



# ΕΦΑΡΜΟΓΗ ΕΝΤΟΠΙΣΜΟΥ ΛΟΓΟΚΛΟΠΗΣ

Δημήτρης Γαβρίλης, Κωνσταντία Κακάλη

Βιβλιοθήκη & Υπηρεσία Πληροφόρησης, Πάντειο  
Πανεπιστήμιο

# Στόχος του προγράμματος

- Κατασκευή υπηρεσίας η οποία θα εκτιμά την πιθανότητα λογοκλοπής εργασιών (επιστημονικές / εκπαιδευτικές)
- Η υπηρεσία θα είναι εγκατεστημένη κεντρικά
- Θα μπορεί να γίνεται χρήση από οποιοδήποτε εξουσιοδοτημένο χρήστη / υπηρεσία
- Αυτόματη ενσωμάτωση με συστήματα αποθετηρίου: πλήρης αυτοματοποίηση

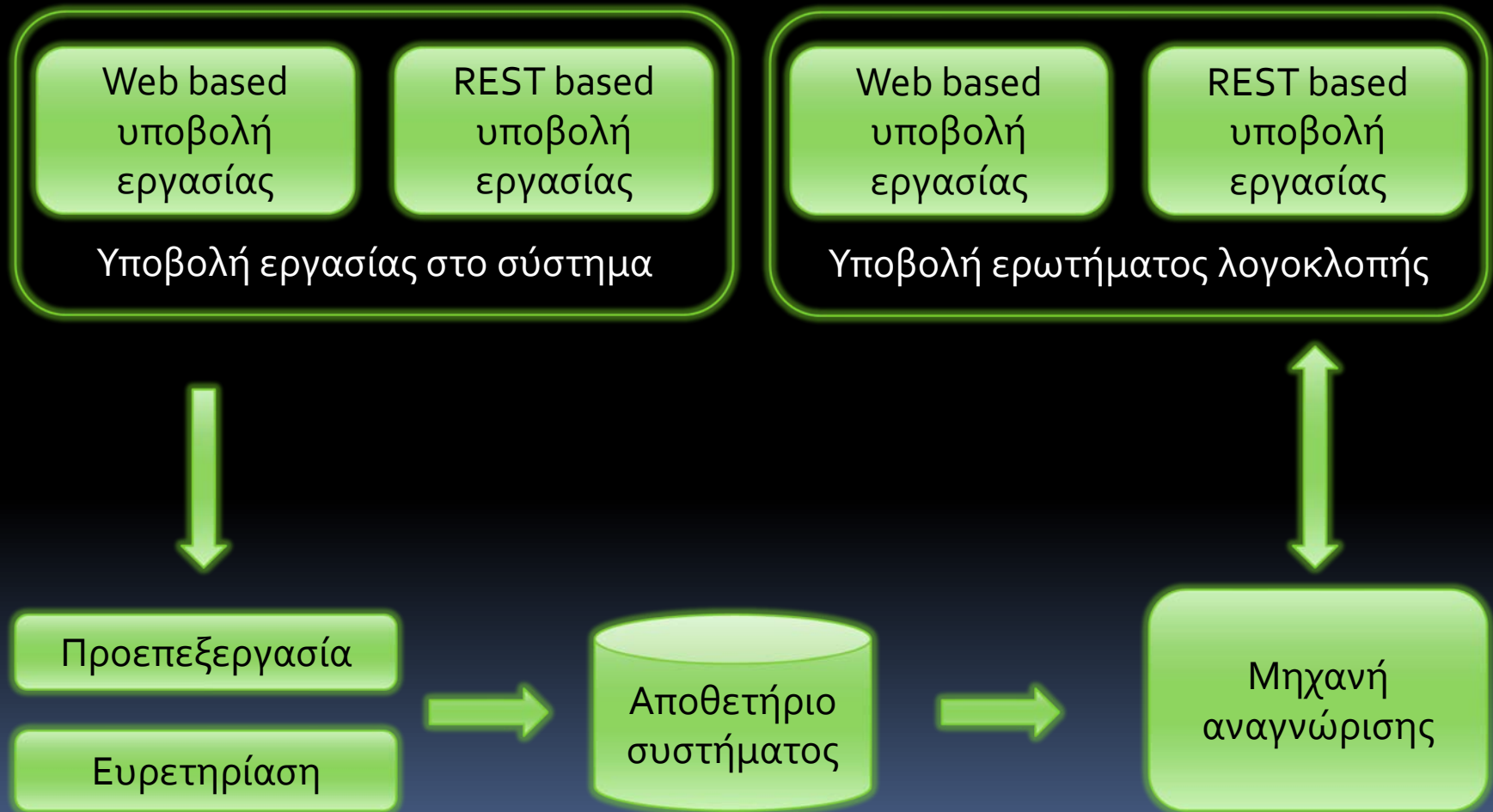
# Προδιαγραφές

- Υποστηριζόμενες γλώσσες: Ελληνικά / Αγγλικά
- Είδη εργασιών: εργασίες, επιστημονικά άρθρα
- Είδος αναγνώρισης: supervised (ύπαρξη υλικού εκπαίδευσης)
- Υλικό εκπαίδευσης: εργασία σε επεξεργάσιμη μορφή (Text, XML) – όχι PDF, Image.
- Χρήση του συστήματος:
  - Χειροκίνητη μέσω web based περιβάλλοντος
  - Αυτόματη με χρήση REST based web service

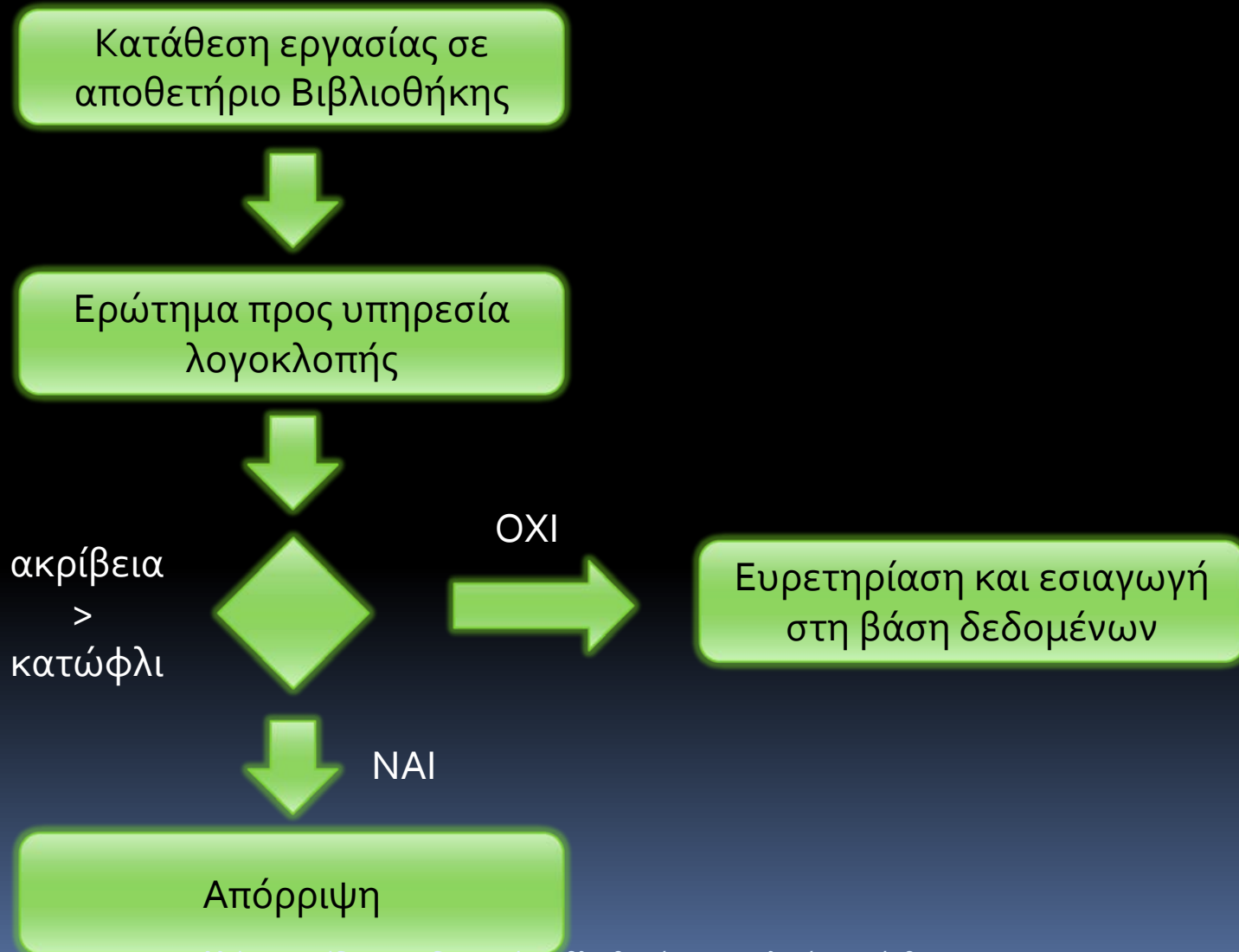
# Μεθοδολογία αναγνώρισης

- Υβριδικός αλγόριθμος
  - Ομοιότητα κειμένου
  - Image/Figure captions
  - Αναφορές
- Για τον βέλτιστο καθορισμό των παραμέτρων του αλγορίθμου απαιτείται:
  - Corpus εκπαίδευσης με θετικά / αρνητικά παραδείγματα.
- Επιτυχία αναγνώρισης εξαρτάται από την ποσότητα και την ποιότητα του υλικού που θα είναι καταχωρημένη.

# Αρχιτεκτονική συστήματος



# Ροή εργασίας



# Συμπεράσματα

- Μια καινοτόμα υπηρεσία
- Θα εξυπηρετεί το σύνολο της ακαδημαϊκής (και όχι μόνο) κοινότητας
- Η απόδοση και ως εκ τούτου η επιτυχία του προγράμματος εξαρτώνται από :
  - Την αποδοχή από την κοινότητα
  - Την ποσότητα και την ποιότητα των εργασιών που θα περιέχει η βάση δεδομένων

Ευχαριστώ για την  
προσοχή σας

