ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
HAROKOPIO UNIVERSITY
.dit/*
Informatics & Telematics

ΕΕΛΛΑΚ
Οργανισμός Ανοιχτών Τεχνολογιών

Αυτόματη εξαγωγή αρμοδιοτήτων και δομής οργανισμών από την Εφημερίδα της Κυβέρνησης

Google Summer of Code

Ηρακλής Βαρλάμης, varlamis@hua.gr
Χαροκόπειο Πανεπιστήμιο: Επίκουρος Καθηγητής, Τμ. Πληροφορικής και Τηλεματικής
ΕΕΛΛΑΚ: Μέλος του ΔΣ, υπεύθυνος για θέματα Ανοιχτής Διακυβέρνησης και Ανοικτών Δεδομένων

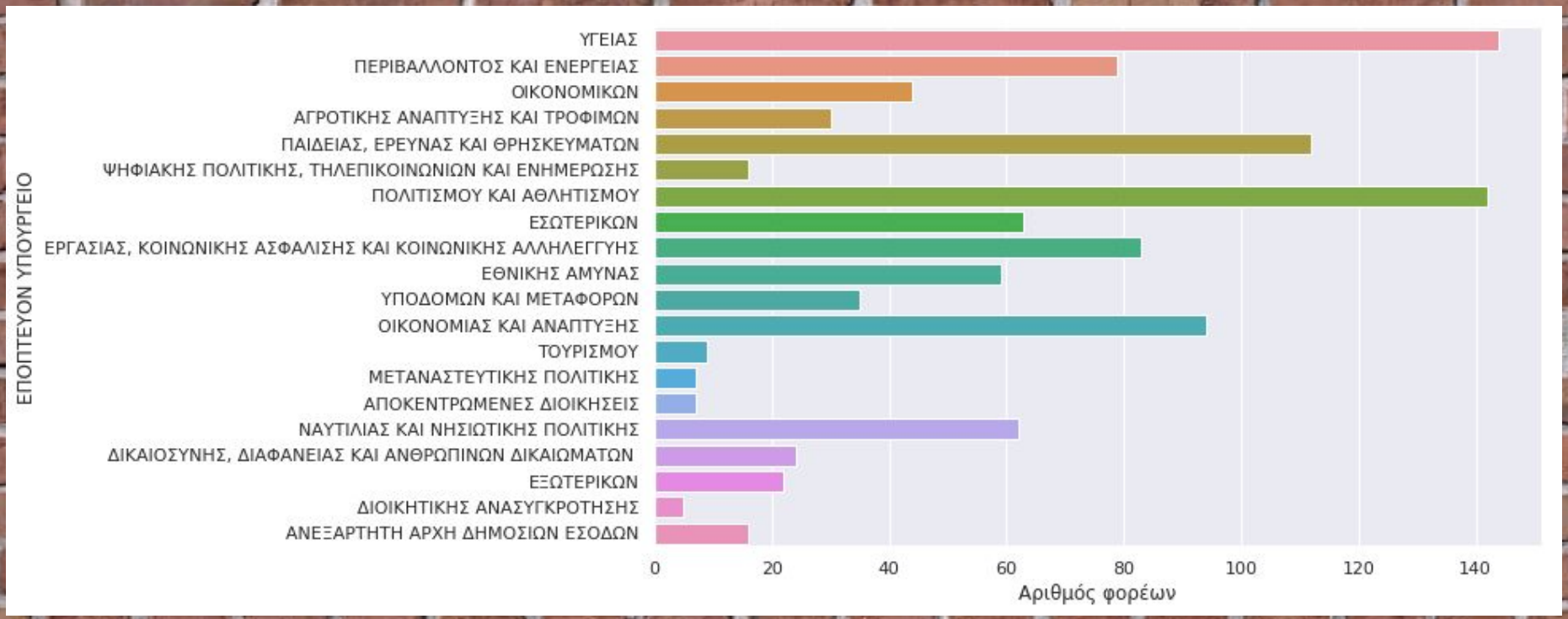
Το πρόβλημα

- Η χαρτογράφηση υπηρεσιών και αρμοδιοτήτων των φορέων της Δημόσιας Διοίκησης

- ΔΙΑΥΓΕΙΑ: ~ 3500 φορείς
- Μητρώο Υπηρεσιών και Φορέων της Ελληνικής Διοίκησης:
  - 5 Ανεξάρτητες Δημόσιες Υπηρεσίες
  - 30 Ανεξάρτητες Αρχές
  - 7 Αποκεντρωμένες Διοικήσεις
  - 348 Οργανισμοί Τοπικής Αυτοδιοίκησης (325 Δήμοι και 13 Περιφέρειες)
  - 18 Υπουργεία στους οποίους υπάγονται άλλοι 1054 φορείς

- Ένα σύνθετο τοπίο υπηρεσιών και αρμοδιοτήτων που δεν αποτυπώνεται με έναν ενιαίο τρόπο στα σχετικά Φ.Ε.Κ.
- Δεν υπάρχουν κάπου συγκεντρωμένα τα σχετικά Φ.Ε.Κ.

Φορείς ανά υπουργείο



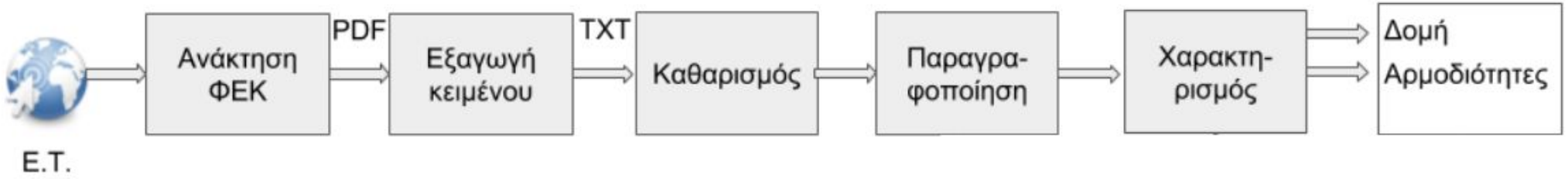
Η ευκαιρία

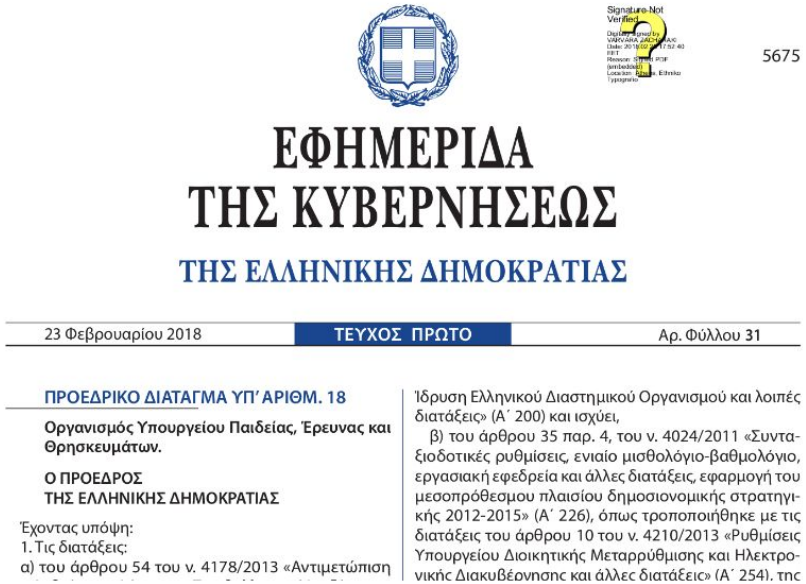
- Google Summer of Code (2018)

- Ανάπτυξη λογισμικού ανοικτού κώδικα για την αυτοματοποίηση της διαδικασίας
  - Εντοπισμού των σχετικών Φ.Ε.Κ.
  - Εξαγωγής της δομής των φορέων (υπηρεσίες)
  - Εξαγωγής των αρμοδιοτήτων τους

- https://github.com/eellak/gsoc2018-GG-extraction
  Προγραμματιστής: ***Chris Karageorgiou Kaneen***
  Μέντορες: Η. Βαρλάμης, Σ. Καπιδάκης, Θ. Καρούνος

Η διαδικασία





5675

ΕΦΗΜΕΡΙΔΑ
ΤΗΣ ΚΥΒΕΡΝΗΣΕΩΣ

ΤΗΣ ΕΛΛΗΝΙΚΗΣ ΔΗΜΟΚΡΑΤΙΑΣ

23 Φεβρουαρίου 2018 | ΤΕΥΧΟΣ ΠΡΩΤΟ | Αρ. Φύλλου 31

ΠΡΟΕΔΡΙΚΟ ΔΙΑΤΑΓΜΑ ΥΠ' ΑΡΙΘΜ. 18

Οργανισμός Υπουργείου Παιδείας, Έρευνας και Θρησκευμάτων.

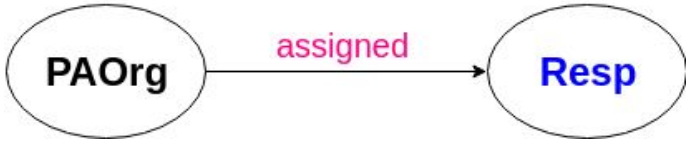
Ο ΠΡΟΕΔΡΟΣ
ΤΗΣ ΕΛΛΗΝΙΚΗΣ ΔΗΜΟΚΡΑΤΙΑΣ

Έχοντας υπόψη:
1. Τις διατάξεις:
α) του άρθρου 54 του ν. 4178/2013 «Αντιμετώπιση

Ίδρυση Ελληνικού Διαστημικού Οργανισμού και λοιπές διατάξεις» (Α΄ 200) και ισχύει,
β) του άρθρου 35 παρ. 4, του ν. 4024/2011 «Συνταξιοδοτικές ρυθμίσεις, ενιαίο μισθολόγιο-βαθμολόγιο, εργασιακή εφεδρεία και άλλες διατάξεις, εφαρμογή του μεσοπρόθεσμου πλαισίου δημοσιονομικής στρατηγικής 2012-2015» (Α΄ 226), όπως τροποποιήθηκε με τις διατάξεις του άρθρου 10 του ν. 4210/2013 «Ρυθμίσεις Υπουργείου Διοικητικής Μεταρρύθμισης και Ηλεκτρονικής Διακυβέρνησης και άλλες διατάξεις» (Α΄ 254), της

RespA

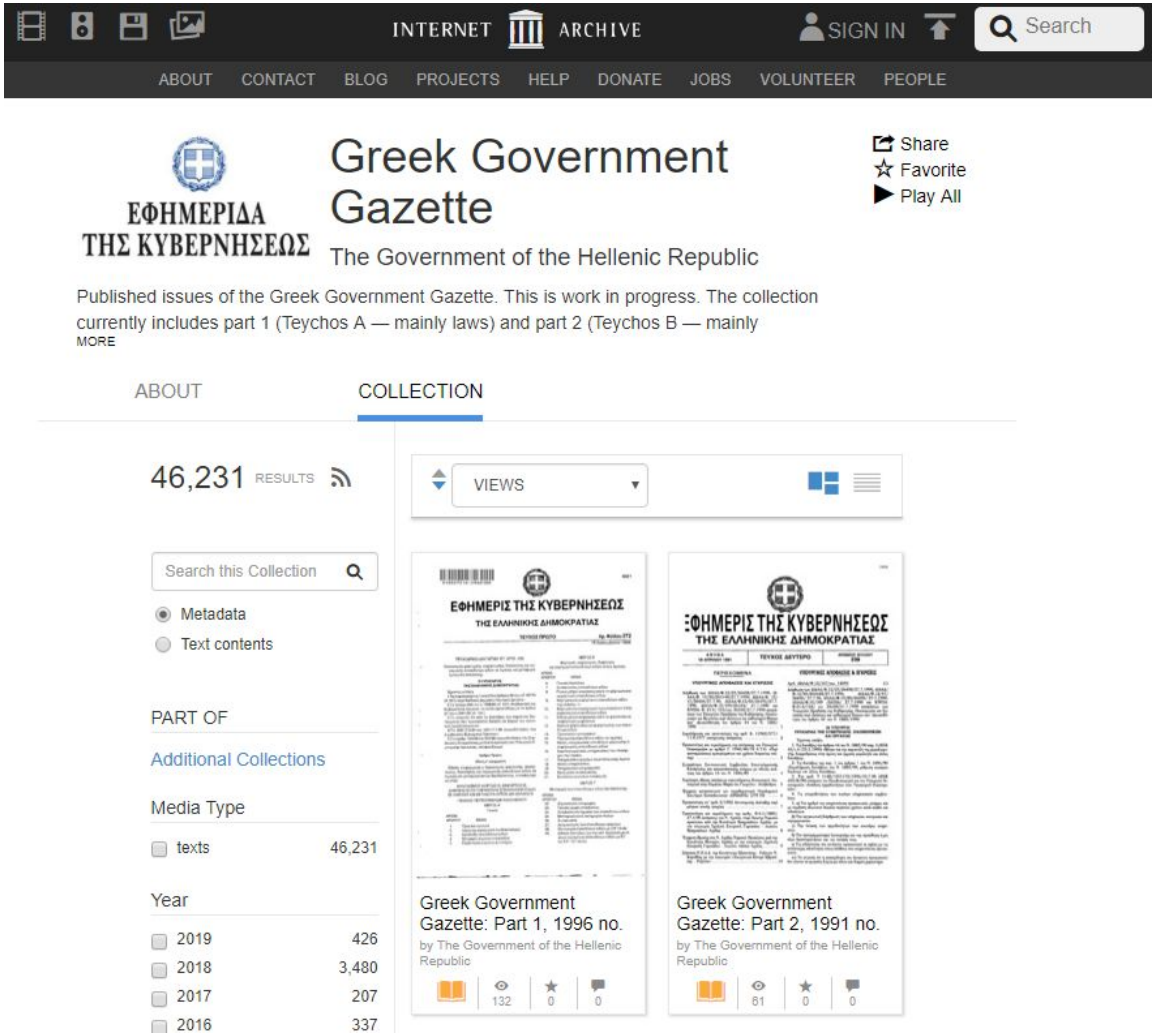
```
{
  "Το Τμήμα Β Παρακολούθησης Προϋπολογισμού και Αναφορών
      "4. Το Τμήμα Β΄ Παρακολούθησης Προϋπολογισμού \nκ
      "α) τη σύνταξη των αποφάσεων έγκρισης και τροποπο
      "β) την υποβολή αιτήματος και σύνταξη αποφάσεων \
      "γ) τις προεγκρίσεις ανάληψης υποχρέωσης σε βάρος
      "δ) τον έλεγχο και την αποστολή στοιχείων των φορ
      "ε) τη μέριμνα για την έγκαιρη κοινοποίηση των εγ
      "στ)  την ενημέρωση του Μητρώου δεσμεύσεων,",
      "ζ) τη βεβαίωση ύπαρξης πίστωσης στον προϋπολογισ
      "η) τη βεβαίωση ύπαρξης πιστώσεων στον προϋπολογι
      "θ) τη σύνταξη των Μνημονίων Συνεργασίας με τους
```



Οι δυσκολίες

- Μεγάλη καθυστέρηση στην ανάκτηση ΦΕΚ από το Ε.Τ.
- Προβλήματα στην ανάκτηση του TXT από ορισμένα PDF
  - αλλοιωμένο κείμενο (κυρίως όπου υπήρχαν πίνακες)
  - θόρυβος (επικεφαλίδες, υποσέλιδα, αριθμοί σελίδων)
- Προβλήματα στην επεξεργασία των πολυεπίπεδων λιστών
- Ετερογένεια
  - Υπάρχει μια αδρή δομή στα ΦΕΚ των περισσοτέρων οργανισμών,
  - Υπάρχουν κάποια βασικά χαρακτηριστικά για τον προσδιορισμό της ανάθεσης αρμοδιοτήτων
  - Δεν υπάρχει καθορισμένο πρότυπο για την περιγραφή ενός οργανισμού

Όλα τα Φ.Ε.Κ. στο Internet Archive



https://archive.org/details/greekgovernmentgazette

Όλα τα ΦΕΚ ανεβαίνουν αυτόματα και έχουν URI

Η δομή του Φ.Ε.Κ.

ΔΟΜΗ → **ΑΡΜΟΔΙΟΤΗΤΕΣ** → **ΣΤΕΛΕΧΩΣΗ**

Αρ. Φύλλου 1

7 Ιανουαρίου 2008

Άρθρο 1
Αποστολή του Ε.Κ.Δ.Δ.Α.

Αποστολή του Εθνικού Κέντρου Δημόσιας Διοίκησης και Αυτοδιοίκησης (Ε.Κ.Δ.Δ.Α.), το οποίο είναι νομικό πρόσωπο δημοσίου δικαίου, υπό την εποπτεία του Υπουργού Εσωτερικών, Δημόσιας Διοίκησης και Αποκέντρωσης (π.δ. 356/1997 - Α΄ 240), αποτελεί ο διαρκής

....

Άρθρο 2
Όργανα Διοίκησης

Το Ε.Κ.Δ.Δ.Α. διοικείται από το Διοικητικό Συμβούλιο και τον Γενικό Γραμματέα του Ε.Κ.Δ.Δ.Α., σύμφωνα με τις εκάστοτε ισχύουσες διατάξεις, οι οποίες αφορούν στην δομή και λειτουργία του Ε.Κ.Δ.Δ.Α.

....

Άρθρο 3
Διάρθρωση Υπηρεσιών

Η διάρθρωση των Υπηρεσιών του Εθνικού Κέντρου Δημόσιας Διοίκησης και Αυτοδιοίκησης (Ε.Κ.Δ.Δ.Α.) είναι η ακόλουθη:

Α. Διοικητικές Υπηρεσίες:
1. Διεύθυνση Προγραμματισμού και Ανάπτυξης.
2. Διεύθυνση Διοικητικού.
3. Διεύθυνση Οικονομικού.
4. Νομική Υπηρεσία.

....

ΚΕΦΑΛΑΙΟ Β΄
ΑΡΜΟΔΙΟΤΗΤΕΣ

Άρθρο 4
Διεύθυνση Προγραμματισμού και Ανάπτυξης

1. Η Διεύθυνση έχει τις αρμοδιότητες των Τμημάτων που τη συγκροτούν, και τα οποία είναι:
α. Τμήμα Σχεδιασμού και Ανάπτυξης.
β. Τμήμα Εφαρμογών Πληροφορικής και Τεχνικής Υποστήριξης.
γ. Τμήμα Τεκμηρίωσης και Βιβλιοθήκης.
δ. Τμήμα Συγχρηματοδοτούμενων Προγραμμάτων.
2. Αρμοδιότητες του Τμήματος Σχεδιασμού και Ανάπτυξης.
Στις αρμοδιότητες του Τμήματος αυτού υπάγονται:
α. Τα θέματα τα σχετικά με τη σύσταση, κατάργηση, μεταφορά, οργάνωση και λειτουργία των υπηρεσιών του Ε.Κ.Δ.Δ.Α.
β. Τα θέματα τα σχετικά με τη μελέτη και την εφαρμογή διοικητικών μεθόδων και διαδικασιών για τη συστηματική και ορθολογική οργάνωση των πάσης φύσεως εργασιών του Ε.Κ.Δ.Δ.Α.
γ. Η μέριμνα για την ανάπτυξη και εφαρμογή συστήματος εξουσιοδοτήσεων για τις διοικητικές και οικονομικές λειτουργίες του Ε.Κ.Δ.Δ.Α.

....

ιι. Η μέριμνα για την αναπαραγωγή του αναγκαίου εκπαιδευτικού υλικού και προετοιμασία και διανομή των φακέλων των επιμορφούμενων.
ιαια. Η προετοιμασία προτάσεων για προμήθειες υλικών και εξοπλισμού και η διαχείριση του υλικού του ΙΝ.ΕΠ. (παραλαβή, φύλαξη, εσωτερική κατανομή, καταγραφή και απογραφή κάθε είδους υλικού).

Άρθρο 15
Κατηγορίες-Κλάδοι-Βαθμοί

1. Οι θέσεις του μόνιμου προσωπικού του Εθνικού Κέντρου Δημόσιας Διοίκησης και Αυτοδιοίκησης (Κεντρική Υπηρεσία και Περιφερειακά Ινστιτούτα Επιμόρφωσης) κατανέμονται κατά κατηγορίες και κλάδους ως εξής:
α. Κατηγορία θέσεων Πανεπιστημιακής Εκπαίδευσης (ΠΕ)
β. Κατηγορία θέσεων Τεχνολογικής Εκπαίδευσης (ΤΕ)
γ. Κατηγορία θέσεων Δευτεροβάθμιας Εκπαίδευσης (ΔΕ)
δ. Κατηγορία θέσεων Υποχρεωτικής Εκπαίδευσης (ΥΕ)
2. Κλάδος Πανεπιστημιακής Εκπαίδευσης (ΠΕ)
α. Κλάδος ΠΕ Διοικητικού-Οικονομικού
β. Κλάδος ΠΕ Πληροφορικής

....

Α. Μεταφορά υφιστάμενων κενών οργανικών θέσεων από υφιστάμενους κλάδους σε άλλους υφιστάμενους ή νέους κλάδους της ίδιας ή άλλης κατηγορίας.
1. Στον Κλάδο ΠΕ Διοικητικού-Οικονομικού μεταφέρονται οι εξής θέσεις:
α. Δυο (2) κενές θέσεις του κλάδου ΥΕ Βοηθητικού Προσωπικού (Κεντρική Υπηρεσία) μεταφέρονται στον υφιστάμενο κλάδο ΠΕ Διοικητικού-Οικονομικού (Κεντρική Υπηρεσία).
β. Μια (1) κενή θέση του κλάδου ΥΕ Προσωπικού Καθαριότητος (ΠΙΝΕΠ Νοτίου Αιγαίου) μεταφέρεται στον υφιστάμενο κλάδο ΠΕ Διοικητικού - Οικονομικού (Κεντρική Υπηρεσία).

Τα χαρακτηριστικά

- Χρήση αρίθμησης της παραγράφου
  - άλλοτε αριθμοί, άλλοτε Ελληνικά γράμματα, άλλοτε παρένθεση, άλλοτε τελεία και συχνά εναλλαγές αυτών
  - σπανιότερα, αλλαγή της αρίθμησης εντός της παραγράφου

- Χρήση συγκεκριμένων λέξεων (ή ζευγών) κυρίως στην αρχή της παραγράφου
  - π.χ. αναθέτει, μεριμνά, επιβλέπει, Τμήμα, Υπηρεσία, Διεύθυνση, Σχεδιασμού και Ανάπτυξης, Οικονομικού, θέσεις ΠΕ, θέση ΤΕ

- Συγκέντρωση παραγράφων ανάθεσης αρμοδιοτήτων εντός του ίδιου άρθρου και σε συνεχόμενες θέσεις

Η επέκταση

- ΕΕΛΛΑΚ 2019
- Επέκταση του λογισμικού ανοικτού κώδικα για τον αυτόματο χαρακτηρισμό ΦΕΚ οργανισμών σε επίπεδο παραγράφου με χρήση μηχανικής μάθησης και η ανάπτυξη μηχανισμών για τη διαρκή εκπαίδευση των μοντέλων
- https://github.com/kontopoulos/gsoc2018-GG-extraction
  Προγραμματιστής: ***Γιάννης Κοντόπουλος***
  Μέντορες: Η. Βαρλάμης, Θ. Καρούνος

Βελτιώσεις

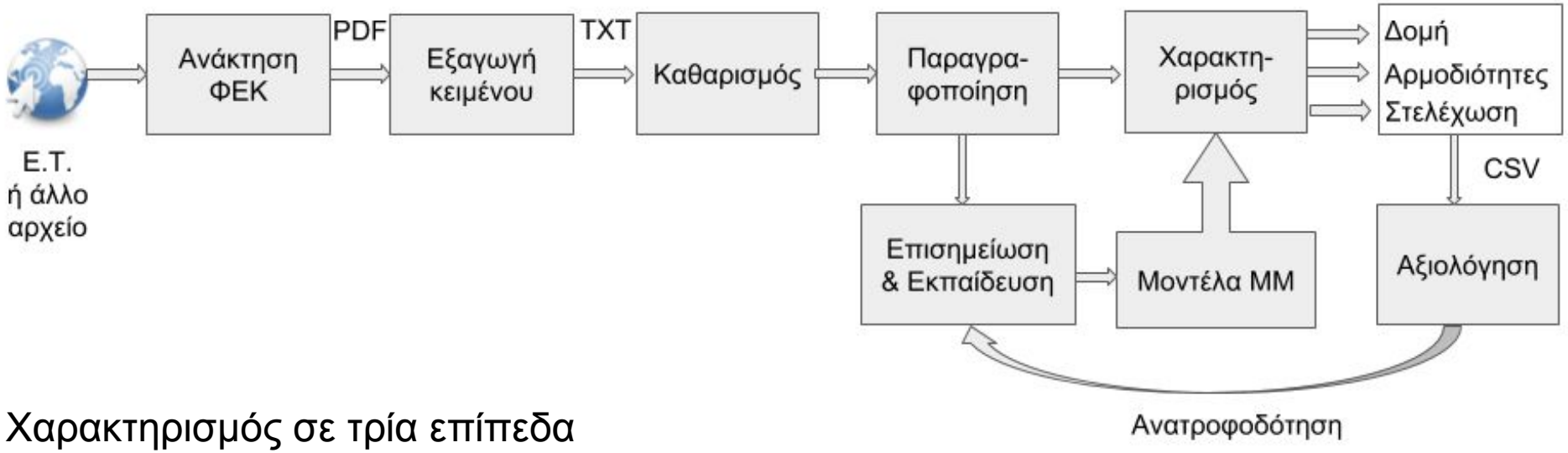
- Βελτίωση της προεπεξεργασίας του κειμένου
  - αφαίρεση stopwords
  - αποκοπή καταλήξεων (stemming)

- Αυτόματη δημιουργία λεξικών (trie indices) με λέξεις ή ζεύγη λέξεων που εμφανίζονται με μεγαλύτερη συχνότητα στην αρχή παραγράφων ενδιαφέροντος απ’ ότι συνολικά, σε δεδομένα εκπαίδευσης.

- Αυτόματη εξαγωγή γνωρισμάτων μορφοποίησης, χρήσης λεξιλογίου, και εκπαίδευση ταξινομητών

- Εξαγωγή χαρακτηρισμών σε CSV αρχεία, δυνατότητα χειροκίνητης διόρθωσης και επανεκπαίδευσης

Κύκλος εκπαίδευσης και χαρακτηρισμού



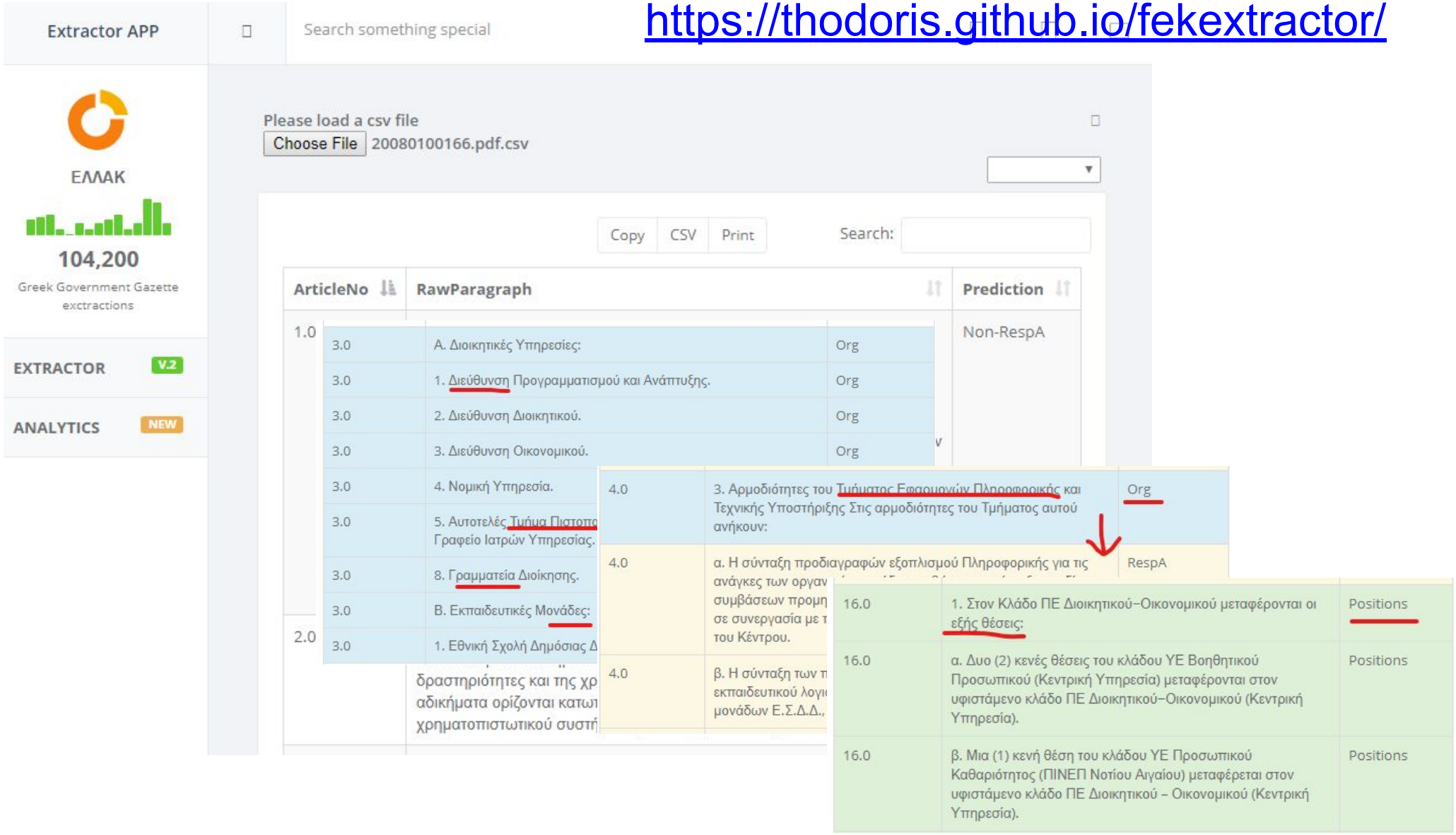
- Χαρακτηρισμός σε τρία επίπεδα
  - Επίπεδο ΦΕΚ: ΦΕΚ οργανισμού ή άσχετο
  - Επίπεδο άρθρου: Άρθρο περιγραφής της δομής του φορέα, ανάθεσης αρμοδιότητας ή περιγραφής της στελέχωσης
  - Επίπεδο παραγράφου: Παράγραφος που ονομάζει μια υπηρεσία του φορέα (Org), παράγραφος που αναθέτει αρμοδιότητες σε υπηρεσία (RespA), παράγραφος που περιγράφει μια ή περισσότερες θέσεις (Position)

Το αποτέλεσμα

- Αναλύθηκαν 492 Φ.Ε.Κ. από το 2008 μέχρι το 2018
- Τα 217 Φ.Ε.Κ που χαρακτηρίστηκαν ως Φ.Ε.Κ. οργανισμών, περιλαμβάνουν
  - ΦΕΚ ίδρυσης οργανισμών
  - ΦΕΚ υπουργείων
  - Τροποποιητικά ΦΕΚ
  - Εφαρμοστικά ΦΕΚ Ευρωπαϊκών οδηγιών
  - ΦΕΚ συγχώνευσης οργανισμών
- Στο σύνολό τους περιλαμβάνουν ανάθεση αρμοδιοτήτων ή αναδιανομή αρμοδιοτήτων σε υπηρεσίες δημοσίων οργανισμών

Επίβλεψη και διόρθωση

https://thodoris.github.io/fekextractor/



Επόμενα βήματα

- Ανάπτυξη ενός πιο ολοκληρωμένου γραφικού περιβάλλοντος (GUI) επισημείωσης και διόρθωσης στα επεξεργασμένα ΦΕΚ
- Αυτόματος εντοπισμός και εξαγωγή επώνυμων οντοτήτων στα ΦΕΚ (Named Entity Recognition), σε επίπεδο διεύθυνσης, τμήματος, ρόλου
- Σύνδεση με το API της Απογραφής (https://hr.apografi.gov.gr/api.html)
- Αποτύπωση της εξαγόμενης πληροφορίας σε ένα πιο ευέλικτο σημασιολογικό πλαίσιο (π.χ. RDF)
- Δημιουργία ενός κοινού template για την περιγραφή φορέων και των αρμοδιοτήτων τους

Ευχαριστώ!

https://opengov.ellak.gr/,
https://opendata.ellak.gr/,
https://odi.ellak.gr/,
https://ellak.gr/

Ηρακλής Βαρλάμης, varlamis@hua.gr