

«Υπολογιστικές Επιστήμες και Τεχνολογίες Δεδομένων, Περιεχομένου και Αλληλεπίδρασης» (ΟΠΣ 5002437)

Υποέργο 1 «Γλωσσικές Τεχνολογίες Ανάλυσης Περιεχομένου και Αλληλεπίδρασης» (Περιφέρεια Αττικής)

Ινστιτούτο Επεξεργασίας του Λόγου/Ε.Κ. «Αθηνά»



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σκοπός

Σκοπός του υποέργου είναι να υπηρετήσει το ερευνητικό σχέδιο του Ινστιτούτου με

- α) εξειδικευμένη έρευνα και ανάπτυξη στις επιμέρους περιοχές της γλωσσικής τεχνολογίας,
- β) δραστηριότητες όπως η ανάπτυξη γλωσσικών υποδομών, εφαρμογών οργάνωσης και διάθεσης γλωσσικών πόρων και τεχνολογιών,
- γ) ένα ευρύ φάσμα δραστηριοτήτων διάχυσης και αξιοποίησης που εκτείνεται από την ενημέρωση του ευρύτερου κοινού και της διεθνούς επιστημονικής κοινότητας μέχρι και την εσωτερική ανακάλυψη τεχνολογιών με ιδιαίτερες προοπτικές εκμετάλλευσης και την περαιτέρω υποστήριξη, διαμόρφωση και εκκόλαψή τους.

Ενότητες Εργασίας

- ΕΕ1: Επεξεργασία Φυσικής Γλώσσας
- ΕΕ2: Τεχνολογίες Φωνής
- ΕΕ3: Φυσική Αλληλεπίδραση και Ενσώματη Επικοινωνία
- ΕΕ4: Τεχνολογίες Μάθησης
- ΕΕ5: Υποδομή Γλωσσικών Πόρων
- ΕΕ6: Δράσεις διάχυσης, προβολής και αξιοποίησης αποτελεσμάτων
- ΕΕ7: Οργάνωση και Διαχείριση

Αποτελέσματα της ΕΕ1

Επεξεργασία Φυσικής Γλώσσας

1. **Ανάπτυξη νέων αρθρωμάτων (modules) ΕΦΓ** που ενσωματώθηκαν στη σουίτα γλωσσικής τεχνολογίας ILSP-NLP:

- συστήματα ανάλυσης γνώμης (sentiment analysis)
- συστήματα ανάλυσης θεματικού περιεχομένου (topic models)

Πλατφόρμα για την παρουσίαση των αποτελεσμάτων σε πρακτικά της ολομέλειας του Ελληνικού Κοινοβουλίου ((4.356 συνεδριάσεις ολομέλειας, 1.063.546 ομιλίες). :

<http://194.177.192.82/presentation/>

2. Σύστημα Μηχανικής Μετάφρασης

Βελτίωση της μεθοδολογίας δημιουργίας συστημάτων μηχανικής μετάφρασης η οποία αναπτύχθηκε στο έργο PRESEMT. Έμφαση δόθηκε στη βελτίωση της μετάφρασης για ζεύγη γλωσσών που περιλαμβάνουν την Ελληνική (EL $\rightarrow$ EN, EN $\rightarrow$ EL & EL $\rightarrow$ DE).

Αποτελέσματα της ΕΕ2

Τεχνολογίες φωνής

1. Ανίχνευση προβλημάτων δυσλεξίας.

Ανάπτυξη ενός ευφυούς συστήματος ανάγνωσης βασισμένου στην αυτόματη αναγνώριση παιδικής φωνής, το οποίο προσπαθεί να ανιχνεύει προβλήματα δυσλεξίας σε άτομα σχολικής ηλικίας προσφέροντας ένα εργαλείο έγκαιρης πρώτης ανίχνευσης.

Δικτυακός τόπος ανίχνευσης δυσλεξίας: <http://www.ilsp.gr/krhpis-dyslexia>

2. Οπτικοακουστική σύνθεση ομιλίας.

Διερεύνηση τεχνικών πολυτροπικής σύνθεσης φωνής με σκοπό την ενσωμάτωσή τους σε τεχνολογία σύνθεσης φωνής από κείμενο, εμπλουτίζοντάς την με τεχνητούς χαρακτήρες (visual agents), μέσω σύνθεσης κινήσεων και εκφράσεων (facial animation) του προσώπου κατά την εκφορά λόγου, σύμφωνα με το κείμενο εισόδου.

Αποτελέσματα της ΕΕ3

Φυσική Αλληλεπίδραση και Ενσώματα Επικοινωνία

1.Τεχνολογίες Νοηματικής Γλώσσας

Επαύξηση της τεχνολογίας σύνθεσης νοηματικού λόγου

- εμπλουτισμός ως προς το περιεχόμενο (κωδικοποίηση λημμάτων ΕΝΓ με σύστημα HamnoSys, έλεγχος)
- πειραματισμός βάσει των παραμέτρων άρθρωσης του νοηματικού λόγου με αξιοποίηση εξειδικευμένου εργαλείου SysBuilder

Αποτέλεσμα ήταν η βελτίωση της εικονικής αναπαράστασης νοηματικού λόγου

<http://sign.ilsp.gr/signilsp-site/index.php/el/sunthesi-noimatwn/>

<http://sign.ilsp.gr/signilsp-site/index.php/el/home/>

2. Τεχνολογίες Ψηφιακών Βοηθών για την Υποστήριξη Πολυτροπικών Διαλόγων

Ανάπτυξη διαλογικού συστήματος για την εξάσκηση μαθητών της αγγλικής γλώσσας στην παραγωγή προφορικού λόγου. Η αλληλεπίδραση χρήστη-υπολογιστή πραγματοποιείται με ψηφιακό χαρακτήρα (Avatar) μέσω φωνής.

Αποτελέσματα ΕΕ4

Τεχνολογίες Μάθησης

1. Έρευνα για την κατανόηση του γραπτού λόγου σε παιδιά σχολικής ηλικίας (βασισμένη στην καταγραφή της οφθαλμοκίνησης)
2. Ανάπτυξη mock-up prototype εφαρμογής για την υποστήριξη της επικοινωνίας στα Ελληνικά σε πραγματικό χρόνο

Αποτελέσματα ΕΕ5

Υποδομή Γλωσσικών Πόρων

1. ΕΘΕΓ

Ποσοτικός και ποιοτικός εμπλουτισμός του περιεχομένου του ΕΘΕΓ

Βελτίωση της διεπαφής πρόσβασης στον ΕΘΕΓ

<http://hnc.ilsp.gr/>

2. ΛΕΞΙΣ: υπολογιστικό λεξικό της Νέας Ελληνικής (μορφολογικό, συντακτικό και σημασιολογικό)

Έλεγχος & εμπλουτισμός σε όλα τα επίπεδα ανάλυσης (Προσθήκη νέων λημμάτων, Σημασιολογικών Μονάδων, σημασιολογικών σχέσεων, σχέσεων συνωνυμίας και αντωνυμίας, Έλεγχος & βελτίωση ορισμών κ.ά)

3. Εννοιολογικό Λεξικό – ΒΔ πολυλεκτικών όρων

Αποτελέσματα ΕΕ 6

Δράσεις διάχυσης, προβολής & αξιοποίησης αποτελεσμάτων

- Πλατφόρμα Υπηρεσιοποίησης γλωσσικών πόρων και υπηρεσιών
- Δράσεις διάχυσης και προβολής αποτελεσμάτων
 - Συμμετοχές σε συνέδρια
 - 20 δημοσιεύσεις σε διεθνή συνέδρια και περιοδικά
 - Συνεργασίες με το Πανεπιστήμιο του Μαγδεμβούργου στη Γερμανία, το KU-LOEVEN και Ghent University στο Βέλγιο, Πανεπιστήμιο του Αμβούργου στη Γερμανία και το Πανεπιστήμιο DePaul, Chicago, ΗΠΑ.
 - Οργάνωση Workshops: «Workshop on Sign Language Technologies», 2017, Αθήνα, «Smart-Access: The 2nd International Workshop on Accessibility and the Smart City: Technological Challenges and Open Accessibility Issues», PETRA 2020 Conference “Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives”, LREC2020 Conference
 - Συμμετοχή με παρουσίαση αποτελεσμάτων σε εθνικές εκδηλώσεις (ASF, PSF, TIF, RN...)
- Δράσεις μεταφοράς και αξιοποίησης αποτελεσμάτων (Οργάνωση πιλοτικής πλατφόρμας εσωτερικής ανακάλυψης και Διαγωνισμός Σχεδίων Αξιοποίησης Ερευνητικών Αποτελεσμάτων INSPEAR)

Δείκτες υποέργου

Αποτελέσματα: Αριθμητικά στοιχεία

- Δημοσιεύσεις σε διεθνή συνέδρια/περιοδικά: 20
- Α/Μ έκτακτου προσωπικού: 329
- Α/Μ προσωπικού (τακτικού και έκτακτου): 335
- Αριθμός ατόμων έκτακτου προσωπικού: 27
- Συνολικός αριθμός ατόμων: 44
- Αριθμός τεχνολογιών/συστημάτων που αναπτύχθηκαν: 7
- Αριθμός γλωσσικών πόρων που εμπλουτίστηκαν: 3

Επιστημονικοί Υπεύθυνοι

- Χαρίλαος Παπαγεωργίου, Διευθυντής Ερευνών (ΕΕ1-1)
- Γεώργιος Ταμπουρατζής, Διευθυντής Ερευνών (ΕΕ1-2)
- Στυλιανός Μπακαμίδης – Ιωάννης Δολόγλου, Διευθυντές Ερευνών (ΕΕ2-1)
- Νάσος Κατσαμάνης, Κύριος Ερευνητής (ΕΕ2-2, ΕΕ3-2)
- Ελένη Ευθυμίου, Διευθυντής Ερευνών (ΕΕ3-1)
- Ιωάννα Αντωνίου, Διευθυντής Ερευνών (ΕΕ4)
- Μαρία Γαβριηλίδου, ΕΛΕ Α' (ΕΕ5-1),
- Πένυ Λαμπροπούλου ΕΛΕ Α' (ΕΕ5-2),
- Αγγελική Φωτοπούλου, Διευθύντρια Ερευνών (ΕΕ5-3)
- Τάσος Πατρικάκος (ΕΕ6)
- Νικόλαος Βογιατζής, Γεώργιος Σάντσες (ΕΕ6-3)
- Γρηγόριος Σταϊνχάουερ, Διευθυντής Ερευνών, Υπεύθυνος Υποέργου (ΕΕ7)

Συνεργαζόμενοι ερευνητές

ΠΟΝΤΙΚΗ ΜΑΡΙΑ (ΕΕ1)
ΠΑΠΠΑΣ ΔΗΜΗΤΡΗΣ (ΕΕ1)
ΚΟΤΙΤΣΑΣ ΣΩΤΗΡΗΣ(ΕΕ1)
ΣΤΑΥΡΟΠΟΥΛΟΣ ΠΕΤΡΟΣ (ΕΕ1)
ΒΑΣΙΛΕΙΟΥ ΜΑΡΙΝΑ (ΕΕ1)
ΚΟΥΛΙΕΡΑΚΗΣ ΙΩΑΝΝΗΣ (ΕΕ1)
ΓΟΥΛΑΣ ΘΕΟΔΩΡΟΣ (ΕΕ1, ΕΕ3)
ΧΑΝΙΩΤΗ ΕΛΕΝΗ (ΕΕ2)
ΣΙΔΗΡΟΠΟΥΛΟΣ ΝΙΚΟΛΑΟΣ (ΕΕ2, ΕΕ6)
ΕΥΘΥΜΙΟΥ ΓΙΩΡΓΟΣ (ΕΕ2,ΕΕ3)
ΨΩΜΑΣ ΒΑΣΙΛΕΙΟΣ (ΕΕ2, ΕΕ3)
ΜΑΣΤΡΟΓΙΑΝΝΟΠΟΥΛΟΣ - ΓΚΕΓΚΑΣ ΔΗΜΗΤΡΗΣ (ΕΕ2, ΕΕ3)
ΒΑΚΑΛΟΠΟΥΛΟΥ ΑΝΝΑ (ΕΕ3, ΕΕ5)
ΔΗΜΟΥ ΑΘΑΝΑΣΙΑ-ΛΗΔΑ (ΕΕ3)
ΚΑΡΙΩΡΗΣ ΠΑΝΑΓΙΩΤΗΣ (ΕΕ3, ΕΕ4)
ΒΑΣΙΛΑΚΗ ΚΙΚΗ (ΕΕ3)
ΣΤΑΜΟΥΛΗ ΣΠΥΡΙΔΟΥΛΑ (ΕΕ3, ΕΕ4)
ΜΠΟΤΙΝΗ ΠΑΡΑΣΚΕΥΗ (ΕΕ4)
ΦΛΟΥΔΑ ΧΡΙΣΤΙΝΑ (ΕΕ4)
ΓΙΑΝΝΟΥΤΣΟΥ ΟΛΓΑ (ΕΕ5)
ΓΙΟΥΛΗ ΠΑΡΑΣΚΕΥΗ (ΕΕ5)
ΚΟΥΛΑΦΕΤΗΣ ΙΩΑΝΝΗΣ (ΕΕ6)
ΑΛΤΑΝΟΠΟΥΛΟΣ ΙΩΑΝΝΗΣ (ΕΕ6)
ΒΟΓΙΑΤΖΗΣ ΝΙΚΟΛΑΟΣ (ΕΕ6)
ΛΙΒΑ ΑΘΑΝΑΣΙΑ (ΕΕ6)
ΣΑΝΤΣΕΣ ΓΕΩΡΓΙΟΣ ΑΝΔΡΕΑΣ (ΕΕ6)
ΠΑΛΗΟΣ ΚΟΣΜΑΣ (ΕΕ2, ΕΕ3)

ΕΥΧΑΡΙΣΤΙΕΣ

Οι εργασίες του υποέργου 1 που αναφέρθηκαν στις προηγούμενες διαφάνειες υλοποιήθηκαν στο πλαίσιο της Πράξης «Υπολογιστικές Επιστήμες και Τεχνολογίες Δεδομένων, Περιεχομένου και Αλληλεπίδρασης» (MIS 5002437) που εντάσσεται στη Δράση «ΔΡΑΣΗ ΣΤΡΑΤΗΓΙΚΗΣ ΑΝΑΠΤΥΞΗΣ ΕΡΕΥΝΗΤΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΩΝ ΦΟΡΕΩΝ» και χρηματοδοτείται από το Επιχειρησιακό Πρόγραμμα «Ανταγωνιστικότητα, Επιχειρηματικότητα και Καινοτομία» στο πλαίσιο του ΕΣΠΑ 2014-2020, με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης (Ευρωπαϊκό Ταμείο Περιφερειακής Ανάπτυξης).

