

Μηχανική της Πληροφορίας σε Λατινογράμματα Κείμενα - «Ανάπτυξη υπολογιστικής τεχνολογίας για την υποστήριξη της collatio και των apparatus κατά την κριτική έκδοση corpus»

Η κριτική έκδοση που προκύπτει από περισσότερους των δύο μαρτύρων εδράζεται ως σήμερα στην επανάληψη κανονικοτήτων, φιλολογικών δηλαδή ενεργειών, βάσει συστήματος κανόνων τους οποίους εφαρμόζει ο ειδικευμένος φιλόλογος. Κατά τη διαδικασία αυτή ιδιαίτερα επίπονο είναι το στάδιο της collation, της αντιπαραβολής των μαρτύρων, και οδηγεί συχνά στην *eliminatio*, τη μόνη ως σήμερα εφικτή πρακτική που επιτρέπει την ανάγνωση και μεταγραφή επιλεγμένου αριθμού μαρτύρων. Στις περιπτώσεις κειμένων που παραδίδονται από corpus πολλών μαρτύρων, η κριτική έκδοση αποτελεί ιδιαίτερα χρονοβόρα και μη απολύτως κανονικοποιημένη διαδικασία.

Η υπολογιστική γλωσσολογία συνεπικουρούμενη από την επιστήμη της πληροφορικής μπορεί πλέον να αναπτύξει τα τεχνολογικά εργαλεία που είναι σε θέση να εκπαιδεύουν βάσει συγκεκριμένων κανόνων, με στόχο τη βελτίωση της απόδοσής τους ώστε να υποβοηθήσουν τον άνθρωπο σε επίπονες επαναληπτικές εργασίες. Μέχρι σήμερα έχουν γίνει προσπάθειες δημιουργίας τέτοιων εργαλείων, χωρίς όμως να υπάρξει ολοκληρωμένη ηλεκτρονική πλατφόρμα για την κριτική έκδοση και τον υπομνηματισμό κειμένων.

Η διδακτορική διατριβή θα προσπαθήσει να πληρώσει αυτό το κενό α) αξιοποιώντας συστηματικά τα υπάρχοντα εργαλεία, β) αναπτύσσοντας νέα, και γ) σχεδιάζοντας υπολογιστικές μεθόδους που θα είναι σε θέση να παραγάγουν μετρήσιμο ποιοτικά και ποσοτικά αποτέλεσμα. Έμφαση αναμένεται να δοθεί στην αυτοματοποιημένη collatio και τον υπομνηματισμό, συμπεριλαμβανομένης και της ιδανικής στίξης του παραχθισμένου κειμένου από πλήρες corpus μαρτύρων, χωρίς αριθμητικό περιορισμό.

Developing intelligent Software tool for supporting the process of automated “collatio” and production of “apparatus criticus”.

The critical version of more than two witnesses is based on actions regarding the repetition of norms, which are defined, by a system of rules applied by the specialist philologist. The comparison of witnesses during the stage of “collation” is a particularly painful task and often leads to “*eliminatio*”, the only feasible practice that allows reading and transcribing a selected number of witnesses. In the case of texts resulted by a corpus of many witnesses, the critical edition is a particularly time-consuming and not completely defined process.

Computational linguistics, which involve computer science and modern technologies, can now develop tools that are capable of “training” following specific rules thus improving their performance in order to help scientists and researchers reducing painstaking repetitive work. Several attempts have been made to create such tools, yet there is no complete online platform for critical publication and creation of “*apparatus criticus*”.

The thesis will try to fill this gap a) by systematically using a set of existing tools, b) by developing new tools, and c) by designing computational methods that will be able to produce measurable qualitative and quantitative results. The thesis will focus on the process of automated “collatio” and the creation of “*apparatus criticus*” considering the case of proper text punctuation using as source a complete corpus of witnesses without numerical limitation.