NLP and Public Engagement: the Case of the Italian School Reform

Tommaso Caselli*, **Giovanni Moretti**◇, **Rachele Sprugnoli**◇♮, **Sara Tonelli**◇,
Damien Lanfrey♣, **Donatella Solda Kutzmann**♣

Vrije Universiteit Amsterdam*, FBK◇, University of Trento♮, MIUR♣
De Boelelaan 1105 1081 HV Amsterdam NL, Via Sommarive, 18 38122 Povo, Viale Trastevere, 76/a 00153 Roma
t.caselli@vu.nl, {moretti;sprugnoli;satonelli}@fbk.eu, {damien.lanfrey,donatella.soldakutzmann}@istruzione.it

Abstract

In this paper we present PIERINO (PIattaforma per l'Estrazione e il Recupero di INformazione Online), a system that was implemented in collaboration with the Italian Ministry of Education, University and Research to analyse the citizens' comments given in #labuonascuola survey. The platform includes various levels of automatic analysis such as key-concept extraction and word co-occurrences. Each analysis is displayed through an intuitive view using different types of visualizations, for example radar charts and sunburst. PIERINO was effectively used to support shaping the last Italian school reform, proving the potential of NLP in the context of policy making.

Keywords: Text Mining, Information visualization, E-government

1.   Introduction

The engagement of citizens in policy making has become an important challenge for public administrations in the last years since current digital technologies could be adapted to promote forms of *direct democracy* by means of online surveys, asking for the citizens' opinion on key issues related to public life. However, the full potential of state-of-the-art technologies in collecting and analyzing such data has not been exploited yet. Online surveys limited to yes/no questions are somehow easily processed but more elaborated questionnaires, involving idea gathering and proposals collection from citizens, even if easily implemented through online forms, raise costs for an effective analysis. Indeed, successful public consultations - especially on topics high on the agenda, can end up with thousands of free-text answers, which cannot be analyzed manually. To make the engagement process effective, reward citizens' efforts, and produce useful outcomes for the administration, such amount of information needs to be automatically processed to extract content that can support civil servants and politicians in taking decisions. Besides, the extracted content must be displayed in an intuitive way, since it should be also accessed by non-technical users.

Research in the automatic analysis of open answers has mainly focused on data from the marketing and advertising fields (Yamanishi and Li, 2002) with few notable studies in the area of e-government (Ku and Leroy, 2014). The literature reports the use of different techniques. For example, supervised approaches (Giorgetti et al., 2003) and active learning (Patil and Ravindran, 2015) are employed to perform survey coding, that is matching open-ended answers with a short description associated to a code, thus converting qualitative information (text) into a quantitative format (code). In addition, text clustering and topic modeling methods are applied to explore information contained in open answers, e.g. to summarize and classify them (Wang and Mulrow, 2014) and to identify the intentions of survey participants (Rosell and Velupillai, 2008).

In this work, we present the first attempt to apply state-of-the-art Italian Natural Language Processing (NLP) techniques to the outcome of an online public consultation for policy making, with the goal to integrate citizens' contributions in the Italian school reform. The work is a joint initiative among the Digital Humanites Group at Fondazione Bruno Kessler (FBK), the Vrije Universiteit Amsterdam (VUA) and the Italian Ministry of Education, Universities and Research (MIUR). The goal of the work is the automatic analysis of linguistic data contained in the answers given by the participants to the public consultation La Buona Scuola[1] (commonly referred to in social media as #labuonascuola).

#labuonascuola has been among the biggest on-line public consultations ever realized in Europe. Intended to gather ideas, opinions and evidence on issues related to the ongoing consisted in an on-line platform with a dedicated survey and several discussion areas, a section for position papers and a calendar of more than 2000 local live debates, whose results have been uploaded to the consultation platform. The survey included a set of multiple-choice and some open questions, to which people could freely answer using natural language. Open questions were for example *What features should teachers have?* or *How would you change the school cycles?*. The questionnaire was online for 2 months, from 15th of September to 15th of November 2014, and could be accessed by the general public, with respondents providing some anagraphical data while keeping their replies anonymous.

The school reform has been highly debated by Italian public opinion and citizens, as also shown by the high number of contributes, ideas, replies and local live debates. Individual participants in the online questionnaire amounted to 130,000 people. Overall, more than 775,000 free text comments were provided by the users. While answers to closed questions were quite easy to process and summarize, the large amount of free texts represented a big challenge for MIUR personnel. Indeed, the amount of data prevented them from a manual analysis due to time, money,

---

[1] https://labuonascuola.gov.it

and errors, which can impact on the final recommendations to policy makers. To address this problem, we developed a system called PIERINO (PIattaforma per l'Estrazione e il Recupero di INformazione Online), which analyzed the free text answers and made them available in a user-friendly format as described in Section 2. The information collected was used to assist policy makers in the development of the new bill for the Italian school system: details on this point are given in Section 3. In addition, a version of PIERINO have been developed to deal with a specific topic related to school, i.e. learning disabilities: this version is described in Section 4. Finally, Section 5. presents conclusions and future work

2. System Description

In this Section we describe data pre-processing together with PIERINO structure and functionalities. Examples taken from the #buonascuola use case and screenshots are provided.

2.1. Data Selection and Preparation

The platform takes in input a list of natural language comments/answers, grouped by question. In our use case, the analysis focused on a set of 10 questions, which were considered as highly relevant by MIUR collaborators. Overall, we processed more than 270,000 free text comments, containing almost 5,000,000 tokens.

Each group of comments is first processed with TextPro[2] (Pianta et al., 2008), an NLP suite performing several linguistic analyses, and with Mallet[3] (McCallum, 2002), a system for the extraction of topic models. Then, we implemented a Web interface to easily navigate the data and query them. The analyses and the visualization strategies were discussed with MIUR and improved iteratively to meet their requirements, which were basically:

- for each question, provide a concise representation of the main topics mentioned in the replies;

- provide an interface which can be intuitively used without the need of technical support;

- allow users to easily make queries to the database;

- give the possibility to see the source comments and not just the summarized information.

These requirements were addressed in different views provided by PIERINO, as detailed in the following subsection. We rely on Highcharts[4] to present the most common chart types (i.e. bar and line charts), while the most interactive and custom data-driven visualizations (i.e. network of co-occurrences) are displayed with d3.js[5] (Bostock et al., 2011).

2.2. PIERINO functionalities

PIERINO includes five types of analysis and views. The platform functionalities are generic and can be applied to any public consultation dataset with textual data. A limited interaction with experts and promoters of the consultation is first necessary in order to set some parameters and for refining the results. It is important to stress out that PIERINO does not provide "the answers" but facilitates the analysis by experts of large amounts of textual data to support decisions. Below we report each functionality of PIERINO by using #labuonascuola data as an example.

The first is the **key-concept based view**, that displays for each question the 20 top-relevant concepts mentioned in the citizens' answers. These were extracted with the first version of KD (Moretti et al., 2015), a highly parametrizable key-concept recognizer[6]. We assigned a higher weight to longer concepts (between 2 and 4 tokens), in order to give a higher relevance to specific concepts and discard generic ones (usually unigrams). The ranked list of key-concepts is displayed through a sunburst, as shown in Figure 1. In this example, related to the question "What is missing in what you learn at school?", we observe a clear prevalence of specific subjects such as civics (first ranked), foreign languages and Italian.

The second section is devoted to the analysis of **verbs**, aimed to establish what actions citizens associate with the various issues presented in the consultation. PIERINO allows users to select a question of interest and see what are the most common verbs displayed in a bar chart. In all the answers, the high frequency of the verbs *"dovere"* (must) and *"fare"* (do) stands out as an index of the decisiveness and commitment with which users have expressed their opinions and the urgency they feel. Figure 2 shows that in the answers to the question "Which features of teachers are important to you?", beyond obvious verbs such as *"insegnare"* (teach) and *"sapere"* (know), there are many verbs encoding participation and relationship such as *"coinvolgere"* (involve), *"trasmettere"* (instill), *"relazionarsi"* (relate), *"interagire"* (interact).

The third section analyzes the degree of overlap between the content of the responses to individual questions and a list of **keywords** provided by MIUR, in order to verify the correspondence between the issues on the political agenda and the outcome of the consultation. We report an example in Figure 3, related to the question about how to train the teachers of the future. The analysis visualized as a radar chart shows a strong overlap between the outcome of the survey and the keywords "public competition", "evaluation" and "certification".

The fourth section of PIERINO includes a **targeted search engine**, which allows users to look for a term or expression of interest in the participants' responses, and see the most common words co-occurring with it. An example is displayed in Figure 4. Co-occurrence selection may be refined by specifying various parameters, including the context in which to search (right context only, left context only or both) and its width (from from 1 to 20 tokens). Besides, the user can search only for co-occurring
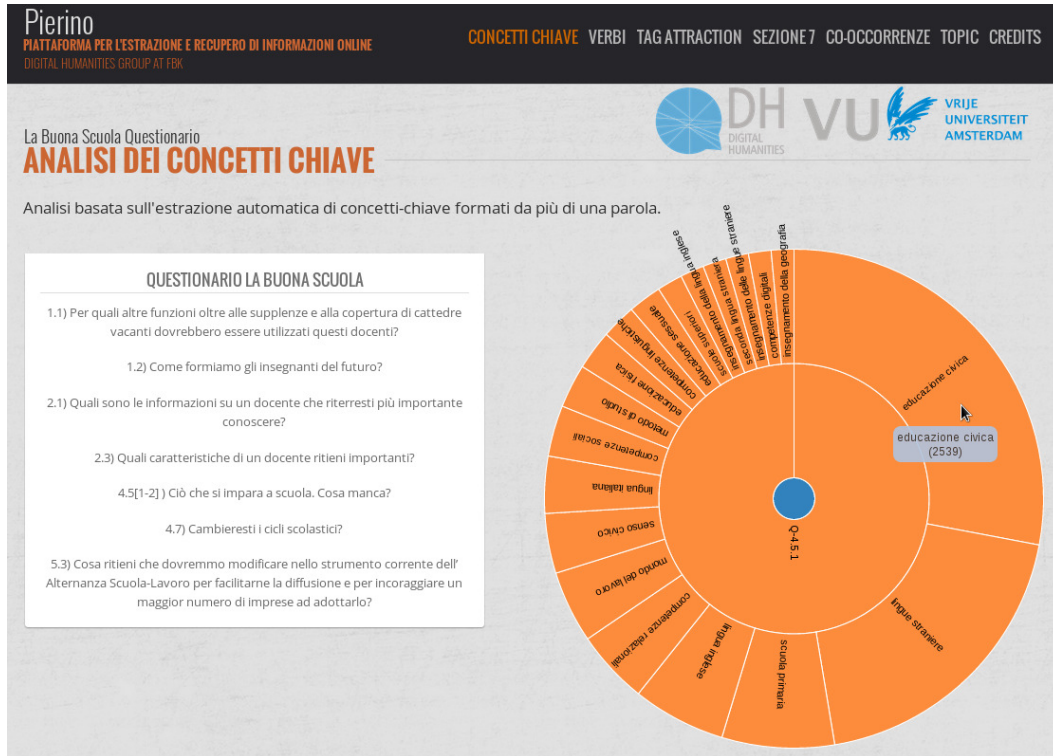
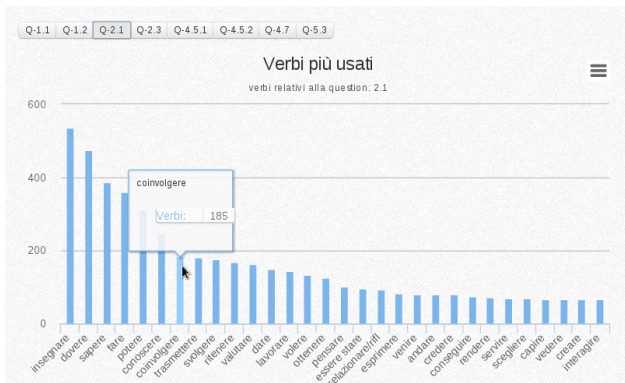Figure 1: Visualization of key-concepts for the question: "What is missing in what you learn at school?").

Figure 2: Most used verbs in the answers to "Which teachers' characteristics are important to you?".

Figure 3: Overlap between the content of the question "How do we train the teachers of the future?" and keywords provided by MIUR

verbs, nouns and adjectives, or for different part-of-speech combinations. Examples of part-of-speech patterns are noun+adjective, e.g. "*alunni stranieri*" (foreign students), and noun+preposition+noun, e.g. "*corso di recupero*" (remedial class). Once the search is completed, users can click on the co-occurrence nodes to display the citizens' comments in which the terms of interest appear. Additionally, the nodes can be colored by polarity (positive or negative), to show whether citizens associated a positive or negative opinion to the given term. The polarity scores associated with a word are extracted from an Italian prior polarity lexicon (Sprugnoli et al., 2015).

Finally, **topics** extracted with Mallet are used to show the most discussed issues as expressed by single words (e.g. "merit") and groups of words (e.g. "work-market"). Topics are displayed as a network (Fig. 5). This highlighted,
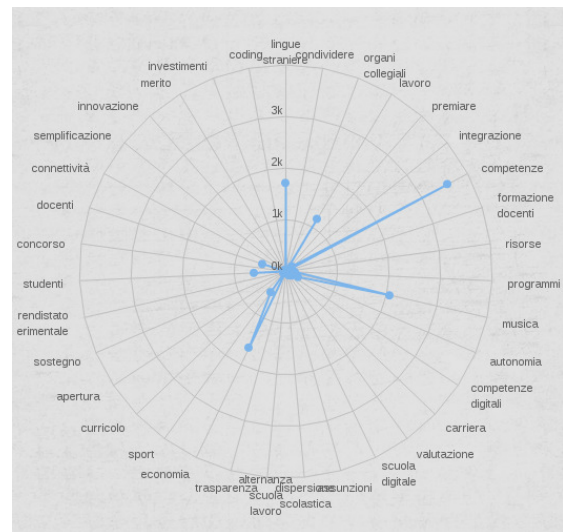
for instance, that respondents stressed the need to put more weight on the training and continuing education of teachers but also on the involvement of families.

3. Impact on Prospect Users

PIERINO can substantially advance the practice of open government. The mere availability of tools like this allows Governments and public administration to organize public consultations that are more stratified and that are able to collect finer data for policy design and implementation.

403

Figure 4: Networks of Co-occurrences: (a) Co-occurrences of the word "lingua" (language) in the answers to question "Whats missing in what you learn at school?"; (b) Co-occurrences of the expression "scuola media" (middle school) for question "How would you reshape school cycles?"
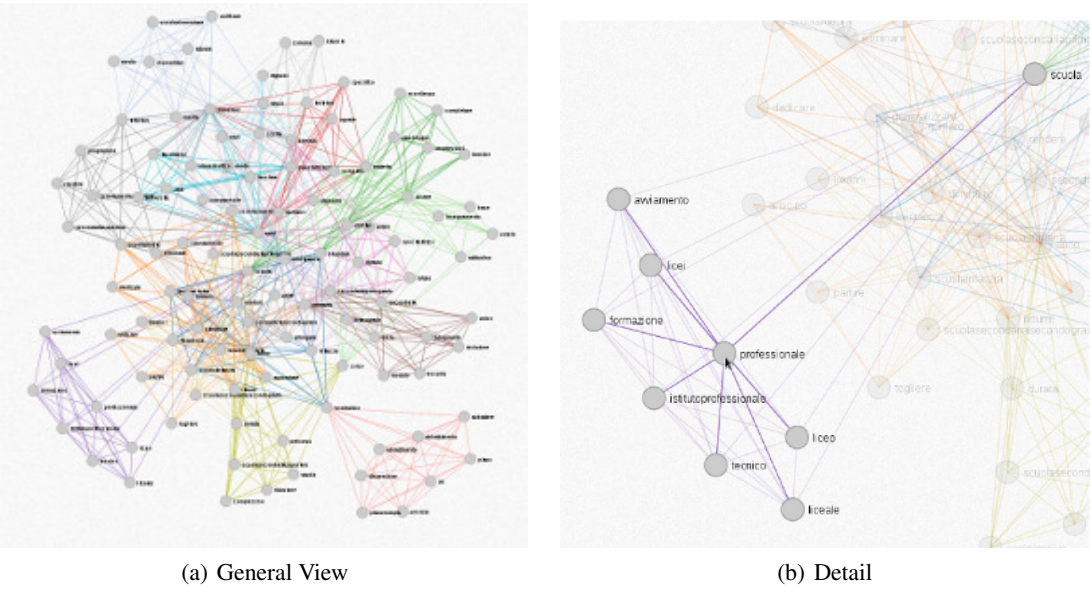

(a) General View

(b) Detail

Figure 5: Topic Network

One of the major impact of PIERINO for prospect users is the enhancement of the quality of decision-making collecting an entire new class of information, referred to natural language expressions, particularly crucial for complex and non-linear policy problems. Those problems tend to be characterized by a wide continuum of options, that cannot be collected by traditional survey methods. For example, a question regarding the possibility of reforming the Italian education cycle (primary, lower secondary and upper secondary education) allowed the MIUR to understand that possible answers with apparent widespread agreement (the possibility to reform and possibly eliminate lower secondary education, merging primary and lower secondary years) proved to be less agreed than other options, more technical and less evident in public debate (such as, for ex-

ample, the possibility of reforming the first two years of upper secondary education, as in the German education system). The combined use of key-concepts (Figure 1) and co-occurrence networks (Figure 4) allowed the Ministry to understand this and other positions on the subject matter and avoid to subscribe to the easier option expressed by public debate.

An additional example of impactful use of PIERINO is provided by the analysis of the answers to the question concerning the permanent hiring of 100,000 new teachers and, particularly, to 50,000 new functional teachers. Over 130,000 textual answers to the question provided a very clear address to policy-makers: at primary school level, additional teachers should be used to promote the coexistence of two teachers in classrooms with over 20 pupils; at

secondary school level, instead, additional teachers should be used to support students with poor performances in the afternoon, with particular respect to schools in areas with lower socio-economic indicators. These and other distinct addresses by the public have been directly embedded in the law finally approved in July 2015.

In many cases, it was the combination between key-concepts and co-occurrence data to provide the clearest reading: while key-concepts analysis provides the general framing to citizens answers, co-occurrence networks allowed to investigate the granularity and nuances of answers, a priceless combination of information for policy-makers.

4. Use case: PIERINO-Dyslexia

A platform version called PIERINO-Dyslexia has been developed to explore participants' comments related to learning disorders on the occasion of the 2015 European Dyslexia Awareness Week.

Although the questionnaire did not include specific questions on learning disability (in Italian DSA, *Disturbi Specifici di Apprendimento*), this theme has been mentioned by several participants to the survey in relation to different subjects such as the training of teachers or the allocation of additional resources. Therefore, we exploited the functionalities of PIERINO to analyze the contexts in which DSA were mentioned and their relevance with respect to the other topics of the survey.

Starting with all the answers given by the citizens, we automatically extract a list of 62,200 key-concepts ordered by relevance. Among these, we checked the ranking of the concepts related to DSA. The term DSA is at position 397, demonstrating the attention to this issue paid by parents and teachers. Concepts with a similar weight are more general expressions such as "*difficoltà*" (difficulty), "*soldi*" (money) and "*scuola superiore*" (high school).

Citizens have discussed learning disabilities in different contexts, which can be analyzed by using the targeted search engine based on co-occurrences. For example, by looking at expressions co-occurring with the acronym "dsa", it is possible to see that participants suggested methods to help the inclusion of dyslexic students in every-day learning activities such as the introduction of occupational therapy in schools and the use of concept maps.

The analyses carried out show that the survey participants are aware of the impact of learning disabilities on the school organization, and propose effective strategies to address them. As an example, answering the question "In which activities should non-tenured teachers be employed apart from teaching?", several participants suggest to employ them in supporting classes with children with learning difficulties after an adequate training. Another example is given by many answers to the question "What features should teacher have?" that highlight the importance of empathy, of knowing how to relate with children with learning difficulties and their families, and how to manage classes with students having different problematic issues. In general, respondents demonstrate a thorough knowledge of learning disabilities, and in particular of dyslexia, as shown also by the use of technical terminology. Indeed, participants could differentiate between various types of disorders and, de-

spite knowing the difficulties of schools and teachers, they want to propose concrete strategies for managing these phenomena.

5. Conclusion and Future Work

We presented PIERINO, the first attempt to integrate NLP and data visualization in the Italian policy making process. The NLP modules in PIERINO have all been benchmarked against standard datasets and successfully applied to other text types such as the citizens answers. Its analyses contributed to the decisions which informed the recently passed bill on the Italian school reform.

On the basis of the results obtained through the analysis of public participation results, the Ministry has decided to use an adapted version of PIERINO to a new stream of texts. In this case, texts are of administrative nature, and relate to school ex-post impact assessment on products and services offered by commercial parties through their CSR campaigns.

We are planning to extend the functionalities of the platform so as to have a finer-grained understanding of citizens' suggestions and concerns. For instance, useful features may include the possibility to cluster answers per content and to perform comparisons based on supporting and contrasting ideas or on the sources. Modality (how the message is conveyed), negation scope and sentiment analysis should be investigated and improved. Another important aspect for policy makers is the possibility to perform real-time analysis of the aggregated data by region, as some specific issues may vary based on the geographical area. Although this piece of information is currently not present in the platform, it can be easily retrieved from the respondents metadata and integrated in the system.

6. Acknowledgements

7. Bibliographical References

Bostock, M., Ogievetsky, V., and Heer, J. (2011). D³ data-driven documents. *Visualization and Computer Graphics, IEEE Transactions on*, 17(12):2301–2309.

Giorgetti, D., Prodanof, I., and Sebastiani, F. (2003). Automatic coding of open-ended surveys using text categorization techniques. In *Proceedings of Fourth International Conference of the Association for Survey Computing*, pages 173–184.

Ku, C.-H. and Leroy, G. (2014). A decision support system: Automated crime report analysis and classification for e-government. *Government Information Quarterly*, 31(4):534–544.

McCallum, A. K. (2002). Mallet: A machine learning for language toolkit.

Moretti, G., Sprugnoli, R., and Tonelli, S. (2015). Digging in the Dirt: Extracting Keyphrases from Texts with KD. In *Proceedings of the Second Italian Conference on Computational Linguistics CLiC-it 2015*, page 198. Accademia University Press.

Patil, S. and Ravindran, B. (2015). Active Learning Based Weak Supervision for Textual Survey Response Classification. In *Computational Linguistics and Intelligent Text Processing*, pages 309–320. Springer.

Pianta, E., Girardi, C., and Zanoli, R. (2008). The TextPro Tool Suite. In *Proceedings of the sixth international conference on Language Resources and Evaluation (LREC 2008)*. Citeseer.

Rosell, M. and Velupillai, S. (2008). Revealing Relations between Open and Closed Answers in Questionnaires through Text Clustering Evaluation. In *LREC*.

Sprugnoli, R., Tonelli, S., Marchetti, A., and Moretti, G. (2015). Towards Sentiment Analysis for Historical Texts. *Digital Scholarship in the Humanities*.

Wang, F. and Mulrow, E. (2014). Analyzing Open-Ended Survey Questions Using Unsupervised Learning Methods. In *Proceedings of the Survey Research Methods Section of the American Statistical Association (JSM 2014)*.

Yamanishi, K. and Li, H. (2002). Mining open answers in questionnaire data. *Intelligent Systems, IEEE*, 17(5):58–63.