

GlossAPI

Alexandros Melidis

Director Open Technologies Alliance - GFOSS

glossAPI as Digital Commons

Mission: To create an open (source, data, models, governance) high-quality language model in Greek.

digiCommons contri:

- Promotes open access to linguistic data sets
- Supports community driven development, transparency and participation

<https://github.com/eellak/glossAPI>

README

Στόχος της ΕΕΛΛΑΚ είναι η ανάπτυξη ενός Ελληνικού γλωσσικού μοντέλου ανοιχτού λογισμικού του glossAPI(a Greek Open Source LLM), που ο κώδικας του θα διατίθεται με την άδεια ανοιχτού λογισμικού [EUPL](#) και τα δεδομένα του με την άδεια [Creative Commons BY-SA](#).

 **Τρέχουσα δράση: Καταγραφή - αποτίμηση ανοιχτών πηγών κειμένου στα Ελληνικά**

 Συμβουλευτείτε το CONTRIBUTING.md για να συνεισφέρετε στην  συγκέντρωση και προτεραιοποίηση συνόλων κειμενικών δεδομένων στα Ελληνικά.

 Επισκεφτείτε το [wiki](#) για να δείτε ή να τροποποιήσετε τις καταγεγραμμένες πηγές.

 Έχετε μια ιδέα που δεν βλέπετε στο αποθετήριο; Θέλετε να προτείνετε μια διόρθωση;  Ανοίξτε ένα [Issue!](#)

 **NEW**: Δοκιμάστε τον [λεξικό κατατεμαχιστή \(παράδειγμα χρήσης\)](#) και  δώστε μας ανατροφοδότηση στα [Issues!](#)

 **Διαβάστε την** [αρθρογραφία μας για τα γλωσσικά μοντέλα και τις διεθνείς εξελίξεις](#)

[Νευρωνικά Δίκτυα και Μηχανική Μάθηση](#)

[Ανοιχτός Κώδικας και Προηγμένα Γλωσσικά Νευρωνικά Δίκτυα](#)

[Τί ξέρουμε για τις ραγδαίες εξελίξεις στην OpenAI...](#)

Releases

No releases published

Packages

No packages published

Contributors 2

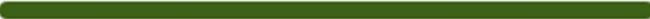


ninagial Nina Gial



tgkarounos Theodoros G. Karounos

Languages

 ● TeX 100.0%

glossAPI journey

- The idea: glossAPI initiated to address the lack of open LLMs in Greek
- Human annotation: strong emphasis on criteria driven preparation of training material and continues quality control

Key processes

- Inventory and prioritization of data sources
- Automated data acquisition and cleaning
- Compilation of training and evaluation data sets

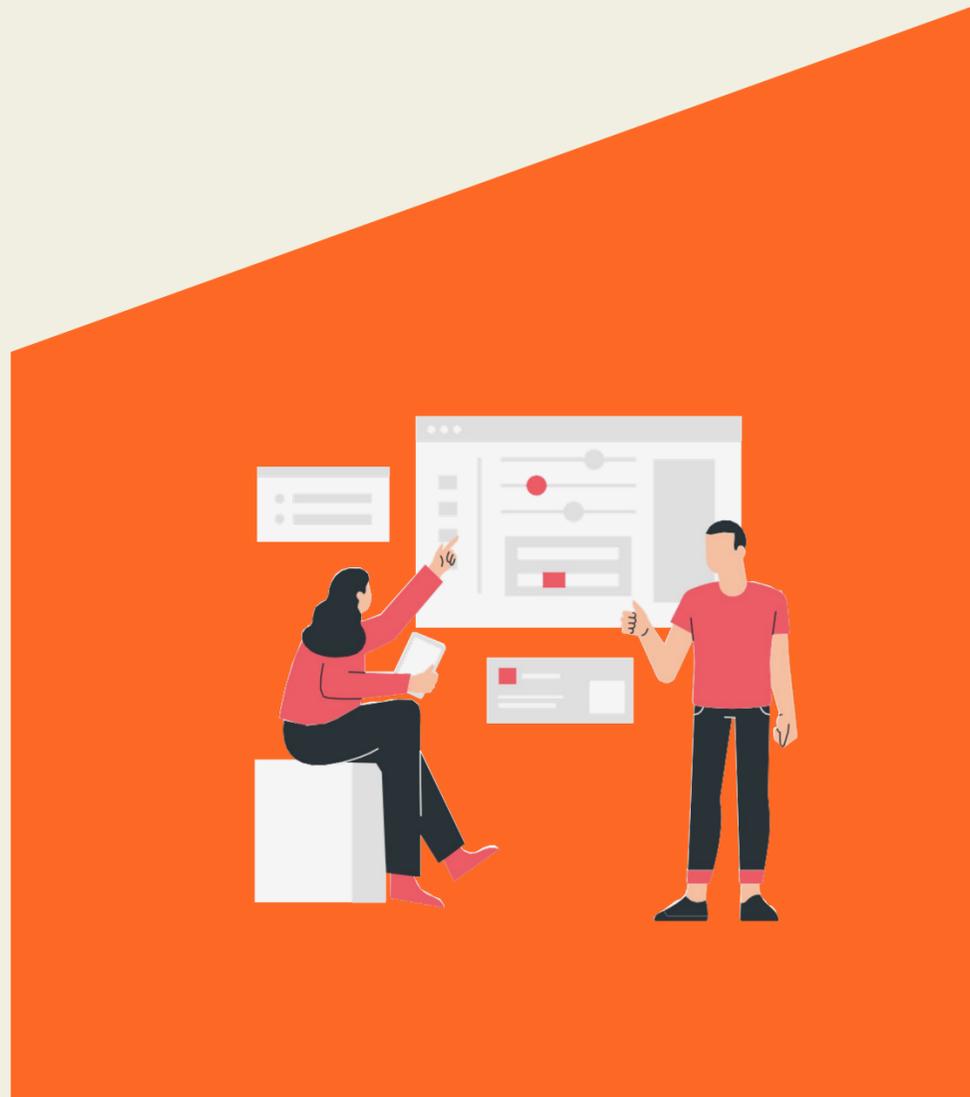
Funding and governance

- Combination of self funding and Digital Europe
- Community driven approach (meet ups, element room, articles, presentations)
- Transparent decision making with key stakeholders
- Regular updates (newsletter) and version controlled protocols to ensure consistency and reproducibility

Future potential

- Broaden the scope and open up dialogue with similar efforts across the oss AI community
- Integration with similar open data- open science initiatives
- Active participation in relevant EU policy discussions - agenda setting and peer pressure for digital commons AI development
- Advocacy for inclusive and ethical AI development in multistakeholder fora such as OGP action plan

thank you



almel@eellak.gr