

## 1. ΓΕΝΙΚΑ

ΔΙΔΑΣΚΩΝ	ΙΩΑΝΝΗΣ ΤΖΙΤΖΙΚΑΣ		
ΕΞΑΜΗΝΟ ΔΙΔΑΣΚΑΛΙΑΣ	ΧΕΙΜΕΡΙΝΟ ΕΑΡΙΝΟ		
ΣΧΟΛΗ	ΘΕΤΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΠΙΣΤΗΜΩΝ		
ΤΜΗΜΑ	ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ		
ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ	ΠΡΟΠΤΥΧΙΑΚΟ / ΜΕΤΑΠΤΥΧΙΑΚΟ		
ΚΩΔΙΚΟΣ ΜΑΘΗΜΑΤΟΣ	HY-463	ΕΞΑΜΗΝΟ ΣΠΟΥΔΩΝ	5 <sup>ο</sup> -8 <sup>ο</sup>
ΤΙΤΛΟΣ ΜΑΘΗΜΑΤΟΣ	ΣΥΣΤΗΜΑΤΑ ΑΝΑΚΤΗΣΗΣ ΠΛΗΡΟΦΟΡΙΩΝ		
ΑΥΤΟΤΕΛΕΙΣ ΔΙΔΑΚΤΙΚΕΣ ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ	ΕΒΔΟΜΑΔΙΑΙΕΣ ΩΡΕΣ ΔΙΔΑΣΚΑΛΙΑΣ		ΠΙΣΤΩΤΙΚΕΣ ΜΟΝΑΔΕΣ
Διαλέξεις και φροντιστήρια	6		6
ΤΥΠΟΣ ΜΑΘΗΜΑΤΟΣ	Επιστημονικής Περιοχής Ε6		
Υποβάθρου , Γενικών Γνώσεων, Επιστημονικής Περιοχής, Ανάπτυξης Δεξιοτήτων			
ΠΡΟΑΠΑΙΤΟΥΜΕΝΑ ΜΑΘΗΜΑΤΑ:	HY-240		
ΓΛΩΣΣΑ ΔΙΔΑΣΚΑΛΙΑΣ και ΕΞΕΤΑΣΕΩΝ:	Ελληνική. Ενδέχεται να γίνεται και στην Αγγλική εάν υπάρχει ενδιαφέρον από αγγλόφωνο ακροατήριο.		
ΤΟ ΜΑΘΗΜΑ ΠΡΟΣΦΕΡΕΤΑΙ ΣΕ ΦΟΙΤΗΤΕΣ ERASMUS	ΝΑΙ		
ΗΛΕΚΤΡΟΝΙΚΗ ΣΕΛΙΔΑ ΜΑΘΗΜΑΤΟΣ (URL)	<a href="http://www.csd.uoc.gr/~hy463">http://www.csd.uoc.gr/~hy463</a>		

## 2. ΜΑΘΗΣΙΑΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

### Μαθησιακά Αποτελέσματα

Περιγράφονται τα μαθησιακά αποτελέσματα του μαθήματος οι συγκεκριμένες γνώσεις, δεξιότητες και ικανότητες καταλλήλου επιπέδου που θα αποκτήσουν οι φοιτητές μετά την επιτυχή ολοκλήρωση του μαθήματος.

Συμβουλευτείτε το Παράρτημα Α

Περιγραφή του Επιπέδου των Μαθησιακών Αποτελεσμάτων για κάθε ένα κύκλο σπουδών σύμφωνα με Πλαίσιο Προσόντων του Ευρωπαϊκού Χώρου Ανώτατης Εκπαίδευσης

Περιγραφικοί Δείκτες Επιπέδων 6, 7 & 8 του Ευρωπαϊκού Πλαισίου Προσόντων Διά Βίου Μάθησης και Παράρτημα Β

Περιοληπτικός Οδηγός συγγραφής Μαθησιακών Αποτελεσμάτων

Τα Συστήματα Ανάκτησης Πληροφοριών (Information Retrieval systems) επιτρέπουν την πρόσβαση σε μεγάλους όγκους πληροφοριών αποθηκευμένων με τη μορφή κειμένου, φωνής, video, ή σε σύνθετη μορφή όπως Ιστοσελίδες. Οι γνωστές μηχανές αναζήτησης (π.χ. google, bing) εντάσσονται σε αυτήν την κατηγορία. Σκοπός των συστημάτων αυτών είναι η ανάκτηση μόνο εκείνων των εγγράφων που είναι συναφή με αυτό που αναζητεί ο χρήστης. Για να το επιτύχουν πρέπει να αντιμετωπίσουν την αβεβαιότητα ως προς το τι πραγματικά αναζητεί ο χρήστης και ποιο το θέμα ενός εγγράφου.

Σκοπός του μαθήματος είναι η εισαγωγή στην περιοχή των συστημάτων ανάκτησης πληροφοριών (information retrieval systems) και η εξέταση των θεωρητικών και πρακτικών ζητημάτων που σχετίζονται με την σχεδίαση, υλοποίηση και αξιολόγηση τέτοιων συστημάτων.

Το μάθημα θα επιτρέψει στους φοιτητές:

- να κατανοήσουν τη θεωρητική βάση των καθιερωμένων μοντέλων ανάκτησης (Boolean, Vector-space, Probabilistic, Logical models),
- να κατανοήσουν τη δυσκολία παράστασης και ανάκτησης εγγράφων, εικόνων, ομιλίας, κλπ.,
- να μάθουν να υλοποιούν και να αξιολογούν IR συστήματα,
- να κατανοήσουν τους καθιερωμένους τρόπους ευρετηρίασης και ανάκτησης του Παγκόσμιου Ιστού,
- να κατανοήσουν πως άλλες τεχνικές από το χώρο της επεξεργασία φυσικής γλώσσας, τεχνητής νοημοσύνης και αλληλεπίδρασης ανθρώπου-μηχανής, σχετίζονται με την Ανάκτηση Πληροφοριών,
- να γνωρίσουν διάφορους αλγόριθμους και συστήματα.

#### Γενικές Ικανότητες

Λαμβάνοντας υπόψη τις γενικές ικανότητες που πρέπει να έχει αποκτήσει ο πτυχιούχος (όπως αυτές αναγράφονται στο Παράρτημα Διπλώματος και παρατίθενται ακολούθως) σε ποια / ποιες από αυτές αποσκοπεί το μάθημα;

Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών

Προσαρμογή σε νέες καταστάσεις

Λήψη αποφάσεων

Αυτόνομη εργασία

Ομαδική εργασία

Εργασία σε διεθνές περιβάλλον

Εργασία σε διεπιστημονικό περιβάλλον

Παράγωγή νέων ερευνητικών ιδεών

Σχεδιασμός και διαχείριση έργων

Σεβασμός στη διαφορετικότητα και στην πολυπολιτισμικότητα

Σεβασμός στο φυσικό περιβάλλον

Επίδειξη κοινωνικής, επαγγελματικής και ηθικής υπευθυνότητας και

ευαισθησίας σε θέματα φύλου

Άσκηση κριτικής και αυτοκριτικής

Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης

- Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών
- Προσαρμογή σε νέες καταστάσεις
- Λήψη αποφάσεων
- Αυτόνομη εργασία
- Ομαδική εργασία
- Εργασία σε διεθνές περιβάλλον
- Εργασία σε διεπιστημονικό περιβάλλον
- Παραγωγή νέων ερευνητικών ιδεών
- Σχεδιασμός και διαχείριση έργων
- Άσκηση κριτικής και αυτοκριτικής
- Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης

### 3. ΠΕΡΙΕΧΟΜΕΝΟ ΜΑΘΗΜΑΤΟΣ

Σκοπός του μαθήματος είναι η εισαγωγή στην περιοχή των συστημάτων ανάκτησης πληροφοριών και η εξέταση των θεωρητικών και πρακτικών ζητημάτων που σχετίζονται με την σχεδίαση, υλοποίηση και αξιολόγηση τέτοιων συστημάτων.

Οργάνωση Περιεχομένου

#### Εισαγωγή

- Τι είναι η Ανάκτηση Πληροφοριών / Βασικές έννοιες
- Ιστορική αναδρομή

#### Μοντέλα Ανάκτησης Πληροφοριών

- Boolean
- Διανυσματικό

- Πιθανοκρατικό
- Εναλλακτικά μοντέλα

#### **Αξιολόγηση Αποτελεσματικότητας Συστημάτων Ανάκτησης Πληροφορίας**

- Ακρίβεια και Ανάκληση
- Εναλλακτικά μέτρα
- Συλλογές αναφοράς και αξιολόγηση συστημάτων

#### **Γλώσσες Επερώτησης για Ανάκτηση Πληροφοριών**

- Λέξεις κλειδιά
- Λογικές επερωτήσεις
- Επερωτήσεις συμφραζομένων
- Επερωτήσεις φυσικής γλώσσας
- Δομημένες επερωτήσεις

#### **Προχωρημένες Λειτουργίες Επερώτησης**

- Ανάδραση συνάφειας
- Επέκταση επερώτησης
- Αυτόματη τοπική/καθολική ανάλυση

#### **Ευρετηριασμός, Προεπεξεργασία και Οργάνωση Αρχείων Κειμένου**

- Λέξεις αποκλεισμού (stopwords), stemming (στελέχωση κειμένου) , θησαυροί όρων
- Οργάνωση αρχείων (ανεστραμμένα αρχεία, δένδρα καταλήξεων, αρχεία υπογραφών)
- Στατιστικά κειμένου
- Συμπίεση κειμένου

#### **Αναζήτηση σε Κείμενα**

- Knuth-Morris-Pratt
- Boyer-Moore
- Αυτόματο καταλήξεων (suffix automaton)
- Φράσεις και εγγύτητα

#### **Ομαδοποίηση Εγγράφων (Clustering)**

##### **Ανάκτηση Πολυμέσων**

- Επερωτήσεις Ομοιότητας
- Ευρετηρίαση και Αναζήτηση βάσει Χαρακτηριστικών (Features)
- Χωρικές Δομές Πρόσβασης και Αναζήτηση σε Πολυδιάστατους Χώρους

#### **Παράλληλη και Κατανεμημένη Ανάκτηση Πληροφοριών**

- Αρχιτεκτονικές MIMD και SIMD
- Διαμερισμός συλλογών
- Επιλογή πηγής
- Επεξεργασία επερωτήσεων
- Αρχιτεκτονικές Peer-2-Peer

#### **Τεχνικές μετα-Διάταξης (meta-ranking)**

- Ενοποιημένες και απομονωμένες μέθοδοι
- Παρεμβολή, Ψηφοφορία

#### **Αναζήτηση στον Παγκόσμιο Ιστό**

- Ιστορική αναδρομή
- Ευρετηριασμός ιστοσελίδων
- Διάσχιση του ιστού (crawling)
- Τεχνικές ανάλυσης συνδέσμων (link analysis)

#### 4. ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ - ΑΞΙΟΛΟΓΗΣΗ

<b>ΤΡΟΠΟΣ ΠΑΡΑΔΟΣΗΣ</b> Πρόσωπο με πρόσωπο, Εξ αποστάσεως εκπαίδευση κ.λπ.	Πρόσωπο με πρόσωπο (αίθουσα διδασκαλίας)																									
<b>ΧΡΗΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ</b> Χρήση Τ.Π.Ε. στη Διδασκαλία, στην Εργαστηριακή Εκπαίδευση, στην Επικοινωνία με τους φοιτητές	<ul style="list-style-type: none"><li>• Ηλεκτρονικό ταχυδρομείο και ηλεκτρονική λίστα ταχυδρομείου</li><li>• Ιστοσελίδα μαθήματος στην πλατφόρμα e-learn (moodle) και χρήση των σχετικών δυνατοτήτων της (Φόρουμ ερωταπαντήσεων)</li><li>• Ηλεκτρονική υποβολή ασκήσεων και project</li><li>• Εφαρμογή παιγνιδοποίησης (gamification) μέσω παιχνιδιού για αυτοαξιολόγηση (QuizzerZone)</li></ul> Αυτοματοποιημένοι έλεγχοι για εντοπισμό αντιγραφών																									
<b>ΟΡΓΑΝΩΣΗ ΔΙΔΑΣΚΑΛΙΑΣ</b> Περιγράφονται αναλυτικά ο τρόπος και μέθοδοι διδασκαλίας. Διαλέξεις, Σεμινάρια, Εργαστηριακή Άσκηση, Άσκηση Πεδίου, Μελέτη & ανάλυση βιβλιογραφίας, Φροντιστήριο, Πρακτική (Τοποθέτηση), Κλινική Άσκηση, Καλλιτεχνικό Εργαστήριο, Διαδραστική διδασκαλία, Εκπαιδευτικές επισκέψεις, Εκπόνηση μελέτης (project), Συγγραφή εργασίας / εργασιών, Καλλιτεχνική δημιουργία, κ.λπ.  Αναγράφονται 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	<table><tr><th>Δραστηριότητα</th><th>Φόρτος Εργασίας Εξαμήνου</th></tr><tr><td>Διαλέξεις</td><td>52</td></tr><tr><td>Εργαστήρια / Φροντιστήρια</td><td>8</td></tr><tr><td>Εργαστηριακές (Προγραμματιστικές) Ασκήσεις</td><td>40</td></tr><tr><td>Μελέτη θεωρίας</td><td>30</td></tr><tr><td>Project</td><td>34</td></tr><tr><td>Μελέτη βιβλιογραφίας</td><td>10</td></tr><tr><td>Εξετάσεις (Ασκήσεις, Πρόοδος, Τελικό Διαγώνισμα)</td><td>6</td></tr><tr><td></td><td></td></tr><tr><td></td><td></td></tr><tr><td></td><td></td></tr><tr><td>Σύνολο Μαθήματος</td><td>180</td></tr></table>	Δραστηριότητα	Φόρτος Εργασίας Εξαμήνου	Διαλέξεις	52	Εργαστήρια / Φροντιστήρια	8	Εργαστηριακές (Προγραμματιστικές) Ασκήσεις	40	Μελέτη θεωρίας	30	Project	34	Μελέτη βιβλιογραφίας	10	Εξετάσεις (Ασκήσεις, Πρόοδος, Τελικό Διαγώνισμα)	6							Σύνολο Μαθήματος	180	
Δραστηριότητα	Φόρτος Εργασίας Εξαμήνου																									
Διαλέξεις	52																									
Εργαστήρια / Φροντιστήρια	8																									
Εργαστηριακές (Προγραμματιστικές) Ασκήσεις	40																									
Μελέτη θεωρίας	30																									
Project	34																									
Μελέτη βιβλιογραφίας	10																									
Εξετάσεις (Ασκήσεις, Πρόοδος, Τελικό Διαγώνισμα)	6																									
Σύνολο Μαθήματος	180																									
<b>ΑΞΙΟΛΟΓΗΣΗ ΦΟΙΤΗΤΩΝ</b> Περιγραφή της διαδικασίας αξιολόγησης  Γλώσσα Αξιολόγησης, Μέθοδοι αξιολόγησης, Διαμορφωτική ή Συμπερασματική, Δοκιμασία Πολλαπλής Επιλογής, Ερωτήσεις Σύντομης Απάντησης, Ερωτήσεις Ανάπτυξης Δοκιμίων, Επίλυση Προβλημάτων, Γραπτή Εργασία, Έκθεση / Αναφορά, Προφορική Εξέταση, Δημόσια Παρουσίαση, Εργαστηριακή Εργασία, Κλινική Εξέταση Ασθενούς, Καλλιτεχνική Ερμηνεία, Άλλη / Άλλες  Αναφέρονται ρητά προσδιορισμένα κριτήρια αξιολόγησης και εάν και που είναι προσβάσιμα από τους φοιτητές.	<p>Γλώσσα ελληνική. Ενδέχεται να γίνεται και στην Αγγλική εάν υπάρχει ενδιαφέρον από αγγλόφωνο ακροατήριο.</p> <p><b>Βαθμολόγηση:</b> 25% Ασκήσεις (A1 8%, A2: 8%, A3: 9%) + 25% Project + 50% Τελικός (αν ο τελικός ≥ 4.75)</p>																									

#### 5. ΣΥΝΙΣΤΩΜΕΝΗ-ΒΙΒΛΙΟΓΡΑΦΙΑ

-Προτεινόμενη Βιβλιογραφία :  
Επιλογές Συγγραμμάτων στον ΕΥΔΟΞΟ:

1. Βιβλίο [41954965]: Ανάκτηση Πληροφορίας, 2η Έκδοση, Baeza-Yates Ricardo, Ribeiro-Neto Berthier [Λεπτομέρειες](#)
2. Βιβλίο [12532681]: ΕΙΣΑΓΩΓΗ ΣΤΗΝ ΑΝΑΚΤΗΣΗ ΠΛΗΡΟΦΟΡΙΩΝ, CHRISTOPHER D. MANNING, PRABHAKAR RAGHAVAN, HINRICH SCHUTZE [Λεπτομέρειες](#)

-Συναφή επιστημονικά περιοδικά:

Επιλεγμένα άρθρα από τις σειρές συνεδρίων SIGIR και ECIR (κυρίως)