γλῶσσαν νωμᾶν

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΚΕΝΤΡΟ ΕΛΛΗΝΙΚΗΣ ΓΛΩΣΣΑΣ
ΜΕΛΟΣ ΤΗΣ ΕΥΡΩΠΑΪΚΗΣ ΟΜΟΣΠΟΝΔΙΑΣ ΕΘΝΙΚΩΝ ΙΔΡΥΜΑΤΩΝ ΓΙΑ ΤΗ ΓΛΩΣΣΑ (EFNIL)

ΕΣΠΑ 2007-2013
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

# Ψηφιακό corpus νεοελληνικής ποίησης

## *Ανεμόσκαλα*

[www.greek-language.gr/digitalResources](www.greek-language.gr/digitalResources)

Σκαρφαλώνοντας λέξεις όπως μιαν ανεμόσκαλα.

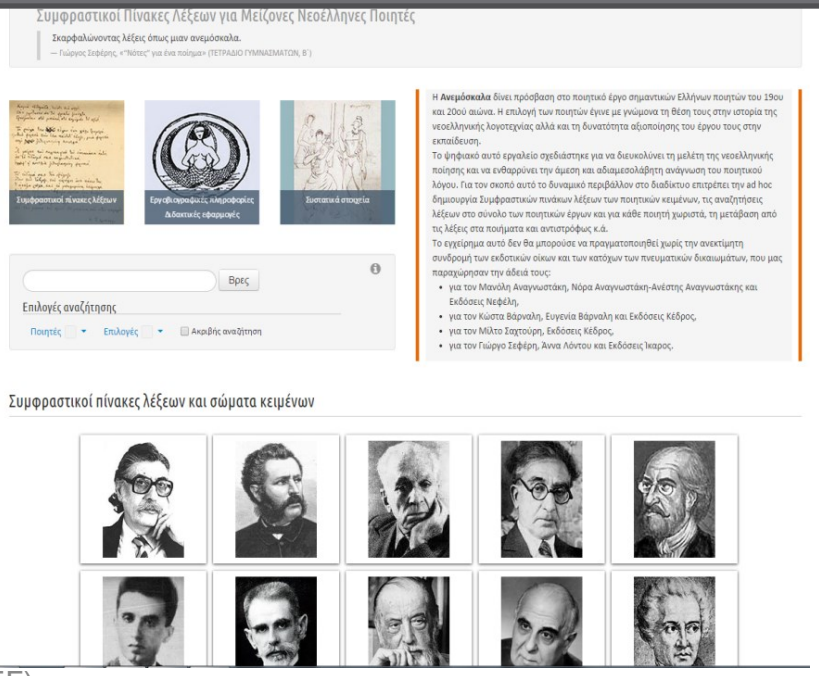
— Γιώργος Σεφέρης, «"Νότες" για ένα ποίημα» (ΤΕΤΡΑΔΙΟ ΓΥΜΝΑΣΜΑΤΩΝ, Β΄)

Στο πλαίσιο ανάπτυξης πόρων για την «Πύλη για την ελληνική γλώσσα και τη γλωσσική διδασκαλία» (www.greek-language.gr)
το ΚΕΓ είχε προχωρήσει σε ψηφιακή επεξεργασία της προηγούμενης *έντυπης έκδοσης του Συμφραστικού Πίνακα Λέξεων* του ποιητικού έργου του Γιώργου Σεφέρη, εντάσσοντάς τον μάλιστα ως γλωσσικό εργαλείο σε ένα ολοκληρωμένο **corpus του σεφερικού έργου, όπου ο αναγνώστης έχει τη δυνατότητα να μεταβεί από τη λέξη στο ποίημα και αντιστρόφως.**

Σήμερα, στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» (ΕΣΠΑ 2007-2013), το ΚΕΓ υλοποιεί ένα νέο ψηφιακό περιβάλλον

**Η *Ανεμόσκαλα* βασίζεται στην προηγούμενη εμπειρία της δημιουργίας του σεφερικού corpus αλλά δεν εκκινεί από μια έντυπη έκδοση (ψηφιοποίηση vs born digital)**

**καινοτόμα χαρακτηριστικά**



Συμβατικά

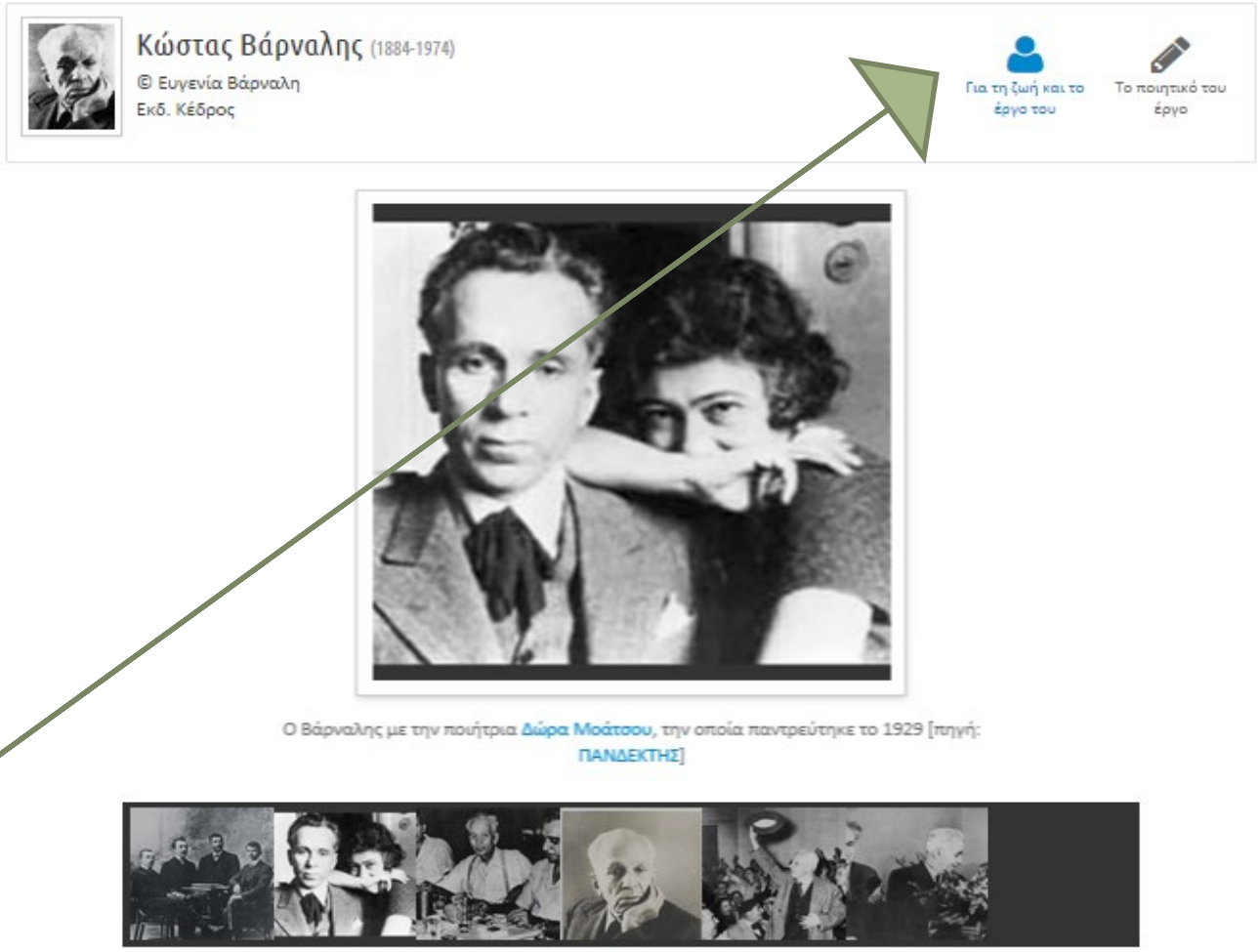
**στοιχεία του ψηφιακού περιβάλλοντος:**

**α) ο συγγραφέας ως άξονας συγκρότησης**

**Χρονολόγιο**

**Εργοβιογραφικό**

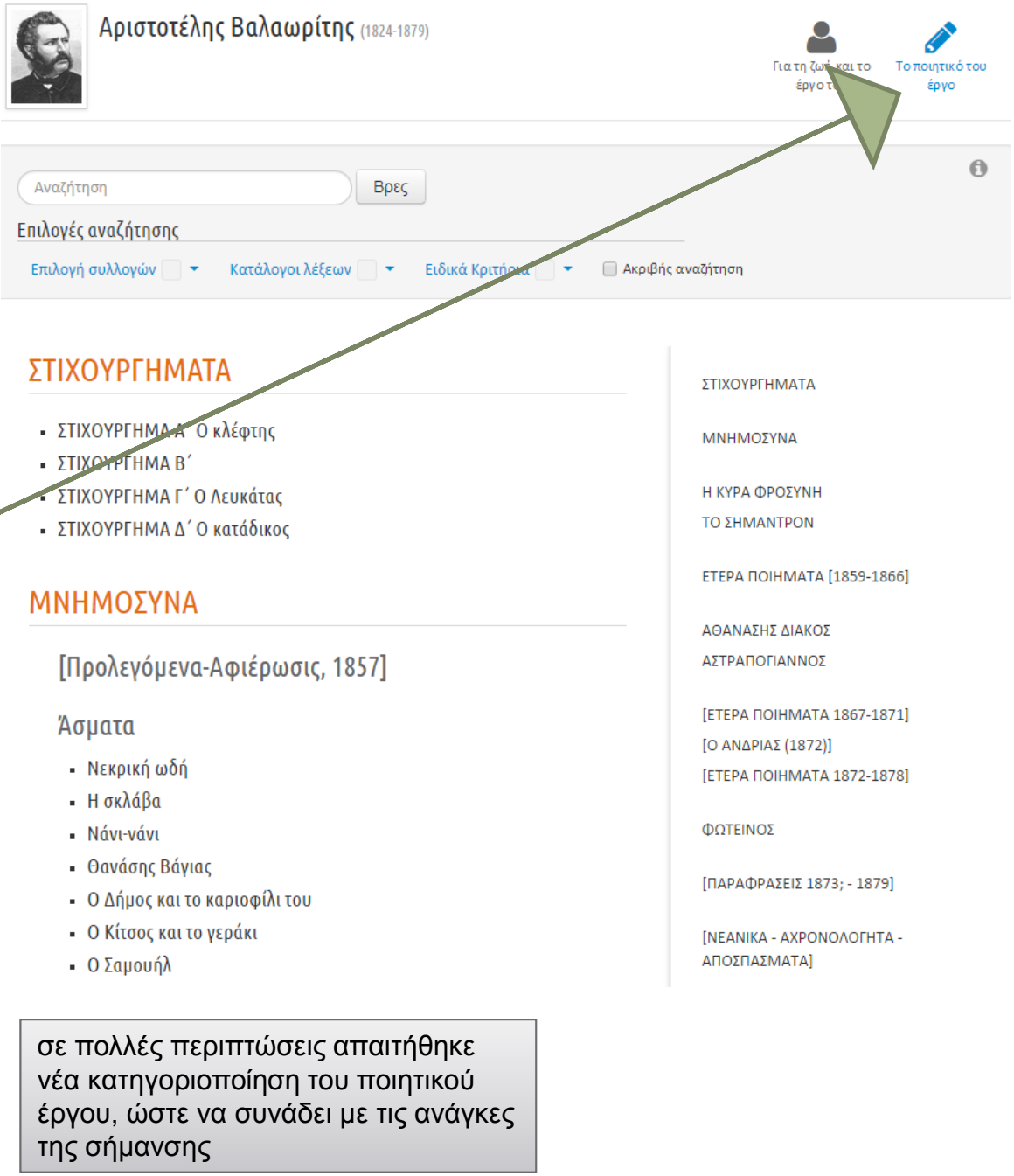
**Εργογραφία-βιβλιογραφία**



**Συμβατικά στοιχεία του ψηφιακού περιβάλλοντος:**

**β) πρόσβαση στο ποιητικό έργο μέσω καταλόγου περιεχομένων**

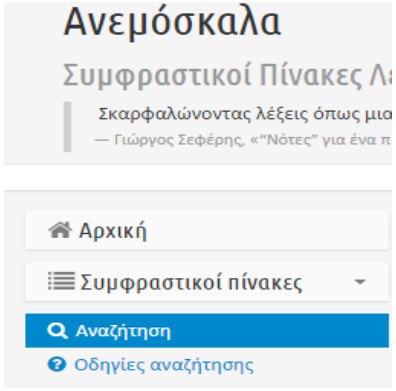
στην περίπτωση ποιητών για τους οποίους δεν έχουμε άδεια δημοσίευσης του πλήρους έργου, ο χρήστης διαθέτει τις ίδιες δυνατότητες αναζήτησης λέξεων, σε επίπεδο συμφραζομένων, χωρίς, ωστόσο, δυνατότητα ανάγνωσης του ποιήματος·
στην περίπτωση αυτή γίνεται λόγος μόνο για συμφραστικούς πίνακες λέξεων και όχι για ολοκληρωμένο σώμα κειμένων.

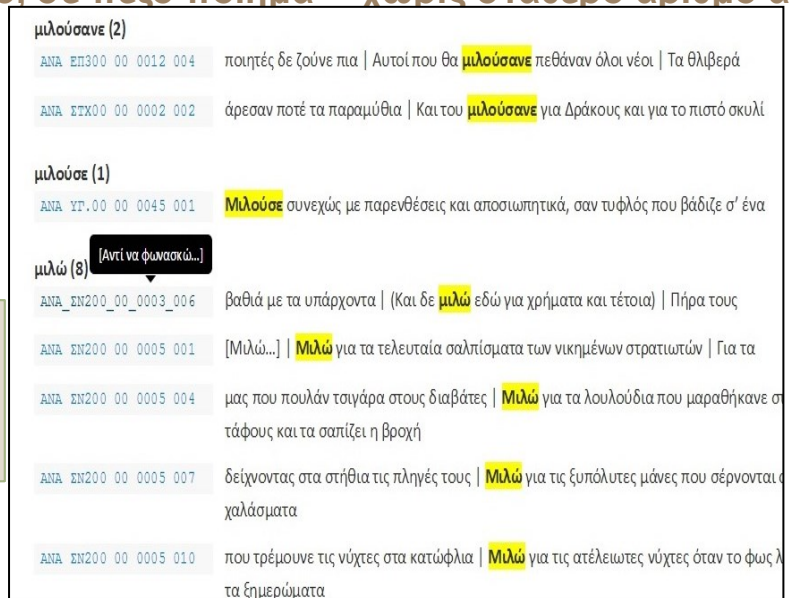



σε πολλές περιπτώσεις απαιτήθηκε νέα κατηγοριοποίηση του ποιητικού έργου, ώστε να συνάδει με τις ανάγκες της σήμανσης

## Οι χρήστες έχουν τη δυνατότητα: 1. απλής αναζήτησης

Η αναζήτηση γίνεται ατονικά
Μπορεί να χρησιμοποιηθεί **αστερίσκος * (στην αρχή, στη μέση ή στο τέλος)** για να αναζητηθούν ομάδες λέξεων που περιέχουν το μόρφημα ή την ακολουθία χαρακτήρων που πληκτρολογείται (βλ. **αναλυτικές οδηγίες αναζήτησης**)
Κάθε λεξικός τύπος καταγράφεται τόσες φορές όσες απαντά, πάντα συνοδευόμενος από τα άμεσα συμφραζόμενά του (context, έκτασης κατά κανόνα ενός στίχου ή μιας αράδας) και από ακριβείς παραπομπές.
**Οι παραπομπές [αριστερά, δίπλα σε κάθε αποτέλεσμα αναζήτησης] επιτρέπουν τη μετάβαση από τη λέξη στο ποίημα (όπου επιτρέπεται λόγω πνευματικών δικαιωμάτων) και δίνουν στοιχεία: =του ποιητή, της συλλογής (ή του «ψευδότιτλου» συλλογής), της σειράς ή υπο-ενότητας της συλλογής (αν υπάρχει, διαφορετικά σημειώνονται μηδενικά), του ποιήματος, αριθμό στίχου (εφόσον ο τύπος συναντάται σε συμφραζόμενα αριθμημένου στίχου ή verset και όχι σε τίτλο, σε πεζό ποίημα —χωρίς σταθερό αριθμό αράδων— κλπ.).**




**Σοφία (για να αναζητήσετε Αγια-Σοφιά, Αγία-Σοφία, δηλαδή τις περιπτώσεις ημιπαραθετικών συνθέτων)*
**δειξε (για να αναζητήσετε π.χ. 'δειξε)*
*ανα*μα (ανάβρασμα, ανάθεμα κλπ.)*

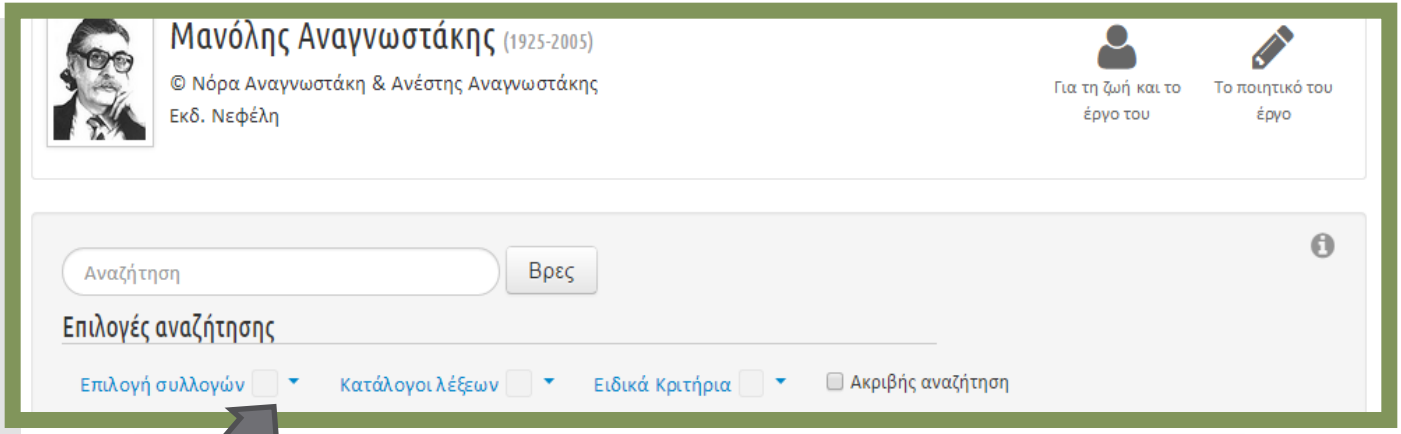
# Οι χρήστες έχουν τη δυνατότητα:

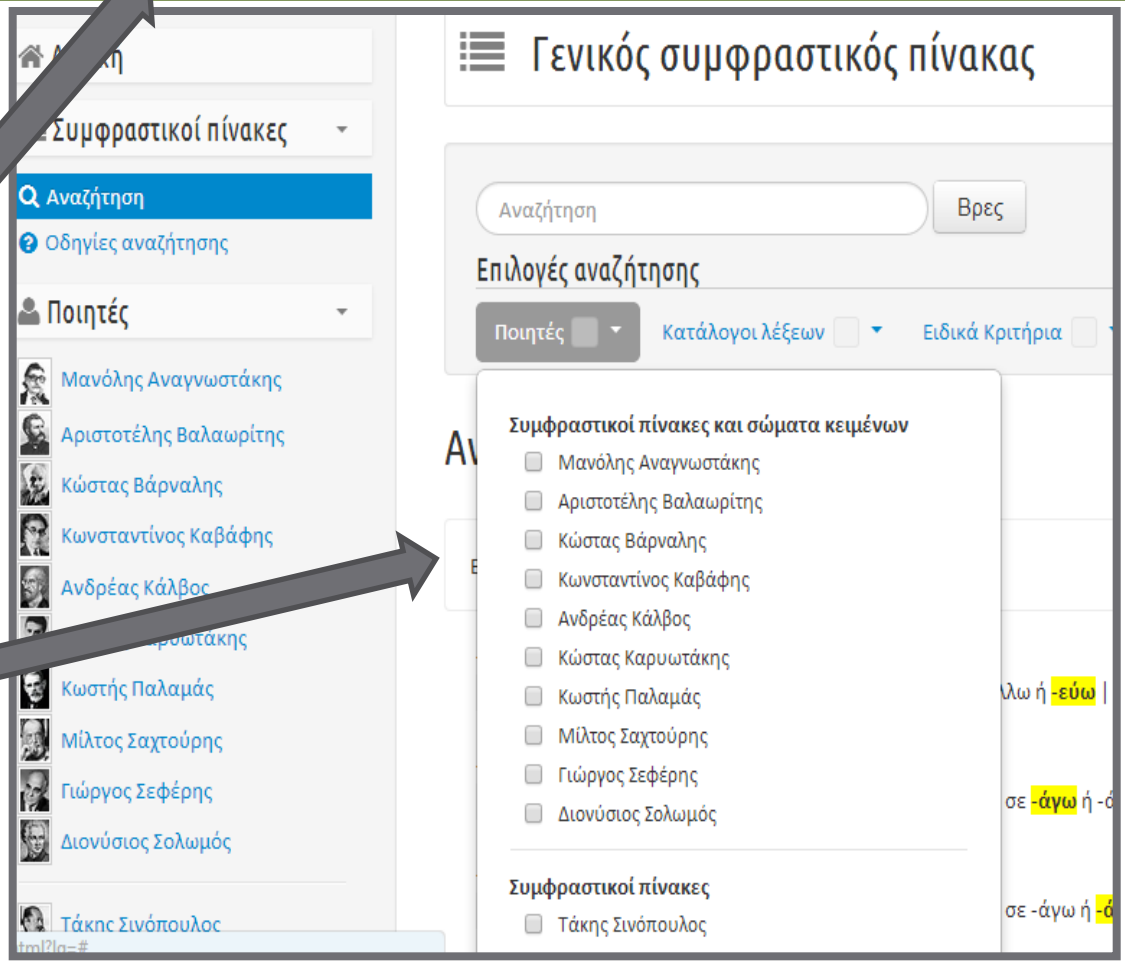
## 1. απλής αναζήτησης λεκτικού τύπου:

**α) σε όλο το ποιητικό έργο ενός ποιητή**

**β) σε συγκεκριμένες συλλογές ενός ποιητή**

**γ) σε έναν ή περισσότερους ποιητής (συνδυαστική αναζήτηση)**



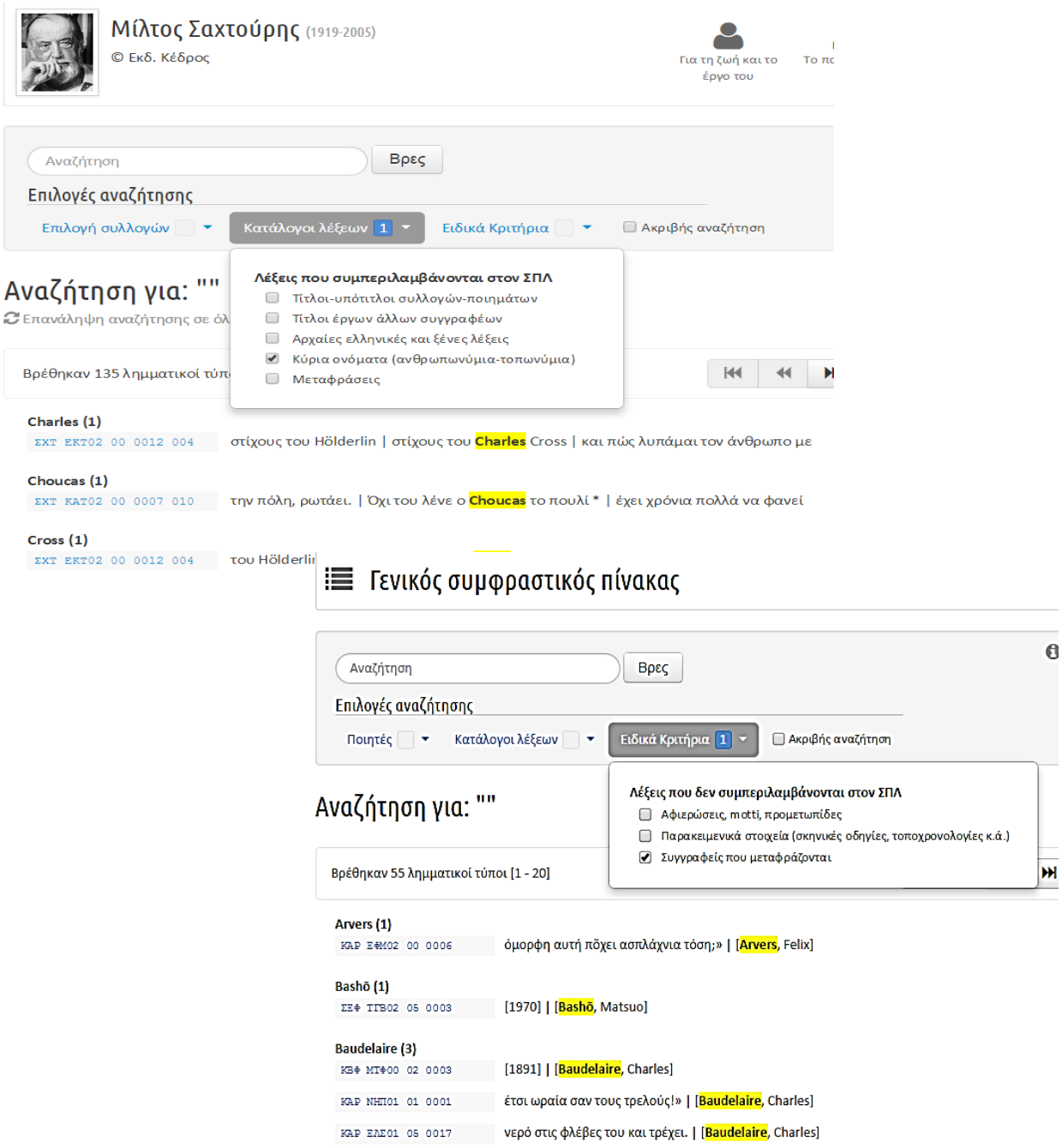



# Οι χρήστες έχουν τη δυνατότητα: 2. αναζήτησης

**α) με περιορισμό σε συγκεκριμένους καταλόγους λέξεων**

Δεν πρόκειται ωστόσο για λημματικούς πίνακες λέξεων ή τυπικά ευρετήρια, αλλά για περιορισμένη αναζήτηση λεξικών τύπων εντός συγκεκριμένων πεδίων.

**β) με ειδικά κριτήρια για λεξικούς τύπους που δεν περιλαμβάνονται στην απλή αναζήτηση του ΣΠΛ**



# Ποιό κείμενο;

# Εκδοτικές αρχές

- Όχι κριτική έκδοση —ένα τέτοιο εγχείρημα προϋποθέτει διαφορετικό χρονικό και μεθοδολογικό πλαίσιο—
- αλλά ένα ψηφιακό corpus με σκόπευση κυρίως **χρηστική**

  προσαρμογή στο ισχύον τονικό και ορθογραφικό σύστημα με βάση συγκεκριμένες προδιαγραφές ανάλογα με το κάθε ποιητικό έργο και τις ιδιαιτερότητές του – εμπλουτισμένο μονοτονικό (π.χ. τονισμός μονοσύλλαβων τύπων της καθαρεύουσας, διατήρηση κράσης κλπ.)

  Π.χ. κι εις ένα ξύλον κρέμεται έν παγωμένον πτώμα (Βαλαωρίτης)

  Ομογενοποίηση ως προς τυπογραφικές, τονικές, ορθογραφικές και άλλες συμβάσεις, χωρίς άλλες διορθωτικές επεμβάσεις πέρα από εμφανή τυπογραφικά σφάλματα ή ήδη διαπιστωμένα παροράματα. Πραγματολογικά λάθη (realia) και ιδιότροπες γραφές του συγγραφέα διατηρούνται - οι επεμβάσεις στην στίξη είναι ελάχιστες.

- Όπου δεν υπάρχει στερεότυπη έκδοση απάντων ενός ποιητή, προχωρούμε και σε εκδοτικές επιλογές ταξινόμησης του ποιητικού έργου.
- Στο corpus περιλαμβάνονται και οι μεταφράσεις, όταν εντάσσονται, κατόπιν επιλογή του ποιητή ή του φιλολογικού επιμελητή, στις ποιητικές συλλογές. Ωστόσο, σημαίνονται με ειδικό τρόπο, ώστε να είναι δυνατή η συμπερίληψη ή μη στην αναζήτηση.

- Επιμέλεια και Σήμανση(tagging)= όχι αυτόματη επεξεργασία με βάση κανόνες αλλά *ερμηνεία. Δουλειά όχι τεχνική αλλά του ψηφιακού φιλολόγου*

εκδοτική προεργασία

[επιστημονικός συντονιστής] *

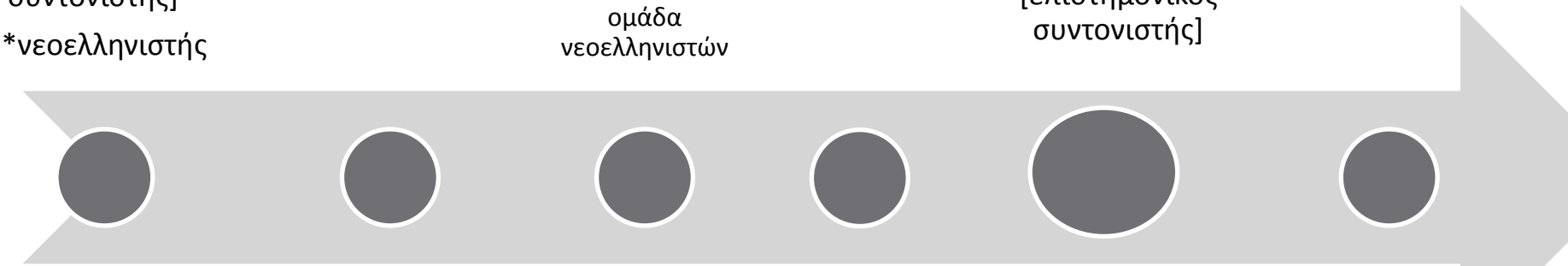
*νεοελληνιστής

Αρχική επιμέλεια (μονοτονισμός - ορθογραφικός εκσυγχρονισμός)

ομάδα νεοελληνιστών

Τελική επιμέλεια - τελική σήμανση

[επιστημονικός συντονιστής]



Πληκτρολόγηση

ομάδα φιλολόγων

Σήμανση αρχική

[γλωσσολόγος υπολογιστικής γλωσσολογίας]

Ψηφιακή επεξεργασία - διορθώσεις

[επιστημονικός συντονιστής σε συνεργασία με τον υπεύθυνο τεχνικής υλοποίησης]

δημιουργείται tagged corpus, δηλαδή είναι ένα ηλεκτρονικό σώμα κειμένων (ΗΣΚ) με γλωσσικές ετικέτες οι οποίες «χαρακτηρίζουν» το κείμενο **Text Encoding Initiative: TEI**
www.tei-c.org

**Ο ΣΠΛ περιλαμβάνει όλους ανεξαιρέτως τους λεξικούς/λεκτικούς τύπους (ΛΤ) που απαντούν στα ποιήματα· πρόκειται για λέξεις-αναφορές/παραπομπές (tokens ή references), και όχι λήμματα.**

Οι ΛΤ καταγράφονται όπως ακριβώς απαντούν

- Μια λεξική μορφή που είναι γραφικά πανομοιότυπη με μια άλλη συμψηφίζεται με αυτήν [εξαίρεση: τα κύρια ονόματα π.χ. Ψαρά – ψαρά]
- Δεν πρόκειται επομένως για λημματολόγιο αλλά για **μορφολόγιο** (=πίνακας των λεξικών μορφών του κειμένου, Peri 1983 και Φιλιππίδου 1986)

## Τι δεν είναι;

- Δεν είναι υπερκείμενο
- Δεν είναι ψηφιακό αρχείο - ψηφιακή (γενετική ή κριτική) έκδοση αρχειακού υλικού
- Δεν περιλαμβάνει στατιστική επεξεργασία γλωσσικών δεδομένων (π.χ. μπορεί να χρησιμοποιηθεί ως ερευνητικό εργαλείο για να καταρτίσει κάποιος τις λέξεις- άπαξ, δεν παρέχει όμως αυτόματα έναν πίνακα, καθώς κάτι τέτοιο θα απαιτούσε λημματοποίηση)

**Αξιοποίηση για (ενδεικτικά):**

- γλωσσολογικές – υφολογικές έρευνες (λ.χ. ιδιωτισμοί, σημασιολογικά πεδία λέξεων κλπ.) σε έναν ή περισσότερους ποιητές, σε ένα ή περισσότερα είδη κειμένων (βλ. πρόσφατη έρευνα για ταύτιση μεταφράσεων του Παπαδιαμάντη με υφομετρικές τεχνικές).
- την ανίχνευση θεμάτων, μοτίβων, εικόνων.

*σε ποιους αφιέρωσε ποιήματά του ο Κ. Παλαμάς;*
*σε ποιους ποιητές απαντούμε πιο συχνά αρχαιόγλωσσα παραθέματα και ποια;*
*ποιο είναι το κοινό λεξιλόγιο Βάρναλη-Σολωμού ή πώς συνομιλεί τελικά ο Βάρναλης με τη σολωμική ποίηση;*
*ποιες οι λέξεις-άπαξ που αξιοποιεί ή/και συνθέτει ο Σαχτούρης;*
*ποια ιστορικά πρόσωπα απαντούν σε στίχους του Βαλαωρίτη;*
*ποια ποιήματα έχουν δηλωμένη χρονολογία συγγραφής το 1940;*

# Προτάσεις διδακτικής αξιοποίησης

- Δ.Ν. Μαρωνίτης: ριζικά επίθετα στον Σολωμό: γλυκός, καλός, μαύρο, χρυσός
- Μαρωνίτης, Δ. Ν. «Οι νέοι στην ποίηση του Κ. Π. Καβάφη», *Κ. Π. Καβάφης, Μελετήματα*
- Γεώργιος Παπαντωνάκης, «Η δυναμική του χρώματος στην ελυτική κοσμογραφία». Μια μελέτη στατιστικής βάσης για τα χρωματικά μοτίβα στον Ελύτη.

- Χρήστος Δανιήλ: Αναζήτηση παράλληλων χωρίων ως προς τις λέξεις *ποιη** και *τραγουδ** προκειμένου να εντοπιστούν και να ανιχνευτούν οι απόψεις που καταγράφονται στα ποιήματα του Καρυωτάκη για την ποίηση.
- Αγγέλα Γιώτη: Επεκτείνοντας την αναζήτηση με τον τύπο μαυρ* στον ΣΠΛ αρχικά του Βάρναλη και στη συνέχεια του Βαλαωρίτη, μπορούμε να εντοπίσουμε τον τρόπο αξιοποίησης του εύρους της σημασίας του συγκεκριμένου επιθέτου σε μεταφορικές χρήσεις και να διαπιστώσουμε τυχόν μετατόπιση από την αξιοποίησή του για απόδοση ηρωικών επεισοδίων της εθνικής ζωής στην λειτουργία του στο πλαίσιο κοινωνικής κριτικής.