

ΙΔΡΥΜΑ	ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ							
ΣΧΟΛΗ	ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ							
ΤΜΗΜΑ	ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ							
ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ	ΜΕΤΑΠΤΥΧΙΑΚΟ - ΔΠΜΣ ΓΛΩΣΣΙΚΗ ΤΕΧΝΟΛΟΓΙΑ							
ΤΙΤΛΟΣ ΜΑΘΗΜΑΤΟΣ	Υπολογιστική Γλωσσολογία και Επεξεργασία Φυσικής Γλώσσας και Επεξεργασία και Διαχείριση Γλωσσικών Πόρων							
ΚΩΔΙΚΟΣ ΜΑΘΗΜΑΤΟΣ	M908	Εξάμηνο	B	ECTS	6			
ΩΡΕΣ ΔΙΔΑΣΚΑΛΙΑΣ	ΘΕΩΡ.		ΦΡΟΝΤ.		ΕΡΓΑΣΤ.			
ΤΥΠΟΣ ΜΑΘΗΜΑΤΟΣ	<p>Επιλέξτε ένα από τα ακόλουθα και διαγράψτε τα υπόλοιπα Βασικό (B) / Επιλογής (E)</p> <table border="1"> <tr> <td>B</td> </tr> <tr> <td>NAI</td> </tr> </table>						B	NAI
B								
NAI								
ΗΛΕΚΤΡΟΝΙΚΗ ΣΕΛΙΔΑ ΜΑΘΗΜΑΤΟΣ (URL)	Διεύθυνση eclass.uoa.gr υποχρεωτική							
ΠΡΟΑΠΑΙΤΟΥΜΕΝΑ ΜΑΘΗΜΑΤΑ	ΟΧΙ							
ΓΛΩΣΣΑ ΔΙΔΑΣΚΑΛΙΑΣ και ΕΞΕΤΑΣΕΩΝ:	ΕΛΛΗΝΙΚΑ							
ΤΟ ΜΑΘΗΜΑ ΠΡΟΣΦΕΡΕΤΑΙ ΣΕ ΦΟΙΤΗΤΕΣ ERASMUS	ΟΧΙ							

<p>ΠΕΡΙΕΧΟΜΕΝΟ ΜΑΘΗΜΑΤΟΣ</p> <p><i>Συμπληρώστε (λεκτική περιγραφή) το περιεχόμενο του μαθήματος αναφέροντας τη δομή και τα θέματα που καλύπτονται.</i></p> <p>Τα αντικείμενα που θα συζητηθούν στο μάθημα παρατίθενται στην συνέχεια. Η παράθεση αντικατοπτρίζει και την δομή του μαθήματος. Τα μαθήματα θα διδαχθούν από τον κ. Χάρη Παπαγεωργίου πλην εκείνων που σαφώς δηλώνεται ότι θα διδαχθούν από την κ. Στέλλα Μαρκαντωνάτου.</p> <ol style="list-style-type: none"> Regular Expressions (Κανονικές εκφράσεις), Text Normalization (Κανονικοποίηση κειμένου) and Edit Distance Classification (Θέματα Κατηγοριοποίησης) και Feedforward neural networks Distributional Semantics & Vector Spaces – Embeddings Language Models (Γλωσσικά μοντέλα) and Recurrent neural networks Contextual Representations & Text Generation (Παραγωγή κειμένου) Sequence Labeling, Named Entity Recognition/Extraction (Εξαγωγή ονοματικών ή ονομασμένων οντοτήτων) & Dependency Parsing (Συντακτική ανάλυση εξαρτήσεων) Argumentation Mining (Αναλυτική Επιχειρηματολογίας & ΕΦΓ) <p>ΣΗΜΑΣΙΟΛΟΓΙΚΟΙ ΚΑΙ ΣΥΝΤΑΚΤΙΚΟΙ ΠΟΡΟΙ (Σ. Μαρκαντωνάτου)</p> <ol style="list-style-type: none"> WordNet, FrameNet BabelNet, Controlled vocabularies (ελεγχόμενα λεξιλόγια) and terminological resources (ορολογικοί πόροι): AAT, LanguaL Treebanks (τράπεζες δένδρων): UPenn, AMR. PARSEME SHARED TASK Semantic Role Labeling (Ετικέτες σημασιολογικών ρόλων), Argument Structure (Ορισματική Δομή) & Event Representations (Αναπαραστάσεις Γεγονότων) Opinion Mining(Εξαγωγή γνώμης) & Sentiment Analysis (Ανάλυση συναισθήματος) Question Answering systems (Συστήματα Ερωταποκρίσεων)

14. Information Extraction (Εξαγωγή πληροφορίας), Common Sense Reasoning & Knowledge Graphs (Γράφοι γνώσης)
15. Ethics, Bias & NLP applications (προαιρετικό)

ΜΑΘΗΣΙΑΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

Περιγράψτε τους στόχους ή/και τα μαθησιακά αποτελέσματα του μαθήματος. Μπορείτε να αναφερθείτε στις επιμέρους κατηγορίες των μαθησιακών αποτελεσμάτων σε επίπεδο γνώσεων, δεξιοτήτων (νοητικών, πρακτικών) και ικανοτήτων. Συμβουλευτείτε το Παράρτημα «Περιληπτικός Οδηγός συγγραφής Μαθησιακών Αποτελεσμάτων» για τη συγγραφή των μαθησιακών αποτελεσμάτων.

Διδακτικοί-Μαθησιακοί Στόχοι -Προσδοκώμενα Μαθησιακά Αποτελέσματα

- Κατανόηση των βασικών μεθόδων, πόρων και εργαλείων που χρησιμοποιούνται στην Επεξεργασία της Φυσικής Γλώσσας.
- Εξοικείωση με μεθόδους δόμησης περιεχομένου στις βάσεις γνώσης
- Υπολογιστικές τεχνικές Επεξεργασίας φυσικής Γλώσσας

Με την επιτυχή ολοκλήρωση του μαθήματος ο φοιτητής/φοιτήτρια θα είναι σε θέση να:

- Αξιοποιήσει βασικούς πόρους στον χώρο της Επεξεργασίας της Φυσικής Γλώσσας.
- Να συμμετέχει στην ανάπτυξη βάσεων γνώσης για διάφορους σκοπούς.
- Να προσαρμόσει και να αναπτύξει μεθοδολογίες ΕΦΓ σε καινοτόμες εφαρμογές

ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ - ΑΞΙΟΛΟΓΗΣΗ

ΤΡΟΠΟΣ ΠΑΡΑΔΟΣΗΣ

Στην τάξη (Πρόσωπο με πρόσωπο)

ΧΡΗΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ

Υποστήριξη μαθησιακής διαδικασίας μέσω της ηλεκτρονικής πλατφόρμας e-class (Παροχή υλικού, Συζητήσεις, Ανακοινώσεις, Ανάθεση εργασιών, Ομάδες φοιτητών)
Επικοινωνία μέσω ηλεκτρονικού ταχυδρομείου
Ζωντανή μετάδοση διαλέξεων

ΟΡΓΑΝΩΣΗ ΔΙΔΑΣΚΑΛΙΑΣ – ΔΙΔΑΚΤΙΚΕΣ ΤΕΧΝΙΚΕΣ

Περιγράφονται αναλυτικά ο τρόπος και τεχνικές διδασκαλίας και αναγράφονται αναλυτικά οι ώρες μελέτης του φοιτητή για κάθε μαθησιακή δραστηριότητα καθώς και οι ώρες μη καθοδηγούμενης μελέτης ώστε ο συνολικός φόρτος εργασίας σε επίπεδο εξαμήνου να αντιστοιχεί στα standards του ECTS (25 ώρες φόρτου εργασίας ανά πιστωτική μονάδα)
Εμπλουτισμένες Διαλέξεις,
Online Διαλέξεις,
Σεμινάρια,
Φροντιστήριο,
Εργαστήριο,
Εργαστηριακή Άσκηση,
Μελέτη & ανάλυση βιβλιογραφίας,
Πρακτική
Εκπόνηση project,
Εκπόνηση ατομικών / ομαδικών εργασιών
Τηλεσυνεργασία (αναφορά σε εργαλεία)
Κλπ

Περιγράφονται αναλυτικά ο τρόπος και τεχνικές διδασκαλίας, αριθμός ατόμων ανά ομάδα κτλ.
Συμπληρώνεται αναλόγως και ο παρακάτω πίνακας.

- Για την ενότητα ΣΗΜΑΣΙΟΛΟΓΙΚΟΙ ΚΑΙ ΣΥΝΤΑΚΤΙΚΟΙ ΠΟΡΟΙ, 25 φοιτητές θα χωριστούν σε ομάδες και κάθε ομάδα θα παρουσιάσει ένα άρθρο ή μια συλλογή σχετικών άρθρων (το μέγεθος των ομάδων αντιστοιχεί στις απαιτήσεις των επιμέρους άρθρων και γι' αυτό ο χωρισμός σε ομάδες θα γίνει σε συνεργασία με την κ. Μαρκαντωνάτου). Οι ομάδες θα παρουσιάσουν και τον πόρο στον οποίο αναφέρεται το άρθρο και θα ετοιμάσουν μικρές ασκήσεις εξοικείωσης των συμφοιτητών τους με τον πόρο. Δύο φοιτητές θα αναλάβουν ατομικά δύο άρθρα και θα εργαστούν όπως και οι ομάδες.
- 4 Εργαστήρια
- 2 Φροντιστήρια
- 4 αναθέσεις Ομαδικών εργασιών

	<ul style="list-style-type: none"> • Εκπόνηση Micro-project 																		
	<table border="1"> <thead> <tr> <th>Δραστηριότητα</th> <th>Φόρτος (ώρες)</th> </tr> </thead> <tbody> <tr> <td>Διαλέξεις</td> <td>32</td> </tr> <tr> <td>MicroProject</td> <td>20</td> </tr> <tr> <td>Φροντιστήριο</td> <td>8</td> </tr> <tr> <td>Εργαστήριο</td> <td>10</td> </tr> <tr> <td>Ομαδική Εργασία</td> <td>40</td> </tr> <tr> <td>Αυτοτελής Μελέτη</td> <td>40</td> </tr> <tr> <td>.....</td> <td>...</td> </tr> <tr> <td>Σύνολο ωρών</td> <td>150</td> </tr> </tbody> </table>	Δραστηριότητα	Φόρτος (ώρες)	Διαλέξεις	32	MicroProject	20	Φροντιστήριο	8	Εργαστήριο	10	Ομαδική Εργασία	40	Αυτοτελής Μελέτη	40	Σύνολο ωρών	150
	Δραστηριότητα	Φόρτος (ώρες)																	
	Διαλέξεις	32																	
	MicroProject	20																	
	Φροντιστήριο	8																	
	Εργαστήριο	10																	
	Ομαδική Εργασία	40																	
	Αυτοτελής Μελέτη	40																	
																	
Σύνολο ωρών	150																		
<p>ΑΞΙΟΛΟΓΗΣΗ ΦΟΙΤΗΤΩΝ <i>Περιγραφή της διαδικασίας αξιολόγησης</i></p> <p>Μέθοδοι αξιολόγησης (Διαμορφωτική ή/και Τελική), Εργαλεία Αξιολόγησης (Δοκιμασία Πολλαπλής Επιλογής, Ερωτήσεις Σύντομης Απάντησης, Ερωτήσεις Ανάπτυξης Δοκιμίων, Επίλυση Προβλημάτων, Γραπτή Εργασία, Εκθεση/Αναφορά, Προφορική Εξέταση, Δημόσια Παρουσίαση, Εργαστηριακή Εργασία, Άλλη/άλλες,) Παροχή ανατροφοδότησης (περιγραφική, μέσω κλίμακας διαβαθμισμένων κριτηρίων)</p>	<p>Περιγράφονται ρητά οι μέθοδοι, τα εργαλεία αξιολόγησης και η παρεχόμενη ανατροφοδότηση αποτελεσμάτων. Συμπληρώνεται αναλόγως και ο παρακάτω πίνακας.</p> <p>Η βαθμολογία των φοιτητών θα προκύψει ως εξής:</p> <ul style="list-style-type: none"> • 60% από τις ομαδικές εργασίες • 37% από το Micro-project & την παρουσίαση άρθρου στο πλαίσιο της Ενότητας ΣΗΜΑΣΙΟΛΟΓΙΚΟΙ ΚΑΙ ΣΥΝΤΑΚΤΙΚΟΙ ΠΟΡΟΙ • 3% από συμμετοχή σε δραστηριότητες και παροχή ανατροφοδότησης 																		
<table border="1"> <thead> <tr> <th>Αξιολόγηση</th> <th>Αριθμός</th> <th>Ποσοστό</th> </tr> </thead> <tbody> <tr> <td>Εκπόνηση microProject</td> <td>1</td> <td>37%</td> </tr> <tr> <td>Εργασίες</td> <td>4</td> <td>60%</td> </tr> <tr> <td>Συμμετοχή</td> <td>1</td> <td>3%</td> </tr> </tbody> </table>	Αξιολόγηση	Αριθμός	Ποσοστό	Εκπόνηση microProject	1	37%	Εργασίες	4	60%	Συμμετοχή	1	3%							
Αξιολόγηση	Αριθμός	Ποσοστό																	
Εκπόνηση microProject	1	37%																	
Εργασίες	4	60%																	
Συμμετοχή	1	3%																	

ΣΥΝΙΣΤΩΜΕΝΗ-ΒΙΒΛΙΟΓΡΑΦΙΑ

- Dan Jurafsky and James H. Martin. 2020-2021 [Speech and Language Processing \(3rd ed. draft\)](#)
- Jacob Eisenstein. 2019 [Natural Language Processing](#)
- Bender, Emily M. and Alex Lascarides. 2019. [Linguistic Fundamentals for Natural Language Processing II: 100 Essentials from Semantics and Pragmatics](#). Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers
- Yoav Goldberg. 2017 [A Primer on Neural Network Models for Natural Language Processing](#)

Για την ενότητα ΣΗΜΑΣΙΟΛΟΓΙΚΟΙ ΚΑΙ ΣΥΝΤΑΚΤΙΚΟΙ ΠΟΡΟΙ

WORDNET

Five papers (first four papers)

<https://wordnetcode.princeton.edu/5papers.pdf>

Jurafsky and Martin 19

<https://web.stanford.edu/~jurafsky/slp3/19.pdf>

FrameNet

<https://framenet.icsi.berkeley.edu/fndrupal/CJFFNintroPPT>

Fillmore Frame Semantics

https://d1wqtxts1xzle7.cloudfront.net/56172451/cognitive-linguistics-basics-readings-dirk-geeraerts.pdf?1522158117=&response-content-disposition=inline%3B+filename%3DCognitive+linguistics+basics+readings+di.pdf&Expires=1609069358&Signature=dqp7Hlp8S7~MaJhNiUwiz-m2qmOSy6M2s71740by-9TFv1RBn4J4xg66uFZFuHIIsUo4cFmcT-JMr9bxtqo1FYuuqv1GGleD~Zs3kMS47Qy5rQaPXQnHv3g~EC190nJhLWc73ag~PcCHKHQt9HKyiTaJ-rwzF1WqAnG0aW6acVhQu34MBJKyykhADgdlxS58ChNRDO4l5fsWoepjKf7dv7TaqAD1v2kh-tZYJHwslwibu3TOUxa9kQiz886TIPGUN--CFqPHVhn4Ez9LTuB4A42nct4EyLkkB7dSFuoxCSvXT1-4~swu0U-CaXbXyomDzoH1eEU1HdAM-GJcW8NTLbg_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA#page=382

Framenet and WordNet as complementary resources

<https://www.aclweb.org/anthology/W09-3021.pdf>

FScore definition

<https://deepai.org/machine-learning-glossary-and-terms/f-score>

BabelNet

Reference BabelNet paper (up to page 235, first 19 pages)

http://wwwusers.di.uniroma1.it/~navigli/pubs/AIJ_2012_Navigli_Ponzetto.pdf

ΘΗΣΑΥΡΟΙ

Stella G. Dextre Clarke. 2001. Thesaural Relationships. In Carol A. Bean and Rebecca Green (Eds) *Relationships in the organization of knowledge*. 2001 Springer Science+Business Media Dordrecht, 37-52

Ελεγχόμενα λεξιλόγια

http://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/

AAT and its company

<https://www.getty.edu/research/tools/vocabularies/lod/index.html>

LanguaI

<https://www.languaI.org/download/Presentation/LanguaIpresentation%202015-10-04.pdf>

UPENN

Classical reference (marcus, sandorini, taylor)

https://www.researchgate.net/profile/Mitchell_Marcus/publication/2873803_The_Penn_Treebank_An_overview/links/00b49538272b5cb5b9000000/The-Penn-Treebank-An-overview.pdf

AMR

<https://amr.isi.edu/a.pdf>

PARSEME SHARED TASK

<file:///C:/Users/marks/AppData/Local/Temp/204-3-1319-1-10-20181105.pdf>