

## Γλωσσική Τεχνολογία: Το πεδίο όπου η γλώσσα συναντά τη σύγχρονη τεχνολογία

Η αλματώδης ανάπτυξη των νέων τεχνολογιών τα τελευταία χρόνια έχει επηρεάσει την ανθρώπινη επικοινωνία μετατρέποντάς την σε πολυτροπική με την αυξανόμενη παρουσία του οπτικού και ηχητικού στοιχείου μέσω της χρήσης πολυμέσων.

Η διάσταση της πληροφορικής επανάστασης, ωστόσο, δεν επιφέρει τον παραγκωνισμό του γλωσσικού στοιχείου· απεναντίας, αναδεικνύει τη γλώσσα ως αποτελεσματικό εργαλείο σε κάθε επίπεδο επικοινωνίας. Τα επιτεύγματα της σύγχρονης τεχνολογίας το αποδεικνύουν.

Η Γλωσσική Τεχνολογία αποτελεί το επιστημονικό πεδίο συνάντησης της Γλωσσολογίας και της Πληροφορικής και έχει ως αντικείμενο την ενσωμάτωση γλωσσικών δυνατοτήτων σε υπολογιστικές εφαρμογές οι οποίες χρησιμοποιούνται σε συστήματα πληροφορικής και επικοινωνιακής τεχνολογίας. Κάποια επιτεύγματα της Γλωσσικής Τεχνολογίας μάς είναι οικεία καθώς τα χρησιμοποιούμε στην καθημερινότητά μας (π.χ. το «έξυπνο» πληκτρολόγιο στα κινητά τηλέφωνα ή η φωνητική αναζήτηση της μηχανής Google), ενώ κάποια άλλα μας είναι λιγότερο γνωστά, επειδή αποτελούν μέρη ευρύτερων εφαρμογών (π.χ. τα φίλτρα που «διαβάζουν» και ταξινομούν αυτόματα τα μηνύματα του ηλεκτρονικού ταχυδρομείου).

Βασικοί τομείς γύρω από τους οποίους αναπτύσσονται γλωσσικές εφαρμογές και υπολογιστικά εργαλεία είναι η αυτόματη ανάλυση και επεξεργασία λόγου, η μηχανική μετάφραση, η σύνθεση και αναγνώριση φωνής, η «εξόρυξη» πληροφοριών από κείμενα και η εκπαιδευτική τεχνολογία. Οι τομείς αυτοί βασίζονται στην επεξεργασία της φυσικής γλώσσας από υπολογιστή και συνδέονται στενά με την αλληλεπίδραση ανθρώπου και μηχανής.

Στο πεδίο της αυτόματης επεξεργασίας και ανάλυσης λόγου ανήκουν υπολογιστικά εργαλεία που έχουν στόχο τον διαχωρισμό προτάσεων και λέξεων, τη μορφολογική ανάλυση κειμένου, τη λημματοποίηση αλλά και τη συντακτική και σημασιολογική ανάλυση κειμένου, εργαλεία τα οποία χρησιμοποιούνται συνήθως ως επιμέρους τμήματα λογισμικών, όπως αυτών της μηχανικής μετάφρασης.

Ως μηχανική μετάφραση νοείται η αυτόματη διαδικασία μετάφρασης μιας γλωσσικής ενότητας (φράσης, πρότασης, κειμένου) από μία γλώσσα σε μία άλλη μέσω υπολογιστικού προγράμματος. Τα σύγχρονα λογισμικά μηχανικής μετάφρασης στηρίζονται σε γραμματική ή στατιστική επεξεργασία γλωσσικών δεδομένων. Σήμερα υπάρχουν πολλές πλατφόρμες μηχανικής μετάφρασης, μερικές από τις οποίες είναι ελεύθερα προσβάσιμες στο διαδίκτυο και ιδιαίτερα δημοφιλείς (Google Translator, Bing Translator). Ως πεδίο, η μηχανική μετάφραση στρέφεται σήμερα προς την επίτευξη υψηλότερης ποιότητας μετάφρασης και την εξάλειψη των αδυναμιών και των σφαλμάτων που παρουσιάζονται στις αυτόματες μεταφράσεις, ενώ πλέον έχει επεκταθεί και στη μηχανική μετάφραση προφορικού λόγου, μέσω της αυτόματης αναγνώρισης ομιλίας.

Οι όροι «σύνθεση» και «αναγνώριση φωνής» συνίστανται στη μηχανική παραγωγή και αναγνώριση της ανθρώπινης ομιλίας. Στην πρώτη περίπτωση ένα κείμενο μετατρέπεται σε ομιλία από ένα υπολογιστικό σύστημα (text-to-speech, TTS), ενώ στη δεύτερη η ανθρώπινη ομιλία είναι αυτή που μετατρέπεται σε κείμενο (speech-to-text, STT). Εφαρμογές σύνθεσης και αναγνώρισης φωνής βρίσκονται σε ηλεκτρονικούς υπολογιστές, κινητά τηλέφωνα και άλλες ηλεκτρονικές συσκευές και περιλαμβάνουν την φωνητική πληκτρολόγηση, τον έλεγχο οικιακών ηλεκτρονικών συσκευών, τη διαδικτυακή αναζήτηση κ.ά., ενώ ενσωματώνονται σε συστήματα αυτόματης παραγωγής κειμένων (π.χ. περιγραφές προϊόντων) και σε συστήματα προφορικών διαλόγων (π.χ. κρατήσεις εισιτηρίων, διάλογοι πλοήγησης σε αυτοκίνητα), αντίστοιχα.

Τα πεδία αυτά είναι πολύ σημαντικά, καθώς προσφέρουν και υποστηρικτικές τεχνολογίες για άτομα με αναπηρίες, επιτρέποντας σε ανθρώπους με προβλήματα όρασης ή δυσκολίες ανάγνωσης να «ακούν» γραπτά κείμενα και σε ανθρώπους με κινητικά προβλήματα να «γράφουν» όσα προφέρουν και εκφωνούν, βελτιώνοντας με αυτόν τον τρόπο την καθημερινή ζωή αυτών των ανθρώπων και διευκολύνοντας την ισότιμη προσβασιμότητά τους στις συνθήκες που διαμορφώνει η κοινωνία της πληροφορικής.

Ο τομέας της «εξόρυξης» ή εξαγωγής πληροφορίας σχετίζεται με την αναγνώριση –μέσα από δεδομένα– εξειδικευμένης πληροφορίας και την ανάκτησή της με αυτόματο τρόπο. Κυρίως στοχεύει στην αυτόματη και μηχανική αναγνώριση ονοματικών οντοτήτων (ονόματα ανθρώπων, τοπωνύμια, ονόματα οργανισμών κ.λπ.), χρονικών πληροφοριών (ημερομηνίες) και γεγονότων, καθώς και των σχέσεων που τα συνδέουν, μέσα από ηλεκτρονικά κείμενα. Τα αποτελέσματα της «εξόρυξης» πληροφορίας χρησιμοποιούνται από ευρύτερες εφαρμογές, όπως η αυτόματη παραγωγή περίληψης κειμένου, τα συστήματα ερωταποκρίσεων (π.χ. αναζήτηση συγκεκριμένων πληροφοριών σε ιστοσελίδες) και η μηχανική μετάφραση. Στον ερευνητικό άξονα του συγκεκριμένου πεδίου βρίσκονται σήμερα η «εξόρυξη» γνώμης και η ανάλυση και αξιολόγηση συναισθήματος μέσα από κείμενα του διαδικτύου (blogs, κοινωνικά δίκτυα).

Τέλος, ως «εκπαιδευτική τεχνολογία», υπό το πρίσμα της γλωσσικής τεχνολογίας, νοείται η τεχνολογία που συμβάλλει στη γλωσσική εκπαίδευση μέσω πολυτροπικών εκπαιδευτικών εργαλείων και ψηφιακών γλωσσικών δραστηριοτήτων, με στόχο τη βελτίωση της γλωσσικής εκμάθησης και την επίλυση προβλημάτων στη διδασκαλία είτε της μητρικής είτε μιας ξένης γλώσσας. Σε αυτό το πλαίσιο αναπτύσσονται μέθοδοι και εργαλεία (εκπαιδευτικός σχεδιασμός, οπτικοακουστικά διδακτικά υλικά, επιμορφωτικά υλικά, εκπαιδευτικός εξοπλισμός, διαγνωστικά εργαλεία, συλλογή και επεξεργασία γλωσσικών πόρων, δημιουργία ηλεκτρονικών λεξικογραφικών εργαλείων κ.λπ.) που υποβοηθούν τη γλωσσική διδασκαλία, αξιολογούν και αντιμετωπίζουν μαθησιακές δυσκολίες και διαταραχές λόγου των παιδιών και ενισχύουν την εξατομικευμένη μάθηση, ενώ παράλληλα μειώνουν τον χρόνο εκμάθησης της ξένης γλώσσας.

Η Γλωσσική Τεχνολογία είναι βασικό συστατικό της ψηφιακής κοινωνίας. Από τη μηχανική μετάφραση έως τη σύνθεση και αναγνώριση φωνής, οι εφαρμογές και οι υπηρεσίες της βελτιώνουν την καθημερινότητά μας, διευρύνουν τη γνώση μας για τη γλώσσα και τους υπολογιστικούς μηχανισμούς, αλλά, πάνω απ' όλα, προσφέρουν στους ανθρώπους τη δυνατότητα να επικοινωνούν πιο αποτελεσματικά με τους υπολογιστές τους και, κυρίως, μεταξύ τους.

*Αρχοντούλα Μεντή*