Position Analysis of Internet Public Opinions -
A Case Study of Nuclear Energy Policy in Taiwan

Center for Survey Research Conference (CSR 2021)

Academia Sinica, Taiwan

September 2-3, 2021

Naiyi Hsiao[1]*

KyleYulunKuo[2]

Hen-Hsuan Huang[3]

Ying-Che Tang[4]

[1] Associate Professor, Department of Public Administration & Taiwan E-Governance Research Center, National Chengchi University. * Contact email: nhsiao@nccu.edu.tw
[2] Doctoral Candidate, Department of Public Administration, National Chengchi University
[3] Assistant Professor, Department of Computer Science, National Chengchi University.
[4] Master, Department of Computer Science, National Chengchi University.

1. Introduction

Nuclear power has been an important source of energy supply in Taiwan for the last few decades. However, the nuclear power plants will be gradually decommissioned. The existing nuclear power plants are going to reach the end of legally licensed lifetime between 2018 and 2025. In addition, the decommission of the nuclear power plants, which is highly technical and meanwhile value dependent, has significant and sustained impact in terms of benefit and risk perception (Slovic[5], 2001; Slovic, Fischhoff & Lichtenstein[6], 1982; Rosa & Freudenberg[7], 1993). Previous studies have found that the general public believe that they are at risk, especially when the policy communications are not transparent in time. Those beliefs then turn into public attitudes against new energy policy (Rubin、Chowdhury & Amlot[8], 2012; Tan, Barnett, Stolz & Links[9], 2011).

With the widespread adoption of digital technology and social media, the internet public opinions (IPO) have received more attention in terms of how they are collected and analyzed. While nearly 90% of the people in Taiwan surfing the internet through mobile phones[10], the rapid and massive IPO data collection and analysis are expected to identify the real-time public sentiments, positions/stances towards specific policy issues. Collecting and analyzing IPO timely and responding actively to the public have become one of the critical governance

---

5 Slovic, P. (Ed.). (2001). ***Smoking: Risk, perception, and policy.*** Sage publications.

6 Slovic, P., Fischhoff, B., & Lichtenstein, S. (1982). Why study risk perception?. ***Risk analysis, 2***(2), 83-93.

7 Rosa, E. A., Freudenberg, W. R. (1993). The historical development of public reactions in nuclear power: implications for nuclear waste policy. Public reactions to nuclear waste: Citizen's views of repository sitting. ***Duke University Press, 32***, 63.

8 Rubin, G. J., Chowdhury, A. K., & Amlôt, R. (2012). How to communicate with the public about chemical, biological, radiological, or nuclear terrorism: a systematic review of the literature. ***Biosecurity and bioterrorism: biodefense strategy, practice, and science, 10***(4), 383-395.

9 Tan, C. M., Barnett, D. J., Stolz, A. J., & Links, J. M. (2011). Radiological incident preparedness: planning at the local level. ***Disaster medicine and public health preparedness, 5***(S1), S151-S158.

10 National Development Council (2020). ***Digital opportunity survey.*** Taipei, Taiwan.

challenges. The previous research has, nevertheless, indicated that IPO analysis remains improved concerning the extent to which the netizens sentiments and positions/stances correlate with each other, and more importantly, why the netizens approve or disapprove a specific public policy issue. Unlike the public sentiments towards general goods or services, positive/negative sentiments do not necessarily correlate with approval/disapproval attitudes/positions/stances of most public policy issues usually more multi-faced and complicated[11] contingent upon benefit and risk as well as relevant cost, condition and context in the diverse perspectives of policy stakeholders.

Rather than sentiment analysis, taking the controversial nuclear energy policy in Taiwan, the study aims to experiment with the procedure where the domains-specific human coders assist the position/stance models based on opinion mining using natural language processing (NLP). Reflecting upon the human-machine collaboration experience, the study attempts to explore how IPO analysis can be applied more constructively in public policy analysis.

2. Proposing IPO Analysis with Human-Machine Collaboration

IPO analysis takes advantage of the existing public comments available on the internet and semantics analysis algorithms enabling a large amount of text/semantic data can be processed simultaneously. Table 1 below proposes the detailed human-machine interaction process[11] that "idealistically" facilitate policy/issue-specific IPO analysis. Firstly, the policy deliberation team and facilitators (Fs) collect texts (T) and data (D) concerning policy/context/issues from multiple online/offline (O2O) sources as specified. The results also help identify the stakeholders (Ss) and domain experts (Es) to be involved in the

[11] Hsiao, N., Liao, Z., & Chen, D. (2018). From naive expectation to realistic progress – Government applications of big data on public opinions mining. ***Big data in computational social science and humanities***. Springer.

subsequent offline deliberation. The deliberation team should pay more attention to online texts and data. The efforts may be properly collaborated with the domain experts with diverse positions who will also participate in the deliberative meetings. Besides, the unexpected stakeholders may be identified when online texts and public opinions are available and browsed preliminarily.

The deliberation team may then work with technical providers of data analytics and machine learning to analyze online texts and to extract public sentiments, positions, arguments and evidences (SPAE). The next two steps (with red stars) substantiate how facilitators, domain experts, and technical experts collaborate to enhance the quality of machine learning and data analytics. The literature has called the iterative process augmented intelligence (AI) to distinguish its uniqueness to integrate domain expertise with automated algorithms. The AI results then hopefully contribute to the preparation and implementation of offline deliberation (Stage 4), which has often been criticized for its lack of online and updated evidences. Given sufficient resources, the results of Stage 4 and 5 may go simultaneously both on online and offline forums (O2O) and stimulate more public discussions. The O2O discussions, coupled with the previous human-machine collaboration and improvement, also ensure that the AI iterations produce and accumulate constructive outcome for subsequent deliberation and policy-making.

Table 1. The iterative stages for human-machine collaboration

Stage	Purpose	Description & Roles
1	Identify policy context, issues, & stakeholders	• Deliberation team/facilitators (Fs) collect texts (T) and data (D) concerning policy/context/issues from multiple online/offline (O2O) sources • Identify stakeholders (Ss) & domain experts (Es)
2	Analyze T/D & extract SPAE by data analytics & machine learning	Fs analyze T/D related to all issues and extract public sentiments, positions, arguments & evidences (SPAE) by data analytics & machine learning algorithms/platforms
3	Revise SPAE from Es	Fs invite Es to explore/explain/revise the previous SPAE results
Repeat Stages 2-3 until SPAE quality good enough		
4	Conduct O2O deliberation	• Demonstrate the prev. SPAE results online • Invite O2O stakeholders for further discussions • Conduct offline deliberation
5	Provide offline deliberation online	• Collect results from offline deliberation *Repeat Stage 2-3 until SPAE quality good enough* • *Demonstrate & invite online discussions
Repeat Stage 2-5 until SPAE quality good enough		

3. Position Analysis by Natural Language Processing

Opinion mining, based on natural language processing (NLP), facilitates the survey for public opinions with a new way. Compared with the traditional offline and online surveys, opinion mining provides the real-time information that reflects public opinions, preferences, stances, and sentiments in a large-scale with a lower latency[12]. Earlier work in opinion mining mainly relies on lexical information, not only using keywords for crawling relevant articles and comments from social network platforms, but also applying a sentiment lexicon for analyzing the polarity of opinions in individual articles/comments[13]. In recent years, learning-based approaches show their effectiveness in sentiment analysis. In particular, a

[12] Blitzer, J., Dredze, M., & Pereira, F. (2007, June). Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In ***Proceedings of the 45th annual meeting of the association of computational linguistics*** (pp. 440-447).

[13] Taboada, M., Brooke, J., Tofiloski, M., Voll, K., &Stede, M. (2011). Lexicon-based methods for sentiment analysis. ***Computational linguistics***, *37*(2), 267-307.

number of models based on deep neural networks generally achieve the performances greater than 80% of F-scores in a variety of datasets in different languages[14].

Experimenting the first three stages in Table 1, we employs deep neural-network models as a tool for automatically analyzing the public opinions, particularly their positions or stances, towards the nuclear energy issue in Taiwan. The high-performance neural network models usually rely on labeled data for training the model in a supervised manner. Popular datasets include Amazon product reviews, restaurant reviews, and hotel reviews. In this work, however, no existing labeled data is available for the issue about the location of nuclear waste. Thus, the first step of this work is to create a dataset in a moderate size, and then a neural-network model based on BERT[15], the dominating pre-trained text-encoder, is trained with the dataset to predict the positions (favor/against/unknown) of public comments from the social media.

3.1 Data Annotation

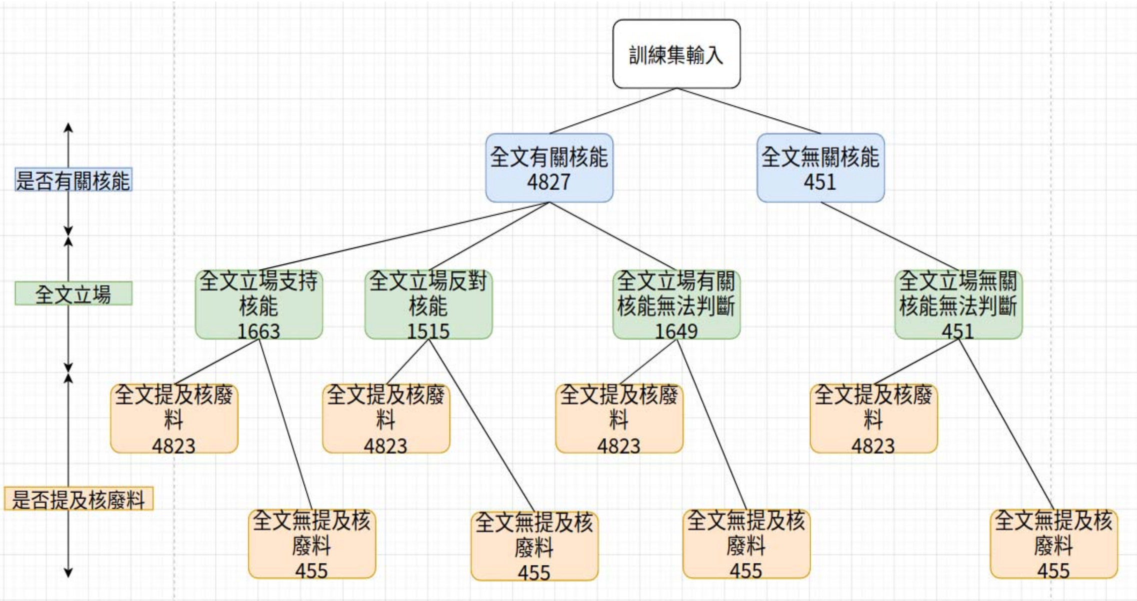
The annotation scheme should be defined before annotation. We organized a committee for determining the full annotation scheme. At first, the coder/annotator has to decide whether the article/comment is relevant to the issue about nuclear power in Taiwan. For a relevant article/comment, the annotator will label the author's stance about nuclear power, i.e. favor or against nuclear power. Finally, the annotator will also label whether the specific topic of nuclear waste is exactly mentioned in the article/comment. A group of native speakers are coached. The labeled data are further manually cleaned. Duplicated articles and noisy data are truncated from the final dataset. The statistics of the resultant dataset is shown in Table 2.

---

[14] Hoang, M., Bihorac, O. A., &Rouces, J. (2019). Aspect-based sentiment analysis using bert. In ***Proceedings of the 22nd Nordic Conference on Computational Linguistics*** (pp. 187-196).

[15] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Table 2. Statistics of the resultant dataset



According to the annotation scheme in Table 2, we propose three models for the three levels of the labels. Firstly, a model is trained to predict if a given article/comment is relevant to nuclear power. For a relevant article/comment, a second model is trained to predict the stance of the authors. The output is one of the three classes, including favor, against, and unknown. Thirdly, the last model is trained to predict if the topic "nuclear waste" is mentioned in the given article/comment. All the three models perform the tasks of sentence classification, a most typical NLP task constructed by the pre-trained BERT text encoder. BERT is a 12- or 24-layer Transformer network that converts a sequence of text into a matrix that is able to represent the information of the original input text. The large Transformer network is pre-trained with two self-supervision tasks with large-amount of data. Previous studies show that the BERT model can achieve decent performance in many NLP tasks without a large-scale dataset for fine-tuning.

Totally 4,500 instances are used for the model training, and 778 instances are held-out for evaluation. The performance metrics include Precision, Recall, and F-score, which are the most widely-used metrics in NLP. They are defined as follows.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$

$$F - score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

As shown in Table 3, True Positive is the number of positive instances that are correctly predicted as positive, False Positive is the number of negative instances that are incorrectly predicted as positive, and False Negative is the number of positive instances that are correctly predicted as Negative. F-score is the harmonic mean of Precision and Recall.

Table 3. Predicted class and Actual class cross-table

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	True Positive (TP)	False Negative (FN) **Type II Error**
	Negative	False Positive (FP) **Type I Error**	True Negative (TN)

3.2 Results for Position/Stance Analysis

Table 4 below shows the performance of the first classification task. Our model achieves an F-score of 0.9178. For the second task, stance detection, our model achieves an F-score of 0.6582. As shown in Table 5, the model is difficult to distinguish the unknown from favor/against stance. Our model achieves a very high performance for detecting whether the topic of nuclear waste is mentioned in the article/comment, achieving an F-score of 0.9563 as indicated in Table 6.

Table 4. Model performance of the first classification task (whether related to nuclear issue or not)

Accuracy	0.9717
Macro-F1	0.9178
Precision	0.9337
Recall	0.9033

模型預測 / 正確標記	有關核能	無關核能
有關核能	693	8
無關核能	14	63

Table 5. Model performance of the second classification task (nuclear policy stance - favor/against/unknown)

模型預測 / 正確標記	支持核能	反對核能	全文有關核能無法判斷	全文無關核能無法判斷
支持核能	169	25	55	0
反對核能	43	118	59	4
全文有關核能無法判斷	43	48	133	4
全文無關核能無法判斷	4	3	7	63

Accuracy	0.6208
Macro-F1	0.6582
Precision	0.6679
Recall	0.6517

Table 6. Model performance of the second classification task (whether nuclear waste mentioned or not)

正確標記 \ 模型預測	有提及核廢料	無提及核廢料
有提及核廢料	688	5
無提及核廢料	8	77

Accuracy	0.9832
Macro-F1	0.9563
Precision	0.9637
Recall	0.9493

4. Discussions and Concluding Remarks

Experimental results (Table 2-6) of the first three stages in Table 1 indicate that all performance indicators can reach 0.9 or above in predicting relevance based on specific keywords. However, machine judgment of the public position/stance (favor/against/unknown) appears most unsatisfactory especially when the predicted position of public opinions is unknown (Table 7). As evidenced by some studies[11], position/stance analysis compared with sentiments analysis of internet public opinions deserves more attention from both academic and research communities particularly seriously utilized as public policy communication and deliberation. The experimental results, as proposed in Table1, do not only call for more iteration of the first three stages but also support the importance of the stage 4 and 5 entailing offline quantitative and qualitative evidence and analysis. Extracting public sentiments and positions can also contribute to public policy communication and deliberation by conducting further qualitative arguments and evidences underlying the public attitudes.

Table 7. Model performance (%) of the second classification task (nuclear policy stance - favor/against/unknown)

Predicted->	Favor	Against	Unknown	Total
Actual favor	169	25	55	249
Actual against	43	118	59	220
Actual unknown	43	48	133	224
Total	255	191	247	693
Predicted->	Favor	Against	Unknown	Total
Actual favor	67.9%	10.0%	22.1%	100.0%
Actual against	19.5%	53.6%	26.8%	100.0%
Actual unknown	19.2%	21.4%	59.4%	100.0%
Total	36.8%	27.6%	35.6%	100.0%
Predicted->	Favor	Against	Unknown	Total
Actual favor	66.3%	13.1%	22.3%	35.9%
Actual against	16.9%	61.8%	23.9%	31.7%
Actual unknown	16.9%	25.1%	53.8%	32.3%
Total	100.0%	100.0%	100.0%	100.0%

The similar methodological reflection has long been explored and practiced by the proponents of mixed methods research[16]. The process proposed in Table 1 is essentially mixed-method as the quantitative-oriented IPO analysis by NLP (stage 2) is embedded in the qualitative-oriented deliberation iterated at stages 1-3 and stages 1-5. Additionally, offline and online sources of public comments and coding such as telephone surveys, focus group interviews, and crowd-sourced annotation may be properly incorporated for better understanding of public opinions[17]. It calls for future implication for inter-disciplinary research and practice.

---

[16] Creswell, J. W., & Plano Clark, V. (2007). ***Designing and conducting mixed methods research***. Thousand Oaks, CA: Sage.

[17] Chen, D. Y., Hsiao, N., Liao, Z. P., & Chen, K. (2016). ***Improving government big data analysis and public policy implementation***. A research report of Taiwan E-governance Research Center & National Development Council. (in Chinese)