

Ημερίδα: Ο φιλόλογος ως δημιουργός, διαχειριστής και χρήστης ψηφιακού περιεχομένου

Ποιος φοβάται το ψηφιακό περιεχόμενο στη Νεοελληνική Φιλολογία;

Σιχάνη Άννα-Μαρία

Περίληψη

Οι ψηφιακές τεχνολογίες όχι μόνο διευκολύνουν την αναζήτηση, τον εντοπισμό και την πρόσβαση στο πρωτογενές υλικό της έρευνάς μας, αλλά κυρίως μάς χορηγούν εργαλεία και μεθόδους προκειμένου να ανανεώσουμε ριζικά τη φυσιογνωμία και την ποιότητα των επιστημονικών μας πρακτικών κι ερωτημάτων. Στην παρουσίαση αυτή, με αφορμή αντίστοιχες εξελίξεις στον διεθνή χώρο των Ψηφιακών Ανθρωπιστικών Σπουδών και πιο ειδικά των Ψηφιακών Λογοτεχνικών Σπουδών (Digital Humanities / Digital Literary Studies), θα επιδιώξω να κάνω μια σύντομη κι ενδεικτική παρουσίαση ψηφιακών εργαλείων και τεχνικών με τα οποία το επιστημονικό αντικείμενο της Νεοελληνικής Φιλολογίας δύναται να εμπλουτιστεί και ταυτόχρονα να συζητήσω μια σειρά από δεξιότητες που ο σημερινός νεοελληνιστής φιλόλογος καλείται να αναπτύξει προκειμένου να χρησιμοποιήσει αυτές τις τεχνολογίες όχι μόνο ως χρήστης αλλά και ως παραγωγός πια ψηφιακού περιεχομένου. Πώς ασχολούμαστε, λοιπόν, στο ψηφιακό παράδειγμα πια με τους κλάδους της βιβλιογραφίας, της εκδοτικής, της λογοτεχνικής ιστορίας και της ερμηνείας της λογοτεχνίας και, επιπλέον, με ποιους τρόπους μπορούμε να (εκ)συγχρονίσουμε τα παραδοσιακά μοντέλα της έρευνας, της διδασκαλίας και της δημοσίευσης στη Νεοελληνική Φιλολογία;

(Ξανα)Γνωρίζοντας τη Νεοελληνική Φιλολογία

Στο πρώτο εξάμηνο του προπτυχιακού προγράμματος διδασκαλίας κάθε σχολής Φιλολογίας στην Ελλάδα οι φοιτητές υποχρωτικά πρέπει να παρακολουθήσουμε και κατόπιν να εξεταστούμε σε ένα μάθημα που λέγεται *Εισαγωγή στη Νεοελληνική Φιλολογία*. Στο μάθημα αυτό ερχόμαστε όλοι μας για πρώτη φορά αντιμέτωποι με ορισμένα από τα κύρια ‘υπαρξιακά ερωτήματα’ στην επιστημονική κι επαγγελματική καριέρα κάθε νεοελληνιστή, τα οποία συστήνουν τον ‘οδικό χάρτη’ (roadmap) του πεδίου και ταυτόχρονα το περιεχόμενο των βασικών εγχειριδίων (reference books) του κλάδου¹:

- Τι είναι η Νεοελληνική Φιλολογία;
- Ποιο είναι το αντικείμενο της ΝΕΦ; Με τι ασχολούμαστε;
- Ποιοι είναι οι επιμέρους κλάδοι της ΝΕΦ;
- Ποιες είναι οι βασικές μεθοδοι με τις οποίες καλούμαστε να εξοικειωθούμε;
- Ποια είναι τα εργαλεία τα οποία χρησιμοποιούμε;

Ψηφιακές τεχνολογίες και Ανθρωπιστικές Επιστήμες

Την ίδια στιγμή, προσπαθώντας να ανιχνεύσουμε και να (συ)στήσουμε την ερευνητική μας πρακτική ανάμεσα σ αυτά τα ερωτήματα, διαπιστώνουμε πως κάτι εδώ και καιρό έχει αλλάξει: η εισαγωγή της υπολογιστικής και κατόπιν ψηφιακής τεχνολογίας επιφέρει μια σειρά από

¹ Βλ. ενδεικτικά τα Μαστροδημήτρης Π.Δ., *Εισαγωγή στη νεοελληνική φιλολογία*, Δόμος, Αθήνα, 2005, Αγγελάτος Δ., *Η Άλφα Βήτα του νεοελληνιστή. Οδηγός για το εισαγωγικό μάθημα στην επιστήμη της νεοελληνικής φιλολογίας*, Gutenberg - Γιώργος & Κώστας Δαρδανός, Αθήνα, 2011.

καθοριστικές αλλαγές, όχι μόνο στην καθημερινότητά μας αλλά και στην ακαδημαϊκή πραγματικότητα, στον χώρο των Φιλολογικών και εν γένει των Ανθρωπιστικών Σπουδών. Για να περιγράψω τη σημερινή συνθήκη συνθηματικά και τηλεγραφικά: το πρωτογενές υλικό (μπορεί να) είναι πια ποιοτικά διαφορετικά προσβάσιμο, οι μέθοδοι και τα εργαλεία που διαθέτουμε και χρησιμοποιούμε χορηγούν πλέον αναλυτικά διαφορετικές δυνατότητες, ενώ τα κανάλια επικοινωνίας, συνεργασίας κι ενημέρωσης έχουν πολλαπλασιαστεί.

Η σύνδεση της τεχνολογίας με τις ανθρωπιστικές σπουδές και την έρευνα είναι βέβαια μια αρκετά παλιά υπόθεση: το ρεύμα των **Humanities Computing** έχει επίσημη πρώτη στα 1949 με τον Roberto Busa και τον Index Thomisticus, έναν αυτοματοποιημένο πίνακα λέξεων των έργων του Θωμα Ακινάτη, σε συνεργασία με την IBM.

Έκτοτε, έχουν αναπτυχθεί κι αναπτύσσονται διαρκώς πλήθος υπολογιστικά εργαλεία και τεχνολογίες που αφορούν από την επεξεργασία δεδομένων, τεχνικές ψηφιοποίησης εικόνας μέχρι γλώσσες προγραμματισμού κι επισημείωσης αλλά και εξειδικευμένα προγράμματα λογισμικού, που στόχο έχουν όχι μόνο την οργάνωση, επεξεργασία, διατήρηση κι οπτικοποίηση της πληροφορίας αλλά κυρίως την ανανέωση της φυσιογνωμίας της έρευνας και της γνώσης στον χώρο των ανθρωπιστικών σπουδών. Ό,τι από το 2004 κι εξής περιγράφουμε ως **Ψηφιακές Ανθρωπιστικές Σπουδές/Digital Humanities**² συμπυκνώνει, ως ακαδημαϊκός κλάδος και ως επιστημονική συμπεριφορά, το σύνολο αυτών των πρακτικών, μεθοδολογιών, εργαλείων που συνιστούν και συστήνουν μια νέα διεπιστημονική, υπολογιστικά προσανατολισμένη δραστηριότητα και γνώση αλλά και τα θεσμικά πλαίσια γύρω απ αυτή.

Νεοελληνική Φιλολογία και Ψηφιακό Περιεχόμενο: τρεις απορίες

Σήμερα, ωστόσο, η τεχνολογία, οι υπολογιστές και τα ψηφιακά μέσα αποτελούν πλέον την απαραίτητη συνθήκη της καθημερινότητάς μας, ορίζοντας, έτσι, τη δημιουργία αλλά και τη χρήση ψηφιακού περιεχομένου ως την απόλυτη καθημερινή πρακτική, γύρω από τις προϋποθέσεις και τις εκδοχές της οποίας ωστόσο σπάνια έχουμε απόλυτη συνείδηση.

Ποια είναι όντως η σχέση της Νεοελληνικής Φιλολογίας ως επιστημονικού κλάδου με το ψηφιακό περιεχόμενο; Άραγε είμαστε ή χρειάζεται να είμαστε εξοικειωμένοι και σε ποιον βαθμό με τις ψηφιακές τεχνολογίες ως παραγωγοί, διαχειριστές και δημιουργικοί χρήστες του ψηφιακού περιεχομένου;

Στην ανακοίνωσή μου, λοιπόν, θα επιχειρήσω να ισορροπήσω μεταξύ θεωρίας και πρακτικής, να θέσω και να επιχειρήσω απαντήσω σε 3 βασικά ερωτήματα:

² Ενδεικτικά σημειώνω τα εισαγωγικά *A Companion to Digital Humanities*, ed. Susan Schreibman, Ray Siemens, John Unsworth. Oxford: Blackwell, 2004, <http://www.digitalhumanities.org/companion/>; Willard McCarty, *Humanities Computing*, Palgrave Macmillan, 2005; *Defining Digital Humanities. A Reader*, Melissa Terras, Julianne Nyhan and Edward Vanhoutte (eds.), Ashgate, 2013.

- Πώς μετατρέπουμε την (πρωτογενή/αναλογική) επιστημονική πληροφορία σε ψηφιακό περιεχόμενο και γιατί η διαδικασία αυτή είναι καίρια προκειμένου να μπορούμε στη συνέχεια να επεξεργαζόμαστε, να δημοσιεύουμε και να χρησιμοποιούμε το ψηφιακό περιεχόμενο, ανοίγοντας ες αεί τον ψηφιακό κύκλο ζωής της πληροφορίας;

Υποστηρίζω πως η συνειδητή προσέγγιση της δημιουργίας και της φυσιογνωμίας του ψηφιακού περιεχομένου - με άλλα λόγια, η 'τεχνική ακτινογραφία' του (τι είναι, πώς φτιάχτηκε, ποιες λειτουργικότητες διαθέτει) - είναι μια **στρατηγικής σημασίας επιστημονική θέση και συμπεριφορά** ψηφιακά ενημερωμένων/προσανατολισμένων φιλολόγων.

- Τι "νέο" φέρνουν το ψηφιακό περιεχόμενο και οι ψηφιακές τεχνολογίες στα χέρια αλλά και στην επιστημονική συμπεριφορά του νεοελληνιστή ;

Θα επιχειρήσω να προτείνω τηλεγραφικά κι ενδεικτικά, με αναφορές άλλοτε σε ελληνικά κι άλλοτε σε ξένα παραδείγματα, πώς καλούμαστε στην πράξη να ξαναδούμε τους κλάδους/στρατηγικές της βιβλιογραφίας, της ανάλυσης κι ερμηνείας, της λογοτεχνικής ιστορίας αλλά και της εκδοτικής, δημιουργώντας και χρησιμοποιώντας ψηφιακή πληροφορία και ψηφιακά εργαλεία.

Επιχειρώ με αυτή μου τη συνολική θεώρηση του κλάδου να δείξω πως οι ψηφιακές τεχνολογίες δεν αποτελούν τόσο ρήξη ή ασυνέχεια αλλά μάλλον ανανέωση και εμπλουτισμό της υφιστάμενης επιστημονικής/επιστημολογικής παράδοσης του κλάδου. Με άλλα λόγια, μάς επιτρέπουν να τροποποιήσουμε, να επεξεργαστούμε με πιο αποτελεσματικό τρόπο παλιά ερευνητικά ερωτήματα ή να θέσουμε νέα που δεν είχαμε καν σκεφτεί πριν εντός του επιστημονικού μας πεδίου ή και σε πιο διεπιστημονικό πλαίσιο.

- Ποιες είναι, τέλος, οι προϋποθέσεις που πρέπει να λάβουμε υπόψιν μας για τη δημιουργία ψηφιακού περιεχομένου ;

1.Από την επιστημονική πληροφορία στο ψηφιακό περιεχόμενο

Ας ξεκινήσουμε από τα βασικά: Τι εννοούμε με τον όρο ψηφιακό περιεχόμενο στη φιλολογία;

Όταν κάνουμε λόγο για ψηφιακό περιεχόμενο περιγράφουμε την ψηφιακή πληροφορία (binary code) που διατίθεται σε ποικιλία μορφών. Η ψηφιακή πληροφορία, ωστόσο, δεν υπάρχει "από μόνη της": πρέπει κάποιος -όλοι εμείς- και με κάποιον τρόπο να τη δημιουργήσει, να την επεξεργαστεί και να τη δημοσιεύσει. Και είναι ακριβώς σε αυτό το σημείο που εγκαινιάζεται η συνάντηση των ψηφιακών τεχνολογιών και του ψηφιακού περιεχομένου, οι δυο βασικές εννοιες του τίτλου της ημερίδας.

Η ψηφιακή πληροφορία μπορεί να αντιστοιχεί σε και να προέρχεται από ένα φυσικό ανάλογό της (physical/ analog counterpart) ή μπορεί να πρόκειται για εκ γενετής ψηφιακή πληροφορία. Οι ψηφιακές τεχνολογίες, συνεπώς, είτε πρόκειται για τη χρήση ενός λογισμικού ή ενός ψηφιακού

μέσου, είτε για μια μορφή γλώσσας προγραμματισμού και μεταφοράς της πληροφορίας, ουσιαστικά χρησιμοποιούνται αφενός για τη για τη μετατροπή και αναπαράσταση της πρωτογενούς πληροφορίας σε βασικούς τύπους ψηφιακής πληροφορίας (σε binary code) και αφετέρου σε δεύτερο στάδιο για την περαιτέρω επεξεργασία και δημοσίευση της ψηφιακής πλέον πληροφορίας. Η εκάστοτε τεχνολογία που θα χρησιμοποιήσουμε τόσο για να μετατρέψουμε την αναλογική πληροφορία σε μια μορφή αναγνώσιμη από τον υπολογιστή όσο και στη συνέχεια για να την επεξεργαστούμε ή να τη δημοσιεύσουμε, υπαγορεύεται από παραμέτρους όπως η φύση του (αναλογικού ή στη συνέχεια του ψηφιακού) υλικού, η απαιτούμενη εκπαίδευση-τεχνογνωσία, ο χρόνος και το κόστος της διαδικασίας, αλλά **κυρίως** από την προσδοκώμενη ή επιθυμητή χρήση του, τι, δηλαδή, θέλουμε να κάνουμε με αυτό το υλικό, εφαρμόζοντας εκ νέου μια ποικιλία από ψηφιακά εργαλεία ανάλογα κάθε φορά με την ερευνητική ανάγκη (βλ. πίνακες 1&2. Οι ψηφιακές τεχνολογίες και τα εργαλεία συνιστούν και χορηγούν, λοιπόν, ένα νέο, ικανό ρεπερτόριο ερευνητικών μεθόδων και στρατηγικών.

Πίνακας 1.

Πίνακας 1.

αναλογικό υλικό	τεχνολογία, εργαλείο	ψηφιακό περιεχόμενο	χρήση/στόχος
φυσικό αντικείμενο ως οπτική πληροφορία (φωτογραφίες, κειμενικά τεκμήρια ως εικόνες)	ψηφιακή φωτογράφιση (digital imaging) <i>τριδιάστατο αντικείμενο</i> <i>ευπαθές αντικείμενο</i> χρήση σαρωτή (scanning) <i>σελίδα</i>	ψηφιακό φωτογραφικό τεκμήριο (digital images) (φωτογραφίες, σκαναρισμένα κειμενικά τεκμήρια)	πρόσβαση συντήρηση υψηλή ανάλυση σύνδεση με άλλα δεδομένα
κείμενο ως περιεχόμενο	λογισμικό οπτικής αναγνώρισης χαρακτήρα (OCR) <i>ευανάγνωστο εντυπο</i> <i>μεγάλος ογκος</i> κωδικοποίηση χαρακτήρων (character encoding-ASCII) γλώσσα μεταφοράς της πληροφορίας (SGML, XML-TEI) <i>επισημαίνει γλωσσολογικές, σημασιολογικές πληροφορίες</i>	ψηφιακό κείμενο (electronic texts) σώματα κειμένων πχ doc, txt, xml, pdf	ανάλυση επεξεργασία σύνδεση με άλλα δεδομένα
άθροισμα πληροφοριών	relational database (σχεσιακή βάση δεδομένων) XML <i>ως μεταγλώσσα για δομημένη πληροφορία</i>	οργανωμένο σύνολο δεδομένων	επεξεργασία σύνδεση με άλλα δεδομένα
άθροισμα των παραπάνω	συνδυασμός των παραπάνω	οργανωμένο σύνολο ψηφιακών τεκμηρίων - ψηφιακός πόρος	άθροισμα των παραπάνω

Πίνακας 2.

κείμενο

επιθυμητό προϊόν	τύπος ψηφιοποίησης/ψηφιακή τεχνολογία
μεταδεδομένα και εικόνες σελίδων	χρήση σαρωτή
αναζήτηση πλήρους κειμένου (σε εικόνες σελίδων)	οπτική αναγνώριση χαρακτήρων (OCR)
αναζήτηση πλήρους κειμένου με εντοπισμό λέξεων σε εικόνες σελίδων	οπτική αναγνώριση χαρακτήρων (OCR) με συντεταγμένες (coordinates)
μεταγραφή του κειμένου	διορθωμένο OCR και βασική κωδικοποίηση δομής (structural markup)
δυνατότητα σύνθετης επεξεργασίας	διορθωμένο OCR και κωδικοποίηση σε επίπεδο (deep encoding)

Το πρωτογενές ψηφιακό περιεχόμενο, λοιπόν, δεν (πρέπει να) είναι αναγκαστικά και πάντα το τελικό προϊόν: αντίθετα ο προορισμός του είναι ουσιαστικά η επανάχρησή του στο ψηφιακό περιβάλλον, η (εξ αεί) δυνατότητα επεξεργασίας του με διαφορετικά ψηφιακά εργαλεία, ο συνδυασμός του με άλλα ψηφιακά δεδομένα και η δημιουργία νέων ψηφιακών προϊόντων κοκ.

Στην κουλτούρα αυτή της επανάχρησης και τροποποίησης του ψηφιακού περιεχομένου, αυτό που πρέπει να επισημάνουμε και να υπενθυμίζουμε διαρκώς είναι ότι πρέπει να ακολουθούνται **ανοιχτά πρότυπα, πρωτόκολλα**, κανόνες καλής πρακτικής και να καθορίζονται οι άδειες πνευματικής ιδιοκτησίας, προκειμένου η ψηφιακή πληροφορία να μπορεί να επεξεργαστεί περαιτέρω, να οπτικοποιηθεί μέσα από μια σειρά από εξειδικευμένα αναλυτικά ψηφιακά εργαλεία και μεθόδους, να οργανωθεί και να δημοσιευθεί, να επεκταθεί, να εμπλουτιστεί στο μέλλον.

2.Χρησιμοποιώντας ψηφιακό περιεχόμενο στη Νεοελληνική Φιλολογία

Έχοντας δει παραπάνω τις βασικές τεχνολογίες για να μετατρέψουμε την επιστημονική σε ψηφιακή πληροφορία, στη συνέχεια θα επιχειρήσω να περιγράψω, τηλεγραφικά κι ενδεικτικά, πώς, χρησιμοποιώντας ορισμένα ψηφιακά εργαλεία και τεχνολογίες -κλιμακούμενης

δυσκολίας- προκειμένου να επεξεργαστούμε ή να δημοσιεύσουμε το ψηφιακό πια περιεχόμενο, καλούμαστε να ξαναδούμε στο ψηφιακό παράδειγμα τους βασικούς κλάδους έρευνας της Νεοελληνικής Φιλολογία, δηλαδή τους κλάδους της βιβλιογραφίας, της εκδοτικής, της ανάλυσης κι ερμηνείας αλλά και της λογοτεχνικής ιστορίας.

Τι μπορεί να σημαίνει βιβλιογραφία στο ψηφιακό περιβάλλον: από την παραπομπή στη διασύνδεση

Καθώς η έννοια της βιβλιογραφίας περιγράφει τη στοχευμένη/εμπρόθετη συγκέντρωση πηγών ή δεδομένων, μια ηλεκτρονική βιβλιογραφία ουσιαστικά αποτελείται από μια μορφή βάσης δεδομένων που επιτρέπει την οργάνωση και τη δυναμική διασύνδεση των επιμέρους τμημάτων της δομημένης μετα-πληροφορίας, τη αναζήτηση και επανάχρησή της, πολλαπλασιάζοντας ουσιαστικά την μια βιβλιογραφική πληροφορία σε πλήθος άλλων δυνάμει πληθυντικών πληροφοριών, καθιστώντας κάθε τμήμα της παραπομπής δυνάμει νέα πληροφορία.

Εκτός από τη δημιουργία βάσεων δεδομένων με βιβλιογραφικές εγγραφές, υπάρχουν πια πολλά λογισμικά οργάνωσης βιβλιογραφικών πηγών (πχ [zotero](#) , [endnote](#) , [mendeley](#)) τα οποία χορηγούν τη δυνατότητα να οργανώσει κανείς τη ψηφιακή βιβλιοθήκη, να δημιουργήσει θεματικές βιβλιογραφίες με το σύνολο των πηγών του, να πραγματοποιήσει παραμετροποιημένες αναζητήσεις, να μοιραστεί με μια ομάδα (πχ συναδέλφους, μαθητές), να χρησιμοποιήσει τις παραπομπές μέσα στη μελέτη του, να επιλέξει τη μορφή, το στυλ της παραπομπής κοκ.

Τι μπορεί να σημαίνει λογοτεχνική ιστορία στο ψηφιακό περιβάλλον: από την εξιστόρηση στην οπτικοποίηση

Προκειμένου για την ιστορική μελέτη και παρακολούθηση της εξέλιξης μιας εθνικής λογοτεχνίας, ενός λογοτεχνικού είδους, αλλά και πιο ειδικά ενός μοτίβου (σχολές ρεύματα στη λογοτεχνική σκηνή κλπ), είτε διατρέχουμε κι αντιπαραβάλλουμε κειμενικά δείγματα, είτε χρησιμοποιούμε ένα σύνολο δευτερογενών ή αρχειακών πηγών (περιοδικά, εφημερίδες, αρχεία).

Η μετατροπή των πρωτογενών ετερογενών αυτών πηγών σε ψηφιακό περιεχόμενο όχι μόνο διευκολύνει την πρόσβαση κι εποπτεία τους αλλά κυρίως επιτρέπει την επεξεργασία, ανάλυση αλλά και οργάνωσή τους με τέτοιο τρόπο ώστε να μπορέσουμε να αναδείξουμε και να υποστηρίξουμε ερευνητικές υποθέσεις γύρω από τη ιστορική εξέλιξη κι ανάγνωση όψεων του λογοτεχνικού φαινομένου.

Για παράδειγμα το προτζεκτ [Mapping the republic of letters](#) από το πανεπιστήμιο του Stanford ουσιαστικά είναι μια σύνθετος ψηφιακός πόρος που βασίζεται σε στα ψηφιακά κείμενα των επιστολών σημαντικών διανοητών του Διαφωτισμού, επισημειωμένα με τέτοιον τρόπο ώστε να μετατρέπεται σε ψηφιακή πληροφορία, εκτός από το ίδιο το κείμενο, και μεγάλος όγκος λεπτομερούς δευτερογενούς σημασιολογικής πληροφορίας όπως παραλήπτως (εθνικότητα, φύλο, τάξη), ημερομηνία, τόποι αποστολής και παραλαβής κλπ. Κατόπιν, το σύνολο των ψηφιακών κειμένων τροφοδοτεί μια βάση δεδομένων όπου οργανώνεται το σύνολο της ψηφιακής πληροφορίας, η οποία στη συνέχεια δύναται να συνδυαστεί και οπτικοποιηθεί με διαφορετικούς τρόπους χρησιμοποιώντας ποικιλία τεχνολογιών όπως συστήματα γεωχωρικών δεδομένων, δίνοντας έτσι μια ή μάλλον πολλές εικονιστικές αφηγήσεις ή αναπαραστάσεις γύρω από τα δίκτυα αλληλογραφίας και την ιστορία τους.

Υπάρχουν ακόμη και πιο απλοί τρόποι να αναπαραστήσει κανείς την εξέλιξη στον χρόνο και σε συνδυασμό με τον χώρο με χρονολόγια ή χωροχρονολόγια, όπου μπορεί να εντάσσει ποικίλο ψηφιακό περιεχόμενο (ψηφιακές φωτογραφίες, δεδομένα κλπ), χρησιμοποιώντας εύχρηστα εργαλεία (πχ [timelines](#), [timemapper](#)).

Τι μπορεί να σημαίνει λογοτεχνική ερμηνεία, κριτική και ανάλυση στο ψηφιακό περιβάλλον: από το close reading στο distant / [differential](#) reading

Αν η λογοτεχνική θεωρία μας χορηγεί μια σειρά από τρόπους και κριτικά λεξιλόγια προκειμένου να αναλύσουμε και να ερμηνεύσουμε τα λογοτεχνικά κείμενα, τα ψηφιακά εργαλεία και η εφαρμογή τους στο ψηφιακό περιεχόμενο μάς επιτρέπουν μια διαφορετική γκάμα και ποιότητα αναγνώσεων: από την επεξεργασία σωμάτων κειμένων και την εξαγωγή ή εξόρυξη πληροφοριών (data mining), την ποσοτικοποίηση των κειμενικών αναλύσεων και την εξ αποστάσεως ανάγνωση μεγάλου όγκου ψηφιακών κειμενικών δεδομένων.

Στο προτζεκτ [Mining the Dispatch](#), για παράδειγμα, χρησιμοποιώντας την τεχνική του topic modelling και το ειδικό εργαλείο mallet στο αναζητησιμ α (OCR) ψηφιοποιημένο σώμα της αμερικανικής εφημερίδας Dispatch, διερευνώνται οι ιστορικές συνέχειες ή ασυνέχειες στον αμερικανικό εμφύλιο μέσα από την παρακολούθηση της εξέλιξης θεμάτων στη διάρκεια του χρόνου, όπως η δουλεία, ο εθνικισμός κλπ .

Το Versioning Machine είναι ένα πρόγραμμα επεξεργασίας και παράλληλης παρουσίασης των εκδοχών ενός κειμένου εφόσον αυτές είναι κωδικοποιημένες σε XML γλώσσα και σύμφωνα με τις προδιαγραφές του προτύπου (TEI Guidelines), επιτρέποντας την αντιπαραβολή έως και 5 εκδοχών του κειμένου ταυτόχρονα.

Τι μπορεί να σημαίνει έκδοση λογοτεχνικού κειμένου στο ψηφιακό περιβάλλον: από το ένα κείμενο στο ρευστό κείμενο

Η εκδοτική επεξεργασία ενός λογοτεχνικού κειμένου είναι συνήθως μια σύνθετη και χρονοβόρα διαδικασία, που περιλαμβάνει μια ποικιλία σταδίων όπως η συγκεντρωση διαθέσιμου υλικού, η κριτική αποκατάσταση & μεταγραφή ανάλογα με το στοχο/φυσιογνωμία της έκδοσης (κριτική, χρηστική, διπλωματική, γενετική), πριν την τελική συχνά κοστοβόρα έντυπη δημοσίευση.

Μια ψηφιακή έκδοση κειμένου δύναται να συμπεριλάβει, να οργανώσει και να επεξεργαστεί μεγάλο όγκο αλλά και διάφορα είδη ψηφιακού περιεχομένου, από υψηλής ανάλυσης ψηφιακά αντιγραφα των εκδοχών/μαρτυρών, ακριβείς ψηφιακές μεταγραφές, δηλαδή ηλεκτρονικά κείμενα που έχουν τη δυνατότητα ενσωμάτωσης και οπτικοποίησης κριτικών παρατηρήσεων, αλλά και ετερογενές συνοδευτικό υλικό όπως πίνακες λέξεων, γλωσσάρια, λεξικά κλπ.

Το ψηφιακό project της κριτικής έκδοσης των περίπου 1.100 αδημοσίευτων χειρογράφων της Jane Austen, περιλαμβάνει ψηφιακές φωτογραφίες υψηλής ανάλυσης ακρίβειας των χειρογράφων μαζί με πλήρεις διπλωματικές μεταγραφές τους, επιμένοντας κυρίως στη σήμανση κι οπτικοποίηση με τη τεχνολογία javascript των εναλλακτικών γραφών ή παραλλαγών (variants) που χρησιμοποιεί η μυθιστοριογράφος, ενώ την έκδοση συμπληρώνουν η τεκμηρίωση και παράθεση όλων των εκδοτικών σταδίων, προβλημάτων και αποφάσεων.

Το πρότζεκτ [Beckett Digital Manuscript Project](#) λειτουργεί ως μια ψηφιακή γενετική έκδοση χειρογράφων δυο δίγλωσσων (αγγλικά και γαλλικά) έργων του Beckett, στοχεύοντας στη συγκρότηση του γενετικού φακέλου. Έτσι, στην ψηφιακή πλατφορμα ο γενετικός φάκελος κάθε έργου περιλαμβάνει οργανωμένα σε μια βάση δεδομένων ψηφιακές φωτογραφίες του συνόλου των χειρογράφων (digital facsimiles), καθένα από τα οποία συνοδεύεται από την ηλεκτρονική μεταγραφή χρησιμοποιώντας XML κωδικοποίηση, η οποία επιτρέπει αφενός μια τοπογραφική (topographic) και μια γραμμική (linear) οπτικοποίηση της μεταγραφής, κι αφετέρου μια εποπτεία της χρονικής διαδοχής και της δυναμικής σύγκρισης των διαφορετικών παραλλαγών τόσο στα στάδια αναθεώρησης όσο και των μεταφραστικών παραλλαγών.

Το πρότζεκτ [Transcribe Bentham](#) αποσκοπεί στη χρηστική έκδοση του τεράστιου όγκου των χειρογράφων του άγγλου πολιτικού φιλοσόφου Jeremy Bentham: το ηλεκτρονικό περιβάλλον φιλοξενεί με μια βάση δεδομένων όπου έχουν οργανωθεί οι υψηλής ανάλυσης φωτογραφίες μαζί με τις ακριβείς διπλωματικές μεταγραφές σε κωδικοποίηση XML, οι οποίες δημιουργούνται από τους ίδιους τους χρήστες, οι οποίοι ωστόσο δεν απαιτείται να έχουν τεχνικές γνώσεις καθώς χρησιμοποιούν ένα “έξυπνο λογισμικό” (του τύπου you type what you see, YTWYS) που μεταφράζει το απλό πληκτρολογημένο κείμενο σε TEI XML ηλεκτρονικό κείμενο. Πρόκειται

για ένα πρότζεκτ πληθοπορισμού (crowdsourcing), όπου απλοί χρήστες συμμετέχουν στην εκδοτική διαδικασία, το τελικό αποτέλεσμα της οποίας ελέγχεται από επιστήμονες.

3. Προϋποθέσεις για τη δημιουργία ψηφιακού περιεχομένου

Κλείνοντας το δικό μου κείμενο και ξεκινώντας ουσιαστικά την συζήτηση γύρω από το ψηφιακό περιεχόμενο, νομίζω πως αξίζει να υπενθυμίσουμε τις βασικές προϋποθέσεις στις οποίες είναι δυνατόν να ανοίξουμε σήμερα τη συζήτηση γύρω από το ψηφιακό περιεχόμενο

Παράγω, αναπαράγω, διαχειρίζομαι, χρησιμοποιώ ψηφιακό περιεχόμενο σημαίνει:

1. ακριβής στοχοθεσία και επιλογή κατάλληλων ψηφιακών εργαλείων και τεχνολογιών
2. χρήση ανοικτών προτύπων κωδικοποίησης και μεταφοράς της πληροφορίας (protocols, standards)
3. υιοθέτηση καλών πρακτικών (best practices)
4. χρήση ανοιχτών εργαλείων και τεχνολογιών (open-source software)
5. επαναχρησιμοποίηση περιεχομένου - ο ρόλος των ερευνητικών υποδομών
6. πνευματικά δικαιώματα & άδειες ανοιχτού λογισμικού / περιεχομένου/ δεδομένων
7. πειραματισμός & εξοικείωση, ανάπτυξη δεξιοτήτων
8. διεπιστημονικότητα
9. συνεργασία & επικοινωνία
10. δημιουργικότητα, φαντασία, καινοτομία, τόλμη.

Σας ευχαριστώ πολύ!