με τις αντίστοιχες σημασιολογικές κλάσεις του EDM. Με τον τρόπο αυτό συλλέγεται απευθείας και σε ένα βήμα όλη η απαραίτητη πληροφορία από το Semantics.gr και αποφεύγεται έτσι η επικοινωνία με το εξωτερικό αυτό σύστημα σε επίπεδο εγγραφής, η οποία θα δημιουργούσε σημαντικό network overhead και θα καθυστερούσε δραματικά τον χρόνο ολοκλήρωσης της διαδικασίας. Το JSON έγγραφο φυλάσσεται στην μνήμη καθ' όλη την διάρκεια της διαδικασίας και χρησιμοποιείται κατά την διάρκεια της ροής από το processor EnrichmentProcessor για τον εμπλουτισμό/ομογενοποίηση κάθε εγγραφής. Για τον σκοπό αυτό το JSON έγγραφο ευρετηριάζεται καταλλήλως ώστε η αναζήτηση πληροφορίας να είναι αποδοτική.

Στο επόμενο βήμα το MetadataRecordItemReader αντικείμενο διαβάσει εγγραφές μεταδεδομένων με σελιδοποιημένο τρόπο από το σύστημα Harvester. Για κάθε MetadataRecord που διαβάζεται, ένας STAX XML parser διαβάσει το πεδίο metadata (το οποίο περιέχει ως XML την OAI-PMH RECORD εγγραφή μεταδεδομένων) και δημιουργεί τον αντίστοιχο EDM-internal γράφο (κλάση PrimaryMetadataIngestionProcessor). Για τον σκοπό αυτό χρησιμοποιείται η υπηρεσία Unmarshaling του EDM Storage API η οποία μπορεί να μετατρέψει ένα EDM-rubic, ESE XML ή DC-DS-XML έγγραφο σε γράφο αποτελούμενο από EDM-internal DTOs. Στην συνέχεια όλα τα rdf:about URIs των EDM στοιχείων ProvidedCHO, Aggregation και Proxy του γράφου μετασχηματίζονται καταλλήλως, σύμφωνα με συγκεκριμένες οδηγίες οι οποίες κωδικοποιούνται σε ειδικό properties αρχείο, ώστε να μετατραπούν σε έγκυρα και ενεργά URIs, προκειμένου να επιτευχθεί η διάθεση των μεταδεδομένων ως Linked Data (processor FixEdmURIs). Στην συνέχεια ο processor EnrichmentProcessor εμπλουτίζει κάθε εισερχόμενη εγγραφή με τις κατάλληλες σημασιολογικές οντότητες σύμφωνα με τους κανόνες αντιστοίχισης και τις σημασιολογικές οντότητες (όροι λεξιλογίου) του JSON εγγράφου που μεταφορτώθηκε αρχικά από το σύστημα Semantics.gr. Τέλος ο ItemWriter PrimaryEdmIngestionItemWriter εισάγει (ή διαγράφει) ανά ομάδες των 100 στοιχείων τις εγγραφές μεταδεδομένων στο EDM Storage System (ως EDM-internal γράφους).

Η υλοποίηση της ροής εισαγωγής εικόνων προεπισκόπησης ακολουθεί παρόμοια λογική. Η πηγή αυτή τη φορά είναι ψηφιακά αρχεία που θα συλλεχθούν ή συλλέχθηκαν στο παρελθόν από το Σύστημα

Συγκομιδής Harvester. Σημειώνεται ότι η μετατροπή σε εικόνα προεπισκόπησης πραγματοποιείται από το ίδιο το Σύστημα Συγκομιδής μέσω υπηρεσίας που αυτό παρέχει ώστε να μειωθεί το network overhead.

3.4.3 To Web layer

Το web layer υλοποιεί την διεπαφή ιστού. Αποτελείται από 1) την δημόσια διαδικτυακή πύλη (Portal) 2) το διαχειριστικό περιβάλλον (management GUI) που επιτρέπει σε έναν εξουσιοδοτημένο χρήστη από το προσωπικό του ΕΚΤ να εκτελεί όλες τις διαχειριστικές λειτουργίες του συσσωρευτή, και 3) την υπηρεσία ιστού ΟΑΙ-ΡΜΗ μέσω της οποίας τα μεταδεδομένα διατίθενται για συγκομιδή σε τρίτα συστήματα (όπως ο συσσωρευτής Europeana).

Για τις δύο εφαρμογές ιστού του SearchCulture.gr, δηλαδή την διαδικτυακή πύλη και το διαχειριστικό περιβάλλον, χρησιμοποιήθηκε το προγραμματιστικό πλαίσιο Spring MVC [14] και κατ' επέκταση, υιοθετήθηκε το προγραμματιστικό υπόδειγμα Μοντέλο-Όψη-Χειριστής (model-view-controller, MVC) [31]. Το προγραμματιστικό αυτό υπόδειγμα διαιρεί την εφαρμογή ιστού σε τρία μέρη, στα μοντέλα (models), τα οποία μοντελοποιούν τα δεδομένα ως java POJOS (ουσιαστικά πρόκειται για τα DTOs που παρουσιάστηκαν στην ενότητα 3.4.2), στις όψεις (views), οι οποίες παρουσιάζουν την πληροφορία στον χρήστη (JSP σελίδες, JSP tags) και στους χειριστές (controllers) οι οποίοι χειρίζονται την είσοδο του χρήστη (δηλαδή τις GET, PUT, POST και DELETE HTTP αιτήσεις) και αποτελούν υλοποιήσεις της διεπαφής Controller του Spring MVC. Κατά κανόνα, για κάθε βασική DTO οντότητα πεδίου δημιουργήθηκαν ξεχωριστοί controllers (που χειρίζονται το σύνολο των σχετικών αιτήσεων εκτελώντας τις κατάλληλες κλήσεις στο Service layer) και ξεχωριστές JSP σελίδες και tags. Παράλληλα, χρησιμοποιήθηκαν διάφορες front-end τεχνολογίες, με κυριότερες το προγραμματιστικό πλαίσιο JQuery [29] και την εργαλειοθήκη Bootstrap [30]. Ο Πίνακας VI συγκεντρώνει όλες τις τεχνολογίες που χρησιμοποιήθηκαν στο web layer για την ανάπτυξη των δύο εφαρμογών ιστού.

Πίνακας VI. Τεχνολογίες που χρησιμοποιήθηκαν στο Web layer για την ανάπτυξη της διαδικτυακής πύλης και του διαχειριστικού περιβάλλοντος

Κατηγορία	Προγραμματιστικό πλαίσιο/εργαλειοθήκη
1 Προγραμματιστικό πλαίσιο δημιουργίας java εφαρμογών ιστού (server-side)	Spring MVC 4.0.2 [14]
2 Προγραμματιστικό πλαίσιο για ταυτοποίηση και αυθεντικοποίηση χρήστη στο διαχειριστικό περιβάλλον (server-side)	Spring Security 3.2.3 [35]
3 Markup γλώσσα δημιουργίας δυναμικών ιστοσελίδων (server-side)	JavaServer Pages Standard Tag Library (JSTL) 2.0 [33]
4 Προγραμματιστικό πλαίσιο για δημιουργία περιγραμμάτων (templates)	Apache Tiles 3.0.3 [32]
5 Επαναχρησιμοποίηση αυτόνομων τμημάτων ιστοσελίδων (server-side)	JSP custom tags 2.0 [33]
6 Προγραμματιστικό πλαίσιο για front-end λειτουργικότητα (client-side)	JQuery 2.1.3 [29]
7 Εργαλειοθήκη για front-end λειτουργικότητα και εμφάνιση (client-side)	Bootstrap 3.3.6 [30]

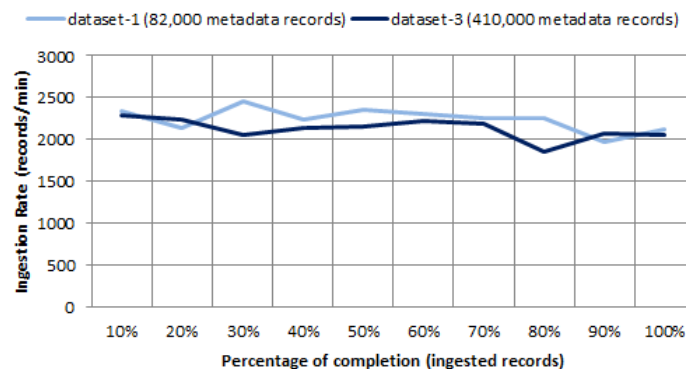
8	Εικονίδια	Font awesome 4.6.3 [34]
9	Διαγράμματα - πίτες	Google Charts (στο διαχειριστικό περιβάλλον) [36] Highcharts 4.2.4 (διαδικτυακή πύλη) [26]
10	Διαδραστική δυναμικά σύννεφα ετικετών	jQCloud 2.0.1 [25]

Για την ανάπτυξη της ΟΑΙ-ΡΜΗ υπηρεσίας ιστού χρησιμοποιήθηκε η βιβλιοθήκη Enhanced OAI Server [23] που υλοποιήθηκε από το ΕΚΤ¹. Η βιβλιοθήκη αυτή είναι μία Servlet εφαρμογή ιστού η οποία χρησιμοποιείται για την εύκολη υλοποίηση του ΟΑΙ-ΡΜΗ πρωτοκόλλου. Βασίζεται στην βιβλιοθήκη OAI-Cat [27] που παρέχει τις κατάλληλες Java διεπαφές για την υλοποίηση του ΟΑΙ-ΡΜΗ πρωτοκόλλου και στο προγραμματιστικό πλαίσιο Biblio Transformation Engine (BTE) [28] το οποίο επίσης αναπτύχθηκε από το ΕΚΤ και χρησιμοποιείται για την υλοποίηση ρών δεδομένων.

3.5 Πειραματική Αξιολόγηση Συσσωρευτή

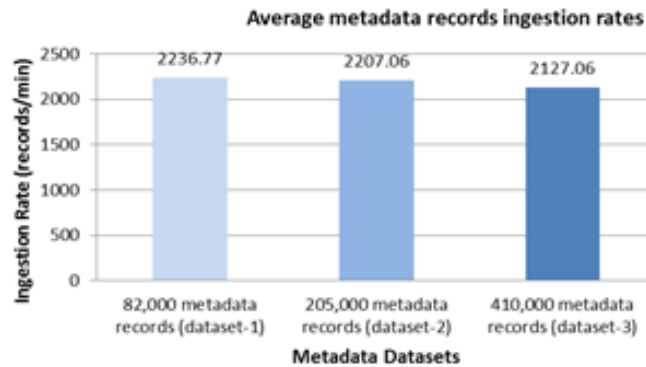
Στην ενότητα αυτή παρουσιάζεται η πειραματική αξιολόγηση της απόδοσης του συστήματος Συσσωρευτή αλλά και της βελτίωσης της ευρεσιμότητας του περιεχομένου που επετεύχθη με τον σημασιολογικό εμπλουτισμό των τύπων.

Για την αξιολόγηση της απόδοσης του συσσωρευτή, χρησιμοποιήθηκαν τα τρία σύνολα μεταδεδομένων που περιγράφει ο Πίνακας II τα οποία εισήχθησαν αρχικά στο Σύστημα Συγκομιδής.



Σχήμα 87. Εξέλιξη του ρυθμού εισαγωγής εγγραφών μεταδεδομένων σε μια διαδικασία συσσώρευσης ως προς το ποσοστό ολοκλήρωσής της

¹ Από τους προγραμματιστές Κώστα Σταμάτη και Κώστα Κουτσοιράκη



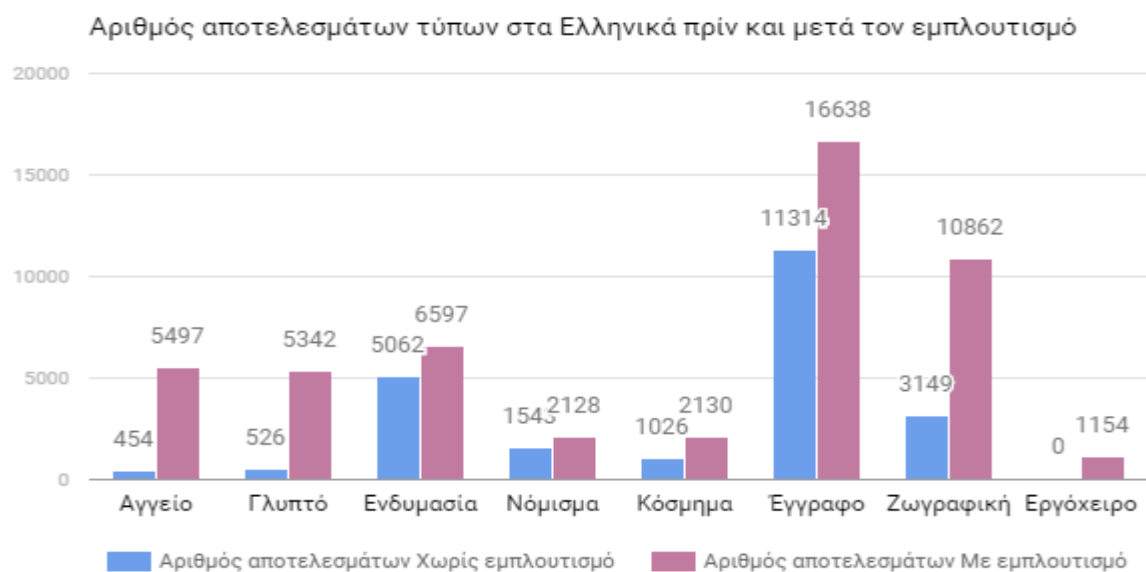
Σχήμα 88. Μέσοι ρυθμοί εισαγωγής εγγραφών μεταδεδομένων για διαφορετικά μεγέθη συλλογών

Το διάγραμμα στο Σχήμα 87 παρουσιάζει την εξέλιξη της ροής ανάκτησης από το σύστημα συγκομιδής Harvester και αποθήκευσης των εγγραφών μεταδεδομένων ως EDM στο σύστημα αποθήκευσης EDM μεταδεδομένων του συσσωρευτή κατά την διάρκεια δύο ολοκληρωμένων διαδικασιών εισαγωγής μεταδεδομένων για τα σύνολα dataset-1 και dataset-3. Παρατηρούμε ότι ο ρυθμός παραμένει σχεδόν σταθερός γεγονός που αποδεικνύει ότι το σύστημα Συσσωρευτή κλιμακώνει ιδιαίτερα ικανοποιητικά. Στο Σχήμα 88 παρουσιάζονται οι μέσοι ρυθμοί εισαγωγής για τα τρία διαφορετικά σύνολα μεταδεδομένων.

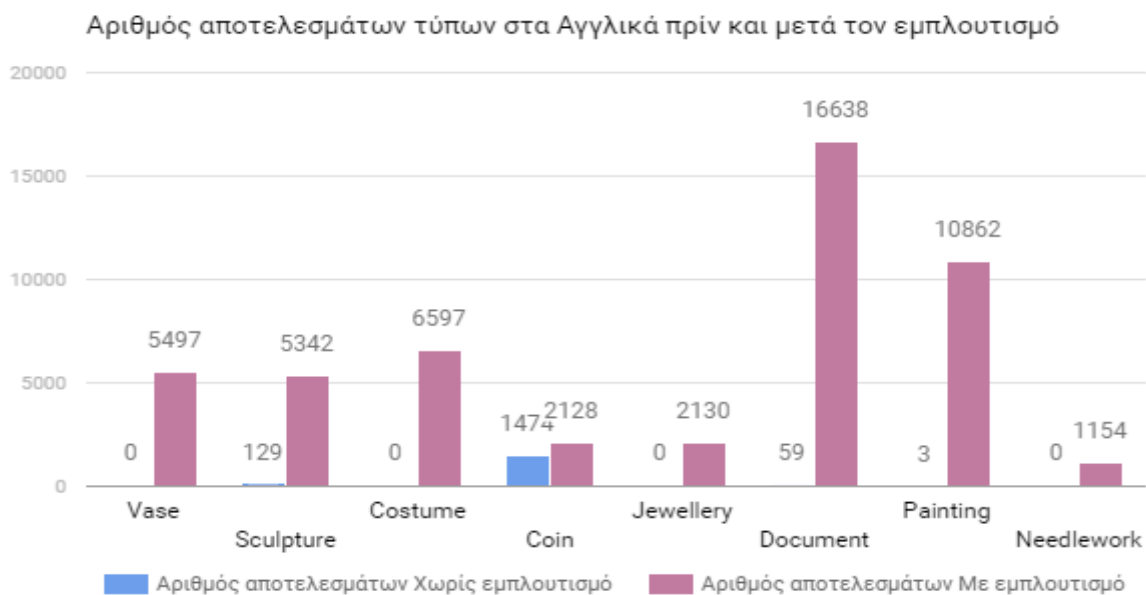
Το ΕΚΤ εμπλουτίζει και ομογενοποιεί το περιεχόμενο που συσσωρεύει βάσει ενός πρότυπου δίγλωσσου ιεραρχικού λεξιλογίου τύπων που δημιούργησε. Τα τεκμήρια του SearchCulture.gr εμπλουτίστηκαν με ένα νέο πεδίο «ΕΚΤ Τύπος» που περιλαμβάνει τις αναφορές στους όρους του λεξιλογίου. Έτσι, 157,000 τεκμήρια ταξινομήθηκαν σε ένα συμπαγές και ισορροπημένο σύνολο 128 τύπων αποκτώντας σημασιολογική συνοχή με θετικό αντίκτυπο στην ευρεσιμότητα, στην παρουσίαση και στην πολυγλωσσικότητα του περιεχομένου. Επιπλέον, το περιεχόμενο του SearchCulture.gr διατίθεται με αυτό τον τρόπο ως Διασυνδεδεμένα Δεδομένα.

Στο Σχήμα 89 παρουσιάζεται ένα διάγραμμα στο οποίο συγκρίνεται η οτελεσματικότητα της αναζήτησης 8 αντιπροσωπευτικών τύπων πριν και μετά τον σημασιολογικό εμπλουτισμό / ομογενοποίηση. Ανάλογα είναι τα αποτελέσματα και για τους υπόλοιπους τύπους τεκμηρίων στους οποίους ταξινομήθηκε το περιεχόμενο.

Στην αγγλική έκδοση, ο χρήστης μπορεί να επιλέξει/αναζητήσει τους ίδιους τύπους στα αγγλικά. Στο Σχήμα 90 παρουσιάζεται το αντίστοιχο διάγραμμα στο οποίο συγκρίνεται η αποτελεσματικότητα της αναζήτησης των ίδιων 8 αντιπροσωπευτικών τύπων στα αγγλικά πριν και μετά τον σημασιολογικό εμπλουτισμό / ομογενοποίηση. Εδώ τα αποτελέσματα είναι ακόμα πιο εντυπωσιακά.



Σχήμα 89. Σημασιολογικός εμπλουτισμός: Βελτίωση ευρεσιμότητας



Σχήμα 90. Σημασιολογικός εμπλουτισμός: Βελτίωση πολυγλωσσικής ευρεσιμότητας

Κατάλογος Σχημάτων

Σχήμα 1. Ο κύκλος ζωής του ψηφιακού περιεχομένου στις υποδομές του EKT : συλλογή, πιστοποίηση, εμπλουτισμός, συσσώρευση-διάχυση και ασφαλής διαφύλαξη.	8
Σχήμα 2. Συνολική αρχιτεκτονική των συστημάτων του EKT για την συλλογή περιεχομένου (Harvester), τον έλεγχο συμμόρφωσης με προδιαγραφές (Validator Back-end / Front-end), την συσσώρευση/ενιαία αναζήτηση (SearchCulture.gr), την ασφαλή διαφύλαξη (Preservator Back-end / Front-end) και τον σημασιολογικό εμπλουτισμό (Semantics.gr).	13
Σχήμα 3. Οι δύο βασικοί στόχοι ενός συσσωρευτή και ενδεικτικοί τρόποι ικανοποίησής τους.	14
Σχήμα 4. Τα κυριότερα λειτουργικά χαρακτηριστικά του SearchCulture.gr.	16
Σχήμα 5. Η αρχική σελίδα της εφαρμογής ιστού του Συστήματος Συγκομιδής.	21
Σχήμα 6. Φόρμα συμπλήρωσης παραμέτρων για την εκκίνηση νέας διαδικασίας συγκομιδής.	22
Σχήμα 7. Μηνύματα σφαλμάτων για μη έγκυρες παραμέτρους εκκίνησης νέας διαδικασίας συγκομιδής.	22
Σχήμα 8 Παρακολούθηση της προόδου μιας εν εξελίξει διαδικασίας σε πραγματικό χρόνο.	23
Σχήμα 9. Προβολή λεπτομερών στατιστικών διαδικασίας συγκομιδής.	23
Σχήμα 10. Εμφάνιση πληροφορίας αποσφαλμάτωσης για διαδικασία που απέτυχε.	24
Σχήμα 11. Προβολή μεταδεδομένων που συλλέχθηκαν στα πλαίσια μιας διαδικασίας συγκομιδής.	24
Σχήμα 12. Σελίδα διαχείρισης διαδικασιών συγκομιδής Ψηφιακών Αρχείων.	25
Σχήμα 13. Φόρμα εκκίνησης νέας διαδικασίας συγκομιδής Ψηφιακών Αρχείων από ZIP αρχείο.	26
Σχήμα 14. Φόρμα εκκίνησης νέας διαδικασίας συγκομιδής Ψηφιακών Αρχείων από ανάγνωση URL από πεδία μεταδεδομένων.	26
Σχήμα 15. Λεπτομέρειες και στατιστικά μιας διαδικασίας συγκομιδής Ψηφιακών Αρχείων.	27
Σχήμα 16. Ενέργειες για διαδικασίες που είτε απέτυχαν είτε ολοκληρώθηκαν με αποτυχίες.	27
Σχήμα 17. Προβολή σφαλμάτων συλλογής ψηφιακών αρχείων.	28
Σχήμα 18. Συνολική Αρχιτεκτονική του Συστήματος Συγκομιδής Ψηφιακού Περιεχομένου.	34
Σχήμα 19. Η εξέλιξη της ροής εισαγωγής κατά την διάρκεια μιας διαδικασίας συγκομιδής μεταδεδομένων ως προς το ποσοστό εξέλιξης αυτής.	35
Σχήμα 20. Η διαδικτυακή πύλη του SearchCulture.gr - Αρχική σελίδα.	43
Σχήμα 21. Η αρχική σελίδα του SearchCulture.gr όπως φαίνεται από κινητά τηλέφωνα.	44
Σχήμα 22. Διαδραστική πίτα με τις μεγαλύτερες συλλογές.	44
Σχήμα 23. Ενημερωτικό πλαίσιο για τον σημασιολογικό εμπλουτισμό που διεξάγει το EKT.	45
Σχήμα 24. Ενημερωτικό πλαίσιο για τους ελέγχους ποιότητας/πιστοποίηση που διεξάγει το EKT.	45
Σχήμα 25. Προαιρετική επιλογή αποθετηρίου.	46
Σχήμα 26. Συγκεκριμενοποίηση αναζήτησης (α) και αυτόματη συμπλήρωση όρου αναζήτησης (β).	46
Σχήμα 27. Σύνθετη αναζήτηση.	46
Σχήμα 28. Προαιρετική επιλογή EKT Τύπου.	47
Σχήμα 29. Προαιρετική επιλογή άδειας χρήσης ψηφιακού περιεχομένου.	47
Σχήμα 30. Αποτελέσματα αναζήτησης.	48
Σχήμα 31. Αποτελέσματα αναζήτησης όπως φαίνονται από οθόνη κινητού τηλεφώνου.	48
Σχήμα 32. Φιλτράρισμα αποτελεσμάτων (Facet Queries).	49
Σχήμα 33. Καρτέλα τεκμηρίου.	50
Σχήμα 34. Καρτέλα τεκμηρίου όπως φαίνεται από οθόνη κινητού τηλεφώνου.	51
Σχήμα 35. Καρτέλα τεκμηρίου σε παράθυρο – πλοήγηση σε προηγούμενα / επόμενα τεκμήρια.	51
Σχήμα 36. RDF απεικόνιση τεκμηρίου - Διάθεση ως Linked Data.	52
Σχήμα 37. Παρουσίαση των φορέων και των συλλογών τους.	53
Σχήμα 38. Περιήγηση – EKT τύποι.	54
Σχήμα 39. Πολυγλωσσικότητα: Η αρχική σελίδα του SearchCulture.gr στα Αγγλικά.	55
Σχήμα 40. Είσοδος στο διαχειριστικό περιβάλλον.	56
Σχήμα 41. Βασικό μενού επιλογών.	56

Σχήμα 42. Η σελίδα με τους παρόχους.....	57
Σχήμα 43. Φόρμα αποθετηρίου φορέα – Κάρτα γενικών στοιχείων.....	58
Σχήμα 44. Φόρμα αποθετηρίου φορέα – Κάρτα καταχώρησης τύπων μεταδεδομένων.....	58
Σχήμα 45. Φόρμα αποθετηρίου φορέα – ΟΑΙ Σύνολα ανά dc:type και ανά τύπο ψηφιοποίησης.....	59
Σχήμα 46. Φόρμα αποθετηρίου φορέα – Στοιχεία για εισαγωγή μεταδεδομένων	60
Σχήμα 47. Φόρμα αποθετηρίου φορέα – Κάρτα δεσμευμένων λεξιλογίων/θησαυρών.	60
Σχήμα 48. Διαχείριση περιεχομένου αποθετηρίου	61
Σχήμα 49. Διαδικασίες Εισαγωγής Μεταδεδομένων	62
Σχήμα 50. Φόρμα εκκίνησης νέας διαδικασίας εισαγωγής για πολλαπλά ΟΑΙ σύνολα	63
Σχήμα 51. Νέα διαδικασία εισαγωγής χωρίς καθορισμό συγκεκριμένου ΟΑΙ συνόλου.....	64
Σχήμα 52. Εισαγωγή μεταδεδομένων που έχουν ήδη συλλεχθεί στο παρελθόν από το Σύστημα Συγκομιδής.	64
Σχήμα 53. Έλεγχος ορθότητας παραμέτρων εκκίνησης εισαγωγής	65
Σχήμα 54. Προβολή λεπτομερών στατιστικών διαδικασίας εισαγωγής.....	65
Σχήμα 55. Εμφάνιση μηνυμάτων για αποσφαλμάτωση για διαδικασίες που απέτυχαν.....	66
Σχήμα 56. Αναφορά εγγραφών μεταδεδομένων που απέτυχαν.....	66
Σχήμα 57. Προβολή EDM εγγραφών που εισήχθησαν στα πλαίσια μιας διαδικασίας εισαγωγής.....	67
Σχήμα 58. Semantics.gr: φόρμα εγγραφής αποθετηρίου	68
Σχήμα 59. Semantics.gr: φόρμα ορισμού αντιστοιχίσεων πεδίου αποθετηρίου – εισαγωγή βασικών στοιχείων....	69
Σχήμα 60. Semantics.gr: η φόρμα αντιστοιχίσεων πεδίου αποθετηρίου αμέσως μετά την συγκομιδή τιμών.....	70
Σχήμα 61. Semantics.gr: χειροκίνητη αντιστοίχιση με όρο λεξιλογίου	71
Σχήμα 62. Semantics.gr: προσθήκη φίλτρων	72
Σχήμα 63. Πραγματικό παράδειγμα εμπλουτισμού αποθετηρίου του SearchCulture.gr. Το αποθετήριο πρωτοτύπως είχε μόνο 6 διαφορετικούς τύπους χωρίς πολυγλωσσική πληροφορία. Μετά τον εμπλουτισμό, τα τεκμήρια κατανεμήθηκαν σε 55 δίγλωσσους τύπους.	73
Σχήμα 64. Το Λεξιλόγιο Τύπων Ψηφιακού Περιεχομένου του ΕΚΤ μέσα από την διαδικτυακή πύλη του Semantics.gr	73
Σχήμα 65. Σελίδα διαχείρισης διαδικασιών εισαγωγής Αρχείων Προεπισκόπησης	74
Σχήμα 66. Φόρμα εκκίνησης νέας διαδικασίας εισαγωγής Αρχείων Προεπισκόπησης.....	75
Σχήμα 67. Αρχεία που απέτυχαν να εισαχθούν ως εικόνες προεπισκόπησης	75
Σχήμα 68. Σελίδα σημασιολογικών οντοτήτων της κατηγορίας 'Άνθρωποι/Οργανισμοί' (edm:Agent).....	77
Σχήμα 69. Αντίστοιχες σελίδες σημασιολογικών οντοτήτων για skos:Concept, edm:Place και edm:Timespan	77
Σχήμα 70. Προβολή μιας σημασιολογικής οντότητας ως RDF.....	78
Σχήμα 71. Φόρμα εισαγωγής/επεξεργασίας σημασιολογικής οντότητας της κατηγορίας 'Άνθρωποι/Οργανισμοί' (edm:Agent).....	78
Σχήμα 72. Φόρμα εισαγωγής/επεξεργασίας σημασιολογικής οντότητας της κατηγορίας 'Τόποι' (edm:Place)	79
Σχήμα 73. Φόρμα επεξεργασίας EDM τεκμηρίου - Καρτέλα 'Τεκμήριο' (edm:ProvidedCHO)	80
Σχήμα 74. Φόρμα επεξεργασίας EDM τεκμηρίου - Καρτέλα 'Συλλογή Ψηφιακών Πόρων' (edm:Aggregation).....	81
Σχήμα 75. Φόρμα επεξεργασίας EDM τεκμηρίου - Καρτέλα 'Ψηφιακοί Πόροι' (edm:WebResource)	82
Σχήμα 76. Ανέβασμα αρχείου κατά την δημιουργία νέου Ψηφιακού Πόρου (edm:WebResource)	82
Σχήμα 77. Επιλογή πεδίων και σειράς εμφάνισής τους στη σελίδα παρουσίασης τεκμηρίου	83
Σχήμα 78. Επιλογή πεδίων και σειράς εμφάνισής τους στα αποτελέσματα αναζήτησης.....	83
Σχήμα 79. Παραμετροποίησης φόρμας επεξεργασίας εγγραφής τεκμηρίου	84
Σχήμα 80. Αρχιτεκτονική του Συστήματος Συσσωρευτή/Ενιαίας Αναζήτησης SearchCulture.	86
Σχήμα 81. Διάγραμμα οντοτήτων συσχετίσεων της σχεσιακής βάσης του SearchCulture.gr.	89
Σχήμα 82: UML διάγραμμα κλάσεων οντοτήτων πεδίου Σχεσιακής Βάσης.....	90
Σχήμα 83. Η τελική λύση: το EDM Storage Implementation	91
Σχήμα 84. Αποθήκευση EDM μεταδεδομένων σε collections της Mongo και εικόνων προεπισκόπησης και ψηφιακών αρχείων σε GridGS collections της Mongo	92

Σχήμα 85. Αφαιρετικό UML διάγραμμα κλάσεων οντοτήτων πεδίου Mongo Βάσης (Αποθήκευση EDM μεταδεδομένων).....	92
Σχήμα 86. Spring Batch ροή εισαγωγής Μεταδεδομένων στο EDM Storage	95
Σχήμα 87. Εξέλιξη του ρυθμού εισαγωγής εγγραφών μεταδεδομένων σε μια διαδικασία συσσώρευσης ως προς το ποσοστό ολοκλήρωσής της.....	98
Σχήμα 88. Μέσοι ρυθμοί εισαγωγής εγγραφών μεταδεδομένων για διαφορετικά μεγέθη συλλογών	99
Σχήμα 89. Σημασιολογικός εμπλουτισμός: Βελτίωση ευρεσιμότητας.....	100
Σχήμα 90. Σημασιολογικός εμπλουτισμός: Βελτίωση πολυγλωσσικής ευρεσιμότητας	100

Κατάλογος Πινάκων

Πίνακας I. Οι Λειτουργικές Απαιτήσεις του Συστήματος Συγκομιδής.....	18
Πίνακας II. Χαρακτηριστικά δεδομένων που χρησιμοποιήθηκαν στα πειράματα	35
Πίνακας III. Οι Λειτουργικές Απαιτήσεις για την δημόσια διαδικτυακή πύλη του SearchCulture.gr.....	38
Πίνακας IV. Οι Λειτουργικές Απαιτήσεις για το Σύστημα Συσσώρευσης και το Διαχειριστικό Περιβάλλον του SearchCulture.gr	39
Πίνακας V. Οι κυριότερες οντότητες πεδίου και τα αντίστοιχα αντικείμενα διαχείρισης στο Persistent και στο Service Layer του SearchCulture.gr.....	94
Πίνακας VI. Τεχνολογίες που χρησιμοποιήθηκαν στο Web layer για την ανάπτυξη της διαδικτυακής πύλης και του διαχειριστικού περιβάλλοντος	97

Δημοσιεύσεις

- Ioanna Ourania Stathopoulou, Haris Georgiadis, Vangelis Banos, Panagiotis Stathopoulos, Nikos Houssos and Evi Sachini:
An Open Cultural Digital Content Infrastructure, International Conference on Digital Libraries (DL'2014), London, 2014.
- Haris Georgiadis, Vangelis Banos, Ioanna Ourania Stathopoulou, Panagiotis Stathopoulos, Nikos Houssos and Evi Sachini:
Ensuring the quality and interoperability of open cultural digital content: System architecture and scalability, 5th International Conference on Information, Intelligence, Systems and Applications (IISA'2014), Chania, 2014.
- Nikos Houssos, Panagiotis Stathopoulos, Ioanna-Ourania Stathopoulou, Vangelis Banos, Haris Georgiadis, Evi Sachini:
Interoperability specifications and an automatic validation tool as a key factor for the support of a repository ecosystem, 9th International Conference on Open Repositories (OR'2014), Helsinki, 2014.
- Haris Georgiadis, Agathi Papanoti, Maria Paschou, Alexandra Roubani, Dimitra Pelekanou, Despoina Chardouveli, Evi Sachini:
Semantics.gr: A self-improving service to repositories and aggregators for massively enriching their content, MTSR 2016: DHC Workshop (to be published in DHC 2016 Workshop proceedings)

Αναφορές

- [1] Παναγιώτης Σταθόπουλος, Νίκος Χούσος, και Γεώργιος Σταύρου, “Προδιαγραφές και χαρακτηριστικά διαλειτουργικότητας για ανοικτό ψηφιακό περιεχόμενο.” Mar-2013. <http://helios-eie.ekt.gr/EIE/handle/10442/8887>
- [2] Πλαίσιο Συνεργασίας Πρόσκλησης 31 και 31.2 <http://www.digitalplan.gov.gr/resource-api/dipla/contentObject/Υποδειγμα-Plaisioy-Synergasias-Prosklhshs-31.2/content>
- [3] Introducing REPOX, a tool to manage metadata spaces, <http://pro.europeana.eu/blogpost/introducing-repox-a-tool-to-manage-metadata-spaces>
- [4] Πλατφόρμα λογισμικού CRIS <http://metrics.ekt.gr/el/cris>
- [5] Open Archives Initiative Protocol for Metadata Harvesting <http://www.openarchives.org/pmh>
- [6] ESE Documentation: <http://pro.europeana.eu/web/guest/ese-documentation>
- [7] EDM Definition: <http://pro.europeana.eu/documents/900548/4b190b0f-9adc-4a65-bc09-3a9b16e45827>
- [8] RDF και RDFS ([http://www.w3.org/2000/01/rdf-schema#/
/](http://www.w3.org/2000/01/rdf-schema#/))
- [9] Linked Data <http://linkeddata.org>
- [10] Το τρέχον «εσωτερικό» EDM schema της Europeana: <http://www.europeana.eu/schemas/edm/EDM-INTERNAL.xsd>
- [11] Expressing Dublin Core Description Sets using XML (DC-DS-XML), available at <http://dublincore.org/documents/dc-ds-xml/>, date issued: 2008-09-01
- [12] Expressing Dublin Core metadata using the Resource Description Framework (RDF), available at <http://dublincore.org/documents/dc-rdf/>, date issued: 2008-01-14.
- [13] Spring Framework <http://spring.io/>
- [14] Spring Web MVC Framework <http://docs.spring.io/spring/docs/current/spring-framework-reference/html/mvc.html>
- [15] Spring Batch <http://docs.spring.io/spring-batch/>
- [16] Data access object Pattern, from Wikipedia, the free encyclopedia http://en.wikipedia.org/wiki/Data_access_object
- [17] Streaming API for XML (StAX) (<http://en.wikipedia.org/wiki/StAX>)
- [18] Apache Solr <http://lucene.apache.org/solr/>
- [19] Morphia, a lightweight type-safe library for mapping Java objects to/from MongoDB Morphia (<https://github.com/mongodb/morphia/wiki>)

- [20] D. Jayathilake, A. Fernando, C. Sooriarachchi, “Thinking Big to Deal with Big Data: A Practical Insight into MongoDB,” in Software Developer’s Journal, May 2012
- [21] Faceted Search, http://en.wikipedia.org/wiki/Faceted_search
- [22] Mongo DB, GridFS, <http://docs.mongodb.org/manual/core/gridfs/>
- [23] Fowler, Martin (2010). Data Transfer Object (DTO). Patterns of Enterprise Application Architecture. Retrieved from <http://martinfowler.com/eaaCatalog/dataTransferObject.html>
- [24] EKT / Enhance OAI-PMH Service, a Java Servlet web application that implements the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) v2.0
<https://github.com/EKT/EnhancedOAIServer>
- [25] jQCloud, <http://mistic100.github.io/jQCloud/>
- [26] Highcharts, <http://www.highcharts.com/>
- [27] The OAIcat Open Source Software (OSS) projec, <http://www.oclc.org/research/themes/data-science/oaicat.html>
- [28] The Biblio-Transformation-Engine Java Framework, <https://github.com/EKT/Biblio-Transformation-Engine>
- [29] The JQuery JS framework, <https://jquery.com/>
- [30] Bootstrap, HTML, CSS, and JS framework <http://getbootstrap.com/>
- [31] The Model–view–controller software architectural pattern,
<https://en.wikipedia.org/wiki/Model%E2%80%93view%E2%80%93controller>
- [32] The Apache tiles framework, <https://tiles.apache.org/>
- [33] Official tutorial: The Java EE 5 Tutorial, Chapter 5, JavaServer Pages Technology,
<http://download.oracle.com/javase/5/tutorial/doc/bnagx.html>
- [34] Font awesome, The iconic font and CSS toolkit,
http://fontawesome.io/?utm_source=hackernewsletter
- [35] Spring Security, <http://projects.spring.io/spring-security/>
- [36] Google charts, <https://developers.google.com/chart/>
- [37] Dublin Core Metadata Initiative Metadata Terms
<http://dublincore.org/documents/2012/06/14/dcmi-terms/?v=terms>



Ενιαία Αναζήτηση
σε ανοικτό ψηφιακό περιεχόμενο



ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ
www.ekt.gr