

Τίτλος Μαθήματος: Επεξεργασία Σωμάτων Κειμένων

- **Κωδικός μαθήματος:**
- **Τύπος μαθήματος:**
- **Επίπεδο μαθήματος (Προπτυχιακό/Μεταπτυχιακό):**
- **Έτος σπουδών:**
- **Εξάμηνο σπουδών:**
- **Αριθμός πιστωτικών μονάδων (ECTS):**
- **Αριθμός διδακτικών μονάδων:**
- **Όνομα διδάσκοντος/ διδασκόντων :**
- **Περιεχόμενο μαθήματος:**

Τα τελευταία χρόνια, τα Σώματα Κειμένων, μεγάλες συλλογές πραγματικού γραπτού ή προφορικού λόγου σε ηλεκτρονική μορφή, έχουν σημαντική συμβολή στη μελέτη τομέων της γλωσσολογίας, όπως στη μορφολογία, σύνταξη, σημασιολογία, πραγματολογία, ανάλυση λόγου, λεξικογραφία, ορολογία, κειμενογλωσσολογία, υφογλωσσολογία, κοινωνιογλωσσολογία, ιστορική και συγκριτική γλωσσολογία, εφαρμοσμένη γλωσσολογία, διαλεκτολογία, κ.ά. Λόγω της αξιοπιστίας, ακρίβειας και ταχύτητας που προσφέρει η επεξεργασία της γλώσσας με τη χρήση του υπολογιστή σε πραγματικό γλωσσικό υλικό (ειδικά μεγάλου μεγέθους του οποίου η επεξεργασία θα ήταν αδύνατη αλλιώς), οι εφαρμογές της Επεξεργασίας Σωμάτων Κειμένων απλώνονται σε επιστημονικά πεδία κι εκτός της Γλωσσολογίας, όπως στην Εκπαίδευση, Ψυχολογία, Φιλολογία, Επικοινωνία Πολιτική, Δικαστική Γλωσσολογία, στις Επιστήμες των Μέσων Μαζικής Ενημέρωσης, κ.ά. Η γλώσσα μπορεί να μελετηθεί πλέον από την πραγματική της χρήση και όχι με υποθέσεις, εκτιμήσεις και προκαταλήψεις. Το μάθημα αποτελείται από το θεωρητικό και εφαρμοσμένο μέρος. Στο θεωρητικό, οι φοιτητές/ήτριες εισάγονται στους στόχους και στις εφαρμογές της μελέτης της Γλώσσας με τη χρήση Σωμάτων Κειμένων (ΣΚ), στα είδη ΣΚ, στη σχεδίαση και κατασκευή ΣΚ, στην επισημείωση ΣΚ, στην κανονικοποίηση συχνοτήτων φαινομένων της γλώσσας. Ιδιαίτερη έμφαση δίνεται στο εφαρμοσμένο μέρος, στο οποίο οι φοιτητές/ήτριες εισάγονται στο λειτουργικό σύστημα Linux και μαθαίνουν τις βασικές και κατάλληλες εντολές για τα βασικά προϊόντα επεξεργασίας ΣΚ.

- **Επιδιωκόμενα μαθησιακά αποτελέσματα:**

Με την επιτυχή ολοκλήρωση αυτού του μαθήματος ο φοιτητής/η φοιτήτρια θα πρέπει να:

- Γνωρίζει τι είναι σώμα κειμένων (ΣΚ).
- Γνωρίζει ποια είναι η αναγκαιότητα για τη χρήση ΣΚ.

- Γνωρίζει τις κατηγοριοποιήσεις των ΣΚ.
- Γνωρίζει για την κατασκευή ενός ΣΚ.
- Γνωρίζει τι είναι η κανονικοποίηση και πώς επιτυγχάνεται.
- Γνωρίζει ΣΚ της Ελληνικής Γλώσσας τα οποία διατίθενται ελεύθερα στο διαδίκτυο.
- Γνωρίζει τι είναι το λειτουργικό σύστημα Linux και την καταλληλότητά του για επεξεργασία ΣΚ.
- Γνωρίζει το βασικό περιβάλλον του λειτουργικού συστήματος Linux.
- Γνωρίζει τι είναι το «τερματικό» και να χρησιμοποιεί βασικές εντολές του Linux.
- Μπορεί να προγραμματίσει σε περιβάλλον Linux για την Παραγωγή Λίστας Συχνοτήτων Λέξεων από ΣΚ.

Με το μάθημα «Επεξεργασία Σωμάτων Κειμένων» ο φοιτητής /η φοιτήτρια ενός Θεωρητικού Πανεπιστημιακού Τμήματος αποκτά ικανότητα να διαχειρίζεται θέματα και προβλήματα με τυπικό και αλγοριθμικό τρόπο, προσόν που αποκτάται με την παρακολούθηση μαθημάτων τα οποία σχετίζονται με τις Θετικές Επιστήμες.

Επιπλέον, ο φοιτητής /η φοιτήτρια Πανεπιστημιακής Σχολής Ανθρωπιστικών Επιστημών, με το πέρας του μαθήματος, αποκτά γνώση των πρώτων βημάτων στον υπολογιστικό προγραμματισμό.

- **Συνιστώμενη βιβλιογραφία προς μελέτη:**

α) Εγχειρίδια του μαθήματος:

Φραντζή, Κ. 2012. Εισαγωγή στην Επεξεργασία Σωμάτων Κειμένων. Αθήνα: Ίων.

Γούτσος, Δ. 2015. Εισαγωγή στη Γλωσσολογία Σωμάτων Κειμένων. Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα - Αποθετήριο "Κάλλιπος".

Μικρός, Γ.. 2009. Ποσοτική Ανάλυση της Κοινωνιογλωσσολογικής Ποικιλίας - Θεωρητικές και Μεθοδολογικές Προσεγγίσεις. Αθήνα : Μεταίχμιο.

β) Συμπληρωματική βιβλιογραφία:

Baker, Paul (2009) Contemporary Corpus Linguistics (Contemporary Studies in Linguistics). Continuum International Publishing Group Ltd.

Biber, Douglas, Conrad, Susan and Randi Reppen, 1998. Corpus linguistics: investigating language structure and use. Cambridge University Press.

Cheng, Winnie (2011) Exploring Corpus Linguistics: Language in Action (Routledge Introductions to Applied Linguistics). Routledge.

Hyland, Ken (2013) Corpus Applications in Applied Linguistics. Bloomsbury Academic.

Kennedy, Graeme (1998) An Introduction to Corpus Linguistics. Longman.

McEnery, Anthony, Xiao, Richard and Yukio Tono, 2006. Corpus-based Language Studies: an advanced resource book. Routledge, 2006.

McEnery, Anthony (2011) Corpus Linguistics: Method, Theory and Practice (Cambridge Textbooks in Linguistics). Cambridge University Press

O’Keeffe, Anne (2012) The Routledge handbook of corpus linguistics (Routledge Handbooks in Applied Linguistics). Routledge.

Stubbs Michael (1996) Text and Corpus Analysis: Computer Assisted Studies of Language and Institutions (Language in Society). John Wiley & Sons.

Teubert Wolfgang (2007) Corpus Linguistics: A Short Introduction. Continuum.

- **Διδακτικές και μαθησιακές μέθοδοι:**

Τα μαθήματα γίνεται με διαλέξεις κατά τις οποίες οι φοιτητές/ήτριες έχουν ενεργή συμμετοχή σε θέματα των οποίων οι τρόποι αντιμετώπισης συζητούνται κατά τη διάρκεια των μαθημάτων.

- **Μέθοδοι αξιολόγησης/ βαθμολόγησης:**

Ο τελικός βαθμός είναι αποτέλεσμα και των παρακάτω:

- 1 Εξέταση στο τέλος του εξαμήνου.
- 2 Εργασία (παράδοση στο τέλος του εξαμήνου).

- **Γλώσσα διδασκαλίας:** Ελληνική/ Αγγλική

- **Τρόπος παράδοσης μαθήματος (με φυσική παρουσία, εξ αποστάσεως):** Η παράδοση των μαθημάτων γίνεται με φυσική παρουσία.