

26 Αναγνώριση παραπομπών σε ελληνικά νομικά κείμενα με χρήση κανόνων

Τσιμπούρης Χαράλαμπος, Κυριάκος Σγάρμπας

ΠΕΡΙΛΗΨΗ

Στην παρούσα ανακοίνωση παρουσιάζεται η κατασκευή αυτόματου συστήματος αναγνώρισης παραπομπών σε ελληνικά νομικά κείμενα, η οποία αναπτύχθηκε υπό την αιγίδα της Επιτροπής Ερευνών του Πανεπιστημίου Πατρών, στα πλαίσια της βασικής έρευνας «Καραθεοδωρής». Η χρήση παραπομπών σε νομικά κείμενα αποτελεί επί μέρους, αλλά αναπόσπαστο στοιχείο μελέτης των ελληνικών νομικών κειμένων. Η ανάγκη έχει προκύψει από το γεγονός ότι η αναγνώριση παραπομπών αποτελεί αρχικό βήμα για μετέπειτα διεργασίες πλοήγησης και ανάλυσης των κειμένων, από χρήστες ή από αυτοματοποιημένα συστήματα επεξεργασίας, όπως δημιουργία δικτύου συσχέτισης, εύκολη πλοήγηση, αλλά και αυτόματη ενημέρωση μετά από τροποποιητικές νομοθετικές διατάξεις. Αν και έχουν γίνει προσπάθειες σε άλλες γλώσσες, η ανάλυση τόνισε την έλλειψη που υπάρχει στον συγκεκριμένο τομέα έρευνας για την ελληνική γλώσσα. Για την σωστή μελέτη και εμπειριστατωμένη δημιουργία αυτοματοποιημένου συστήματος, έγινε κατασκευή ομάδας κανονικών εκφράσεων που θα μπορούν να καλύψουν όσο το δυνατόν περισσότερες περιπτώσεις, δημιουργώντας κατάλληλη βάση δεδομένων για αυτό το σκοπό. Συγκεκριμένα, επιλέχθηκε μία λίστα από 14 αρχεία με νομικό περιεχόμενο ποικίλης προέλευσης και προχωρήσαμε στην καταγραφή όλων των παραπομπών που βρέθηκαν στα επιλεγμένα αρχεία, προς άλλα νομικά κείμενα, ή παραπομπή σε άλλο άρθρο ή παράγραφο του ίδιου κειμένου. Με την ολοκλήρωση της καταγραφής, δημιουργήθηκε η κατάλληλη βάση δεδομένων με 796 καταχωρήσεις παραπομπών πάνω στην οποία βασίστηκε εξ' ολοκλήρου η δημιουργία κανόνων. Παράλληλα, γίνεται καταγραφή συντομογραφιών για νομοθεσία, κώδικες, νομοθετήματα, πολιτικά δικαστήρια, ποινικά δικαστήρια, διοικητικά δικαστήρια, και ευρωπαϊκά δικαστήρια καθώς αποτελούν αναπόσπαστο κομμάτι σε ένα ευρύ φάσμα νομικών παραπομπών. Η μελέτη έδειξε την έλλειψη προτυποποίησης που επικρατεί ως γενικευμένο φαινόμενο στα ελληνικά νομικά κείμενα γεγονός το οποίο βασίζεται στην ποικίλης μορφής γραφή παραπομπών, όπως γίνεται και σαφές από την αντίστοιχη βάση δεδομένων που κατασκευάστηκε για αυτό το σκοπό.

Reference identification, in greek law texts with use of regular expressions

Tsimpouris Charalampos, Kyriakos Sgarbas

ABSTRACT

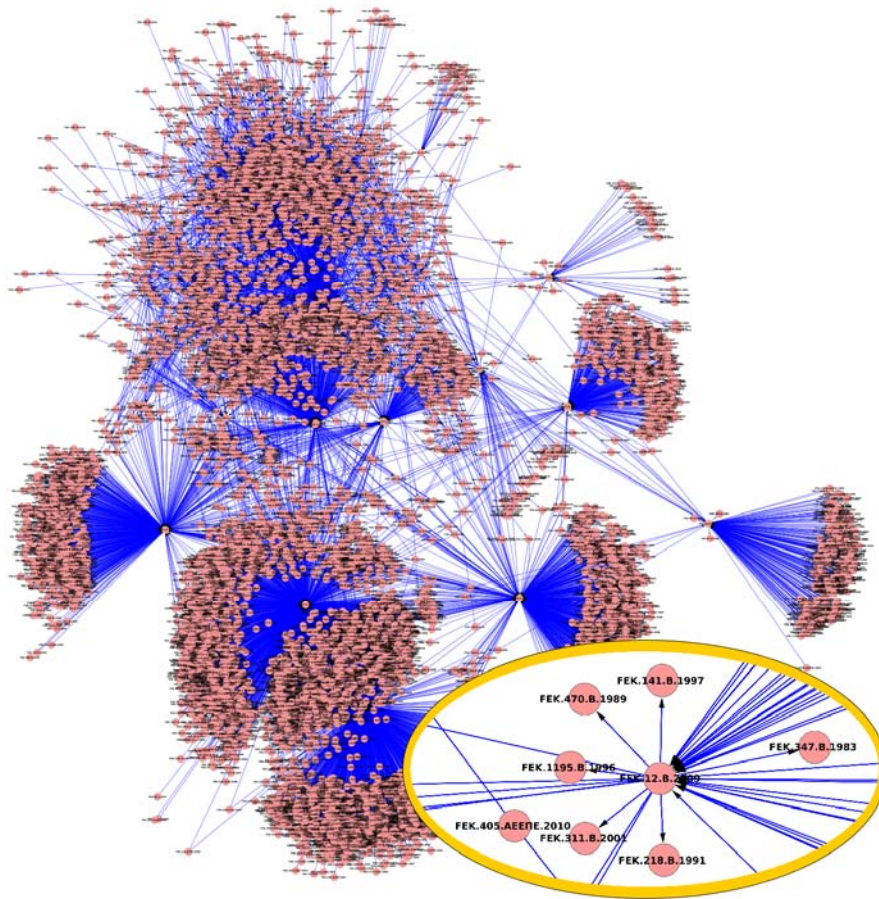
In this paper, we present the construction of an autonomous identification referncing platform in greek law texts, which was developed with the support of the Research Program Caratheodory, 2010-2013, of the Research Committee of the University of Patras, Greece. Legal citation is needed in every step of studying legal texts, as it is used both by human and legal parsers for easy traversal through legal databases, creating connection graphs and even maintaining existing laws. Even there has been development in other languages, the analysis showed the lack in this area of expertise, in Greek legal texts. In order to proceed with the construction of the autonomous system, we created a group of regular expressions, that can identify as most as possible of a list of occurrences, after creating a

database for this purpose. In particular, we selected 14 documents, with different kind of legal texts, and exported manually all citations to other legal documents. A database was constructed with 796 citations, on which all regular expressions were based. The analysis showed the significant lack of standardization, as a generalized issue among all kind of legal documents, based at the fact of various ways of writing citations.

0 Εισαγωγή

Σκοπός της παρούσας ανακοίνωσης είναι η παρουσίαση της μελέτης και ανάπτυξης αυτόματου συστήματος αναγνώρισης παραπομπών σε ελληνικά νομικά κείμενα. Αποφάσεις συλλογικών οργάνων, βασίζονται σε επίσημα νομικά κείμενα (Νόμοι, Υπουργικές αποφάσεις, και άλλα), που περιλαμβάνει μία αρκετά δύσκολη διαδικασία λήψης υλικού, συντήρησής του, αλλά και κατανόησής του. Κεντρική πηγή αποτελεί κάθε φορά τα γνωστά “ΦΕΚ” (Φύλλο Εφημερίδας της Κυβερνήσεως) που αποτελούν την επίσημη μέθοδο δημοσίευσης ανακοινώσεων, νομοθεσίας, συνθηκών, εκθέσεις δικαστηρίων, προκηρύξεις προσωπικού, διεθνής συμφωνίες, οικονομικές καταστάσεις και πολλά άλλα. Περιέχουν έτσι μια πληθώρα νομικών πληροφοριών για νομικούς ερευνητές, εφημερίδες και νομικές υπηρεσίες [1].

Για παράδειγμα, στην περίπτωση “*Για τους όρους και τη διαδικασία καταβολής του ως άνω επιδόματος ισχύουν τα οριζόμενα με τις υπ’ αριθμ. Γ1α 440/7.2.1991 (ΦΕΚ 90 Β΄) και Π3δ/οικ. 1078/ 19.3.1997 (ΦΕΚ 241 Β΄) κοινές αποφάσεις των Υπουργών Εθνικής Οικονομίας, Οικονομικών και Υγείας, Πρόνοιας και Κοινωνικών Ασφαλίσεων και με τα άρθρα 2 έως 9 της υπ’ αριθμ. Π1α/Γ.Π.οικ.60227/19.5.2006 (ΦΕΚ 645 Β΄) κοινής απόφασης των Υπουργών Οικονομίας και Οικονομικών και Υγείας και Κοινωνικής Αλληλεγγύης.*” ο αναγνώστης θα πρέπει να αναφερθεί σε 3 διαφορετικά ΦΕΚ ούτως ώστε να συνεχίσει τη μελέτη του.



Εικόνα 1: Συνδέσεις όλων των γειτονικών κόμβων που σχετίζονται με το
ΦΕΚ 2317/B/2008

Οι πολλαπλές παραπομπές, οδηγούν σε πολύπλοκο γράφο συνδέσεων μεταξύ των ΦΕΚ, όπως φαίνεται ενδεικτικά για ένα συγκεκριμένο ΦΕΚ στην εικόνα 1. Η πληθώρα νόμων αλλά και οι αμφίσημες αναφορές μας οδηγούν στην απαραίτητη δημιουργία κατάλληλων εργαλείων.

1 Σχετικές εργασίες

Στην προσπάθεια αναζήτησης υπάρχουσας ελληνικής βιβλιογραφίας σχετικά με τον τρόπο γραφής αναφορών προς νομικά κείμενα, δεν κατέστη δυνατόν να βρεθεί κάποια μελέτη.

Επίσης, αναζήτηση τυποποίησης στον Ελληνικό Οργανισμό Τυποποίησης (ΕΛΟΤ) [2] σχετικά με τον καθιερωμένο τρόπο γραφής αναφορών σε νομικά κείμενα, δεν έδωσε σαφή αποτελέσματα. Η μόνη προσπάθεια ομογενοποίησης του τρόπου γραφής βρέθηκε μόνο στις οδηγίες γραφής αναφορών σύμφωνα με την Αμερικανική Ψυχολογική Εταιρεία (American Psychological Association, APA), η οποία μέσα από μία ενδελεχής λίστα περιπτώσεων κυβερνητικών εγγράφων προτείνει συγκεκριμένο τρόπο γραφής, ενώ τα συνοδεύει με αντίστοιχα παραδείγματα [3, 4]. Αντίστοιχος κανονισμός έχει δημιουργηθεί και από την Εταιρεία σύγχρονης γλώσσας (Modern Language Association, MLA) [5] και την βιβλιοθήκη του Κογκρέσου των Ηνωμένων Πολιτειών [6].

2 Εργαλεία και Βάση δεδομένων

Για την πληρέστερη μελέτη της υπάρχουσας κατάστασης έπρεπε αρχικά να γίνει συγκέντρωση του κατάλληλου υλικού αλλά και η κατασκευή κατάλληλου εργαλείου για την εύκολη διαχείρισή του. Συγκεκριμένα, έγινε τυχαία επιλογή 14 εγγράφων νομικού περιεχομένου (230 σελίδες ελεύθερο κείμενο από 11 ΦΕΚ, 1 Προεδρικό διάταγμα, 1 νομικό έγγραφο, 1 υπουργική απόφαση) και από διαφορετικές ημερομηνίες (1990-2012), ενώ στην συνέχεια έγινε πλήρης καταγραφή όλων των παραπομπών που βρέθηκαν μέσα στα κείμενα, προς άλλα νομικά κείμενα ή προς το ίδιο κάθε φορά. Οι 796 παραπομπές καταγράφηκαν σε κατάλληλη βάση δεδομένων (MySQL) η οποία κατασκευάστηκε για το συγκεκριμένο σκοπό.

Παράλληλα, η δημιουργία κατάλληλου εργαλείου (Εικόνα 2) σε μορφή διαδραστικής ιστοσελίδας (PHP και HTML με χρήση AJAX) και χρήση φυλλομετρητή, βελτίωσε δραματικά σε χρόνο την διαδικασία δημιουργίας κανόνων, δοκιμάζοντας εύκολα και γρήγορα πιθανές κανονικές εκφράσεις βοηθώντας με την ποιοτική και ποσοτική οπτική αναπαράσταση των αποτελεσμάτων.

3 Μεθοδολογία

Αρχικά, έγινε βασικός διαχωρισμός των παραπομπών ανάλογα με το είδος νομικού κειμένου στο οποίο παραπέμπουν. Συγκεκριμένα, καταλήξαμε στις εξής κατηγορίες: α) Εφημερίδα της Κυβερνήσεως (ΦΕΚ), β) Προεδρικό Διάταγμα (ΠΔ), γ) Νόμος, δ) Υπουργική απόφαση και ε) Βασιλικό διάταγμα. Είναι σαφές πως υπάρχουν και άλλες κατηγορίες νομικών κειμένων όπως αποφάσεις δικαστηρίων, Ευρωπαϊκές οδηγίες, Πράξεις νομοθετικού περιεχομένου κ.α. Παρόλαυτα, θεωρήσαμε σκόπιμο να μην επεκταθούμε στην παρούσα φάση, ούτως ώστε να μελετήσουμε την υπάρχουσα κατάσταση σε ένα μέρος του προβλήματος. Ακολουθεί η λίστα με τις κανονικές εκφράσεις στις οποίες καταλήξαμε μετά τη μέθοδο των δοκιμών για κάθε μία από τις προαναφερθείσες κατηγορίες.

ΦΕΚ:

1. /(?<fek>(Φ\.?Ε\.?Κ\.)?~? ?(?<arithmos>[0-9]+) *\/
*(τ\.)?(?<tipos>[ΑΒΒΓΔ]´?´?´?´|Δ\.?Ε\.?Β\.?Ι\.)|(Α\.?ΕΙ\.?Δ\.)|(Α\.
?Σ\.?Ε\.?Π\.)|(Α\.?Ε-
Ε\.?Π\.?Ε\.)|(Δ\.?Δ\.?Σ\.)|(Ο\.?Π\.?Κ\.)|(Υ\.?Ο\.?Δ\.?Δ\.)|(Α\.?Α
\.?Π\.)?) *\/ *(?<imerominia>[0-9]{1,2}[.---][0-9]{1,2}[.---][0-
9]{2,4}|[0-9]{2,4})/ui
2. /(?<fek>(?(?<imerominia>[0-9]{1,2}[.---][0-9]{1,2}[.---][0-9]{2,4}|[0-
9]{2,4}) *\\((Φ\.?Ε\.?Κ\.)?~? ?(?<arithmos>[0-9]+)
?\/?(τ\.)?(?<tipos>[ΑΒΒΓΔ]´?´?´?´|Δ\.?Ε\.?Β\.?Ι\.)|(Α\.?ΕΙ\.?Δ\.)|(Α\.
?Σ\.?Ε\.?Π\.)|(Α\.?Ε-
Ε\.?Π\.?Ε\.)|(Δ\.?Δ\.?Σ\.)|(Ο\.?Π\.?Κ\.)|(Υ\.?Ο\.?Δ\.?Δ\.)|(Α\.?Α
\.?Π\.)?) *\\)/ui
3. /(?<fek>(?(?<imerominia>[0-9]{1,2}[.---][0-9]{1,2}[.---][0-9]{2,4}|[0-
9]{2,4}) *(<[^\>]*)) *\\((Φ\.?Ε\.?Κ\.)?~?
?(τ\.)?(?<tipos>[ΑΒΒΓΔ]´?´?´?´|Δ\.?Ε\.?Β\.?Ι\.)|(Α\.?ΕΙ\.?Δ\.)|(Α\.
?Σ\.?Ε\.?Π\.)|(Α\.?Ε-
Ε\.?Π\.?Ε\.)|(Δ\.?Δ\.?Σ\.)|(Ο\.?Π\.?Κ\.)|(Υ\.?Ο\.?Δ\.?Δ\.)|(Α\.?Α
\.?Π\.)?) *\\/(?<arithmos>[0-9]+)/ui
4. /(?<fek>(Φ\.?Ε\.?Κ\.)?~? ?(?<arithmos>[0-9]+)\/(?<imerominia>[0-
9]{1,2}[.---][0-9]{1,2}[.---][0-9]{2,4}|[0-9]{2,4}),? *τ\.
*(?<tipos>[ΑΒΒΓΔ]´?´?´?´|Δ\.?Ε\.?Β\.?Ι\.)|(Α\.?ΕΙ\.?Δ\.)|(Α\.?Σ\.?Ε
\.?Π\.)|(Α\.?Ε-
Ε\.?Π\.?Ε\.)|(Δ\.?Δ\.?Σ\.)|(Ο\.?Π\.?Κ\.)|(Υ\.?Ο\.?Δ\.?Δ\.)|(Α\.?Α
\.?Π\.)?)/ui
5. /(?<fek>(Φ\.?Ε\.?Κ\.)?~? ?(?<arithmos>[0-9]+)\/? *(τ\.)?
*(?<tipos>[ΑΒΒΓΔ]´?´?´?´|Δ\.?Ε\.?Β\.?Ι\.)|(Α\.?ΕΙ\.?Δ\.)|(Α\.?Σ\.?Ε
\.?Π\.)|(Α\.?Ε-
Ε\.?Π\.?Ε\.)|(Δ\.?Δ\.?Σ\.)|(Ο\.?Π\.?Κ\.)|(Υ\.?Ο\.?Δ\.?Δ\.)|(Α\.?Α
\.?Π\.)?) *(τη\.) *(?<imerominia>[0-9]{1,2}[.---][0-9]{1,2}[.---][0-
9]{2,4}|[0-9]{2,4})/ui

Προεδρικό διάταγμα:

1. /(?<pd>Π\.?Δ\.?[.-]*(?<arithmos>[0-9]*)\/(?<imerominia>[0-
9]{1,2}[.---][0-9]{1,2}[.---][0-9]{2,4}|[0-9]{2,4})/ui
2. /(?<pd>((?<arithmos>[0-9]*)\/)?(?<imerominia>[0-9]{1,2}[.---][0-
9]{1,2}[.---][0-9]{2,4}|[0-9]{2,4}) *Π\.?Δ\.?(\./τος)?/ui
3. /(?<pd>Π\.?Δ\.?(\./τος)?((?<arithmos>[0-9]*)\/)?(.{0,4}
)?(?<imerominia>[0-9]{1,2}[.---][0-9]{1,2}[.---][0-9]{2,4}|[0-
9]{2,4})/ui

Νόμος:

1. / (?<nomos>(Α\.)?Ν\.?([όο]μο[ιυ]|Δ\.?(\./τος)?[.-]*(?<arithmos>[0-
9]+)\/(?<imerominia>[0-9]{1,2}[.---][0-9]{1,2}[.---][0-9]{2,4}|[0-
9]{2,4})/ui

Υπουργική απόφαση:

1. /(?<apofasi>(Υ\.?Α\.?|Αποφ)[.-]*(?<kw dikos>(.{0,4}[0-9]+[\/
]?)*)\/(?<imerominia>[0-9]{1,2}[.---][0-9]{1,2}[.---][0-9]{2,4}|[0-
9]{2,4})/ui

- 2. αμμομησικές εισακές περιημένων μριου, του μμ-3997/24
- 3. του Ν-2545/97
- 4. της περιπτ.α της παρ.2 αρθ-3
- 5. του μμ-159/97 περτ
- 6. του ελέγχου κατά το Αρθ-12
- 7. Οι δραστηριότητες που αναφέρονται στα Αρθ-3 και Αρθ-4 του Ν-1650/86 (ΦΕΚ-160/Α/86)
- 8. του ΠΔ/5-12-79 (ΦΕΚ-707/Α/79)
- 9. όπως τροποποιήθηκε με το ΠΔ/31-3-87 (ΦΕΚ-303/Α/87)
- 10. των Αρθ-5 και Αρθ-6 του ΠΔ/23-2-87 (ΦΕΚ-166/Α/87)
- 11: προσαρτώνται στο Ν-3105/03 (ΦΕΚ-29/Α/03)
- 12: κατά την παρ.δ.α του Αρθ-41
- 13: στο Ν-3190/03 (ΦΕΚ-249/Α/03) κατά την παρ.7 του Αρθ-20 αυτού.
- 14: περιγράφονται στις διατάξεις της παρ.1 του Αρθ-2 με στοιχεία α, β, γ, δ και ε.
- 15: που υπάρχουν στις διατάξεις του Ν-1559/85 (ΦΕΚ-135/Α/85)
- 16: έχουν αντικατασταθεί από τις διατάξεις του Ν-2244/94 (ΦΕΚ-168/Α/94)
- 17: σύμφωνα με το Ν-210/73 (ΦΕΚ-277/Α/73) και το Ν-2115/93 (ΦΕΚ-15/Α/93)
- 18: σύμφωνα με το Αρθ-39 της Υγ.Δίς.ΑΙβ/8577/83 (ΦΕΚ-526/Β/83)
- 19: επιτρέπεται και από τις διατάξεις του Αρθ-4 του Ν-3741/29 (ΦΕΚ-4/Α/29)
- 20: (Ε.Σ.Βλέπε Αποφ-Φ15/7816/616/05 ΦΕΚ-542/Β/22-4-05
- 21: της παρ.1 του Αρθ-5
- 22: σύμφωνα με το Ν-2773/99 (ΦΕΚ-286/Α/99)
- 23: στις διατάξεις του παρόντος νόμου και του Αρθ-41 (περπτ.ιθ και περπτ.κ) του Ν-2218/94 (ΦΕΚ-90/Α/94)
- 24: κυκλοποιήθηκαν με το ΠΔ-410/95 (ΦΕΚ-231/Α/95 (ε.σ αρθ-24))
- 25: στις διατάξεις του παρόντος νόμου και του Αρθ 52 της Υγ.Δίς.ΑΙβ/8577/83 (ΦΕΚ-526/Β/83)
- 26: της περπτ.β της παρ.1 του Αρθ-2,
- 27: στις εξαιρέσεις του Αρθ-3
- 28: που αναφέρεται στην παρ.5 του άρθρου αυτού.
- 29: που υπερβαίνουν τα όρια της παρ.28 του Αρθ-1 της κοινης υπουργικής Αποφ-3329/15-2-89 (ΦΕΚ-132/Β/89)
- 30: των δραστηριοτήτων της παρ.1
- 31: εφαρμόζονται οι διατάξεις της παρ.1 του Αρθ-24.
- 32: (Ε.Σ.Βλέπε Αποφ-Φ15/7816/617/4-4-05 ΦΕΚ-542/Β/22-4-05 Προθεσμίες
- 33: που τίθενται στην παρ.2β του Αρθ-6.
- 34: μονάδες που αναφέρονται στην παρ.2 του άρθρου αυτού.
- 35: βελτίωση της παρ.8α του Αρθ-4.
- 36: (Ε.Σ.Βλέπε Αποφ-Φ15/7816/616/05 ΦΕΚ-542/Β/22-4-05 Απαλλαγή από άδεια
- 37: βελτισίσεις του Ν-3325/05)
- 38: σύμφωνα με την κοινή υπουργική Αποφ-ΗΗ/13727/724/5-8-03 (ΦΕΚ-1087/Β/03)
- 39: προφύλακτων της 16-8-23, σύμφωνα με το ΠΔ/2-3-81 (ΦΕΚ-138/Α/81)
- 40: απόσπληκν της παρ.1γ του Αρθ-2
- 41: πλην των περπτ.ββ, περπτ.δδ
- 42: αεριών και περπτ.εε.
- 43: Οι διατάξεις του Ν-1650/86 όπως
- 44: ορίζονται στην υπουργική απόφαση της παρ.3.
- 45: εφαρμόζονται οι διατάξεις του παρ.3 και παρ.4 του Αρθ-4.
- 46: (Ε.Σ.Βλέπε Αποφ-Φ15/7816/615/05 ΦΕΚ-542/Β/22-4-05
- 47: εγκαταστάσεων του Ν-3325/05)
- 48: (Ε.Σ.Βλέπε Αποφ-Φ15/8101/624/05 ΦΕΚ-552/Β/25-4-05 Διαδικασία

```
Show only (7)? : [^.*]
Test regular expression (7)? : [^<feko>(0|1|2|3|4|5|6|7|8|9|10|11|12|13|14|15|16|17|18|19|20|21|22|23|24|25|26|27|28|29|30|31|32|33|34|35|36|37|38|39|40|41|42|43|44|45|46|47|48)]

(/?<pd>Π?7Δ(.?)?/ro<?7(0-9)(0,5) etos ... [ ]
(/?<feko>(ΦΕΚ)?7(0-9)-/v fek tp etos ... [ ]
/?<nomos>- ar nomou etos ... [ ]
(/?<apofasi>(Υ.Δ.)?/Αποφ/ [^?<kwidikos>... [ ]
(/?<arthros>/Αρθ<?7(0-9)-> [ ]
(/?<ar nomou>N-(0-9)-> [ ]
(/?<ar nomou>N<?7(0-9)-> [ ]
(/?<basiliko diatagma>B.7Δ.7.../??<kwidikos>... [ ]
(/?<diataksi>?<?<tipos>-(0-9) [^?<kwidikos>... [ ]
(/?<cedafio>ε&[α|β|γ|δ|ε|στ|ζ|η|θ|ι|κ|λ|μ|ν|ξ|ο|π|ρ|ς|σ|τ|υ|φ|χ|ψ|ω|... [ ]
(/?<etos>-(0-9)(2-4)) [ ]
(/?<etos>-(0-9)(1,2)-[0-9](1,2)-[0-9](2,3)) [ ]
(/?<feko>(Φ.7)7K.7?7?<arithmos>-(0-9)) [ ]
(/?<feko>?<?<imerominia>-(0-9)(1,2) [^-[0-9](1,2) [ ]
(/?<feko>?<?<imerominia>-(0-9)(1,2) [^-[0-9](1,2) [ ]
(/?<feko>(Φ.7)7K.7?7?<arithmos>-(0-9)) [ ]
(/?<feko>(Φ.7)7K.7?7?<arithmos>-(0-9) [ ]
(/?<feko>(Α.7)?N.7(0-9) [ ]
(/?<paragrafos>... [ ]
(/?<paragrafos>... [ ]
(/?<paragrafos>... [ ]
(/?<pd>Π.7Δ.7(7ro<?7(0-9) [ ]
(/?<pd>Π.7Δ.7(7ro<?7(0-9) [ ]
(/?<periptwsi>peript. [ ])
```

Εικόνα 2: Πλατφόρμα για δοκιμή κανονικών εκφράσεων, με οπτική υποβοήθηση αποτελεσμάτων και καταγραφή σε βάση δεδομένων

Βασιλικό διάταγμα:

- 1. / (?<basiliko diatagma>B\ .?Δ\ .? [\-]) * ((?<kwidikos> (. { 0 , 4 } [0 - 9] + [\ /] ?) *) \ * / * ((?<imerominia> [0 - 9] { 1 , 2 } [. -] [0 - 9] { 1 , 2 } [. -] [0 - 9] { 2 , 4 }) | [0 - 9] { 2 , 4 }) * , ? * (Αρθ \ . [? \ -] * ((?<arthro> [0 - 9] +)) ?) / ui

Λόγω της μεγάλης πολυπλοκότητας των κανόνων, χρησιμοποιήθηκαν κανονικές εκφράσεις με παραγωγή επίσημείωσης (Named Capturing Groups) ούτως ώστε, εφόσον ο κανόνας ενεργοποιείται, να γίνεται αυτόματος διαχωρισμός των επί μέρους στοιχείων παράλληλα με τον αλγόριθμο ανάλυσης.

Κατασκευάστηκαν, επίσης, κανόνες για την αναγνώριση παραπομπών μέσα στο τρέχον νομικό κείμενο, όπως για παράδειγμα “...ρυθμίζονται, αναλυτικά, τα θέματα διάρθρωσης, στελέχωσης, αρμοδιοτήτων που αναφέρονται στις Διευθύνσεις Ανάπτυξης των ΝΑ, καθώς και τα θέματα υλικοτεχνικής υποδομής του ελέγχου κατά το Αρθ-12 και κάθε άλλη αναγκαία λεπτομέρεια...”, για την μεγαλύτερη πληρότητα των περιπτώσεων μελέτης.

4 Αποτελέσματα

Διάφορα προβλήματα εμφανίστηκαν κατά τη διαδικασία μελέτης και κατασκευής των κανονικών εκφράσεων. Η λίστα που ακολουθεί καταγραφεί τα πιο χαρακτηριστικά παραδείγματα που παρουσιάστηκαν σε μεγαλύτερο βαθμό, χωρίς να έχει κάποια σημασία η σειρά με την οποία έχουν καταγραφεί, κυρίως σε αναφορές ΦΕΚ, και δευτερευόντως σε άλλα νομικά κείμενα.

- *Ελλιπής ή λανθασμένη γραφή αναφοράς ΦΕΚ:* Σε πολλές περιπτώσεις δεν έχει σημειωθεί ο τόμος του φύλλου της κυβερνήσεως, παράδειγμα “ΦΕΚ 26/9–2–2007”, ενώ σε άλλες περιπτώσεις έχει γίνει λάθος συμπλήρωση, όπως για παράδειγμα “ΦΕΚ-93Α/Α/61” (επανάληψη τόμου) και “16/17–3–1950 (Α΄ 82)” (εννοεί ότι αναφέρεται σε ΦΕΚ αλλά αγνοεί το αντίστοιχο ακρωνύμιο). Επίσης, σε μεγάλο βαθμό δεν γίνεται αναφορά του χρόνου/ημερομηνίας, αλλά εννοείται από τον αντίστοιχο νόμο ή προεδρικό διάταγμα, όπως για παράδειγμα “π.δ. 33/2009 (ΦΕΚ Α΄ 50)”, χωρίς απαραίτητα να υπάρχει σαφής λύση στο συγκεκριμένο πρόβλημα λόγω ενδιάμεσων κειμένων, όπως για παράδειγμα “...του ν. 3852/2010 «Νέα Αρχιτεκτονική της Αυτοδιοίκησης και της Αποκεντρωμένης Διοίκησης -Πρόγραμμα Καλλικράτης.» (ΦΕΚ 87 Α΄)...”. Σε κάθε περίπτωση, έγινε καταγραφή ακόμα και περιπτώσεων που είναι εντελώς ξένες από οτιδήποτε έχει γίνει σύνηθες στα νομικά κείμενα, όπως για παράδειγμα “π.7001/4544/14.10.77 ΚΟΙΝΗΣ ΑΠΟΦΑΣΕΩΣ ΤΩΝ ΥΠΟΥΡΓΩΝ... (φεκ 418.1.11.77, Τ. Δ΄)” (πιθανόν εννοεί ΦΕΚ Δ/418/1-11-77).
- *Ελλιπής αναφορά πρωτοκόλλου προεδρικού διατάγματος:* Ουκ ολίγες φορές, γίνεται αναφορά σε προεδρικό διάταγμα του οποίου χαρακτηριστικό στοιχείο αποτελεί μόνο η ημερομηνία, όπως για παράδειγμα “ΠΔ/23-2-87”.
- *Λανθασμένη γραφή ή χρήση συμβόλων:* Επανειλημμένα, έγινε καταγραφή αναφοράς σε ΦΕΚ με χρήση του αγγλικού χαρακτήρα Β (κωδικός 66) στη θέση του ελληνικού χαρακτήρα Β (κωδικός 914), γεγονός το οποίο δεν επηρεάζει την ανάγνωση από τον οποιονδήποτε που μελετά το νομικό κείμενο, αλλά μπορεί να καταστρέψει σε μεγάλο βαθμό την εξαγωγή αποτελεσμάτων από αυτοματοποιημένο σύστημα αναγνώρισης συγκεκριμένων κωδικών, αν δεν μπορεί να παρακάμψει τέτοιου τύπου προβλήματα. Επίσης, υπάρχει ασάφεια ως προς το κατάλληλο σύμβολο χρήσης αποστρόφου. Συγκεκριμένα, παρατηρήθηκε χρήση των συμβόλων ['] (κωδικός 900), ['] (κωδικός 8217) και ['] (κωδικός 39).

- *Διαφορετική σειρά των επί μέρους στοιχείων:* Αν κωδικοποιήσουμε τα επί μέρους στοιχεία αναφοράς σε ΦΕΚ με τις εξής κωδικές ονομασίες {ΤΕΥΧΟΣ}, {ΗΜΕΡΟΜΗΝΙΑ}, {ΑΡΙΘΜΟΣ} χωρίς να μας απασχολήσει πως ακριβώς έχει καταγραφεί κάθε στοιχείο ή ο χαρακτήρας σύνδεσής τους (δλδ. κάθετος, τελεία, κόμμα, πλήθος κενών), η μελέτη έδειξε τις παρακάτω παραλλαγές:

1. ΦΕΚ {ΑΡΙΘΜΟΣ} {ΤΕΥΧΟΣ} {ΗΜΕΡΟΜΗΝΙΑ}
2. ΝΟΜΟΣ {ΗΜΕΡΟΜΗΝΙΑ} ΦΕΚ {ΑΡΙΘΜΟΣ} {ΤΕΥΧΟΣ}
3. ΝΟΜΟΣ {ΗΜΕΡΟΜΗΝΙΑ} ΦΕΚ {ΤΕΥΧΟΣ} {ΑΡΙΘΜΟΣ}
4. ΦΕΚ {ΑΡΙΘΜΟΣ} {ΗΜΕΡΟΜΗΝΙΑ} {ΤΕΥΧΟΣ}
5. ΦΕΚ {ΑΡΙΘΜΟΣ} {ΤΕΥΧΟΣ} {ΗΜΕΡΟΜΗΝΙΑ}

Σε αυτό το σημείο, πρέπει να σημειωθεί το γεγονός ότι το πλήθος των παραπάνω παραλλαγών ταυτίζεται με το πλήθος των κανονικών εκφράσεων που κατασκευάσαμε για την αναγνώριση παραπομπών σε ΦΕΚ (βλέπε κεφάλαιο 3, Μεθοδολογία), καθώς δεν κατέστη δυνατή η δημιουργία μοναδικής κανονικής έκφρασης που να μπορεί να καλύψει παράλληλα όλες τις παραπάνω παραλλαγές.

- *Έλλειψη αναφοράς έτους:* Στις περιπτώσεις οι οποίες δεν γίνεται αναφορά έτους σε παραπομπές ΦΕΚ, γίνεται προσπάθεια να αναγνωρισθεί το έτος από την εμφάνιση νόμου ακριβώς πριν (βλέπε κανόνες 2 και 3). Παρόλαυτα, υπάρχουν περιπτώσεις οι οποίες ακόμα και έτσι, υπάρχει αδυναμία πλήρους ταυτοποίησης του ΦΕΚ, όπως για παράδειγμα “...του ν. 3852/2010 «*Νέα Αρχιτεκτονική της Αυτοδιοίκησης και της Αποκεντρωμένης Διοίκησης -Πρόγραμμα Καλλικράτης.*» (ΦΕΚ 87 Α’)...”.
- *Αδυναμία αποκωδικοποίησης πληροφορίας:* Η διαφοροποίηση μηχανογράφησης κάθε υπηρεσίες του κράτους, ή ακόμα και του ίδιου υπουργείου, αναγκάζει και σε διαφορετικό τρόπο πρωτοκόλλησης και αρχειοθέτησης όλων εγγράφων που διαχειρίζεται. Ενδεικτικό παράδειγμα αποτελεί η εξής αναφορά “...*Με την υπ’ αριθμ. ΥΠΑΙΘΠΑ/ΓΔΔΥ/ΔΙΟΙΚ/ΤΔΠΙΔ/132904/12956/11036/3829/27-12-2012 απόφαση του Υπουργού Παιδείας...*” .
- *Λανθασμένη αρίθμηση:* Επίσης, παρατηρήθηκαν αναφορές στην ίδια πηγή αλλά με διαφορετικό σύμβολο αρίθμηση. Για παράδειγμα, στην αναφορά “*Υγ_Δξς-ΑΙβ/8577/83*” ο κωδικός πρωτοκόλλου είναι *ΑΙβ* εν’ αντιθέσει με “*Υγ_Δξς-Α1β/8577/83*” που ο κωδικός πρωτοκόλλου είναι *Α1β*.

Τέλος, με τυχαία επιλογή ΦΕΚ, βάλαμε σε πλήρη λειτουργία του κανόνες και έγινε εκ νέου καταμέτρηση των αποτελεσμάτων. Συγκεκριμένα, έγινε αναγνώριση 33% αναφορών σε νόμους (αναγνωρίστηκαν 3 από σύνολο 9), 66% σε ΦΕΚ (10 από 15), 92% σε ΠΔ (11 από 12) ενώ καμία αναγνώριση δεν επιτεύχθηκε από τις 40 σε άλλου είδους έγγραφα (όπως αποφάσεις συλλόγων, αναφορές στο ίδιο ΦΕΚ, κ.α.)

5 Συμπεράσματα

Στη συγκεκριμένη μελέτη έγινε προσπάθεια περιγραφής της υπάρχουσας κατάστασης που επικρατεί στα νομικά κείμενα. Είναι σαφές, πως η έλλειψη προτυποποίησης όσον αφορά τον τρόπο γραφής από όλα τα εμπλεκόμενα μέρη (νομοθέτης, υπηρεσίες) οι οποίες συγγράφουν κείμενα νομικού περιεχόμενου, δεν ακολουθούν κάποια συγκεκριμένη μεθοδολογία, γεγονός το οποίο οδηγεί σε δυσκολία αυτόματης χαρτογράφησης της πληροφορίας στο σύνολό της. Τα αποτελέσματα, αν και υπάρχουν μεγάλα περιθώρια βελτίωσης, μας δίνουν θετικές ενδείξεις ούτως ώστε να συνεχίσουμε στην ίδια κατεύθυνση και την περαιτέρω βελτιστοποίηση του συστήματος, χωρίς να παραμερίζεται το γεγονός κατάλληλης τυποποίησης με παράλληλη ενημέρωση και κατάρτιση του κοινού.

6 Υποστήριξη

Η συγκεκριμένη εργασία υποστηρίχθηκε υπό την αιγίδα της Επιτροπής Ερευνών του Πανεπιστημίου Πατρών, στα πλαίσια της βασικής έρευνας «Καραθεοδωρής».

7 Βιβλιογραφία

- [1] Beatrice A. Tice, *Foreign Official Gazettes: Solving a Collection conundrum*, 97 *Law Library Journal* 299 (2005)
- [2] Ελληνικός οργανισμός τυποποίησης, <http://www.elot.gr/> (Τελευταία προσπέλαση Ιούνιος 2013)
- [3] Οδηγίες γραφής αναφορών σε κυβερνητικά έγγραφα σύμφωνα με η Αμερικανική Ψυχολογική Εταιρεία (American Psychological Association, APA), http://www.unk.edu/uploadedFiles/academics/library/gov_doc/about/Citing%20Government%20Documents%20-%20APA%20Style%202010.pdf, (Τελευταία προσπέλαση Ιούνιος 2013)
- [4] Οδηγίες γραφή αναφορών σε κυβερνητικά έγγραφα σύμφωνα με το πανεπιστήμιο του Michigan <http://apu.libguides.com/content.php?pid=54832&sid=401524>, (Τελευταία προσπέλαση Ιούνιος 2013)

- [5] Citing Government Information Sources Using MLA University of Nevada, Reno Libraries, http://knowledgecenter.unr.edu/help/manage/government_cite.aspx (Τελευταία προσπέλαση Ιούνιος 2013)
- [6] Citation Guide: U.S. Congressional Documents Library of Congress, <http://memory.loc.gov/ammem/amlaw/lwcite.html> (Τελευταία προσπέλαση Ιούνιος 2013)

Χαράλαμπος Τσιμπούρης

Υπ. Διδάκτορας, Τμήμα Ηλεκτρολόγων Μηχανικών, Πανεπιστήμιο Πατρών
Τμ. Ηλ.Μηχ & Τεχν. Υπολ, Εργαστήριο Ενσύρματης Τηλεπικοινωνίας
Ηλ-ταχ.: xtsimpouris@upatras.gr

Δρ. Μηχ. Κυριάκος Σγάρμπας

Επ. Καθηγητής, Τμήμα Ηλεκτρολόγων Μηχανικών, Πανεπιστήμιο Πατρών
Τμ. Ηλ.Μηχ & Τεχν. Υπολ, Εργαστήριο Ενσύρματης Τηλεπικοινωνίας
Ηλ-ταχ.: sgarbas@upatras.gr