

Θέμα: Συστήματα Συλλογής και Αυτόματης Σύνοψης Ειδήσεων με Χρήση Σύγχρονων Μεθόδων Επεξεργασίας Φυσικής Γλώσσας. Διερεύνηση Υφιστάμενων Προσεγγίσεων και Ολοκληρωμένων Λύσεων και Πρόταση Σχεδιασμού Υλοποίησης Αντίστοιχου Συστήματος	
Επιβλέπων: Β. Ταμπακάς	e-mail: tampakas@uop.gr
Μέλη:	Ακαδημαϊκό Έτος: 2024-25
<p>Στόχοι Διερεύνηση και ανάλυση της τρέχουσας κατάστασης των τεχνολογιών και συστημάτων (state-of-the-art) που χρησιμοποιούνται στη συλλογή και αυτόματη σύνοψη ειδήσεων, συμπεριλαμβανομένων τόσο των μεθόδων όσο και των λύσεων που υπάρχουν. Ανάπτυξη ενός πρωτοτύπου συστήματος συλλογής ειδήσεων και σύνοψης περιεχομένου (π.χ. για ένα συγκεκριμένο θέμα), το οποίο να αξιοποιεί τεχνικές NLP και μηχανικής μάθησης, προσφέροντας στους χρήστες σύντομες και περιεκτικές περιλήψεις από πολλαπλές έγκυρες πηγές ειδήσεων.</p>	
<p>Αντικείμενο Η παρούσα διπλωματική εργασία επικεντρώνεται στην ανάπτυξη ενός καινοτόμου συστήματος συλλογής και αυτόματης σύνοψης ειδήσεων, καθώς και στην ανάλυση των σύγχρονων τεχνολογιών και συστημάτων (state-of-the-art) που εφαρμόζονται στην κατηγοριοποίηση και την αυτόματη σύνοψη ειδήσεων αλλά και στη βασικό σχεδιασμό και ανάπτυξη ενός τέτοιου συστήματος για μια ειδική κατηγορία ειδήσεων. Αρχικά, θα πρέπει να γίνει διερεύνηση και αξιολόγηση των παραδοσιακών αλλά και πιο σύγχρονων μεθόδων μηχανικής μάθησης (όπως τα νευρωνικά δίκτυα και οι τεχνικές transformers) και επεξεργασίας φυσικής γλώσσας (NLP), καθώς και των διαθέσιμων συστημάτων, τόσο open-source (όπως το Gensim) όσο και εμπορικών πλατφορμών (όπως το Google News). Στη συνέχεια θα ακολουθήσει ο σχεδιασμός (π.χ. ειδικότερο πεδίο εφαρμογής, τεχνικές και αλγόριθμοι, τεχνολογίες υποδομής, σύνολα δεδομένων) και η ανάπτυξη ενός βασικού τέτοιου συστήματος με συγκεκριμένο πεδίο εφαρμογής. Η ανάπτυξη του συστήματος μπορεί για παράδειγμα να επικεντρωθεί στη συλλογή και φιλτράρισμα ειδήσεων από αξιόπιστες πηγές και στην παραγωγή αυτόματων περιλήψεων, προσφέροντας στον χρήστη εύκολη και περιεκτική ενημέρωση. Το σύστημα θα πρέπει να υιοθετήσει τις απαιτήσεις και να θέσει τους κανόνες για αμεσότητα και ακρίβεια, προσφέροντας στους χρήστες τη δυνατότητα να ενημερώνονται γρήγορα για τα βασικά θέματα της επικαιρότητας.</p> <p>Gambhir, M., & Gupta, V. (2017). Recent automatic text summarization techniques: a survey. <i>Artificial Intelligence Review</i>, 47(1), 1-66.</p> <p>Gupta, R., Kumaraguru, P., & Sureka, A. (2016). Characterizing Factual News, Fake News, and Clickbait: A Large-Scale Measurement Study of Social Media and News from Multiple Perspectives. In <i>Proceedings of the 2016 IEEE Conference on Collaboration and Internet Computing</i>.</p>	

Η εργασία περιλαμβάνει:

Μελέτη Βιβλιογραφίας και Εξοικείωση με το Αντικείμενο

Εξόρυξη Δεδομένων:

Ανάπτυξη Μοντέλου, Σχεδιασμός συστήματος και υλοποίηση

Αξιολόγηση: