



www.xenesglosses.eu

Οι περισσότερες ευρωπαϊκές γλώσσες αντιμετωπίζουν τον κίνδυνο ψηφιακής εξαφάνισης σύμφωνα με νέα μελέτη Ευρωπαίων ειδικών της γλωσσικής τεχνολογίας.

Αξιολογώντας το επίπεδο της υποστήριξης που παρέχει η γλωσσική τεχνολογία σε 30 ευρωπαϊκές γλώσσες, οι ειδικοί καταλήγουν ότι η ψηφιακή υποστήριξη για 21 από αυτές είναι ανύπαρκτη ή, στην καλύτερη περίπτωση, ανεπαρκής. Η μελέτη πραγματοποιήθηκε από το META-NET, ένα ευρωπαϊκό δίκτυο αριστείας που απαρτίζεται από 60 ερευνητικά κέντρα σε 34 ευρωπαϊκές χώρες.

Η μελέτη, η οποία καταρτίστηκε από περισσότερους από 200 ειδικούς και καταγράφεται στους 30 τόμους της σειράς των Λευκών Βιβλίων του META-NET (διαθέσιμων σε ηλεκτρονική και έντυπη μορφή), αξιολόγησε για κάθε γλώσσα το επίπεδο ανάπτυξης της γλωσσικής τεχνολογίας σε τέσσερις διαφορετικές περιοχές: αυτόματη μετάφραση, φωνητική αλληλεπίδραση, ανάλυση κειμένου και διαθεσιμότητα γλωσσικών πόρων. Συνολικά 21 από τις 30 γλώσσες κατατάχτηκαν από τους ειδικούς στην τελευταία κατηγορία, με 'μικρή ή καθόλου υποστήριξη', ως προς τουλάχιστον μία από τις παραπάνω περιοχές. Αρκετές γλώσσες, όπως τα ισλανδικά, τα λεττονικά, τα λιθουανικά και τα μαλτέζικα έχουν τη χαμηλότερη βαθμολογία σε όλες τις περιοχές. Στο άλλο άκρο της κατάταξης, μόνο η αγγλική αξιολογήθηκε ότι έχει 'καλή υποστήριξη', ενώ δεν υπάρχει καμία γλώσσα με 'άριστη υποστήριξη'. Ακολουθούν τα γαλλικά, τα γερμανικά, τα ιταλικά, τα ισπανικά και τα ολλανδικά με 'μέτρια υποστήριξη'. Γλώσσες όπως τα ελληνικά, αλλά και τα

βασικά, τα βουλγαρικά, τα καταλανικά, τα ουγγρικά και τα πολωνικά έχουν 'αποσπασματική υποστήριξη' και ως εκ τούτου ανήκουν στις γλώσσες υψηλού κινδύνου.

Σύμφωνα με τους Hans Uszkoreit και Georg Rehm, επιμελητές της μελέτης, «Τα αποτελέσματα της μελέτης μας είναι πολύ ανησυχητικά. Οι περισσότερες ευρωπαϊκές γλώσσες παρουσιάζουν μεγάλες ελλείψεις σε γλωσσικούς πόρους ενώ κάποιες από αυτές είναι τελείως παραμελημένες. Από αυτή την άποψη, είναι δύσκολο να διασφαλιστεί το ψηφιακό τους μέλλον. Υπάρχουν δραματικές διαφορές στην υποστήριξη της γλωσσικής τεχνολογίας ανάμεσα στις διάφορες ευρωπαϊκές γλώσσες και τεχνολογικές περιοχές. Το χάσμα μεταξύ 'μικρών' και 'μεγάλων' γλωσσών ολοένα και διευρύνεται. Πρέπει να εξασφαλίσουμε τον εφοδιασμό των μικρότερων και λιγότερο πλούσιων σε ψηφιακούς πόρους γλωσσών με τις απαραίτητες βασικές τεχνολογίες, αλλιώς οι γλώσσες αυτές είναι καταδικασμένες σε ψηφιακή εξαφάνιση.»

Όπως σημειώνουν οι συγγραφείς της ελληνικής μελέτης «Τα ελληνικά είναι μία από τις γλώσσες με μικρό αριθμό ομιλητών και μικρό όγκο ψηφιακών γλωσσικών πόρων και τεχνολογιών. Η κρισιμότητα της ανάπτυξης της γλωσσικής τεχνολογίας για τα ελληνικά προβάλλει έντονα μόλις κανείς εξετάσει τι ακριβώς είναι και τι προσφέρει η γλωσσική τεχνολογία, τους τρέχοντες περιορισμούς της αλλά και τις τεράστιες δυνατότητες καινοτομίας.»

Η γλωσσική τεχνολογία παράγει λογισμικό που επεξεργάζεται τον προφορικό και τον γραπτό λόγο. Χαρακτηριστικά παραδείγματα λογισμικού γλωσσικής τεχνολογίας είναι οι ορθογραφικοί και συντακτικοί διορθωτές, οι διαδραστικοί προσωπικοί βοηθοί στα έξυπνα τηλέφωνα (όπως η Siri στο iPhone) τα διαλογικά συστήματα μέσω τηλεφώνου, τα συστήματα αυτόματης μετάφρασης, οι μηχανές αναζήτησης στο διαδίκτυο και η συνθετική φωνή σε συστήματα πλοήγησης στα αυτοκίνητα. Τα σύγχρονα συστήματα γλωσσικής τεχνολογίας βασίζονται κυρίως σε στατιστικές μεθόδους που απαιτούν πολύ μεγάλες ποσότητες προφορικών και γραπτών δεδομένων. Ο απαραίτητος όγκος δεδομένων είναι πολύ δύσκολο να βρεθεί ή να δημιουργηθεί όταν πρόκειται για γλώσσες με σχετικά λίγους ομιλητές. Επιπλέον, τα στατιστικά συστήματα γλωσσικής τεχνολογίας έχουν εγγενείς περιορισμούς ως προς την ποιότητά τους, κάτι που είναι εμφανές π.χ. στις –συχνά διασκεδαστικές– λανθασμένες μεταφράσεις που παράγονται από συστήματα μηχανικής μετάφρασης.

Η Ευρώπη έχει καταφέρει να άρει σχεδόν όλα τα σύνορα μεταξύ των χωρών της. Υπάρχει όμως ακόμη ένα σύνορο το οποίο φαίνεται αδιαπέραστο: οι γλωσσικοί φραγμοί εμποδίζουν την ελεύθερη ροή της γνώσης και της πληροφορίας. Καθυστερούν επίσης τον μακροπρόθεσμο στόχο μιας ενιαίας ψηφιακής αγοράς, καθώς παρακωλύουν την ελεύθερη διακίνηση των αγαθών, των προϊόντων και των υπηρεσιών. Αν και η γλωσσική τεχνολογία μπορεί να άρει τους γλωσσικούς φραγμούς μέσω των σύγχρονων συστημάτων μηχανικής μετάφρασης, τα αποτελέσματα της μελέτης του META-NET δείχνουν ξεκάθαρα ότι πολλές από τις ευρωπαϊκές γλώσσες δεν είναι ακόμα έτοιμες. Υπάρχουν σημαντικά τεχνολογικά κενά που οφείλονται κυρίως στην εστίαση της έρευνας και της ανάπτυξης στην αγγλική γλώσσα, στην έλλειψη πολιτικής δέσμευσης και οικονομικών πόρων, αλλά και στην απουσία συγκεκριμένου ερευνητικού και τεχνολογικού οράματος.

Χρειάζεται μια συντονισμένη και μεγάλης κλίμακας ευρωπαϊκή προσπάθεια ώστε να καλυφθούν τα διαπιστωμένα κενά και να δρομολογηθεί η μεταφορά τεχνολογίας στις περισσότερες γλώσσες. Η πρόκληση είναι τεράστια και πρέπει να αντιμετωπιστεί συλλογικά, με τη συμμετοχή της ΕΕ, των κρατών μελών, των συνδεδεμένων χωρών και της βιομηχανίας. Στοιχεία της πρόκλησης είναι το μεγάλο κατά κεφαλή οικονομικό φορτίο για τις μικρότερες γλωσσικές κοινότητες, η απαιτούμενη μεταφορά τεχνολογίας μεταξύ των γλωσσών, η έλλειψη διαλειτουργικότητας των γλωσσικών πόρων, εργαλείων και υπηρεσιών, καθώς και το γεγονός ότι συχνά τα γλωσσικά σύνορα δεν συμπίπτουν με τα πολιτικά. Η Ευρώπη πρέπει να αναλάβει δράση έτσι ώστε να υποστηρίξει τις γλώσσες της στο ψηφιακό περιβάλλον. Οι γλώσσες αποτελούν πολύτιμο συστατικό της πολιτιστικής μας κληρονομιάς και συνεπώς αξίζει να διασφαλίσουμε το μέλλον τους.

Η 26η Σεπτεμβρίου έχει καθιερωθεί από το Συμβούλιο της Ευρώπης ως η Ευρωπαϊκή Ημέρα Γλωσσών, ανάμεσα στους στόχους της οποίας είναι και η προώθηση της γλωσσικής και πολιτιστικής κληρονομιάς της Ευρώπης. Οι δράσεις του META-NET στοιχίζονται απόλυτα με τις προκλήσεις και τις προοπτικές που αντιμετωπίζει η πολιτιστική μας κληρονομιά στην εποχή της πληροφορίας.

Τι είναι η γλωσσική τεχνολογία;

Η γλωσσική τεχνολογία μάς υποστηρίζει ήδη σε καθημερινές δραστηριότητες, όπως η συγγραφή ηλεκτρονικών μηνυμάτων ή η αγορά εισιτηρίων. Τα οφέλη της γλωσσικής τεχνολογίας είναι προφανή σε περιπτώσεις εφαρμογών με τις οποίες π.χ. αναζητούμε ή θέλουμε να μεταφράσουμε ιστοσελίδες, διορθώνουμε ορθογραφικά ένα κείμενό μας, χρησιμοποιούμε το σύστημα ψυχαγωγίας του αυτοκινήτου ή το κινητό μας τηλέφωνο με προφορικές εντολές, ή παίρνουμε οδηγίες που εκφωνούνται από μια εφαρμογή πλοήγησης. Στο άμεσο μέλλον θα είμαστε σε θέση να συνομιλούμε με υπολογιστικά προγράμματα, καθώς και με μηχανές και συσκευές, όπως τα πολυαναμενόμενα ρομπότ με τα οποία σύντομα θα συνυπάρχουμε καθημερινά στο σπίτι ή στη δουλειά. Όπου και να είμαστε, όποτε και να χρειαστούμε πληροφορίες ή άλλου είδους βοήθεια, απλώς θα τις ζητούμε με τον πιο απλό τρόπο: τη γλώσσα μας. Η κατάργηση των επικοινωνιακών φραγμών μεταξύ των ανθρώπων και της τεχνολογίας θα αλλάξει τον κόσμο μας.

Η γλωσσική τεχνολογία αναγνωρίζεται ευρέως στις μέρες μας ως σημαντικός χώρος ανάπτυξης των τεχνολογιών της πληροφορίας. Μεγάλες πολυεθνικές εταιρίες όπως η Google, η Microsoft, η IBM και η Nuance έχουν επενδύσει σημαντικά στον χώρο αυτό. Εκατοντάδες μικρομεσαίες επιχειρήσεις στην Ευρώπη εξειδικεύονται σε εφαρμογές ή υπηρεσίες γλωσσικής τεχνολογίας. Η γλωσσική τεχνολογία βοηθά τους ανθρώπους να συνεργάζονται, να μαθαίνουν, να οργανώνουν την επιχείρησή τους και να μοιράζονται τη γνώση πέρα από τα γλωσσικά σύνορα και ανεξάρτητα από την τεχνολογική κατάρτισή τους.

Η σειρά των Λευκών Βιβλίων του META-NET

Η σειρά των Λευκών Βιβλίων του META-NET «Γλώσσες στην Ευρωπαϊκή Κοινωνία της Πληροφορίας», καταγράφει το επίπεδο της γλωσσικής τεχνολογίας σε 30 ευρωπαϊκές γλώσσες και παρουσιάζει τις προοπτικές αλλά και τους άμεσους κινδύνους. Η σειρά καλύπτει όλες τις επίσημες γλώσσες των ευρωπαϊκών κρατών μελών και αρκετές άλλες γλώσσες που μιλιούνται στην Ευρώπη. Αν και κατά καιρούς έχουν υπάρξει αρκετές

αξιόλογες και περιεκτικές επιστημονικές μελέτες σε θέματα γλώσσας και τεχνολογίας, εντούτοις δεν έχει καταρτιστεί μέχρι στιγμής μια ευρέως κατανοητή παρουσίαση των κύριων ευρημάτων κάθε γλώσσας και των προκλήσεων που καλούνται να αντιμετωπίσουν σε σχέση με μια τεχνολογικά ενήμερη πολύγλωσση Ευρώπη. Η σειρά των Λευκών Βιβλίων του META-NET έρχεται να καλύψει αυτό το κενό, καθώς εξηγεί γιατί οι περισσότερες γλώσσες αντιμετωπίζουν σημαντικά προβλήματα και επισημαίνει τις πιο κρίσιμες ελλείψεις. Στην κατάρτιση των Λευκών Βιβλίων συνεισέφεραν συνολικά πάνω από 200 ειδικοί.

Λευκές Βίβλοι έχουν καταρτιστεί για τις ακόλουθες ευρωπαϊκές γλώσσες: αγγλικά, βασκικά, βουλγαρικά, γαλικιανά, γαλλικά, γερμανικά, δανικά, ελληνικά, εσθονικά, ιρλανδικά, ισλανδικά, ισπανικά, ιταλικά, καταλανικά, κροατικά, λεττονικά, λιθουανικά, μαλτέζικα, νορβηγικά (μπουκμόλ και νινόρσκ), ολλανδικά, ουγγρικά, πολωνικά, πορτογαλικά, ρουμανικά, σερβικά, σλοβακικά, σλοβενικά, σουηδικά, τσεχικά και φινλανδικά. Κάθε Λευκή Βίβλος είναι γραμμένη στη γλώσσα στην οποία αναφέρεται και είναι μεταφρασμένη στα αγγλικά.

Σχετικά με το META-NET και το META

Το META-NET, ένα Δίκτυο Αριστείας που απαρτίζεται από 60 ερευνητικά κέντρα σε 34 ευρωπαϊκές χώρες, έχει ως στόχο να προάγει τα τεχνολογικά θεμέλια μιας πολύγλωσσης ευρωπαϊκής κοινωνίας της πληροφορίας. Συγχρηματοδοτείται από την Ευρωπαϊκή Επιτροπή μέσω τεσσάρων ερευνητικών έργων.

Το META-NET ίδρυσε την Σύμπραξη Τεχνολογίας για την Πολύγλωσση Ευρώπη (Multilingual Europe Technology Alliance – META). Περισσότεροι από 600 οργανισμοί (ερευνητικά κέντρα, πανεπιστήμια, μικρομεσαίες και μεγάλες επιχειρήσεις) από 55 χώρες έχουν ήδη γίνει μέλη της σύμπραξης αυτής.

esos.gr