

Gaining insight from government data

gov.insight

E. Zavitsanos, G. Giannakopoulos, G. Paliouras



Institute of Informatics and Telecommunications, NCSR Demokritos
gov.insight.iit.demokritos.gr



The Target

- Exploit freely available, government data
- for decision support...
- and social studies.
- Demonstrate and apply state-of-the-art technology
- Usable by
 - Social and political analysts and students, Journalists
 - IT researchers and students
 - Lawmakers



The Bow: gov.insight

- Web application
- Data from opengov.gr
- Knowledge from comments on public consultations
- Created by the Software and Knowledge Engineering Lab (SKEL)
- Accessible through gov.insight.iit.demokritos.gr



The Bowstring: Data

- 176 consultations
- 3201 topics (law articles)
- 46190 comments

Last crawl includes Jan 2012 consultations



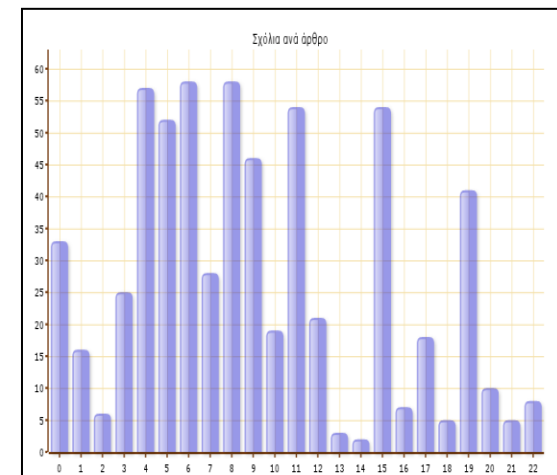
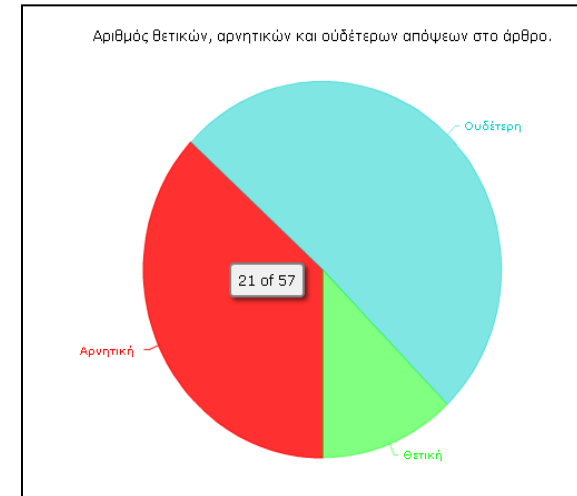
The Arrow: Services

- Navigation and search
- Statistics per **consultation, article** or **search results**
- Sentiment analysis
- Discussion activity quantification
- Trend visualization
- Author-based grouping
- Implied (latent) topic grouping



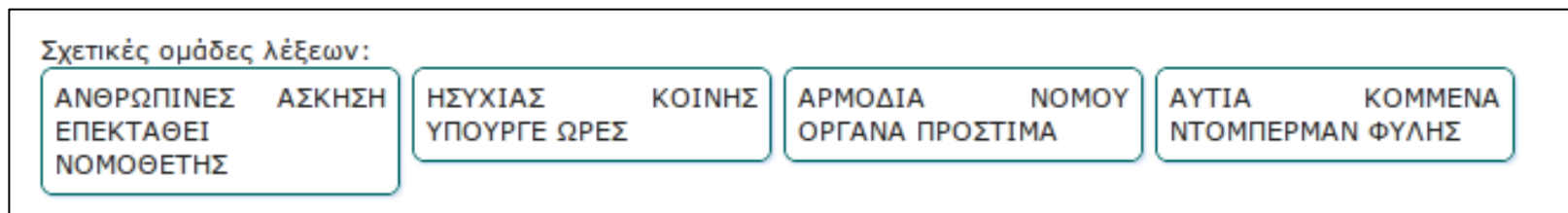
Use cases – (1/2)

- What do users think of a given article?
- Which articles were the most popular or controversial?



Use cases – (2/2)

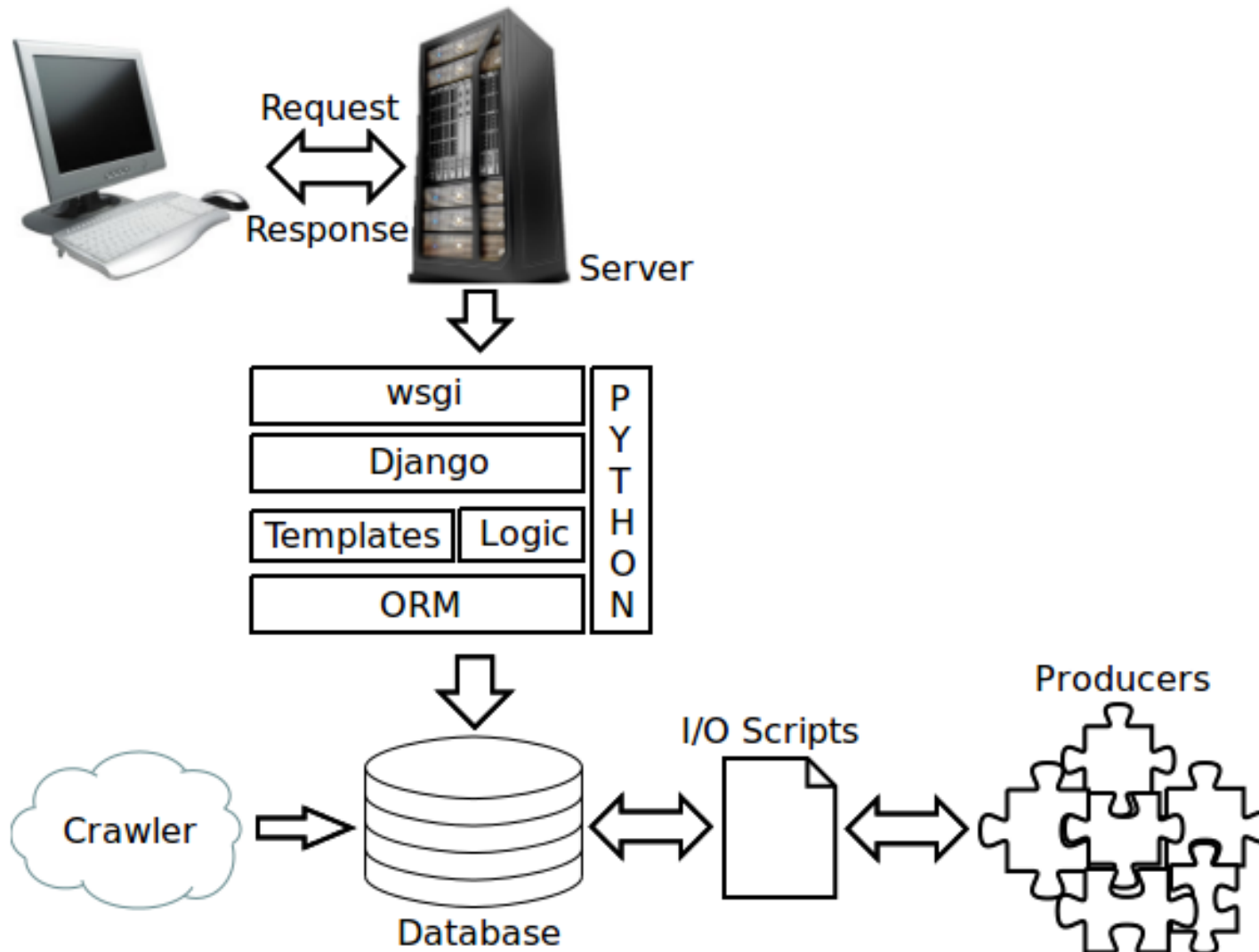
- Which (sub)topics showed up in the discussion?



- What were the trends in term use within a discussion?



Architecture Overview



Science and Technology Tools

- Preprocessing
- Part-Of-Speech tagging
 - Machine Learning (Athens University of Economics and Business)
- Comment Clustering (and implied topics)
 - Hierarchical Dirichlet Processes (SKEL implementation)
- Sentiment Classification
 - Machine Learning (i-sieve)
- Statistics and word cloud (open source tools)



Future work

- Named Entities
- Users
 - Roles and rights
 - Personalization
- More producers
- More statistics

Conclusions

- Extendable text analysis platform
- Open access to high tech solutions
- Dissemination and demonstration of research results
- Cooperation between industry and academia
- Open data exploitation
- **Science for the people**



Questions and Comments

- What would you expect from gov.insight;
- How can it be used?
- What can be improved?

Thank you for your attention
<http://gov.insight.iit.demokritos.gr>

Elias Zavitsanos

George Giannakopoulos

George Paliouras

