

Ενημερωτικός Οδηγός Σπουδών

Ψηφιακά Εργαλεία για τη Διαχρονική Ανάλυση της Γλώσσας

Τίτλος	Ψηφιακά Εργαλεία για τη Διαχρονική Ανάλυση της Γλώσσας
Κωδικός ΚΕΔΙΒΙΜ	1829
Πιστοποιητικό	Πιστοποιητικό Εξειδικευμένης Επιμόρφωσης
Επίπεδο	6 του Ευρωπαϊκού Πλαισίου Προσόντων
Διάρκεια	8 μήνες
Συνολικός χρόνος ενασχόλησης	300 ώρες
Μέθοδος	Ασύγχρονη εξ αποστάσεως, στην πλατφόρμα Open eClass
Γλώσσα	Ελληνικά
Αιτήσεις	έως 26 Οκτωβρίου 2026
Περίοδος υλοποίησης	30 Οκτωβρίου 2026 έως 29 Ιουνίου 2027
Τέλη	500 €, σε έως 4 δόσεις, ή 475 € με εφάπαξ πληρωμή
Ακαδημαϊκός Υπεύθυνος	Νικόλαος Λαβίδας, Καθηγητής Διαχρονικής Γλωσσολογίας

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών (ΕΚΠΑ), Φιλοσοφική Σχολή, Τμήμα Αγγλικής Γλώσσας και Φιλολογίας, Τομέας Γλώσσας-Γλωσσολογίας, Εργαστήριο Διαχρονικής Γλωσσολογίας και Γλωσσικής Αλλαγής AthDGC. Κέντρο Επιμόρφωσης και Διά Βίου Μάθησης (ΚΕΔΙΒΙΜ ΕΚΠΑ).

1. Σκοπός του προγράμματος

Το πρόγραμμα του Κέντρου Επιμόρφωσης και Διά Βίου Μάθησης του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών (ΚΕΔΙΒΙΜ ΕΚΠΑ) παρέχει τις γνώσεις και τις δεξιότητες που χρειάζονται για την κατανόηση της γλωσσικής αλλαγής και για την ανάλυση γλωσσικών δεδομένων σε βάθος χρόνου. Οι εκπαιδευόμενοι εξοικειώνονται με ψηφιακά εργαλεία και υπολογιστικές μεθόδους, με τη στατιστική ανάλυση και μοντελοποίηση γλωσσικών δεδομένων και με την εφαρμογή τους σε σώματα κειμένων, ώστε να μπορούν να διαχειρίζονται ψηφιακά αρχεία και συλλογές, να σχεδιάζουν, να υλοποιούν και να αξιολογούν ερευνητικά έργα και να αναπτύσσουν ψηφιακά εκπαιδευτικά εργαλεία.

Η θεωρητική βάση του προγράμματος είναι ότι ο τρόπος με τον οποίο επισημειώνεται ένα σώμα κειμένων καθορίζει ποια ερωτήματα μπορούν να τεθούν σε αυτό. Η εργασία γίνεται κυρίως σε πραγματικά κείμενα της ελληνικής διαχρονίας, με τη μεθοδολογία του ερευνητικού προγράμματος PROIEL και του διαχρονικού σώματος κειμένων AthDGC (Athens Digital Glossa Chronos). Όσοι το επιθυμούν μπορούν, μετά τον ποιοτικό έλεγχο και με τη σύμφωνη γνώμη τους, να εντάξουν επισημειωμένο δείγμα της εργασίας τους στην επόμενη ανοιχτή έκδοση του σώματος AthDGC.

2. Σε ποιους απευθύνεται και τι προϋποθέτει

Το πρόγραμμα απευθύνεται σε κλασικούς και βυζαντινούς φιλόλογους, σε νεοελληνιστές και σε φιλόλογους της αγγλικής, της γερμανικής, της γαλλικής, της ισπανικής, της ιταλικής και άλλων ινδοευρωπαϊκών γλωσσών, σε εκπαιδευτικούς με ενδιαφέρον για ψηφιακές μεθόδους, σε μεταπτυχιακούς φοιτητές γλωσσολογίας, σε ερευνητές των ψηφιακών ανθρωπιστικών επιστημών και σε επαγγελματίες της γλωσσικής τεχνολογίας. Προϋποθέτει βασικές γνώσεις χειρισμού υπολογιστή, πρόσβαση στο διαδίκτυο και λογαριασμό ηλεκτρονικού ταχυδρομείου. Δεν απαιτείται προηγούμενη εμπειρία στον προγραμματισμό, ο οποίος διδάσκεται από τις βασικές αρχές του.

3. Οργάνωση της εκπαίδευσης

Οι 300 ώρες του προγράμματος αντιστοιχούν κατά μέσο όρο σε 8 έως 9 ώρες την εβδομάδα. Δεν υπάρχουν υποχρεωτικά μαθήματα σε συγκεκριμένες ώρες. Οι καταγεγραμμένες διαλέξεις, τα tutorials, οι ασκήσεις και τα φόρουμ συζητήσεων βρίσκονται στην πλατφόρμα Open eClass του ΚΕ-ΔΙΒΙΜ, όπου δίνονται και οι οδηγίες εγκατάστασης του λογισμικού. Τα βασικά εργαλεία, όπως η Python, η R, το NLTK και το spaCy, είναι ελεύθερο λογισμικό.

Αξιολόγηση

Σκέλος	Βάρος
Πρακτικές ασκήσεις	40%
Ομαδικές εργασίες (projects) ανά ενότητα	20%
Τελική ερευνητική εργασία	40%

Για την επιτυχή ολοκλήρωση απαιτούνται η μελέτη όλου του ασύγχρονου υλικού, η επιτυχής υποβολή όλων των ασκήσεων με βαθμό τουλάχιστον 5 στα 10 και η τελική ερευνητική εργασία.

4. Διδακτικές ενότητες

Διδακτική Ενότητα 1: Θεμέλια Διαχρονικής Γλωσσολογίας (60 ώρες)

Η ενότητα εξετάζει τους μηχανισμούς της γλωσσικής αλλαγής, τις θεωρίες της γραμματικοποίησης και τις φωνητικές και σημασιολογικές μεταβολές, με μελέτες περίπτωσης από την ελληνική και άλλες γλώσσες, και εισάγει τις βασικές έννοιες της γλωσσολογίας των σωμάτων κειμένων (corpus linguistics).

Υποενότητες:

- Εισαγωγή στη γλωσσική αλλαγή και εξέλιξη
- Μηχανισμοί φωνητικής και φωνολογικής αλλαγής
- Γραμματικοποίηση και σημασιολογική μεταβολή
- Γλωσσολογία των σωμάτων κειμένων για διαχρονική ανάλυση

- Μελέτες περίπτωσης από την ελληνική και άλλες γλώσσες

Με την ολοκλήρωση της ενότητας ο/η εκπαιδευόμενος/η μπορεί να:

- αναλύει τους μηχανισμούς φωνολογικής, μορφολογικής και συντακτικής αλλαγής σε διαχρονικά δεδομένα,
- ταξινομεί τύπους γλωσσικών μεταβολών (γραμματικοποίηση, αναλογία, επανανάλυση),
- εφαρμόζει τη συγκριτική μέθοδο για την ανασύσταση προγενέστερων γλωσσικών σταδίων,
- αξιολογεί θεωρίες γλωσσικής εξέλιξης με βάση εμπειρικά δεδομένα,
- κατασκευάζει φυλογενετικά δέντρα γλωσσικής συγγένειας,
- διακρίνει εσωτερικούς και εξωτερικούς παράγοντες γλωσσικής αλλαγής,
- τεκμηριώνει χρονολογικά στάδια γλωσσικής εξέλιξης με φιλολογικά τεκμήρια.

Διδακτική Ενότητα 2: Υπολογιστικές Μέθοδοι και Εργαλεία (60 ώρες)

Η ενότητα διδάσκει προγραμματισμό σε Python και R για γλωσσική ανάλυση, επεξεργασία κειμένων με τις βιβλιοθήκες NLTK και spaCy, κανονικές εκφράσεις (regular expressions) και διαδικτυακή άντληση δεδομένων (web scraping), με πρακτική εφαρμογή σε πραγματικά γλωσσικά δεδομένα.

Υποενότητες:

- Βασικές αρχές Python και δομές δεδομένων για φιλόλογους και γλωσσολόγους
- Επεξεργασία και ανάλυση κειμένων με Python
- Οι βιβλιοθήκες NLTK και spaCy για την επεξεργασία φυσικής γλώσσας
- Διαδικτυακή άντληση δεδομένων και διεπαφές προγραμματισμού εφαρμογών (API) για τη συλλογή γλωσσικών δεδομένων
- Αυτοματοποίηση ροών εργασίας για γλωσσολογικές εργασίες

Με την ολοκλήρωση της ενότητας ο/η εκπαιδευόμενος/η μπορεί να:

- γράφει προγράμματα (scripts) σε Python για την αυτόματη εξαγωγή γλωσσικών δεδομένων,
- χειρίζεται εργαλεία επεξεργασίας φυσικής γλώσσας (NLTK, spaCy) για την ανάλυση ιστορικών κειμένων,
- κατασκευάζει κανονικές εκφράσεις (regex) για την αναζήτηση γλωσσικών μοτίβων,
- διαχειρίζεται βάσεις δεδομένων γλωσσικών πληροφοριών με SQL,
- εφαρμόζει αλγόριθμους τμηματοποίησης σε λεκτικές μονάδες (tokenization) και λημματοποίησης (lemmatization) σε ιστορικά κείμενα,
- αυτοματοποιεί διαδικασίες επισημείωσης (annotation) γλωσσικών δεδομένων,
- επεξεργάζεται δεδομένα από ψηφιακά σώματα κειμένων με υπολογιστικά εργαλεία.

Διδακτική Ενότητα 3: Στατιστική Ανάλυση και Μοντελοποίηση (60 ώρες)

Η ενότητα καλύπτει την περιγραφική στατιστική, τους ελέγχους υποθέσεων και τα μοντέλα παλινδρόμησης για γλωσσικά φαινόμενα, την εφαρμογή μοντέλων χρονοσειρών στη γλωσσική εξέλιξη και την οπτικοποίηση δεδομένων.

Υποενότητες:

- Οπτικοποίηση με ggplot2 και matplotlib
- Περιγραφική στατιστική για γλωσσικά δεδομένα
- Έλεγχοι υποθέσεων και στατιστική ανάλυση
- Μοντέλα παλινδρόμησης για τη γλωσσική εξέλιξη
- Χρονοσειρές και ανάλυση τάσεων

Με την ολοκλήρωση της ενότητας ο/η εκπαιδευόμενος/η μπορεί να:

- εκτελεί στατιστικούς ελέγχους (t-test, chi-square) σε γλωσσολογικά δεδομένα,
- υπολογίζει δείκτες γλωσσικής ποικιλότητας και πολυπλοκότητας,
- δημιουργεί μοντέλα παλινδρόμησης για την πρόβλεψη γλωσσικών αλλαγών,
- οπτικοποιεί γλωσσικές τάσεις με διαγράμματα και χάρτες θερμότητας,
- εφαρμόζει τεχνικές ομαδοποίησης (clustering) σε γλωσσικά φαινόμενα,
- ερμηνεύει στατιστικά αποτελέσματα στο πλαίσιο της γλωσσολογικής θεωρίας,
- αξιολογεί τη στατιστική σημαντικότητα διαχρονικών μεταβολών.

Διδακτική Ενότητα 4: Εφαρμογές σε Ιστορικές Γλώσσες (60 ώρες)

Η ενότητα εστιάζει στην ψηφιακή επεξεργασία αρχαίων και μεσαιωνικών κειμένων: οπτική αναγνώριση χαρακτήρων (OCR) σε παλαιές εκδόσεις, διαχείριση του πολυτονικού, κωδικοποίηση κατά το πρότυπο TEI/XML και δημιουργία ψηφιακών κριτικών εκδόσεων.

Υποενότητες:

- OCR και ψηφιοποίηση παλαιών εκδόσεων
- Διαχείριση του πολυτονικού και ιστορικών γραφών
- Κωδικοποίηση TEI/XML για κριτικές εκδόσεις
- Ανάπτυξη διαχρονικών σωμάτων κειμένων
- Εργαλεία για αρχαία και μεσαιωνικά κείμενα

Με την ολοκλήρωση της ενότητας ο/η εκπαιδευόμενος/η μπορεί να:

- κωδικοποιεί ιστορικά κείμενα κατά το πρότυπο TEI/XML,
- δημιουργεί ψηφιακές κριτικές εκδόσεις με παραλλαγές χειρογράφων,
- εφαρμόζει τεχνικές OCR για την ψηφιοποίηση παλαιών εντύπων,
- κατασκευάζει γλωσσάρια και λεξικά ιστορικών όρων,

- χρησιμοποιεί υπολογιστικά εργαλεία για την επεξεργασία χειρογράφων και για εκδοτικές εργασίες,
- συγκρίνει διαφορετικές εκδοχές κειμένων με αλγόριθμους αντιπαραβολής,
- επεξεργάζεται πολύγλωσσα ιστορικά σώματα κειμένων και παράλληλα κείμενα.

Διδακτική Ενότητα 5: Ερευνητικό Εργαστήριο και Πρακτική Άσκηση (60 ώρες)

Στην τελευταία ενότητα οι εκπαιδευόμενοι αναπτύσσουν πρωτότυπο ερευνητικό έργο, δημιουργούν ανοιχτό σύνολο δεδομένων (dataset) ή εργαλείο, συγγράφουν επιστημονική εργασία και προετοιμάζουν την παρουσίασή της σε συνέδριο.

Υποενότητες:

- Σχεδιασμός ερευνητικού έργου
- Μεθοδολογία και συλλογή δεδομένων
- Ανάπτυξη ανοιχτής συλλογής γλωσσικών δεδομένων ή γλωσσικού τεχνολογικού εργαλείου
- Συγγραφή επιστημονικής εργασίας
- Παρουσίαση και διάχυση των αποτελεσμάτων

Με την ολοκλήρωση της ενότητας ο/η εκπαιδευόμενος/η μπορεί να:

- σχεδιάζει πλήρη ερευνητική πρόταση με μεθοδολογία και χρονοδιάγραμμα,
- υλοποιεί αυτόνομο ερευνητικό έργο από τη συλλογή έως την ανάλυση των δεδομένων,
- συντάσσει επιστημονική εργασία σύμφωνα με τα διεθνή πρότυπα δημοσίευσης,
- παρουσιάζει ερευνητικά αποτελέσματα με επαγγελματικές τεχνικές παρουσίασης,
- δημοσιεύει ερευνητικά δεδομένα σε αποθετήρια ανοιχτής πρόσβασης (Zenodo, GitHub),
- αξιολογεί κριτικά ερευνητικές εργασίες συναδέλφων (peer review),
- συνεργάζεται σε ομαδικά ερευνητικά έργα με κατανομή ρόλων και ευθυνών.

5. Πιστοποίηση και μοριοδότηση

Το πιστοποιητικό εκδίδεται στα ελληνικά και αφορά το πρόγραμμα στο σύνολό του, αφού δεν εκδίδονται χωριστά πιστοποιητικά ανά διδακτική ενότητα. Σύμφωνα με την επίσημη σελίδα του προγράμματος, μοριοδοτείται ως εξής:

- δύο (2) μόρια σε υποψηφίους για διορισμό ως μόνιμοι εκπαιδευτικοί στη γενική εκπαίδευση, στην ειδική αγωγή και ως ειδικό εκπαιδευτικό προσωπικό, ν. 4589/2019 (ΦΕΚ 13/Α/29.1.2019),
- δύο (2) μόρια σε υποψηφίους για διορισμό ως αναπληρωτές εκπαιδευτικοί και ειδικό εκπαιδευτικό προσωπικό μέσω του Ανώτατου Συμβουλίου Επιλογής Προσωπικού (ΑΣΕΠ), ν. 4589/2019 (ΦΕΚ 13/Α/29.1.2019),

- δέκα (10) μόρια σε υποψηφίους για τους κλάδους εκπαιδευτικών των κατηγοριών ΤΕ και ΔΕ της γενικής εκπαίδευσης και της ειδικής αγωγής και εκπαίδευσης, ΥΑ 48190/ΕΙ (ΦΕΚ 1088/Β/2.4.2019),
- ένα (1) έως δύο (2) μόρια στην επιλογή περιφερειακών διευθυντών εκπαίδευσης, διευθυντών πρωτοβάθμιας και δευτεροβάθμιας εκπαίδευσης, προϊσταμένων των Κέντρων Διεπιστημονικής Αξιολόγησης, Συμβουλευτικής και Υποστήριξης, συμβούλων εκπαίδευσης και διευθυντών ή προϊσταμένων των τμημάτων εκπαιδευτικών θεμάτων, ν. 4823/2021 (ΦΕΚ 136/Α/3.8.2021).

Οι όροι κάθε διαδικασίας διορισμού ή επιλογής ορίζονται από την προκήρυξή της.

6. Τέλη παρακολούθησης

Οι κατηγορίες μερικής απαλλαγής, όπως αναγράφονται στην επίσημη σελίδα του προγράμματος, είναι οι εξής:

Τέλη και εκπτώσεις

Κατηγορία	Έκπτωση	Τέλη	Εφάπαξ
Άτομα με ειδικές ανάγκες	50%	250 €	238 €
Άνεργοι, γονείς τέκνων με ειδικές ανάγκες, άτομα με ετήσιο εισόδημα κάτω από 12.000 €	30%	350 €	333 €
Πρόσφυγες και μετανάστες	25%	375 €	356 €
Κάτοικοι περιοχών που έχουν κηρυχθεί σε κατάσταση έκτακτης ανάγκης λόγω φυσικών καταστροφών, εκπαιδευόμενοι που έχουν παρακολουθήσει ή παρακολουθούν επιμορφωτικό πρόγραμμα του ΕΚΠΑ	20%	400 €	380 €
Απόφοιτοι τριτοβάθμιας εκπαίδευσης σχολών Παιδαγωγικών, Ανθρωπιστικών και Θετικών Επιστημών και κάτοχοι διδακτικής/παιδαγωγικής επάρκειας, υπάλληλοι του δημόσιου τομέα	15%	425 €	404 €
Γονείς μονογονεϊκών οικογενειών με ετήσιο εισόδημα κάτω από 15.000 €	10%	450 €	428 €
Άτομα ηλικίας άνω των 65 ετών	5%	475 €	451 €

Για κάθε λεπτομέρεια ισχύει η επίσημη σελίδα του προγράμματος.

7. Αίτηση

Η αίτηση γίνεται σε τρία βήματα:

1. Ανοίξτε την επίσημη σελίδα του προγράμματος στο ΚΕΔΙΒΙΜ ΕΚΠΑ (cce.uoa.gr, «Όλα τα προγράμματα», κωδικός 1829) και πατήστε «Κάνε Αίτηση».
2. Επιλέξτε τον «1ο Κύκλο». Θα μεταφερθείτε στην Κυβέλη, την πλατφόρμα εκπαιδευομένων του ΚΕΔΙΒΙΜ.

3. Συνδεθείτε στον λογαριασμό σας στην Κυβέλη ή δημιουργήστε νέο και ολοκληρώστε την αίτηση. Οδηγίες για τη δημιουργία λογαριασμού και τη σύνδεση υπάρχουν στην αρχική σελίδα της Κυβέλης.

Ο [απευθείας σύνδεσμος της αίτησης](#) στην Κυβέλη (cybele.cce.uoa.gr) οδηγεί στη σύνδεση ή στη δημιουργία λογαριασμού και από εκεί κατευθύνει στην αίτηση για τον πρώτο κύκλο.

8. Επικοινωνία

Σπουδές και περιεχόμενο του προγράμματος: Νικόλαος Λαβίδας, Καθηγητής Διαχρονικής Γλωσσολογίας, Ακαδημαϊκός Υπεύθυνος, nlavidas@enl.uoa.gr, τηλ. 210 7277462.

Αιτήσεις, τέλη, πληρωμές και δικαιολογητικά μερικής απαλλαγής: Γραμματεία ΚΕΔΙΒΙΜ ΕΚΠΑ, secr-cce@uoa.gr, τηλ. 210 3689543 και 210 3689544, Δευτέρα έως Παρασκευή, 09:00 έως 17:00, εκτός επίσημων αργιών.

Κυβέλη και Open eClass: Υποστήριξη Πληροφοριακών Συστημάτων, ithelp-cce@uoa.gr.

Ιστότοπος του προγράμματος: athdgc.github.io/Digital-Diachrony