



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών
— ΙΔΡΥΘΕΝ ΤΟ 1837 —

«ΑΠΟΛΛΩΝΙΣ»

Γλωσσικοί πόροι στο ΕΚΠΑ: Υποδομή, χρήση και εργαλεία

Κλεοπάτρα Φέρλα

ΕΚΠΑ, 01/03/2019

Από την υποδομή *CLARIN* στην υποδομή *ΑΠΟΛΛΩΝΙΣ*

CLARIN

(<https://www.clarin.eu/>)

[About](#) [Participants](#) [Services](#) [Knowledge Base](#) [Events](#) [News](#) [Contact](#) [Applications](#) [Intranet login](#)

CLARIN - European Research Infrastructure for Language Resources and Technology

CLARIN makes digital language resources available to scholars, researchers, students and citizen-scientists from all disciplines, especially in the humanities and social sciences, through single sign-on access. CLARIN offers long-term solutions and technology services for deploying, connecting, analyzing and sustaining digital language data and tools. CLARIN supports scholars who want to engage in cutting edge data-driven research, contributing to a truly multilingual European Research Area. [Read more...](#)



Common Language Resources and Technology Infrastructure



The Latvian FrameNet-annotated text corpus is a balanced, multilayered corpus annotated according to the latest frame inventory of Berkeley FrameNet.



Participating Organizations

The operations, services and centres of the CLARIN infrastructure are provided and funded by the national consortia in the countries that have joined CLARIN ERIC or by associated centres.

[Learn more](#)



Services, Tools and Data

CLARIN provides a wide variety of Services and Data sets

[Learn more](#)

• • • • •

CLARIN

ΜΙΑ ΕΥΡΩΠΑΪΚΗ ΕΡΕΥΝΗΤΙΚΗ ΥΠΟΔΟΜΗ

Το 2012 ιδρύθηκε το CLARIN (= Common Language Resources and Technology Infrastructure) ERIC (= European Research Infrastructure Consortium).

Η κοινοπραξία ανέλαβε την αποστολή δημιουργίας και συντήρησης μιας υποδομής που θα υποστηρίζει την **ανταλλαγή, χρήση και βιωσιμότητα** γλωσσικών δεδομένων και εργαλείων για την έρευνα στις ανθρωπιστικές και κοινωνικές επιστήμες.

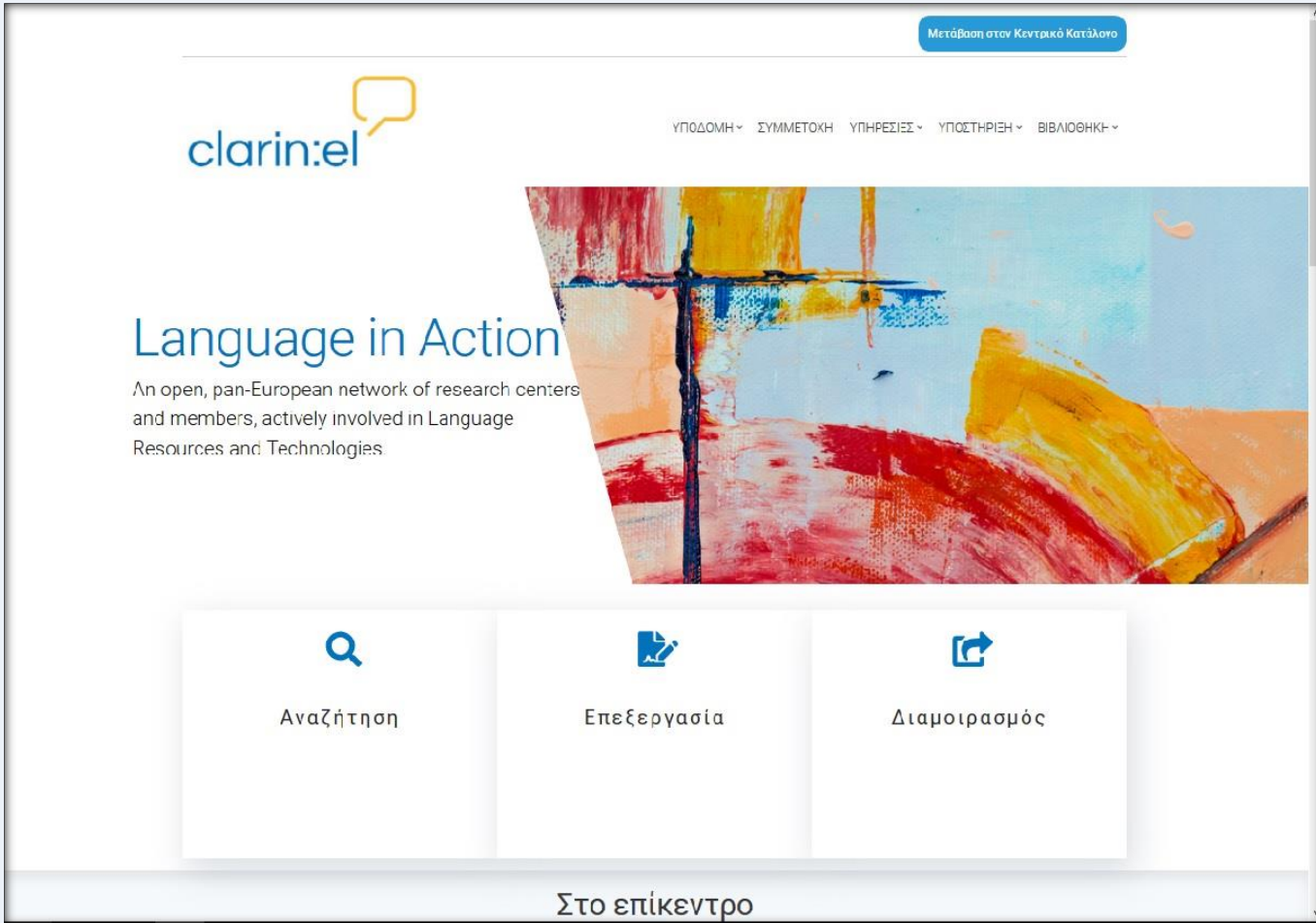
Σήμερα, το CLARIN παρέχει εύκολη και βιώσιμη πρόσβαση σε ψηφιακά γλωσσικά δεδομένα (γραπτά, προφορικά ή και σε πολυτροπική μορφή). Επίσης, προσφέρει στον ερευνητή προηγμένα εργαλεία για να ανακαλύψει, να διερευνήσει, να εκμεταλλευτεί, να σχολιάσει, να αναλύσει ή να συνδυάσει τέτοια σύνολα δεδομένων, **σε όποια χώρα και αν αυτά βρίσκονται.**

CLARIN

Σκοπός της υποδομής είναι, ιδανικά, ένας έλληνας ερευνητής να μπορεί από το γραφείο του με MIA, και μόνη, πιστοποιημένη εγγραφή (single sign-on):

- να ψάξει, να βρει και να πάρει έγκριση ώστε να χρησιμοποιήσει κείμενα από τους συνεργαζόμενους φορείς του εξωτερικού,
- να επιλέξει το ακριβές σύνολο δεδομένων στα οποία θέλει να δουλέψει και να αποθηκεύσει την επιλογή του,
- να «τρέξει» πάνω στην επιλογή του εργαλεία π.χ. σημασιολογικής ανάλυσης από την Αθήνα και στατιστικά εργαλεία από τη Βουδαπέστη,
- να αποθηκεύσει τη διαδικασία και τα αποτελέσματα της ανάλυσης, και τέλος,
- να τα μοιραστεί με συνεργάτες του στο εξωτερικό.

CLARIN:EL
(<http://www.clarin.gr/el>)



CLARIN:EL

Λίγο αργότερα ξεκίνησε και στην Ελλάδα η αντίστοιχη διαδικασία και έτσι δημιουργήθηκε η υποδομή CLARIN:EL, το ελληνικό σκέλος της ευρωπαϊκής υποδομής CLARIN ERIC.

Η Ερευνητική Υποδομή CLARIN:EL λειτουργεί ως ένα κατανεμημένο δίκτυο αποθετηρίων, το οποίο αποτελείται από:

- Ιδρυματικά Αποθετήρια (που δημιουργούνται για κάθε οργανισμό που συμμετέχει στο δίκτυο CLARIN:EL)
- το Αποθετήριο Φιλοξενούμενων Πόρων (ΑΦΠ) και
- τον Κεντρικό Κατάλογο (ή Κεντρικό Συσσωρευτή)

Στο δίκτυο CLARIN:EL συμμετέχουν μέχρι στιγμής 10 Ερευνητικοί Οργανισμοί και Πανεπιστημιακά Ιδρύματα από όλη την Ελλάδα:

ΕΚ «Αθηνά» (φορέας και συντονιστής του δικτύου CLARIN:EL), ΕΔΕΤ, ΔΗΜΟΚΡΙΤΟΣ, ΕΚΠΑ, ΑΡΙΣΤΟΤΕΛΕΙΟ, ΚΕΓ, ΙΟΝΙΟ, ΑΙΓΑΙΟ, ΕΚΚΕ, ΠΑΝΤΕΙΟ.

DARIAH-EU

(<https://www.dariah.eu/>)



Αναπτύσσει, διατηρεί και λειτουργεί μια υποδομή για την υποστήριξη των ερευνητικών πρακτικών που βασίζονται στις ΤΠΕ και υποστηρίζει τους ερευνητές στη χρήση τους για την κατασκευή, την ανάλυση και την ερμηνεία των ψηφιακών πόρων.

DARIAH-GR/ΔΥΑΣ

(<http://www.dyas-net.gr/>)

**δυάς**
Δίκτυο Υποδομών για την Έρευνα στις Ανθρωπιστικές Επιστήμες

GR | EN

DARIAH - EU | GR | FR | DE | NL | SI

ΑρχικήΤαυτότηταΒιβλιοθήκηΕργαλείαΠαρατηρητήριοΚοινότηταΕπικοινωνία

-δυάςλογοι

Εργαστήρι για θέματα αξιοποίησης των ψηφιακών τεχνολογιών προς όφελος της έρευνας στο πεδίο των ανθρωπιστικών επιστημών.

Κτίριο Ερευνητικών Κέντρων της Ακαδημίας Αθηνών, Αναγνωστοπούλου 14, Αίθουσα DARIAH Ισόγειο




Αναζήτηση

Σύνδεση

☐ Remember me?

Ημερολόγιο

February

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

Βιβλιοθήκη



Εδώ θα βρείτε δημόσια έγγραφα, εκθέσεις, αξιολογήσεις, αναφορές, ενημερωτικό υλικό και βιβλιογραφία σχετική με τις εξελίξεις των ψηφιακών εφαρμογών στις ανθρωπιστικές επιστήμες.

Εργαλεία και Μητρώα



Η περιοχή αυτή φιλοξενεί ψηφιακά εργαλεία και μητρώα που αναπτύσσονται ή διατίθενται στο πλαίσιο του έργου DARIAH-GR. Τόσο τα εργαλεία όσο και τα μητρώα διατίθενται σε ανοικτή πρόσβαση κι απευθύνονται σε ερευνητές των ανθρωπιστικών επιστημών καθώς και σε διαχειριστές πολιτιστικών συλλογών (αρχαιονόμους, επιμελητές μουσείων κ.ά.).

Παρατηρητήριο

Κοινότητα

DARIAH-GR/ΔΥΑΣ

Σήμερα το δίκτυο ΔΥΑΣ υλοποιεί τις δράσεις του DARIAH-GR σε συντονισμό με την Ευρωπαϊκή Ερευνητική Υποδομή για τις Τέχνες και τις Ανθρωπιστικές Επιστήμες.

Οι υπηρεσίες της εθνικής υποδομής οργανώνονται σε τέσσερις ενότητες:

- Υπηρεσίες κοινωνίας ψηφιακών πόρων
- Εργαλεία και οδηγίες ανάπτυξης ψηφιακών πόρων
- Υπηρεσίες Ευρωπαϊκής Υποδομής DARIAH
- Παρατηρητήριο Ψηφιακών Ανθρωπιστικών Σπουδών

Οι φορείς που συμμετέχουν είναι:

Ακαδημία Αθηνών (Συντονιστής), ΕΚ «Αθηνά» (με τη ΜΟΨΕ), ΕΚΠΑ (με το Τμήμα Ιστορίας και Αρχαιολογίας), ΕΜΠ (με το ΕΠΙΣΕΥ), ΑΣΚΤ, ΙΤΕ (με το ΕΠΣ)

(<https://apollonis-infrastructure.gr/>)



ΑΠΟΛΛΩΝΙΣ

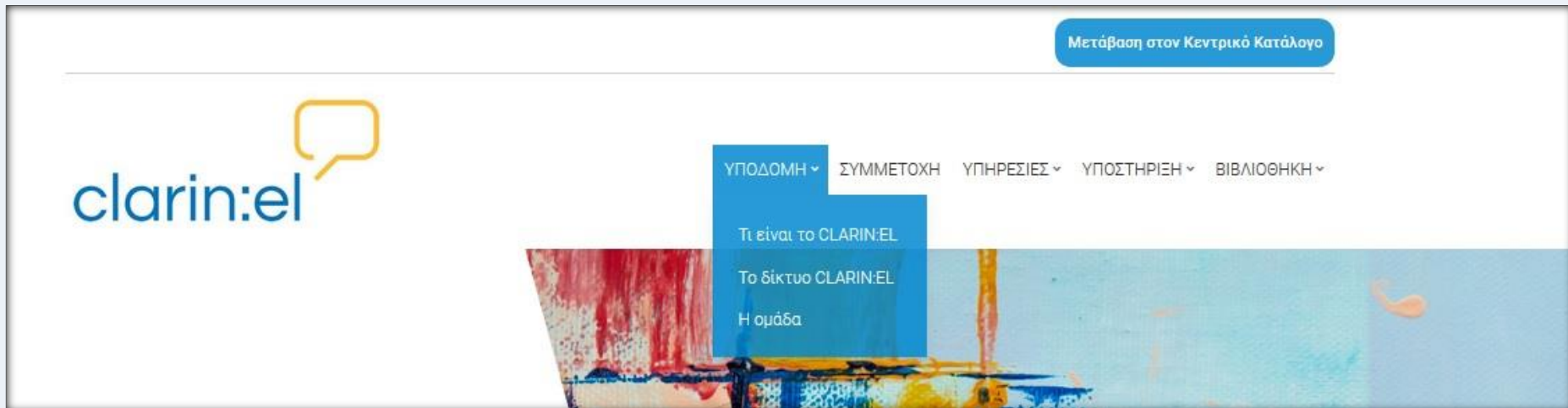
Η ΑΠΟΛΛΩΝΙΣ είναι η εθνική υποδομή που υποστηρίζει και προωθεί τις ψηφιακές ανθρωπιστικές επιστήμες και τέχνες, και τη γλωσσική τεχνολογία και καινοτομία στην Ελλάδα. Εντάσσεται στη Δράση «Ενίσχυση των Υποδομών Έρευνας και Καινοτομίας» και χρηματοδοτείται από το Επιχειρησιακό Πρόγραμμα «Ανταγωνιστικότητα, Επιχειρηματικότητα και Καινοτομία» στο πλαίσιο του ΕΣΠΑ 2014-2020.

Δημιουργήθηκε από τη σύμπραξη του Εθνικού Δικτύου Γλωσσικής Τεχνολογίας clarin:el και του Εθνικού Δικτύου Ψηφιακών Υποδομών για τις Ανθρωπιστικές Επιστήμες DARIAH-GR/ΔΥΑΣ, τα οποία είναι συνιστώσες των αντίστοιχων ευρωπαϊκών υποδομών CLARIN και DARIAH.

Στην παρούσα φάση εξέλιξης της υποδομής περιλαμβάνονται δράσεις περαιτέρω ανάπτυξης και εμπέδωσης των υπηρεσιών των δύο συνιστωσών υποδομών, μέσα σε ένα πλαίσιο διαλειτουργικότητας και αμοιβαίας ενίσχυσης, ενώ συνεχίζεται η συμμετοχή της Ελλάδας στις αντίστοιχες ευρωπαϊκές υποδομές.

CLARIN: EL

Η δομή του κόμβου



CLARIN: EL

Η δομή: Συμμετοχή



ΥΠΟΔΟΜΗ > **ΣΥΜΜΕΤΟΧΗ** ΥΠΗΡΕΣΙΕΣ > ΥΠΟΣΤΗΡΙΞΗ > ΒΙΒΛΙΟΘΗΚΗ >

Συμμετοχή

 > ΣΥΜΜΕΤΟΧΗ



Συμμετοχή στο δίκτυο CLARIN:EL

Μάθετε τα πάντα για το πώς μπορείτε να συμμετάσχετε και εσείς στο δίκτυο CLARIN:EL και πραγματοποιήστε την εγγραφή σας!

Η Ερευνητική Υποδομή CLARIN:EL σας δίνει την ευκαιρία να συμμετάσχετε στο εθνικό δίκτυο Γλωσσικών Πόρων και Τεχνολογιών στην Ελλάδα! Μοιραστείτε τα δικά σας γλωσσικά δεδομένα, εργαλεία και διαδικτυακές υπηρεσίες γλωσσικής επεξεργασίας και επωφεληθείτε της δυνατότητας πρόσβασης σε εκατοντάδες πόρους άλλων, σε εθνικό και ευρωπαϊκό επίπεδο.



Για ιδιώτες

Πληροφορίες για ιδιώτες που επιθυμούν να εγγραφούν στην Υποδομή.

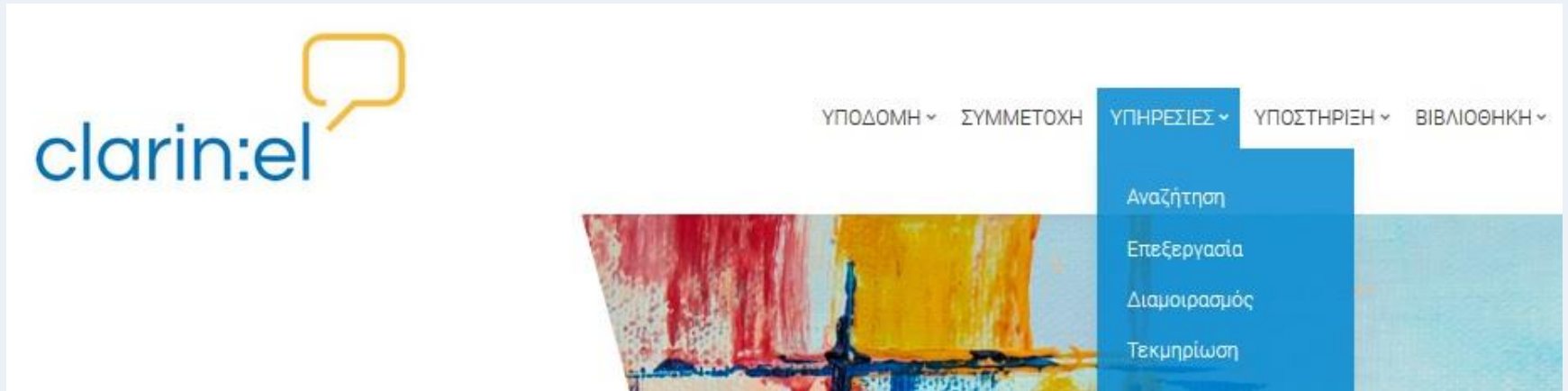


Για οργανισμούς

Πληροφορίες για οργανισμούς που επιθυμούν να γίνουν μέλη του Δικτύου.

CLARIN: EL

Η δομή του κόμβου



CLARIN: EL

Η δομή: Υπηρεσίες



CLARIN: EL

Υπηρεσίες: Αναζήτηση

Η Υπηρεσία Αναζήτησης CLARIN:EL περιλαμβάνει τη διαδικασία απλών ή σύνθετων αναζητήσεων πόρων, εργαλείων και υπηρεσιών στον Κεντρικό Κατάλογο CLARIN:EL. Πιο συγκεκριμένα, ο χρήστης από τον Κεντρικό Κατάλογο του CLARIN:EL έχει τη δυνατότητα να αναζητήσει τον ή τους πόρους που επιθυμεί με δύο τρόπους:

Χρήση λέξεων-κλειδιών

Αναζήτησε πόρους στον Κεντρικό Κατάλογο CLARIN:EL

Ο χρήστης μπορεί γράφοντας λέξεις-κλειδιά στο πεδίο της αναζήτησης πόρων στην αρχική σελίδα του Κεντρικού Καταλόγου CLARIN:EL να ανακτήσει τους πόρους που πληρούν τις προϋποθέσεις που έθεσε.

Χρήση διαθέσιμων φίλτρων

Αναζήτησε πόρους στον Κεντρικό Κατάλογο CLARIN:EL

Σε περίπτωση που ο χρήστης θέλει να περιορίσει τον αριθμό των αποτελεσμάτων της αναζήτησής του, έχει τη δυνατότητα να χρησιμοποιήσει τα διαθέσιμα φίλτρα που του προσφέρει η Υποδομή CLARIN:EL. Επιλέγοντας το κουμπί search (χωρίς να συμπληρώσει κάτι στο πεδίο της αναζήτησης) ή εναλλακτικά πατώντας πάνω στο κουμπί browse (από το κεντρικό μενού) στην αρχική σελίδα του Κεντρικού Καταλόγου CLARIN:EL, οδηγείται αυτόματα στο συνολικό κατάλογο όλων των πόρων, εργαλείων και υπηρεσιών που διατίθενται μέσω των Ιδρυματικών Αποθετηρίων και του Αποθετηρίου Φιλοξενούμενων Πόρων της Υποδομής CLARIN:EL. Στα αριστερά βρίσκεται η λίστα με όλα τα διαθέσιμα φίλτρα που μπορεί να ο χρήστης να επιλέξει προκειμένου να περιορίσει την αναζήτησή του, συνεπώς και τα αποτελέσματα που θα πάρει. Μπορεί για παράδειγμα να περιορίσει την αναζήτησή του με βάση τη γλώσσα (**Language**) στην οποία θέλει να είναι ο πόρος, τον τύπο του πόρου (**Resource Type**) ή το μέσον αυτού (**Media Type**) ή να φιλτράρει τα αποτελέσματα ανάλογα με τη διαθεσιμότητα (**Availability**) και την άδεια του πόρου (**Licence**), την θεματική περιοχή (**Domain**) αλλά και το ιδρυματικό αποθετήριο (**Source Repository**) από το οποίο προέρχεται ο πόρος που αναζητάμε.



225846

QUERIES



400

CORPORA



35

TOOLS

ab

46

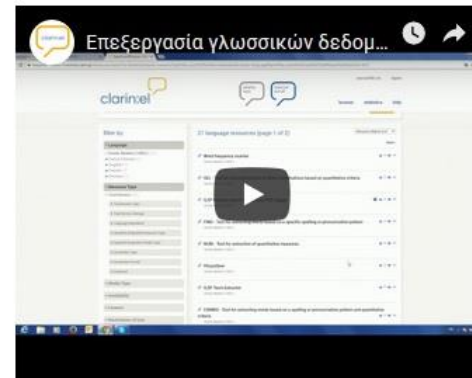
LEXICAL
RESOURCES

CLARIN: EL

Υπηρεσίες: Επεξεργασία

Η **Ερευνητική Υποδομή CLARIN:EL** υλοποιεί, παρέχει, υποστηρίζει και συντηρεί εργαλεία και διαδικτυακές υπηρεσίες γλωσσικής επεξεργασίας. Στον **Κεντρικό Κατάλογο CLARIN:EL** ο χρήστης έχει τη δυνατότητα να ανακαλύψει και να κάνει χρήση των υπηρεσιών γλωσσικής επεξεργασίας, οι οποίες περιλαμβάνουν μεταξύ άλλων υπηρεσίες **λεκτικής ανάλυσης**, υπηρεσίες **αναγνώρισης λέξεων, ορίων προτάσεων και μερών του λόγου**, υπηρεσίες **μορφολογικής και συντακτικής ανάλυσης**, υπηρεσίες **αναγνώρισης ονοματικών οντοτήτων, ορολογίας, άποψης και συναισθήματος** καθώς και πολλές ακόμα υπηρεσίες αναλυτικής κειμένου.

Η πρόσβαση στις διαδικτυακές υπηρεσίες γλωσσικής επεξεργασίας CLARIN:EL είναι **ελεύθερη** για όλη την ακαδημαϊκή και ερευνητική κοινότητα, τη βιομηχανία, αλλά και το ευρύ κοινό εφόσον ο χρήστης έχει προηγουμένως πραγματοποιήσει την εγγραφή του στην Υποδομή CLARIN:EL. Περισσότερες πληροφορίες σχετικά με τη διαδικασία εγγραφής στην Υποδομή CLARIN:EL μπορείτε να βρείτε [εδώ](#).



Προτεινόμενα εργαλεία & υπηρεσίες CLARIN:EL

ILSP Dependency parser

Συντακτικός αναλυτής

[Δείτε τον πόρο στον Κεντρικό Κατάλογο CLARIN:EL](#)

Ο ILSP Dependency Parser είναι ένας συντακτικός αναλυτής για την ελληνική γλώσσα. Ο αναλυτής αυτός αναπαριστά αυτόματα τη συντακτική δομή ελληνικών προτάσεων, αναγνωρίζοντας για κάθε λέξη τη λέξη που αποτελεί την "κεφαλή" της, αλλά και τη σχέση της κάθε λέξης με τη λέξη-κεφαλή. Έτσι, στην πρόταση **Οι αστρονόμοι ανακάλυψαν ένα νέο φεγγάρι** αναγνωρίζει, μεταξύ άλλων, ότι οι λέξεις **αστρονόμοι** και **φεγγάρι** αποτελούν το υποκείμενο και το αντικείμενο της λέξης **ανακάλυψαν**. Ο ILSP Dependency Parser έχει αναπτυχθεί με βάση το **Greek Dependency Treebank**, έναν γλωσσικό πόρο του ΙΕΛ/ΕΚ "Αθηνά" που ανέρχεται σε περίπου 175 χιλιάδες λέξεις, χειρωνακτικά επισημειωμένες στο επίπεδο της σύνταξης.

GrNE-Tagger

Επισημειωτής Ονοματικών Οντοτήτων

[Δείτε τον πόρο στον Κεντρικό Κατάλογο CLARIN:EL](#)

Ο Επισημειωτής Ονοματικών Οντοτήτων GrNE-Tagger επεξεργάζεται κείμενα και αναγνωρίζει αυτόματα και ταξινομεί τις Ονοματικές Οντότητες (κύρια ονόματα) που αυτά περιέχουν στις ακόλουθες πέντε κατηγορίες: **Άτομο/Πρόσωπο, Τοποθεσία, Οργανισμός, Ονόματα κτιρίων ή/και άλλων ανθρώπινων κατασκευών, Γεωπολιτική Οντότητα**. Ο Επισημειωτής Ονοματικών Οντοτήτων GrNE-Tagger αναπτύχθηκε και συντηρείται από το ΙΕΛ/ΕΚ "Αθηνά".

CLARIN: EL

Υπηρεσίες: Διαμοιρασμός

Διαμοιρασμός πόρων στο CLARIN:EL

Από Οργανισμούς-μέλη του δικτύου CLARIN:EL

[Μάθετε περισσότερα για τη συμμετοχή Οργανισμού στο Δίκτυο CLARIN:EL](#)

Οι Οργανισμοί που επιθυμούν να διαμοιραστούν τους πόρους τους μέσω της Ερευνητικής Υποδομής CLARIN:EL πρέπει πρώτα να γίνουν μέλη του δικτύου CLARIN:EL σύμφωνα με τη διαδικασία που περιγράφεται εδώ. Αφού ο Οργανισμός υπογράψει το Καταστατικό του δικτύου, το νέο μέλος είναι έτοιμο να αποθέσει τους πόρους του στην Υποδομή CLARIN:EL είτε μέσω του δικού του Ιδρυματικού Αποθετηρίου είτε μέσω του Αποθετηρίου Φιλοξενούμενων Πόρων.

Από ιδιώτες που συνεργάζονται με Οργανισμό-μέλος του CLARIN:EL με Ιδρυματικό Αποθετήριο

[Μάθετε περισσότερα για τις κατηγορίες και τα δικαιώματα των χρηστών CLARIN:EL](#)

Στην περίπτωση που ανήκετε ή/και συνεργάζεστε με Οργανισμό που είναι μέλος του δικτύου CLARIN:EL και διαθέτει το δικό του Ιδρυματικό Αποθετήριο, ο διαμοιρασμός των πόρων σας γίνεται μέσω αυτού. Στην περίπτωση αυτή, για την είσοδό σας στον Κεντρικό Κατάλογο CLARIN:EL θα πρέπει να χρησιμοποιήσετε τον Ακαδημαϊκό σας λογαριασμό. Εάν αντιμετωπίσετε κάποιο πρόβλημα, μπορείτε να ζητήσετε βοήθεια από τον Τεχνικό Υπεύθυνο (το άτομο που είναι υπεύθυνο για τα τεχνικά θέματα του αποθετηρίου του οργανισμού σας) ή τον Επιστημονικό Υπεύθυνο (το άτομο που είναι υπεύθυνο για τα επιστημονικά και διοικητικά θέματα του αποθετηρίου του οργανισμού σας).

Από ιδιώτες που συνεργάζονται με Οργανισμό-μέλος του CLARIN:EL χωρίς Ιδρυματικό Αποθετήριο

[Μάθετε περισσότερα για τις κατηγορίες και τα δικαιώματα των χρηστών CLARIN:EL](#)

Στην περίπτωση που συνεργάζεστε με Οργανισμό που είναι μέλος του δικτύου CLARIN:EL αλλά δεν έχει δικό του Ιδρυματικό Αποθετήριο, ο διαμοιρασμός των πόρων σας γίνεται μέσω του Αποθετηρίου Φιλοξενούμενων Πόρων. Για τον σκοπό αυτό, χρησιμοποιείτε για την είσοδό σας στον Κεντρικό Κατάλογο CLARIN:EL τον Ακαδημαϊκό σας λογαριασμό (κάνοντας χρήση της Υπηρεσίας Ταυτοποίησης και Πιστοποίησης χρηστών - Academic SSO login), και ζητάτε από τον Επιστημονικό Υπεύθυνο του Οργανισμού σας (το άτομο που είναι υπεύθυνο για τα επιστημονικά και διοικητικά θέματα του Οργανισμού σας στο δίκτυο CLARIN:EL) να σας δώσει τα πρόσθετα δικαιώματα για την τεκμηρίωση των Γλωσσικών Πόρων.

Από ιδιώτες που συνεργάζονται με Οργανισμό-μη μέλος του CLARIN:EL

[Μάθετε περισσότερα για τις κατηγορίες και τα δικαιώματα των χρηστών CLARIN:EL](#)

Καταρχάς διερευνήστε εάν ο Οργανισμός σας σκοπεύει να γίνει μέλος του δικτύου CLARIN:EL και να δημιουργήσει δικό του Ιδρυματικό Αποθετήριο. Εάν ναι, επικοινωνήστε με το άτομο που είναι υπεύθυνο για το Ιδρυματικό Αποθετήριο στον Οργανισμό σας για να σας δώσει τα αντίστοιχα δικαιώματα και, στη συνέχεια, μπορείτε να αποθέσετε τους πόρους σας. Εάν ο Οργανισμός δεν σκοπεύει να συμμετάσχει στο δίκτυο CLARIN:EL, ακολουθήστε τις οδηγίες για τους ιδιώτες-ανεξάρτητους παρόχους που δίνονται παρακάτω.

Από ιδιώτες-ανεξάρτητους παρόχους

[Μάθετε περισσότερα για τις κατηγορίες και τα δικαιώματα των χρηστών CLARIN:EL](#)

Στην περίπτωση που δε συνεργάζεστε με κανένα Οργανισμό, ο διαμοιρασμός των πόρων σας γίνεται μέσω του Αποθετηρίου Φιλοξενούμενων Πόρων. Για το σκοπό αυτό και εφόσον επιθυμείτε να διαμοιραστείτε τους πόρους σας ως ως ιδιώτης-ανεξάρτητος πάροχος, θα πρέπει πρώτα να συμπληρώσετε και να υπογράψετε το Συμφωνητικό Κατάθεσης Πόρων. Στη συνέχεια, παρακαλείστε να επικοινωνήσετε με τον Διευθυντή της Ερευνητικής Υποδομής CLARIN:EL και τον Τεχνικό Συντονιστή, οι οποίοι και θα φροντίσουν να σας δώσουν τα αντίστοιχα δικαιώματα για να μπορείτε να μεταφορτώσετε τους πόρους σας στην Υποδομή CLARIN:EL.

CLARIN: EL

Υπηρεσίες: Τεκμηρίωση

Υπόδειγμα τεκμηρίωσης στο CLARIN:EL Παραδείγματα εγγραφής μεταδεδομένων τεκμηρίωσης στο CLARIN:EL

Για τη διευκόλυνση των χρηστών, έχουμε προετοιμάσει και θέσει στη διάθεσή τους ένα σύνολο υποδειγματικών εγγραφών μεταδεδομένων (με τη μορφή XML αρχείων), οι οποίες είναι συμβατές με το βασικό σχήμα META-SHARE. Τις εγγραφές αυτές μπορούν οι χρήστες να τις μεταφορτώσουν στην ΕΥ CLARIN:EL και να τις χρησιμοποιήσουν ως βάση για να αλλάξουν και να προσθέσουν στοιχεία που αφορούν τους δικούς τους πόρους.

Πιο συγκεκριμένα, διατίθενται εγγραφές (XML αρχεία) για τα ακόλουθα είδη πόρων:

- μονόγλωσσα κειμενικά σώματα
- παράλληλα κειμενικά σώματα
- ορολογικά λεξικά

Περιβάλλον τεκμηρίωσης πόρων CLARIN:EL Χρήση του εργαλείου τεκμηρίωσης

Οι εγγραφές μεταδεδομένων που περιγράφουν τους ΓΠ που διατίθενται μέσω του CLARIN:EL πρέπει να αποθηκευθούν στην Υποδομή CLARIN:EL, διαδικασία που υποστηρίζεται από το **εργαλείο τεκμηρίωσης CLARIN:EL**.

Για να έχετε πρόσβαση στο εργαλείο τεκμηρίωσης θα πρέπει να έχετε λογαριασμό χρήστη με δικαιώματα τεκμηριωτή. Περισσότερες πληροφορίες για τη διαδικασία που θα πρέπει να ακολουθήσετε για τον διαμοιρασμό πόρων θα βρείτε [εδώ](#). Μόλις ενημερωθείτε ότι έχετε την κατάλληλη εξουσιοδότηση, θα μπορείτε να προσθέσετε περιγραφές για τους πόρους σας. Ανάλογα με τον ρόλο χρήστη που θα σας έχει αποδοθεί, ελέγξτε το κατάλληλο [εγχειρίδιο χρήστη](#) για περισσότερες πληροφορίες.

Το μοντέλο τεκμηρίωσης META-SHARE Meta-share 2.0 (Έκδοση 3.0.2)

Για την περιγραφή και τεκμηρίωση των πόρων, η Υποδομή CLARIN:EL χρησιμοποιεί το μοντέλο τεκμηρίωσης META-SHARE.

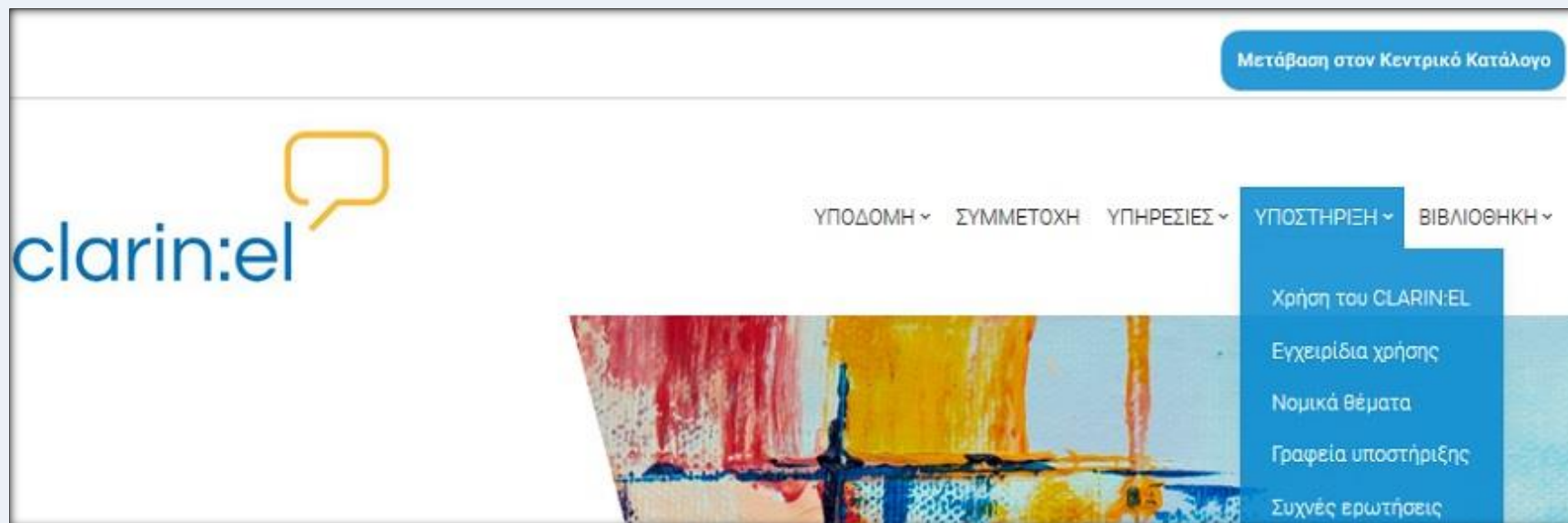
Οι βασικές αρχές του μοντέλου εμπνέονται από το **Component MetaData Infrastructure** (CMDI), που παρέχει ένα πλαίσιο περιγραφής και επαναχρησιμοποίησης έτοιμων υποσυνόλων μεταδεδομένων. Στοιχεία (γνώστα και ως **κατηγορίες δεδομένων**, τα οποία κωδικοποιούν ιδιότητες) που είναι σημασιολογικά συναφή ομαδοποιούνται σε **συστατικά**, τα οποία με τη σειρά τους μπορούν επίσης να ομαδοποιηθούν σε έτοιμα προκατασκευασμένα **προφίλ** για την περιγραφή/τεκμηρίωση συγκεκριμένης κατηγορίας ΓΠ. Με στόχο την προώθηση της επαναχρησιμοποίησης, το CLARIN συντηρεί το **Μητρώο Συστατικών** (Component Registry), όπου οι χρήστες μπορούν να αποθηκεύουν και να διαμοιράζονται συστατικά και προφίλ. Επίσης, μπορούν να συνδέουν στοιχεία και συστατικά που είναι σημασιολογικά παρόμοια με άλλα στοιχεία και συστατικά μέσω του **ISOcat Data Category Registry** (DCR).

Η κεντρική οντότητα στην οντολογία META-SHARE είναι ο **Γλωσσικός Πόρος**, ο οποίος συνδέεται με άλλες περιφερειακές οντότητες μέσω σχέσεων (που κωδικοποιούνται με τη μορφή στοιχείων). Η σύνδεση του ΓΠ με τις περιφερειακές οντότητες επιτρέπει την αναπαράσταση του πλήρους κύκλου ζωής ενός ΓΠ, από τη στιγμή που δημιουργείται μέχρι τη χρήση του. Περιφερειακές οντότητες είναι, ενδεικτικά, τα κείμενα αναφοράς που σχετίζονται με τον ΓΠ (άρθρα, αναφορές, εγχειρίδια χρήσης κτλ.), άτομα και οργανισμοί που συμμετείχαν στη δημιουργία και χρήση των πόρων (δημιουργοί, διαβέτες κτλ.), σχετικά έργα και δραστηριότητες (έργα χρηματοδότησης, δράσεις χρήσης κτλ.), οι άδειες χρήσης κτλ.

Οι ΓΠ ταξινομούνται σε δύο βασικούς άξονες: **είδος πόρου** και **είδος μέσου** (δηλαδή με βάση το μέσο στο οποίο έχει υλοποιηθεί ο ΓΠ).

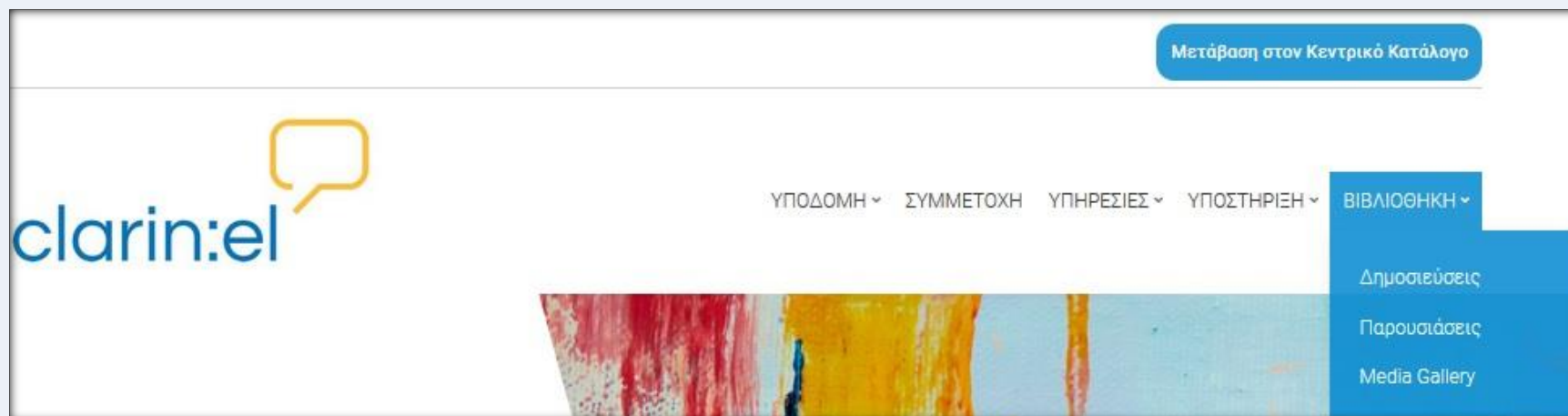
CLARIN: EL

Η δομή



CLARIN: EL

Η δομή



CLARIN: EL

Η δομή: Βιβλιοθήκη → Παρουσιάσεις

 Παρουσίαση της Ερευνητικής Υποδομής CLARIN:EL

2ο διεθνές συνέδριο "Δημιουργική Γραφή"

 Ανοιχτά γλωσσικά δεδομένα: η Υποδομή γλωσσικών πόρων και υπηρεσιών CLARIN:EL (διαφάνειες στα ελληνικά)

12ο διεθνές συνέδριο ελληνικής γλωσσολογίας (ICGL12)

 CLARIN:EL - μια Υποδομή τεκμηρίωσης, διαμοιρασμού και επεξεργασίας γλωσσικών δεδομένων (διαφάνειες στα αγγλικά)

Ημερίδα του CLARIN:EL στη Ρόδο

 Παρουσίαση της συμμετοχής του Πανεπιστημίου του Αιγαίου στην Υποδομή CLARIN:EL

 Παρουσίαση τεχνικών και νομικών θεμάτων & σεναρίων χρήσης

 Γενική παρουσίαση της Υποδομής CLARIN:EL

Ημερίδα του CLARIN:EL στην Κέρκυρα

 Παρουσίαση τεχνικών και νομικών θεμάτων & σεναρίων χρήσης

 Παρουσίαση της συμμετοχής του Ιονίου Πανεπιστημίου στην Υποδομή CLARIN:EL

 Γενική παρουσίαση της Υποδομής CLARIN:EL

Ημερίδα του CLARIN:EL στη Θεσσαλονίκη

 Παρουσίαση της συμμετοχής του ΚΕΓ στην Υποδομή CLARIN:EL

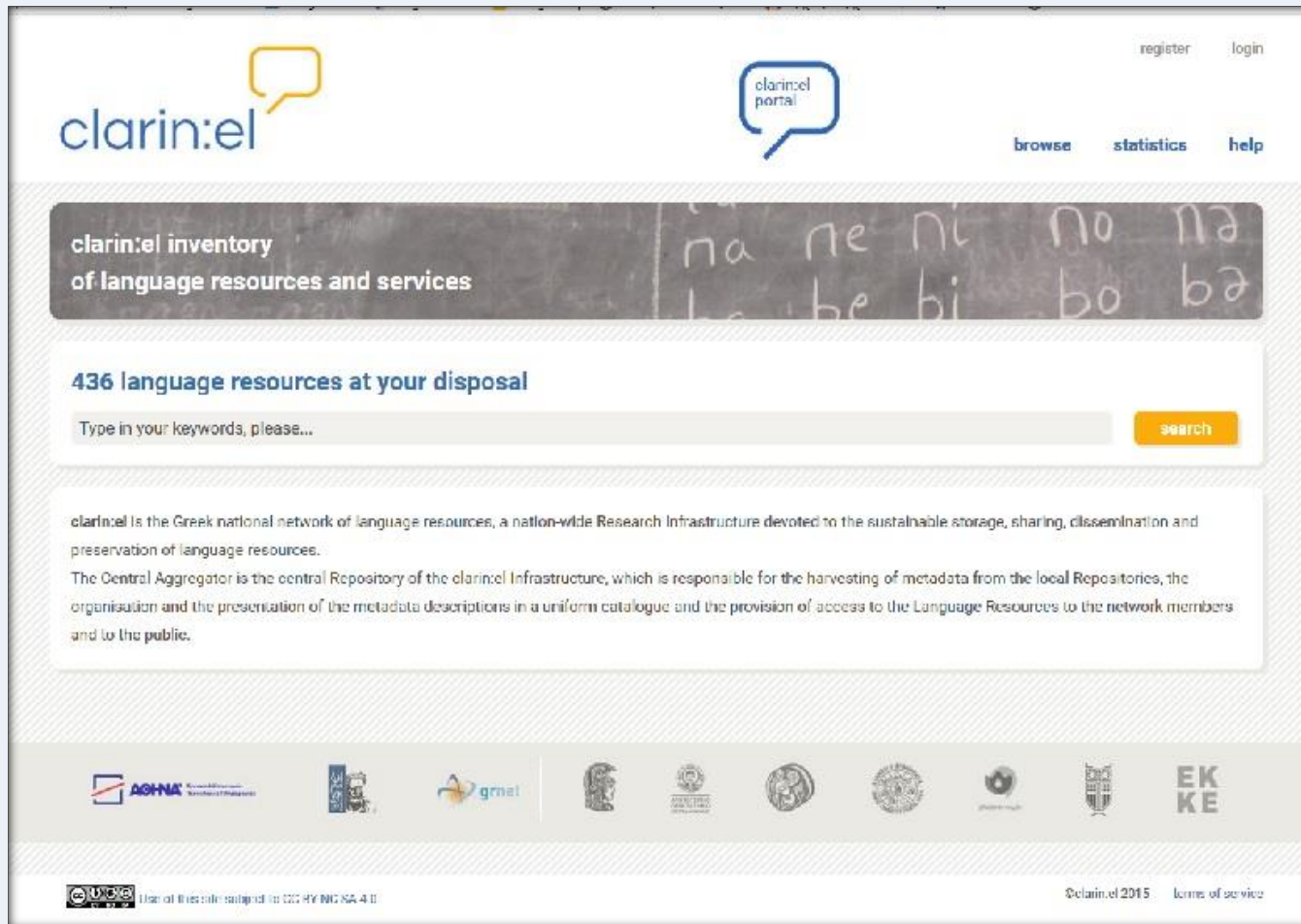
 Παρουσίαση της συμμετοχής του ΑΠΘ στην Υποδομή CLARIN:EL

 Παρουσίαση τεχνικών και νομικών θεμάτων & σεναρίων χρήσης

 Γενική παρουσίαση της Υποδομής CLARIN:EL

CLARIN: EL

Κεντρικός κατάλογος



CLARIN: EL

Είδος πόρων

Το CLARIN EL συγκεντρώνει τόσο δεδομένα όσο και μεταδεδομένα.

Όπου:

- Δεδομένα: το πρωτογενές γλωσσικό υλικό
- Μεταδεδομένα: η τεκμηρίωση των δεδομένων με βάση ένα συγκεκριμένο σχήμα περιγραφής

Η οργάνωση γίνεται σε 3 κατηγορίες:

- Μεταδεδομένα και πόροι στο CLARIN EL
- Μεταδομένα στο CLARIN EL και οι πόροι διαθέσιμοι από αλλού
- Μεταδεδομένα στο CLARIN EL, ενώ οι πόροι δεν είναι ψηφιοποιημένοι ή διαθέσιμοι ηλεκτρονικά

CLARIN: EL

Οργάνωση και χρήστης

- Μεταδεδομένα & πόροι στο CLARIN EL: η βασική κατηγορία
 - το CLARIN EL παρέχει τον πόρο και την τεκμηρίωσή του και άρα ο χρήστης μπορεί
 - ✓ να ψάξει στον κατάλογο με βάση την τεκμηρίωση του πόρου,
 - ✓ να εντοπίσει και να καταφορτώσει τον πόρο ή
 - ✓ να χρησιμοποιήσει κάποια υπηρεσία γλωσσικής επεξεργασίας και να πάρει τα αποτελέσματα στον υπολογιστή του.
 - ✓ Το CLARIN EL φέρει την ευθύνη για τη διάθεση του πόρου, σύμφωνα με τους όρους διάθεσής του όπως καθορίζονται από τον πάροχο.

Παράδειγμα:

CLARIN: EL

Οργάνωση και χρήστης

1. Ο ΠΟΡΟΣ βρίσκεται στο clarin:el και **ΕΙΝΑΙ** επεξεργάσιμος

 **Babiniotis archive**

 142  2  46

other names Αρχείο Μπαμπινιώτη

PID <http://hdl.grnet.gr/11500/UOA-0000-0000-2515-F>

description
The "Babiniotis archive" includes the written transcripts of authentic everyday conversations between students of the Department of Linguistics that took place between 2001 and 2006, of which the original sound archives are not available. The data were transcribed as part of the undergraduate project for Prof. Babiniotis' class on textlinguistics of several years and were subsequently checked and standardized. It includes 618 individual archives with a total 1.747.330 words.

Το "Αρχείο Μπαμπινιώτη" περιλαμβάνει τις γραπτές μεταγραφές αυθεντικών καθημερινών συνομιλιών μεταξύ φοιτητών του Τομέα Γλωσσολογίας που πραγματοποιήθηκαν μεταξύ 2001 και 2006, για τις οποίες δεν υπάρχουν τα ηχητικά αρχεία. Τα δεδομένα μεταγράφηκαν ως μέρος της προπτυχιακής εργασίας στο μάθημα της Κειμενογλωσσολογίας του Καθ. Μπαμπινιώτη επί σειρά ετών και στη συνέχεια ελέγχθηκαν και κανονικοποιήθηκαν. Περιλαμβάνει 618 αρχεία με σύνολο 1.747.330 λέξεις.

[download](#) [edit resource](#) [process](#) [back](#)

Οργάνωση και χρήστης

clarin:el

CLARIN: EL

Οργάνωση και χρήσης

2. Ο ΠΟΡΟΣ βρίσκεται στο clarin:el αλλά **ΔΕΝ** είναι επεξεργάσιμος

Glossary of feminist theory

 63  1  25

other names Γλωσσάρι Φεμινιστικής Θεωρίας

PID <http://hdl.gnnet.gr/11500/UOA-0000-0000-25EB-E>

description

This glossary includes terms from the field of feminist theory and criticism as well as the broader area of gender studies and its contribution to literary criticism. As no such dictionary exists in the Greek language, the terms used in feminist theory and in feminist literary criticism have been so far translated quite inconsistently. The glossary intends to fill this void through the compilation of approximately 100 such terms, contributing at the same time to a research project involving many researchers at the National and Kapodistrian University of Athens, the electronic lexicon Fylopaedia which serves as a self-learning tool for students interested in gender studies.

Το Γλωσσάρι Φεμινιστικής Θεωρίας περιλαμβάνει όρους από το χώρο της φεμινιστικής θεωρίας και κριτικής, καθώς και από το ευρύτερο πλαίσιο της μελέτης του φύλου και της συμβολής της στην κριτική της λογοτεχνίας. Σημειώνεται ότι δεν υπάρχει τέτοιο εργαλείο στην ελληνική και επομένως τα νοήματα που χρησιμοποιούνται στο χώρο της φεμινιστικής θεωρίας και κριτικής της λογοτεχνίας δεν έχουν μέχρι τώρα αποδοθεί με ενιαίο τρόπο. Η ληματοποίηση περίπου 100 τέτοιων όρων αποτελεί τη συμβολή του μαθήματος «Θεωρία Λογοτεχνίας και Λεξικογραφία» σε ένα ερευνητικό έργο, στο οποίο συνεργάζονται πολλές/οί επιστήμονες του ΕΚΠΑ: στη σύνταξη του ηλεκτρονικού λεξικού με τίτλο «Φυλο-Παιδεία» (Fylo-raedia), το οποίο επιδιώκει να χρησιμεύσει ως εργαλείο αυτο-εκπαίδευσης για φοιτήτριες, φοιτητές και όποιες/ους / άλλες/ους ενδιαφερόμενες/ους για το χώρο των σπουδών φύλου.

 download

edit resource

back

CLARIN: EL

Οργάνωση και χρήστης

Distribution	 	Resource Creation
Availability Available - Restricted Use	Bilingual text lexicalConceptualResource	Resource Creator Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών , Interdepartmental M.A. programme 
Licence CC - BY - NC Restrictions: Academic - Non Commercial Use, Attribution Distribution Access/Medium: Downloadable Licensors: National and Kapodistrian University of Athens_School of Philosophy_Faculty of English Language and 	Languages English Greek, Modern (1453 -)	Funding Project Θεωρία Λεξικογραφίας και Εφαρμογές, Lexicography: Theory and Applications Funding Types: Eu Funds, National Funds Project duration: 10/01/2002 - 07/30/2008
Contact Person Maria Sifianou 	Linguality Linguality type: Bilingual	Metadata Created: 12/17/2015 Last Updated: 12/27/2015 Source: http://uoa.clarin.gr
	Text Format Application/pdf	
	Size 100 Entries	
	Character encoding US - ASCII	
	Domains Sociology & Anthropology (DDC301) Language (DDC400)	
	Modalities Written Language	
	Time Coverage Modern time	
	Geographic coverage Greece	

CLARIN: EL

Οργάνωση και χρήστης

- Μεταδεδομένα στο CLARIN EL, αλλά όχι οι πόροι
 - το CLARIN EL παρέχει την τεκμηρίωση του πόρου αλλά όχι τον ίδιο τον πόρο (τον οποίο δεν έχει στα αποθετήριά του)
 - παρέχει απλώς σύνδεσμο για το σημείο όπου βρίσκεται ο πόρος, ο οποίος παρέχεται από άλλη υποδομή ή φορέα
 - το CLARIN EL δεν φέρει καμία ευθύνη για την διαθεσιμότητα του πόρου

Παράδειγμα:

CLARIN: EL

Οργάνωση και χρήστης

3. ΜΟΝΟ τα μεταδεδομένα βρίσκονται στο clarin:el

GLOSSOLOGIA

142 2 59

other names ΓΛΩΣΣΟΛΟΓΙΑ

URLs <http://glossologia.phil.uoa.gr>

PID <http://hdl.gnnet.gr/11500/UOA-0000-0000-257A-E>

description

Glossologia is a journal of general and historical Greek linguistics, focusing on the synchrony and diachrony of the Greek language, as well as publishing studies on theoretical, applied and cross-linguistic research. The journal accepts articles and book reviews in Greek, English, French and German.

Ετήσιο ελληνικό περιοδικό Γενικής και Ιστορικής Γλωσσολογίας. Φιλοξενεί μελέτες και βιβλιοκρισίες στην Ελληνική, Αγγλική, Γαλλική και Γερμανική που αφορούν, κατά προτεραιότητα, τη συγχρονία και τη διαχρονία της Ελληνικής γλώσσας, αλλά και μελέτες γενικότερης θεωρητικής, εφαρμοσμένης και διαγλωσσικής έρευνας.

CLARIN: EL

Οργάνωση και χρήστης

Distribution

Availability

Available - Restricted Use

Licence

CC - BY - NC
Restrictions: Academic - Non Commercial Use, Attribution
Distribution Access/Medium: Accessible Through Interface
Execution location: *hidden*
Licensors:
National and Kapodistrian University of Athens 
Distribution rights holders:
Department of Linguistics 

Contact Person

Dionysios Goutsos 

text 

Multilingual text corpus

Languages

French
Greek, Modern (1453 -)
English
German

Linguality

Linguality type: Multilingual
Multi-linguality type: Parallel

Text Format

Application/pdf

Size

167 Articles

Character encoding

UTF - 8

Domains

Linguistics (DDC410)

Modalities

Written Language

Time Coverage

Modern time

Geographic coverage

Greece

Metadata

Created: 11/04/2015
Last Updated: 11/22/2018
Source: <http://uoa.clarin.gr>

CLARIN: EL

Είδος πόρων

► Domain

► Geographic Coverage

► Time Coverage

► Publication Status

► Processing Service

▼ Processable

✚ No (12)

✚ Yes (1)

resource type:

Corpus: 

Lexical/Conceptual: 

Tool/Service: 

Language: 

Description: 

media type:

Text: 

English Greek, Modern (1453-)

 **GLOSSOLOGIA** 

English French German Greek, Modern (1453-)

↓ 59  141

 **aformes** 

English Greek, Modern (1453-)

↓ 6  84

 **Corpus of Greek Texts** 

Greek, Modern (1453-)

↓ 23  88

 **French-greek and greek-french lexicon of football terms** 

French Greek, Modern (1453-)

↓ 9  37

 **Diachronic corpus of Greek of the 20th century** 

Greek, Modern (1453-)

↓ 26  64

 **Babiniotis archive** 

Greek, Modern (1453-)

 ↓ 46  143

CLARIN: EL

Οργάνωση και χρήστης

Ο πόρος δεν είναι διαθέσιμος διαδικτυακά ή δεν είναι καν σε ψηφιακή μορφή

Είναι όμως χρήσιμος! Τι γίνεται;

- ΔΕΝ είναι προτεραιότητα του CLARIN EL η συλλογή τέτοιου τύπου υλικού
- Παρόλα αυτά, το CLARIN EL παρέχει την τεκμηρίωση και παραπέμπει στον δημιουργό/διαθέτη του πόρου σε χωριστό κατάλογο "συμπληρωματικού υλικού", ώστε να μην υπάρξει σύγχυση με τους πραγματικούς πόρους

CLARIN: EL

Προδιαγραφές συγκέντρωσης πόρων

- ψηφιοποιημένα κείμενα γραπτού ή προφορικού (μεταγγραμμένου) λόγου
- ηχητικά αρχεία καταγεγραμμένου προφορικού λόγου
- βίντεο
- λεξικά/γλωσσάρια
- συλλογές κειμένων

CLARIN: EL

Προδιαγραφές συγκέντρωσης πόρων

Γλώσσα

- μονόγλωσσοι (Ελληνικά)
- δίγλωσσοι ή πολύγλωσσοι (που περιλαμβάνουν Ελληνικά)
- προτεραιότητα στα Νέα Ελληνικά
- δεν αποκλείονται παλαιότερες μορφές και διάλεκτοι, απλώς είναι σε χαμηλότερη προτεραιότητα

CLARIN: EL

Προδιαγραφές συγκέντρωσης πόρων

- Μέσο
 - Κείμενο
 - Ήχος
 - Βίντεο
 - Εικόνα
- Κειμενικό είδος / επίπεδο λόγου
 - σε οποιοδήποτε τομέα, επίπεδο λόγου ή κειμενικό είδος
 - λογοτεχνία / επιστήμες / τύπος ...
 - οικείο/επίσημο ύφος
 - άρθρα / μυθιστορήματα / δημόσια έγγραφα / συγγράμματα / επιστολές / αναρτήσεις σε κοινωνικά μέσα δικτύωσης ...

CLARIN: EL

Προδιαγραφές συγκέντρωσης πόρων

- Μορφότυποι κειμενικού υλικού
 - txt, xml, html, doc/rtf, pdf, ppt
 - τελευταία προτεραιότητα: κείμενο σε μορφή εικόνας (tiff, jpg)
 - **δακτυλόγραφο & χειρόγραφο ΟΧΙ**
- Μορφότυποι ηχητικού υλικού & πολυμεσικού υλικού (βίντεο)
 - Ηχητικό υλικό
 - ✓ ηχητικό/ακουστικό αρχείο wav file, Mp3, απλή ορθογραφική μεταγραφή από απομαγνητοφώνηση, φωνητική μεταγραφή σε IPA
 - Πολυμεσικό υλικό
 - ✓ Mp4, avi

CLARIN: EL

Προδιαγραφές συγκέντρωσης πόρων

- Κωδικοποίηση
 - **UTF 8**
- Εικόνα
 - jpg, tiff, pdf (όχι σκαναρισμένα)
 - φωτογραφίες με λεζάντες: **ΝΑΙ**
 - φωτογραφίες σκέτες: **ΌΧΙ**
- Λεξικά και θησαυροί
 - βέλτιστη επιλογή: το περιεχόμενό τους (=τα λήμματα), ώστε να είναι δυνατόν να αποθηκεύονται και να καταφορτώνονται από το CLARIN σε αυτή την περίπτωση
 - ✓ LMF, tbx
 - ✓ txt, csv
 - ✓ xml, SKOS / RDF
 - λιγότερο επιθυμητή επιλογή: διαθέσιμα μέσω διεπαφής, οπότε τα τεκμηριώνουμε και κάνουμε απλή παραπομπή στο url τους
 - έντυπα: **ΟΧΙ**, εκτός αν είναι εύκολα μετατρέψιμα σε ηλεκτρονικά επεξεργάσιμη μορφή

ΕΥΧΑΡΙΣΤΩ!