

Ανάλυση δεδομένων από συνεδριάσεις της ολομέλειας του ελληνικού κοινοβουλίου

Δημήτρης Γκούμας, Μαρία Ποντίκη, Κωνσταντίνα Παπανικολάου &
Χάρης Παπαγεωργίου

ΕΚ ΑΘΗΝΑ/ ΙΕΛ

Exploring the Political Agenda of the Greek Parliament Plenary Sessions
(ParlaCLARIN@LREC2018, 7 May 2018, Miyazaki, Japan)



clarin:el

Ερευνητικό ερώτημα:

- Πως διαμορφώνεται και εξελίσσεται στο πέρασμα του χρόνου η πολιτική ατζέντα;
- Πως δρουν και αντιδρούν τα πολιτικά κόμματα και τα μέλη τους στο πέρασμα του χρόνου;

Ανάλυση θεματικού περιεχομένου (topic modelling) και πολιτικός λόγος:

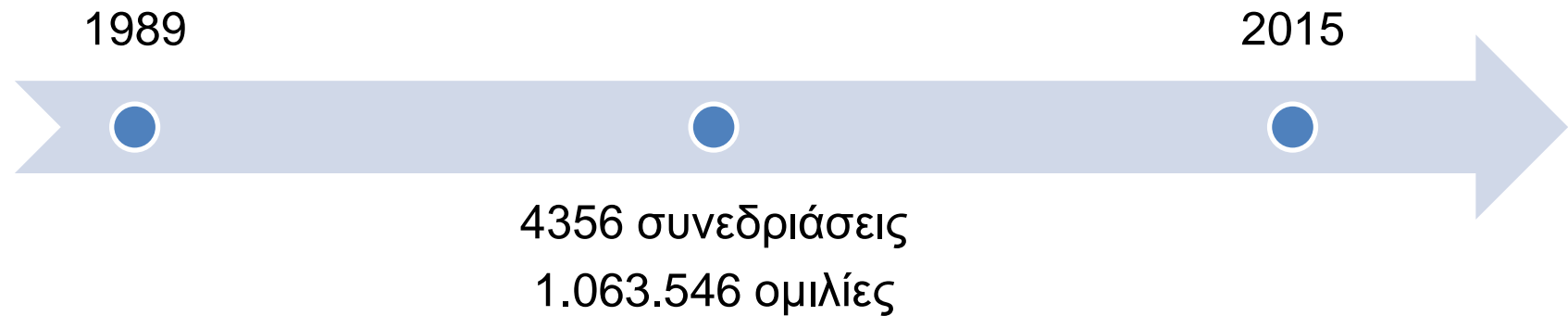
Greene & Gross, 2017 → Ευρωπαϊκό Κοινοβούλιο

Quinn et al., 2010 → Αμερικανική Γερουσία

Grimmer, 2010 → Κογκρέσο (δελτία τύπου)

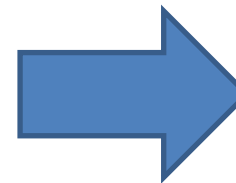
Δεδομένα

- Συνεδριάσεις της ολομέλειας του ελληνικού κοινοβουλίου

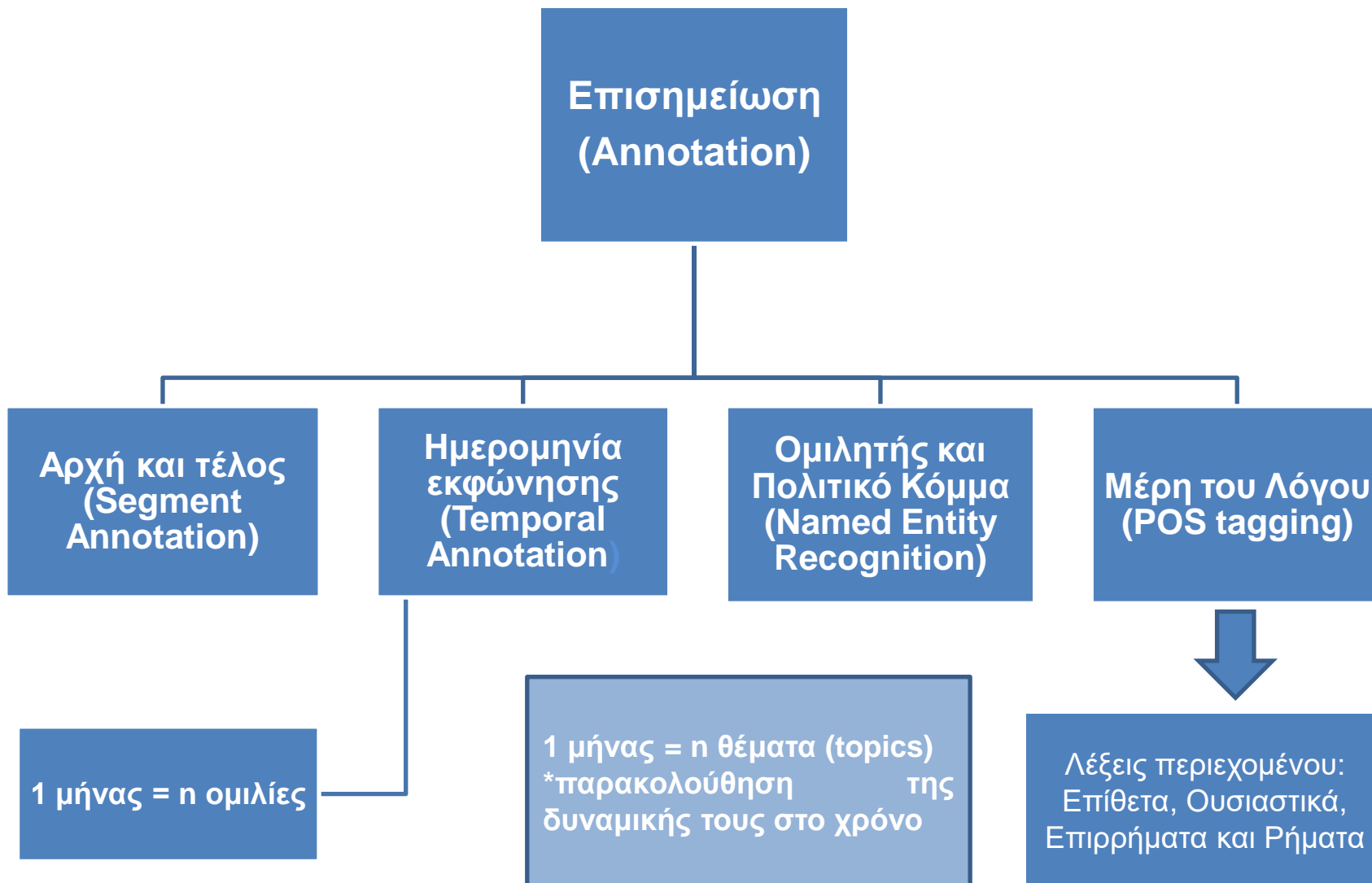


Μορφότυπος
(format)
δεδομένων

{*txt, *doc, *docx, *pdf,
*jpeg}



XML



Ανάλυση θεματικού περιεχομένου (topic modelling)

- TF-IDF σχήμα (weighting scheme)
- Non-negative Matrix Factorization (**NMF**)

Δυναμική **NMF**

Μοντέλο NMF
→ Παράθυρο Μήνα X

- n Θέματα/ Μήνα X

Εφαρμογή της μετρικής Word2Vec
Topic Coherence (Συνοχή Θεμάτων)
(TC-W2V)
Μοντέλο NMF

- Σημασιολογικά συγγενή
θέματα → δυναμικά θέματα
 $DT_1 = \{\text{topic}_1 \text{ topic}_2 \text{ topic}_3\}$

Αποτελέσματα

- **Greek Parliament topic modeling**
- 283 παράθυρα (1 παράθυρο → 1 μήνας)
- 90 δυναμικά θέματα (dynamic topics)
- Παρουσίαση ανά θέμα/ ανά μήνα
- Αναζήτηση με όρους
- Πολιτικά κόμματα/ μέλη κομμάτων
- Εκλογικές περιφέρειες

