# Dynamic stakeholder-associated topic modeling on public concerns in mega infrastructure projects: a case of Hong Kong–Zhuhai–Macao Bridge (HZMB)

Jin XUE[1], Geoffrey Qiping Shen[2], Yiming LI[3], Jiashuo WANG[4], Irfan Zafar[5]

[1]PhD. Candidate, Department of Building and Real Estate, Hong Kong Polytechnic University, Kowloon, Hong Kong. Email: jin.xue@connect.polyu.hk
[2]Chair Professor, Department of Building and Real Estate, Hong Kong Polytechnic University, Kowloon, Hong Kong. Email: bsqpshen@polyu.edu.hk
[3]Research Scientist, College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China, Email: liyiming90@zju.edu.cn
[3]Research Scientist, Department of Research and Development, Atypon System Limited, Oxford, United Kingdom, Email: yli@atypon.com
[4]Research Assistant, Department of Computing, Hong Kong Polytechnic University, Kowloon, Hong Kong. Email: jiashuo.wang@connect.polyu.hk
[5]PhD. Candidate, Department of Building and Real Estate, Hong Kong Polytechnic University, Kowloon, Hong Kong. Email: irfan.zafar@connect.polyu.hk

## Abstract

Public concerns are essential for mega infrastructure project success, meanwhile profoundly influencing on the stakeholder performance. To understand the dynamics of public concerns in the project duration is beneficial for decision-makers and project stakeholders to improve their adaptability in an uncertain project environment. Therefore, the study proposes a dynamic stakeholder-associated topic modeling approach to identify, evaluate, and manage the public concerns by learning the experience from large quantities of unstructured project documents. The approach is composed of three parts, including the Topic Over Time text-mining model, stakeholder relevance scoring system, and managerial map. Based on the managerial map, the proposed "Mirror Z" and "Letter N" strategies guide the decision-maker and project stakeholders to determine the management priority of public concerns considering the public popularity and stakeholder relevance in different project phases. A sixteen-year case of Hong Kong–Zhuhai–Macao Bridge is used to validate the effectiveness of the proposed approach with 1748 official project documents. The dynamics of

16 critical public concerns in the project are revealed and verified by reality. Furthermore, the management guides are summarized for decision-makers, political groups, construction groups, and pressure groups. The study establishes a text-mining based method to provide useful guidelines for future mage infrastructure projects in the view of stakeholder management.

**Keywords**

Public Concern; Topic Modeling; Stakeholder; Hong Kong–Zhuhai–Macao Bridge

## 1. Introduction

Public concern is an informal regulation frequently used by external stakeholders, including government, media, green groups, and concerned citizens (Leonidou et al., 2017). The general public profoundly influences the decision-making process of mega infrastructure projects, either directly or indirectly (Lin et al., 2018). Through an official public consultation, public concerns can directly influence the development of mega infrastructure projects (Boyer et al., 2015). Moreover, public concerns may escalate to indirect appeals with regard to political issues; in this case, it may encourage authorities to interfere with the implementation of mega infrastructure projects (Breukers and Upham, 2015). Previous studies have proved that public concerns heavily influence organizations involved in the projects (Lin et al., 2018). Thus, understanding the relevance between public concerns and stakeholders' participation is of great significance for decision-makers and project stakeholders, since gaining public support through managing related potential risks may benefit them(Rodriguez- Melo and Mansouri, 2011).

Existing studies have performed a static analysis of public concerns through questionnaires, surveys, and interviews; however, they have failed to show the dynamic features involved during the whole project

duration(Lin et al., 2018). Although data from surveys and interviews have revealed patterns of public concerns to some extent, results are still highly dependent on the involved respondents(Vandeweerdt et al., 2016). Respondents' subjective views determine a study's final empirical results. To address the limitation of traditional survey-based research, large quantities of official documents on mega infrastructure projects might be used as a possible dataset to reflect the changeable public concerns in project duration through more objective and convincing empirical data (Zidane et al., 2013). However, few solutions have been provided to explore critical public concerns and their stakeholder relevance from the extensive unstructured text data in the research domain of mega infrastructure projects. Thus, a call for a more efficient approach to analyze official project documents exists.

Given the development of text-mining technology, topic modeling has been utilized as a robust tool to detect core concepts from a considerable amount of texts (Yao et al., 2016). Topic models are established according to various probabilistic theories, including Probability Latent Semantic Analysis (PLSA), Latent Dirichlet Allocation (LDA), and Topic Over Time (TOT) (Li et al., 2015b). TOT model is a dynamic topic model, which does not only explores the contents of core concepts but also reflects the dynamic patterns of concepts (Wang and McCallum, 2006). Therefore, analyzing project documents through the mining of critical issues and presenting the dynamics of various issues in project duration is considered a potential solution. However, given that the current TOT model has few considerations of stakeholders' identifications, it should be further developed to integrate the evaluation function of stakeholders' participation.

The primary purpose of this study is to develop a method that performs dynamic analysis of public concerns and their stakeholder relevance in mega infrastructure projects from unstructured project documents, providing significant management guides from stakeholders' perspectives. This study is divided

into two parts. First, a dynamic stakeholder-associated topic modeling approach (DSTM) is proposed. DSMT is composed of the TOT model, stakeholder relevance scoring system, and managerial map of public concerns. Second, a 16-year longitudinal case study of the Hong Kong-Zhuhai-Macao Bridge (HZMB) is conducted to validate the effectiveness of the proposed approach and reveal the dynamic patterns of public concerns in mega infrastructure projects.

## 2. Background

### 2.1 Public concerns in mega infrastructure projects

Public concerns reflect legitimate demands from public opinions and have been widely discussed concerning public issues, such as environmental protection, technology revolution, and immigration affairs (Carmichael and Brulle, 2017, Wardman and Löfstedt, 2018, McLaren et al., 2017). Previous studies have proved the substantial influence of public concerns toward the decision-making of public policies (Stimson, 2018). Therefore, precautionary measures are recommended for decision-makers to control the risk amplification of public concerns (Wardman and Löfstedt, 2018). Thus, a call for a more robust tool that can guide proactive management exists. One of the shortcomings of existing tools is the lack of effective measurement for public concerns, given that the insufficient and unrepresentative samples in traditional survey-based approach may lead to unreliable findings (Vandeweerdt et al., 2016). To overcome the defect of data source, previous studies have also investigated the strong connection between the congressional voting and public concerns on social issues, wherein the council members always take what their constituents want as primary considerations to determine whether they support the bill or not (Vandeweerdt et al., 2016). Thus, the documents of council meetings may be considered a reliable data source to represent public concerns.

92 Given that most mega infrastructure projects are public-funded, public concerns have significantly
93 affected environmental and social issues concerning mega infrastructure projects (Wang et al., 2016), and
94 thus affecting a wide range of stakeholder behaviors (Li et al., 2013). For instance, public concern influences
95 project planning, which may drive the government to halt and cancel a project, particularly in cases of
96 energy projects (Wu et al., 2019). In the construction stage, public concern drives project stakeholders to
97 adopt social responsibilities (Lin et al., 2018). Moreover, a public concern can be amplified to public
98 opposition by pressure groups (Li et al., 2016). Thus, understanding stakeholder relevance is essential when
99 evaluating the criticalness of public concerns effect on mega infrastructure projects.
100 Given the significance of public concerns, effective public participation has been studied in the
101 field of mega infrastructure projects. According to the ladder of citizen participation proposed by Arnstein
102 (1969), public participation can drive public concern into the decision-making process and achieve
103 democratic and justice governance, legitimate demand, and social justice for mega infrastructure projects
104 (Zhou et al., 2019, Bryson et al., 2013, Fung, 2015). The public participation level varies in different project
105 phases (André et al., 2006), indicating the need for understanding public concerns in the dynamic project
106 environment. To further understand the public concern, interactive participations are considered, including
107 public meetings, public hearings, workshops, negotiation and mediation, and online-based public
108 forums(Lin et al., 2018, Higgs et al., 2008, Rowe and Frewer, 2004). Besides the traditional interactive
109 participation, online public opinions have received popularity among researchers to track public concerns on
110 different social media platforms in the construction industry (Jiang et al., 2016a, Tang et al., 2017). As
111 social media has a significant influence on the daily operation of the construction industry (Azhar et al.,
112 2019), decision-makers use social media platforms to perceive public concerns among the wider population

(Jiang et al., 2016a). To utilize large quantities of online information, the topic modeling approach has been applied to explore significant concepts with the development of text-mining technologies (Tang et al., 2017). However, limitations still exist. First, online views are still probably biased, given that most Internet surfers are commonly composed of younger generations. Second, limited patterns of stakeholder relevance can be obtained due to random and unorganized information found in online communication.

In summary, previous studies have indicated that council documents may reflect public concerns with regard to major events occurring in society. Moreover, considering stakeholder relevance is essential to assess the significance of public concerns. Although topic modeling has been applied to track public concerns in mega infrastructure projects, a call exists for a more robust solution to comprehensively explore the relationship between public concerns and the corresponding stakeholders within the dynamic project duration.

### 2.2 Topic modeling

Topic models are derived from probabilistic graphical models to provide a method that discovers the hidden structure of data (Koller and Friedman, 2009). Traditional topic models such as PLSA (Hofmann, 2001) and LDA (Blei et al., 2003) belong to the Latent Semantic Indexing method. Through indexing, the document-term occurrence matrix can be reduced into low dimensions, which are denoted as latent features. According to these two classical topic models, several topic model-based methods have been proposed and applied in various areas including document classification, recommendation, and transfer learning (Li et al., 2015b, Rubin et al., 2012, Zhuang et al., 2010, Nguyen et al., 2015, Yang et al., 2015, Yao et al., 2018).

Although these models have been widely used in different research domains, unsupervised methods still face challenges concerning dynamic data, which are collected over time (Wang and McCallum, 2006).

LDA model is focused on the co-occurrence of words and their interdependencies but unable to capture the dynamic patterns of the texts (Wang and McCallum, 2006).. Therefore, the time-stamped document cannot be well analyzed by LDA to detect meaningful topics. To overcome the shortcomings of LDA, TOT is developed for exploring the dynamics of the text.

TOT model is established through a continuous distribution over time associated with topics considering word co-occurrence and document timestamps (Wang and McCallum, 2006). Therefore, the time distribution of one topic is determined by the period when a strong word co-occurrence pattern bursts. TOT model has employed beta distribution over time, covering all data and used a Dirichlet distribution to sample each document and word co-occurrence similar to the LDA model (Wang and McCallum, 2006). Thus, the proposed model has been considered as an effective topic model to explore the dynamic pattern of text documents. To measure the validity of the topic model, an automatic coherence measure proposed by Mimno et al. (2011) is used. This measure provides a semantic coherence score based on word co-occurrence among the featured words associated with the latent topic. The measure is used as an alternative method replacing the traditional human judge method to validate the results of the topic model (Chang et al., 2009).

In the research domain of mega infrastructure projects, the topic model has been used as a mining approach to review the literature and analyze online public opinions (Jiang et al., 2016b, Jiang et al., 2016a). However, a static topic model such as LDA remains as the dominating method used to explore information (Jiang et al., 2016c). Thus, a call for a more dynamic model exists to reflect further how information changes within the development of the project. Moreover, existing studies have focused on the mining and interpretation of results in a general way without specifying the information from stakeholders' perspectives.

Therefore, learning through investigating past experiences from existing project documents is essential for project stakeholders. Furthermore, although the data-mining approaches have been extended to extract information from project documents (Le et al., 2018, Liu et al., 2018b), it is still a lack of a robust tool on analysis of text data among unstructured project documents. Hence, it calls for further study of the topic model on a large number of unstructured text project documents.

## 3. Method

### 3.1 Research framework

The framework of the DSTM is presented in Figure 1. First, data were collected from the official project documents. Second, the TOT model was employed to detect the public concerns in the project, through which it was obtained the critical public concerns and the annual trend of each public concern. Third, according to the identified public concerns, we developed a scoring system to measure the link of each public concern and the corresponding stakeholders. Fourth, we generated the stakeholder relevance score for each public concern. Fifth, we designed a managerial map for public concerns in mega infrastructure projects considering the annual popularity and stakeholder relevance of each concern, providing relevant guides for decision-makers and project stakeholders in managing public concerns in different stages of mega infrastructure projects.

**<Figure 1. The framework of DSTM>**

### 3.2 Data collection

Official project documents are used as the primary data sources in conducting dynamic analysis, given that they provide time-stamped information of public concerns in the changeable environment of mega infrastructure projects (Lee et al., 2017). These documents are required to meet two criteria. First, the

content of these documents should be highly relevant to public concerns. Second, these documents should follow the text format given that the text-mining approach will be applied to implement data analysis.

In this study, the official project documents in the council library are chosen as the pool of datasets. Given that the legislative council is an institution wherein the council members who represent the general public participate in the social governance (Vandeweerdt et al., 2016), council documents provide a unique opportunity to reflect the major public concerns in an official way. Moreover, most of the documents of the legislative council are open for the public in a democratic society. For instance, the archives of local major infrastructure projects can be explored in a variety of congress libraries in the United States (https://www.congress.gov/congressional-record), Singapore (https://sprs.parl.gov.sg/search/home), United Kingdom (https://archives.parliament.uk/), Canada (https://lop.parl.ca/sites/PublicWebsite/default/en_CA/), Hong Kong (https://www.legco.gov.hk/general/english/library/index.html), etc. Thus, researchers can access council documents, which contain reliable longitudinal information of public concerns related to mega infrastructure projects.

### 3.3 TOT model

TOT model was developed based on the traditional LDA. TOT model detects the topic of the target text, considering not only the word co-occurrences but also temporal information. Therefore, the TOT model is used to explore the topic distributions in the documents with the timestamp, reflecting the dynamic features of the detected topic (Wang and McCallum, 2006). The graphical representation of TOT, as proposed by Wang and McCallum (2006), is shown in Figure 2, whereas the summary of the notations used in this section is shown in Table 1. The generative process used in Gibbs sampling for parameter estimation is stated as follows:

1. For each topic $z$, draw a multinomial distribution $\emptyset_z$ from a Dirichlet prior β.
2. For each document $d$, draw a multinomial distribution $\theta_d$ from a Dirichlet prior α.
3. For each word in document $d$:
   (a) Draw a topic $z_{di}$ from multinomial $\theta_d$,
   (b) Draw a word $w_{di}$ from multinomial $\emptyset_{z_{di}}$,
   (c) Draw a timestamp $t_{di}$ from Beta $\varphi_{z_{di}}$.

In this study, the setting of hyper-parameter α and β is the same as in TOT suggested by Wang and McCallum (2006), wherein α = 50/T and β = 0.1. Moreover, data preprocessing is conducted with term frequency-inverse document frequency (TF-IDF) filtering before running the TOT model to improve the quality of the text, which has removed some frequent but meaningless words (Ramos, 2003).

**<Table 1 Notations of the symbols in TOT>**

**<Figure 2. Graphic representation of the TOT model for Gibbs sampling>**

After the Gibbs sampling procedure, it is obtained three probability matrixes: document-topic probability matrix, topic-word probability matrix, and topic-time probability matrix. The first one is the probability distribution over the topics for each document. The second one is the probability distribution over the words for each topic. The last one is the probability distribution over time for each topic.

To describe each topic, 15 most probable feature words from the topic-word probability matrix are extracted and listed in descending order of their probabilities related to the topic, which are generated by the TOT model. The threshold setting of the top 15 feature words to interpret the topic is based on two considerations. On the one hand, the number of selected feature words varies in the previous studies with the scale of top 5, top 10, top 15, and top 20, which proves the range of selected feature words is not sensitive to

the research results (Jiang et al., 2016b, Wang and McCallum, 2006, Mimno et al., 2011, Jiang et al., 2016c). On the other hand, the coherence validity test of TOT modeling is suggested to be undertaken with the top 15 feature words (Mimno et al., 2011). Therefore, the top 15 featured words were selected, in consideration of the concept interpretation and the modeling validity test requirement.

To measure the validity of the coherence and typicality of the topics generated by the TOT model, a score called topic coherence is used to assess the results (Mimno et al., 2011). Given the feature word list $V^{(z)}$, topic coherence is calculated for each topic $z$ through the following Formula 1:

$$C\left(Z; V^{(z)}\right) = \sum_{m=2}^{15} \sum_{l=1}^{m-1} log \frac{D\left(v_m^{(z)}, v_l^{(z)}\right)+1}{D\left(v_l^{(z)}\right)}, \quad (1)$$

where $\mathrm{D}(\mathrm{v})$ represents the number of documents the feature word *v* appears, and $\mathrm{D}(v_m, v_n)$ is the number of documents that contain feature words $v_m$ and $v_n$. An absolute value of coherence that is closer to 0 indicates a more representative topic.

### 3.4 Scoring system for stakeholder relevance

When displaying the relative importance of each stakeholder and identified topic *Z*, a relevance score is utilized. Given the topic $Z$ and the timestamp $Y$ and the stakeholder word $v_h$, we calculate the probability of how much $v_h$ contributes to each feature word $v_w$ of topic $Z$ in the documents from timestamp $Y$. The probability is calculated by counting the frequency of paragraphs where stakeholder $v_h$ occurs given the paragraphs containing feature word $v_w$ for all documents from timestamp $Y$ has the feature word $v_w$ occurs. We multiply this probability and the probability of $v_w$ under topic $Z$, which has already been analyzed by the TOT model. Finally, the relevance score is formulated as the sum of the product of these multiplications. Let $\mathrm{S}(v_w, d)$ be the number of paragraphs that $v_w$ appears in the document $d$, $\mathrm{S}(v_m, v_n, \mathrm{d})$ be the number of paragraphs that $v_m$ and $v_n$ appear together in the document

$d$, $D(v,Y)$ be the number of documents containing word $v$ with the timestamp $Y$, and $p(v_w,Z)$ be the probability of feature word $v_w$ of topic $Z$, which is derived from the topic-word probability matrix. The score is defined in the Formula 2 as follows:

$$S(v_h,Y,Z)=\sum_{W=1}^{W}\frac{\sum_{d=1}^{D(Y)}\frac{S(v_w,v_h,d)}{S(v_w,d)}}{D(v_w,Y)}*p(v_w,Z), \quad (2)$$

where D(Y) is the number of documents with timestamp $Y$. In our case study, it is calculated the relevance score of 29 kinds of stakeholders listed in Table 2, which is generated by fast browsing the content of project documents.

**<Table 2. 29 kinds of stakeholders>**

### 3.5 Proposed managerial map for public concerns

A managerial map for public concerns called MPC-Map is designed to show the importance degree of public concerns with a specific timestamp according to the relationship with the popularity of the public concerns and the scores of stakeholder relevance. The vertical axis represents the popularity of each topic using the average annual document correlated probability of the corresponding topic, whereas the horizontal axis represents the relevance of the stakeholders using the sum of stakeholder relevance scores related to each topic. The graph is divided into four blocks by the average of horizontal and vertical values of all points. In Figures 3 and 4, the managerial map is divided into four zones, which classifies the criticalness of public concerns according to four levels. Zone One is composed of public concerns with high popularity and stakeholder relevance. Zone Two includes public concerns with high popularity and limited stakeholder relevance. Zone Three represents the public concerns with low popularity and wide stakeholder relevance. Zone Four has public concerns with little popularity and limited stakeholder relevance.

Two strategies have been proposed based on the managerial map. The first one is for decision-makers, whereas the other one is for each kind of stakeholders. The decision-makers focus more on the trend popularity of public concerns rather than the stakeholder relevance. Thus, addressing the issues receiving high public attention is essential to decision-makers. Each project stakeholder always considers its interest first, therefore paying more attention to the stakeholder relevance. The divergence on the focus between decision-makers and project stakeholders leads to different priority paths on the managerial map when ranking the criticalness of public concerns. For decision-makers, the management priority on public concerns is to address first the high popularity issues from Zone One to Zone Four, which is similar to the mirror image of the letter "Z" shown in Figure 3. For project stakeholders, the management priority on public concerns is to address first the high stakeholder-relevant issues from Zone One to Zone Three to Zone Two to Zone Four, which is similar to the letter "N" as shown in Figure 4. Therefore, the "Mirror Z" and "Letter N" strategies are proposed to determine the management priority of public concerns in mega infrastructure projects for decision-makers and project stakeholders, respectively.

**<Figure 3. "Mirror Z" approach for the management of public concerns>**

**<Figure 4. "Letter N" approach for the management of public concerns>**

## 4. Case study

HZMB is a 55-km cross-boundary mega transport project, connecting Hong Kong, Macau, and Zhuhai—the three major cities located at the Pearl River Delta in China. HZMB was built to enhance the economic and sustainable development of the Greater Bay Region. The project was planned in 2008, and the construction work began on 15 December 2009 on the Guangdong side. Nine years later, the construction work was completed on 6 February 2018 and was opened to the public on 24 October 2018. With the

proposed DSTM, we conducted a dynamic analysis of public concerns in HZMB using 1748 official documents from 2003 to 2018. Through the proposed DSTM approach, we detected the public concerns and their trends in the timeline and drew the managerial map to reflect the importance for each public concern in three different project phases: planning (2003–2009), construction (2010–2017), and handover (2018).

### 4.1 Results of the data collection

As the figure (shown in Appendix 1) shows, we downloaded the project documents related to HZMB from the open-source library of Hong Kong Legislative Council by searching the terms: "Hong Kong–Zhuhai–Macao Bridge" and "HZMB." Consequently, it was obtained the raw documents listed in the figure (shown in Appendix 2), which would be used for further data cleaning.

Given that unrelated contents to the project were found in the collected raw documents, we did a data cleaning by extracting the paragraphs from the raw files with the keywords related to HZMB, ensuring the high relevance between data and the project. As a result, we collected a total of 1748 official documents on HZMB from 2003 to 2018, including the planning, construction, and handover stages. The number distribution of collected documents is shown in the figure (shown in Appendix 3), which presents the number of documents in each year. According to the classical literature of TOT modeling, the model is proved to be valid from the mega dataset of State-of-Union Addresses to the medium dataset of email collections, and the small dataset of conference proceedings, indicating the size of the dataset is not the constraint of the validity of TOT model (Wang and McCallum, 2006). Therefore, the number of documents in the case (1748) is adequate for the modeling research, which is also in line with the topic model research by the similar size of the dataset (Jiang et al., 2016c).

### 4.2 Critical public concerns

Through TOT modeling, the 16-topic model was considered as the most valid model to interpret the documents by the sensitivity analysis according to the topic coherence score shown in Table 3. Therefore, 16 topics on public concerns were explored, as presented in Table 4. Each topic has corresponding top 15 featured words that exhibited the highest relevance with the clustered topic, indicating the concept of the topic. For instance, clustered topic 0 had the highest relevant feature words "align" and "*lantau*," which represent the landing location of HZMB. Moreover, other words such as "tunnel," "environment," and "finance" reflect the considerations in the alignment, thus leading to the concept of topic 0 as "alignment."

In light of the featured words, we interpreted the concept of each topic as follows: alignment, local community, cooperation across governments, environment, immigration and custom service, operation on cross-boundary traffic, local traffic in connective areas, local industry, operation on local traffic, urban development, finance, construction, design, regional economy, and commercial development. The identified topics represent the major public concerns in the project duration on HZMB from 2003 to 2018.

**<Table 3. Distribution of coherences for different topic numbers>**

**<Table 4. Top 15 featured terms for each topic in the 16-topic model>**

### 4.3 Trend of public concerns

Given the consideration of timestamp information in each document by the TOT model, a trend of public concerns is detected, illustrating the dynamic patterns of each public concern in the project duration. The trend of each public concern is shown in Figure 5. The figure reflects the average correlation probability of the documents under each topic in the timeline derived from the document-topic probability matrix of TOT results, reflecting the annual popularity of the topics in the dataset. Generally, public concerns could be divided into three categories due to the various level of popularity for each public concern in HZMB. The

detected dynamic features can be explained with the development history of HZMB, which proves the validity of the proposed stakeholder-associated topic modeling approach. The detailed interpretations are as follows.

**<Figure 5. The trend of public concerns in HZMB>**

First, the most critical public concern is from the cooperation across governments (#2), which attracts high popularity from 2008 to 2014. This period starts from the late planning stage to the early construction stage. Given that the period is at the decision-making phase of HZMB, policies and agreements are reached among three governments: Hong Kong SAR, Macao SAR, and Guangdong Province. Therefore, the performance of governmental cooperation has received substantial public attention and become the only hot topic among public concerns during the given period.

Second, environment (#3), local connectivity (#9), finance (#11), and commercial development (#15) have received medium popularity among critical public concerns. Interestingly, all these four kinds of concerns have received their peaks at the early and late phases of the project duration. First, the environmental issues (#3) have become the focus when the worries of the adequate protection measures on Chinese White Dolphin, one endangered species, have been raised among the public in the early planning stage of HZMB. The issue resurfaces given that the project has been faced with the juridical review of its environmental impact assessment that caused severe project delays in the early construction stage. Besides, the concern of local connectivity (#9) is mainly from the residents in the neighborhood of the project. At the beginning of the planning stage, the voice of simultaneous completion of local connectivity and HZMB has been raised by local residents with the willingness to upgrade the less-developed road system within the areas affected by the projects. However, the local residents' advocating voice turned to worries as the

344 construction of the local connectivity faced project delays when HZMB is close to completion. A similar
345 situation related to commercial development (#15) has occurred, which raises public concern. Given that
346 Hong Kong is an advanced commercial city, one commercial development plan attached with HZMB called
347 "bridgehead economy" has been proposed in the early planning stage, which has gained different opinions
348 from the public. However, the plan has faced fierce debate among the public due to the delay and the
349 technical difficulties at the late construction stage of HZMB. Furthermore, financial arrangement (#11) is
350 another critical public concern. The financial pressure caused by HZMB is widely discussed among public
351 in the initial planning stage, wherein the public worries have escalated due to the severe cost overrun of the
352 project at the end of the construction stage.

353 Third, besides the five groups of public concerns mentioned above, the other 11 types of public
354 concerns attract the third-level popularity, reflecting the dynamics of abundant concerns from the public.

355 **4.4 Managerial map for public concerns**

356 The managerial map of public concerns presents the information on the popularity and stakeholder
357 relevance for each public concern in three different stages: planning (2003–2009), construction (2010–2017),
358 and handover (2018). As shown in Figure 6, the managerial map is provided for decision-makers to learn the
359 management priority of public concerns in the development of HZMB with the "Mirror Z" approach.
360 Besides, Figures 7–9 indicate the management priority of public concerns for a variety of project
361 stakeholders using the "Letter N" approach, including political groups, construction groups, and pressure
362 groups. The instructions on management priority would benefit decision-makers and project stakeholders to
363 understand the criticalness of public concerns and provide the management guide for public concerns in
364 future mega infrastructure projects.

**<Figure 6. Managerial map of public concerns for decision-makers>**

4.4.1 Decision-makers

Decision-makers overview the general stakeholder relevance for each public concern to understand the stakeholder participation hereby determining the management priority based on the MPC-Map by the "Mirror Z" strategy. In Figure 6, the stakeholder relevance in the x-axis represents the involvement level of all project stakeholder groups listed in Table 2. According to the results in Figure 6, the management guide for decision-makers in each project phase is discussed as follows.

4.4.1.1 Planning Stage

In the planning stage, one of the most critical public concerns in Zone One is cooperation across regional governments (#2). Given that the political issues are sensitive to the success of mega infrastructure projects (Flyvbjerg, 2014), government cooperation always attracts much attention from the general public, especially in the cross-boundary transport project, which also influences the interest of stakeholders. Hence, government cooperation is quite significant for maintaining a good working relationship among governments across different regions and keeping the collaboration information transparent toward the public. In response to the broad public and stakeholder focus on the governmental cooperation issue, decision-makers need to disclose timely information to dismiss the public worries on "closed decisions" made by authorities (Wu et al., 2019). The instant and transparent information are also recommended to be disseminated by multiple accesses: press, websites, and social media (Li et al., 2012b, Leung et al., 2013), which could benefit a wide range of stakeholders to receive the latest project progress under the political level to make their corresponding preparations accordingly.

Another critical public concern in Zone One is environmental issues (#3). This study supports the

386 findings of the existing research on environmental issues, which are considered as an ongoing concern
387 among various stakeholders in the early site selection phase (Min et al., 2018). The environmental impact of
388 the project is a serious concern among the public, indicating the criticalness of early public participation in
389 the planning stage (Diduck et al., 2013, Herian et al., 2012). Therefore, the decision-makers are suggested to
390 build a widely accepted process for environmental impact assessment through the frequent public
391 consultations among project stakeholders (Liu et al., 2018a). Besides, the proper public petition mechanism
392 is also required to assist the local residents and green groups in conveying their worries and receiving the
393 feedback from the authorities (Xie et al., 2014).

394 Moreover, commercial development (#15) is a focal public concern in Zone Two with limited
395 stakeholder relevance. The primarily affected stakeholders on this issue are the government and local
396 residents. The precise explanation of positive local economic benefits has been proved to be useful to gain
397 public support (Tummers, 2011, Valentin et al., 2018). In HZMB, one proposal of the development of
398 commercial facilities was initiated by decision-makers in the project planning, which received the warm
399 welcome of local communities.

400 4.4.1.2 Construction Stage

401 In the construction stage, besides the three focal public concerns (government cooperation,
402 environment, and commercial development), another two public concerns have appeared in Zone Two: local
403 connectivity (#9) and finance (#11).

404 The local connectivity (#9) is introduced by the worries of local residents on the potential heavy
405 traffic pressures caused by the completion of HZMB, which escalated when it was closer to the project due
406 date. Since the expert power of technicians and scientists are capable of enhancing the trust between relevant

stakeholders (Leung et al., 2013), the decision-makers are suggested to invite professionals to communicate with the local public to ease their worries.

The financial issue (#11) is caused by the public doubts on financial viability facing the severe cost overrun of HZMB. The number of relevant stakeholders is limited, including the government and contractors. Hence, the focus group meeting is recommended to arrange between relevant stakeholders and public representatives to achieve the collaborations and consensus in the process of applying additional bills to continue the project (Rowe and Frewer, 2005).

4.4.1.3 Handover stage

In the handover stage, the top critical public concerns in Zone One include the local community (#1), local industry (#7), custom and immigration affairs (#4), and design (#13).

As the project is getting closer to the operation, the realization level for the original planning proposal has started to attract substantial public concerns, which introduces the reviews of the promised economic benefits for the local community (#1) and the performance of project design (#13). The increase in public attention has also exerted the pressure for all internal project stakeholders (contractors, subcontractors, suppliers, and consultants) and the government to prepare for a proper explanation of the divergence between the original planning proposal and the actual project toward the public. Therefore, decision-makers need to establish sufficient public communication access (i.e., public hotline and public email) together with all relevant stakeholders to respond to the complaints from the public instantly (Wang, 2001, Rowe and Frewer, 2005).

Another two public concerns are relevant to the local industry (#7) and custom and immigration affairs (#4). The two issues are referred to the technical arrangements of operational management for HZMB.

Therefore, the roundtable negotiation meeting has the potential to address the concerns by inviting all the relevant professionals and public representatives (Xie et al., 2019).

**<Figure 7. Managerial map of public concerns for political groups>**

4.4.2 Political Groups

Political groups hold various political powers in the council system of Hong Kong, including legislative council, district council, the ruling party, and the opposition party. In Figure 7, the stakeholder relevance in the x-axis represents the involvement level of the abovementioned four political groups. According to the results in Figure 7, the management guide for political groups in each project phase is discussed using the "Letter N" strategy.

Compared with the general stakeholder relevance shown in Figure 6, the concern of local connectivity (#9) moves from Zone Two to Zone One in the planning stage, as shown in Figure 7. Thus, the issue of local links has become one of the most critical public concerns for political groups. The development synchronization of local transport systems connecting the HZMB with other urban areas of Hong Kong is one of the significant issues fiercely debated among various political groups. The issue has heavily influenced the final pass of bills related to the project in the planning stage. Therefore, the public-participated advisory committee is recommended to be founded by inviting different political groups and the representatives of local communities to boost the satisfactory development plan of local connectivity with the planning of mega infrastructure projects (Plummer and Taylor, 2013, Webler and Tuler, 2006).

The public concern on the local industry (#7) is located in Zone Three with the highest stakeholder relevance in the construction stage, which makes it the only critical concern located in Zone One in the handover stage. The criticalness of public concerns on the local industry is also validated by the

449 development history of HZMB. The local industry, which mainly refers to the logistics and tourism industry,
450 is a significant issue in the council, given the economic benefits brought by the construction of HZMB in
451 Hong Kong. Therefore, the council has been regularly debating the proposed policies on the development of
452 the local industry. The debate among political groups is initiated at the start of the construction stage and
453 intensified at the handover stage when the project is getting closer to completion and operation. As shown in
454 Figure 7, the issue of the local industry attracts less public attention in the construction stage before it
455 becomes a hot topic among the public in the handover stage. Thus, the construction stage would be a
456 suitable period to hold the roundtable negotiation forums between political groups and the representatives of
457 relevant local industry (i.e., logistics, tourism) (Xie et al., 2014), for reaching the acceptable agreement
458 among relevant stakeholders concerning the development strategy of local industry.

459 **<Figure 8. Managerial map of public concerns for construction groups>**

460 4.4.3 Construction Groups

461 Construction groups are composed of different participants who are directly involved in the
462 construction work of HZMB, including contractors, consultants, subcontractors, suppliers, and workers. In
463 Figure 8, the stakeholder relevance in the x-axis represents the involvement level of the abovementioned
464 five construction groups. According to the results in Figure 8, the management guide for construction groups
465 in each project phase is discussed by the "Letter N" strategy.

466 The public concern of design (#13) has great relevance to construction groups, which is located in
467 Zone Three, with the highest stakeholder relevance during the planning and construction stages. The public
468 popularity of the project design dramatically increases in the handover stage, causing the issue to move to
469 Zone One. The mapping result reflects the close correlation between construction groups and design issues,

given that the project design is highly relevant to the behavior of construction groups in the project (Zhao et al., 2016). Less public attention to design issues in the planning and construction stages would not disturb the activities of construction groups. However, the result indicates that the construction groups should still pay special attention to public doubts on design issues in the handover stage. When the mega infrastructure project is close to the operation, the performance of design becomes a focal issue in the media, attracting mass attention among the general public. In the handover stage of HZMB, the media report of some design errors, such as the unexpected movement of the seawall and falsification of the cement report, has raised severe public concern toward the construction groups. Thus, the active feedback of negative news and reports to the public is essential for construction groups to ease public worries (Wu et al., 2019). Besides, the public exhibition of project achievements, such as the awards and new records gained by project design, is recommended to be held to improve the civic pride on the project (Xie et al., 2019).

Another interesting finding is from the construction issue (#12), which is located in Zone Four and considered the least critical issue for construction groups in the planning and construction stages. The technical professions of construction issues cause less public attention (y-axis). The general public would not focus too much on construction activities due to their lack of engineering knowledge. Moreover, the public tends to blame the decision-makers rather than the construction groups concerning poor construction performance (i.e., cost overrun, time delay), which leads to the low relevance between construction issues and construction groups in the x-axis of the map. In the view of public concerns, the result indicates that the general public does not heavily influence detailed construction activities. However, the construction issue is still a critical concern for construction groups, given that the technical difficulties are the significant features of mega infrastructure projects (Flyvbjerg, 2014).

**<Figure 9. Managerial map of public concerns for pressure groups>**

4.4.4 Pressure Groups

Pressure groups are composed of external stakeholders who pose political oppositions on the project, including environmental groups, fisherman associations, local communities, and media(Li et al., 2015a). In Figure 9, the stakeholder relevance in the x-axis represents the involvement level of the abovementioned four pressure groups. According to the results in Figure 9, the management guide for pressure groups in each project phase is discussed by the "Letter N" strategy.

In the planning stage, the construction (#12) and design (#13) issues (Zone Three) are the two concerns with the highest stakeholder relevance of pressure groups in the managerial map. The pressure groups are afraid of the disturbance brought by the construction of mega infrastructure projects (Li et al., 2012a). Hence, they emphasize the significance of protection measures into the project design and construction to diminish various adverse effects. In HZMB, the pressure groups have negotiated with decision-makers in a variety of social disturbance issues caused by the construction and design, including the noise led by the coming construction works and the inadequate technical measures to protect the involved marine life in the project design. Hence, pressure groups are encouraged to arrange several rounds of public consultation and polling with decision-makers to convey their worries and demands to improve the quality of design and construction (Norton and Hughes, 2017, Xie et al., 2019). Furtherly, the construction (#12) and design (#13) issues are placed in Zone One in the handover stage because of the review of project performance led by pressure groups. The public-participated evaluation team is advised to build for the joint work between decision-makers and pressure groups to analyze the performance of project execution and urge the authorities to make up for the shortcomings timely (Wang, 2001, Xie et al., 2014).

Another finding is the issue of cross-boundary traffic arrangement (#5), which is located in Zone Three in the construction stage and upgraded to Zone One in the handover stage. The result indicates that the issue of the operational arrangement of mega infrastructure projects has already received full attention among pressure groups in the construction stage, although not acquiring much public focus. In HZMB, the arguments on whether the cross-boundary traffic should be right-driving followed by mainland regulation or left-driving maintained with the Hong Kong regulation have already been discussed by media in the construction stage. In the same period, the worries of gas-emission caused by heavy cross-boundary traffic after the completion of HZMB have been presented by local communities and environmental groups. Before these arguments escalate as major public concerns in the handover stage, public hearings and consultations are recommended to organize in the construction stage to address the conflict of operational arrangement between decision-makers and pressure groups (Boudet et al., 2011, Webler and Tuler, 2006).

## 5. Managerial contributions and Implications

### 5.1 Contributions for stakeholder management on public concerns

This study provides a method for decision-makers to undertake stakeholder management on public concerns in mega infrastructure projects with exploring knowledge from unstructured official project documents. Around the public concerns, the proposed dynamic stakeholder-associated topic model (DSTM) benefits the decision-makers to identify stakeholder issues, assess the stakeholder relevance, and obtain the stakeholder management strategies (shown in Figure 10). With the database of unstructured official text documents in the mega infrastructure projects, the decision-makers first use the method of TOT modeling to identify the critical public concerns and their dynamic patterns automatically. Then, a scoring system is provided for decision-makers to assess the relevance between project stakeholders and each public concern

533 in the timeline. Finally, the managerial map for public concerns is presented for decision-makers to conduct
534 the stakeholder management in each project phase with the “Mirror Z” and “Letter N” approaches. The
535 “Mirror Z” approach is useful for decision-makers to determine the managerial priority of public concerns,
536 while the “Letter N” approach is to guide the managerial priority for each specific project stakeholder.

537 **<Figure 10. The framework of Stakeholder management on public concerns >**

538 5.2 Implications for project stakeholders on managing public concerns

539 The managerial implications (shown in Figure 11) are summarized based on the interpretations of
540 research results, which provides the reference for project stakeholders on managing public concerns in the
541 future mega infrastructure projects. The implications (shown in Figure 11) benefit project stakeholders from
542 three aspects. First, it indicates the critical public concerns for stakeholders in each project phase, as all the
543 included concerns are allocated in the first and second criticalness zones based on “Mirror Z” and “Letter N”
544 approaches. Second, the management strategies are proposed for stakeholders to address the challenges
545 brought by public concerns with effective public involvement methods. Third, the precautionary measures
546 are highlighted for stakeholders based on the dynamic analysis of the managerial map, as the corresponding
547 concerns may attract much more public attention and critics in the later project phase. For instance, the
548 roundtable negotiation forums on the development of the local industry are suggested for political groups to
549 arrange in the construction stage, before the concern would upgrade to Zone One in the handover stage
550 stated in 4.4.2. A similar precaution is also provided for pressure groups on the public concern of
551 cross-boundary traffic arrangements. With the precautionary intervention, the implication is helpful for
552 stakeholders to manage public concerns in advance.

553 **<Figure 11. Managerial implications for project stakeholders >**

## 6. Conclusion

Efficient dynamic analysis is a bottleneck of studies on public concerns. Thus, the present study proposes a DSTM approach to fill the gap by using longitudinal text data from the official project documents. The proposed method is composed of three modules. First, the TOT model identifies critical public concerns and automatically tracks its trend within the development of mega infrastructure projects. Second, a stakeholder relevance scoring system is used to assess the relationship between public concerns and project stakeholders in each time point of the project duration. Third, a managerial map of public concerns is developed based on concern trend and stakeholder relevance derived from the first and second modules in three different stages of the project. According to the managerial map, determining the management priority through "Mirror Z" and "Letter N" strategies is beneficial for decision-makers and project stakeholders when facing various kinds of public concerns.

Through the case study with HZMB, 16 critical public concerns have been identified, wherein the popularity trend of each public concern is tracked. The trend has revealed that the cooperation among governments has the highest public attention during the decision-making period. In contrast, public concerns on the environment, local connectivity, finance, and commercial development have two peaks of public popularity in the early and late stages of the project. All the dynamic patterns are proved by the development history of HZMB, which validates the effectiveness of the DSTM approach. Moreover, the managerial maps are provided for decision-makers, political groups, construction groups, and pressure groups to further understand the management priority of public concerns in three different stages of HZMB: planning, construction, and handover.

The present study has two significant contributions. First, a new method is proposed to generate the managerial map by detecting the dynamic patterns and stakeholder relevance of public concerns from large quantities of unstructured text documents in mega infrastructure projects. Second, "Mirror Z" and "Latter N" strategies are proposed to manage public concerns in three stages during the project duration in mega infrastructure projects considering the concern popularity and stakeholder relevance. Therefore, the DSTM approach realizes the dynamic public concern analysis for decision-makers and project stakeholders from identification to evaluation and management. The proposed method could be extended to further investigate the experience of various mega infrastructure projects with open official text documents in different regions. Moreover, the proposed method has a potential application on data available from news websites, social network media, and other text sources, which can be used to further understand different public concerns comprehensively.

## Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request. The data refers to the longitudinal text documents of HZMB which was used to examine the validity of the proposed method.


## Acknowledgement

The research discussed in this paper is fully supported by the National Natural Science Foundation of China (Grant No.71671156).


## Appendix

**<Appendix 1. The example of a database of legislative council records>**

**<Appendix 2. The example of raw documents related to HZMB>**

595 **< Appendix 3. The number distribution of collected documents from 2003 to 2018>**

596 **References**


597 ANDRé, P., ENSERINK, B., CONNOR, D. & CROAL, P. 2006. Public participation international best practice
598 principles. *Special publication series,* 4.
599 ARNSTEIN, S. R. 1969. A ladder of citizen participation. *Journal of the American Institute of planners,* 35**,**
600 216-224.
601 AZHAR, S., RIAZ, Z. & ROBINSON, D. 2019. Integration of social media in day-to-day operations of
602 construction firms. *Journal of Management in Engineering,* 35**,** 06018003.
603 BLEI, D. M., NG, A. Y. & JORDAN, M. I. 2003. Latent dirichlet allocation. *Journal of machine Learning*
604 *research,* 3**,** 993-1022.
605 BOUDET, H. S., JAYASUNDERA, D. C. & DAVIS, J. 2011. Drivers of Conflict in Developing Country
606 Infrastructure Projects: Experience from the Water and Pipeline Sectors. *Journal of Construction*
607 *Engineering and Management,* 137**,** 498-511.
608 BOYER, E. J., VAN SLYKE, D. M. & ROGERS, J. D. 2015. An empirical examination of public involvement in
609 public-private partnerships: Qualifying the benefits of public involvement in PPPs. *Journal of Public*
610 *Administration Research and Theory,* 26**,** 45-61.
611 BREUKERS, S. & UPHAM, P. 2015. Organisational aspects of public engagement in European energy
612 infrastructure planning: the case of early-stage CCS projects. *Journal of environmental planning and*
613 *management,* 58**,** 252-269.
614 BRYSON, J. M., QUICK, K. S., SLOTTERBACK, C. S. & CROSBY, B. C. 2013. Designing public participation
615 processes. *Public administration review,* 73**,** 23-34.
616 CARMICHAEL, J. T. & BRULLE, R. J. 2017. Elite cues, media coverage, and public concern: an integrated path
617 analysis of public opinion on climate change, 2001–2013. *Environmental Politics,* 26**,** 232-252.
618 CHANG, J., GERRISH, S., WANG, C., BOYD-GRABER, J. L. & BLEI, D. M. Reading tea leaves: How humans
619 interpret topic models. Advances in neural information processing systems, 2009. 288-296.
620 DIDUCK, A. P., PRATAP, D., SINCLAIR, A. J. & DEANE, S. 2013. Perceptions of impacts, public participation,
621 and learning in the planning, assessment and mitigation of two hydroelectric projects in
622 Uttarakhand, India. *Land Use Policy,* 33**,** 170-182.
623 FLYVBJERG, B. 2014. What you should know about megaprojects and why: An overview. *Project*
624 *management journal,* 45**,** 6-19.
625 FUNG, A. 2015. Putting the public back into governance: The challenges of citizen participation and its
626 future. *Public Administration Review,* 75**,** 513-522.
627 HERIAN, M. N., HAMM, J. A., TOMKINS, A. J. & PYTLIK ZILLIG, L. M. 2012. Public participation, procedural
628 fairness, and evaluations of local governance: The moderating role of uncertainty. *Journal of Public*
629 *Administration Research and Theory,* 22**,** 815-840.
630 HIGGS, G., BERRY, R., KIDNER, D. & LANGFORD, M. 2008. Using IT approaches to promote public
631 participation in renewable energy planning: Prospects and challenges. *Land Use Policy,* 25**,** 596-607.
632 HOFMANN, T. 2001. Unsupervised learning by probabilistic latent semantic analysis. *Machine learning,* 42**,**
633 177-196.
634 JIANG, H., QIANG, M. & LIN, P. 2016a. Assessment of online public opinions on large infrastructure projects:

A case study of the Three Gorges Project in China. *Environmental Impact Assessment Review,* 61**,** 38-51.
JIANG, H., QIANG, M. & LIN, P. 2016b. Finding academic concerns of the Three Gorges Project based on a topic modeling approach. *Ecological indicators,* 60**,** 693-701.
JIANG, H., QIANG, M. & LIN, P. 2016c. A topic modeling based bibliometric exploration of hydropower research. *Renewable and Sustainable Energy Reviews,* 57**,** 226-237.
KOLLER, D. & FRIEDMAN, N. 2009. *Probabilistic graphical models: principles and techniques*, MIT press.
LE, T., LE, C. & DAVID JEONG, H. 2018. Lifecycle data modeling to support transferring project-oriented data to asset-oriented systems in transportation projects. *Journal of Management in Engineering,* 34**,** 04018024.
LEE, C., WON, J. W., JANG, W., JUNG, W., HAN, S. H. & KWAK, Y. H. 2017. Social conflict management framework for project viability: Case studies from Korean megaprojects. *International Journal of Project Management,* 35**,** 1683-1696.
LEONIDOU, L. C., CHRISTODOULIDES, P., KYRGIDOU, L. P. & PALIHAWADANA, D. 2017. Internal drivers and performance consequences of small firm green business strategy: The moderating role of external forces. *Journal of Business Ethics,* 140**,** 585-606.
LEUNG, M. Y., YU, J. Y. & LIANG, Q. 2013. Improving Public Engagement in Construction Development Projects from a Stakeholder's Perspective. *Journal of Construction Engineering and Management,* 139.
LI, T. H., NG, S. T. & SKITMORE, M. 2012a. Conflict or consensus: An investigation of stakeholder concerns during the participation process of major infrastructure and construction projects in Hong Kong. *Habitat international,* 36**,** 333-342.
LI, T. H., NG, S. T. & SKITMORE, M. 2012b. Public participation in infrastructure and construction projects in China: From an EIA-based to a whole-cycle process. *Habitat international,* 36**,** 47-56.
LI, T. H., NG, S. T. & SKITMORE, M. 2013. Evaluating stakeholder satisfaction during public participation in major infrastructure and construction projects: A fuzzy approach. *Automation in construction,* 29**,** 123-135.
LI, T. H., NG, S. T. T., SKITMORE, M. & LI, N. Investigating stakeholder concerns during public participation. Proceedings of the Institution of Civil Engineers-Municipal Engineer, 2016. ICE Publishing, 199-219.
LI, T. H., THOMAS NG, S. & SKITMORE, M. 2015a. Modeling multi-stakeholder multi-objective decisions during public participation in major infrastructure and construction projects: A decision rule approach. *Journal of Construction Engineering and Management,* 142**,** 04015087.
LI, Y., WEI, B., CHEN, H., JIANG, L. & LI, Z. Cross-domain learning based traditional Chinese medicine medical record classification. 2015 10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), 2015b. IEEE, 335-340.
LIN, H., SUI, Y., MA, H., WANG, L. & ZENG, S. 2018. CEO narcissism, public concern, and megaproject social responsibility: Moderated mediating examination. *Journal of Management in Engineering,* 34**,** 04018018.
LIU, B. S., HU, Y., WANG, A. M., YU, Z. L., YU, J. & WU, X. L. 2018a. Critical Factors of Effective Public Participation in Sustainable Energy Projects. *Journal of Management in Engineering,* 34.
LIU, J., CUI, Z., YANG, X. & SKITMORE, M. 2018b. Experimental investigation of the impact of risk preference on construction bid markups. *Journal of Management in Engineering,* 34**,** 04018003.

MCLAREN, L., BOOMGAARDEN, H. & VLIEGENTHART, R. 2017. News coverage and public concern about immigration in Britain. *International Journal of Public Opinion Research,* 30**,** 173-193.

MIMNO, D., WALLACH, H. M., TALLEY, E., LEENDERS, M. & MCCALLUM, A. Optimizing semantic coherence in topic models. Proceedings of the conference on empirical methods in natural language processing, 2011. Association for Computational Linguistics, 262-272.

MIN, J. H., JANG, W., HAN, S. H., KIM, D. & KWAK, Y. H. 2018. How conflict occurs and what causes conflict: conflict analysis framework for public infrastructure projects. *Journal of Management in Engineering,* 34**,** 04018019.

NGUYEN, D. Q., BILLINGSLEY, R., DU, L. & JOHNSON, M. 2015. Improving topic models with latent feature word representations. *Transactions of the Association for Computational Linguistics,* 3**,** 299-313.

NORTON, P. & HUGHES, M. 2017. *Public Consultation and Community Involvement in Planning: A twenty-first century guide*, Routledge.

PLUMMER, J. & TAYLOR, J. G. 2013. *Community participation in China: Issues and processes for capacity building*, Routledge.

RAMOS, J. Using tf-idf to determine word relevance in document queries. Proceedings of the first instructional conference on machine learning, 2003. Piscataway, NJ, 133-142.

RODRIGUEZ-MELO, A. & MANSOURI, S. A. 2011. Stakeholder engagement: defining strategic advantage for sustainable construction. *Business Strategy and the Environment,* 20**,** 539-552.

ROWE, G. & FREWER, L. J. 2004. Evaluating public-participation exercises: a research agenda. *Science, technology, & human values,* 29**,** 512-556.

ROWE, G. & FREWER, L. J. 2005. A typology of public engagement mechanisms. *Science, Technology, & Human Values,* 30**,** 251-290.

RUBIN, T. N., CHAMBERS, A., SMYTH, P. & STEYVERS, M. 2012. Statistical topic models for multi-label document classification. *Machine learning,* 88**,** 157-208.

STIMSON, J. 2018. *Public opinion in America: Moods, cycles, and swings*, Routledge.

TANG, L., ZHANG, Y., DAI, F., YOON, Y., SONG, Y. & SHARMA, R. S. 2017. Social media data analytics for the US Construction industry: Preliminary study on twitter. *Journal of Management in Engineering,* 33**,** 04017038.

TUMMERS, L. 2011. Explaining the willingness of public professionals to implement new policies: A policy alienation framework. *International Review of Administrative Sciences,* 77**,** 555-581.

VALENTIN, V., NADERPAJOUH, N. & ABRAHAM, D. 2018. Impact of Characteristics of Infrastructure Projects on Public Opinion. *Journal of Management in Engineering,* 34.

VANDEWEERDT, C., KERREMANS, B. & COHN, A. 2016. Climate voting in the US Congress: the power of public concern. *Environmental Politics,* 25**,** 268-288.

WANG, X. 2001. Assessing public participation in US cities. *Public Performance & Management Review***,** 322-336.

WANG, X. & MCCALLUM, A. Topics over time: a non-Markov continuous-time model of topical trends. Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, 2006. 424-433.

WANG, Y., HAN, Q., DE VRIES, B. & ZUO, J. 2016. How the public reacts to social impacts in construction projects? A structural equation modeling study. *International Journal of Project Management,* 34**,** 1433-1448.

WARDMAN, J. K. & LöFSTEDT, R. 2018. Anticipating or accommodating to public concern? Risk amplification and the politics of precaution reexamined. *Risk analysis,* 38**,** 1802-1819.

WEBLER, T. & TULER, S. 2006. Four perspectives on public participation process in environmental assessment and decision making: Combined results from 10 case studies. *Policy Studies Journal,* 34**,** 699-722.

WU, L., JIA, G. & MACKHAPHONH, N. 2019. Case study on improving the effectiveness of public participation in public infrastructure megaprojects. *Journal of Construction Engineering and Management,* 145**,** 05019003.

XIE, L.-L., YANG, Y., HU, Y. & CHAN, A. P. 2014. Understanding project stakeholders' perceptions of public participation in China's infrastructure and construction projects. *Engineering, Construction and Architectural Management*.

XIE, L. L., HU, Y., ZHANG, B., XIA, B., SKITMORE, M. & GAO, X. Y. 2019. Accountability Effects of Public Participation Practice in Chinese Public Projects: Partial Least-Squares Path Analysis. *Journal of Management in Engineering,* 35.

YANG, S., LU, W., YANG, D., YAO, L. & WEI, B. Short text understanding by leveraging knowledge into topic model. Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2015. 1232-1237.

YAO, L., ZHANG, Y., WEI, B., LI, L., WU, F., ZHANG, P. & BIAN, Y. 2016. Concept over time: the combination of probabilistic topic model with wikipedia knowledge. *Expert Systems with Applications,* 60**,** 27-38.

YAO, L., ZHANG, Y., WEI, B., ZHANG, W. & JIN, Z. 2018. A topic modeling approach for traditional Chinese medicine prescriptions. *IEEE Transactions on Knowledge and Data Engineering,* 30**,** 1007-1021.

ZHAO, D., MCCOY, A. P., KLEINER, B. M., MILLS, T. H. & LINGARD, H. 2016. Stakeholder perceptions of risk in construction. *Safety science,* 82**,** 111-119.

ZHOU, Y., HOU, L., YANG, Y., CHONG, H.-Y. & MOON, S. 2019. A comparative review and framework development on public participation for decision-making in Chinese public projects. *Environmental Impact Assessment Review,* 75**,** 79-87.

ZHUANG, F., LUO, P., SHEN, Z., HE, Q., XIONG, Y., SHI, Z. & XIONG, H. Collaborative dual-plsa: mining distinction and commonality across multiple domains for text classification. Proceedings of the 19th ACM international conference on Information and knowledge management, 2010. ACM, 359-368.

ZIDANE, Y. J.-T., JOHANSEN, A. & EKAMBARAM, A. 2013. Megaprojects-Challenges and lessons learned. *Procedia-Social and Behavioral Sciences,* 74**,** 349-357.

**Table 1 Notations of the symbols in TOT**

| Symbol | Description |
|---|---|
| T | Number of topics |
| D | Number of documents |
| $N_d$ | Number of word tokens in document d |
| $\theta_d$ | The multinomial distributions of topics specific to the document d |
| $\emptyset_z$ | The multinomial distributions of words specific to topic z |
| $\varphi_z$ | The beta distributions of time specific to topic z |
| $z_{di}$ | The topic associated with the ith token in document d |
| $w_{di}$ | The ith token in document d |
| $t_{di}$ | The timestamp associated with the ith token in document d |

**Table 2. 29 kinds of stakeholders**

| Stakeholder groups | | | | |
|---|---|---|---|---|
| Hong Kong SAR | Central government | Macao government | Guangdong Province | Shuttle operator |
| Zhuhai government | Coordination group | HZMB Authority | District council | Tender |
| Legislative Council | Ruling Party | Opposition Party | Court | Aviation |
| Contractor | Consultant | Subcontractor | Supplier | Immigration |
| Worker | Environmental group | Local community | Fishermen | Insurance |
| Logistic industry | Tourism industry | Media | Hong Kong Citizen | |

**Table 3. Distribution of coherences for different topic numbers**

| Topic Number | 10 | 12 | 14 | 16 | 18 | 20 | 22 | 24 |
|---|---|---|---|---|---|---|---|---|
| Average Coherence | -48.5053 | -48.9602 | -51.1143 | **-47.1811** | -51.8116 | -48.1022 | -48.8801 | -49.1057 |

**Table 4. Top 15 featured terms for each topic in the 16-topic model**

| Topic | Concept | Feature Words |
|---|---|---|
| #0 | alignment | align, lantau, passing, hope, improv, facil, clearance, infrastructure, engine, delay, transport, plan, tunnel, environment, financ |
| #1 | local community | commun, western, infrastructure, control, demand, tuenmun, hksar, preliminary, mainland, macao, arrang, direct, plan, develop, transport |
| #2 | cooperation across governments | volum, invest, expedit, expediti, reach, nwnt, agreement, examin, coordinationgroup, environmentalimpact, competit, contribut, integr, corridor, economicbenefit |
| #3 | environment | environmentalimpact, eastern, cargo, integr, judgment, contribut, economicbenefit, competit, corridor, examin, landfall, mitig, reach, cheklapkok, cooper |
| #4 | immigration and custom service | control, boundary, tuenmun, plan, environ, problem, increase, hksar, develop, transport, territory, environmentalimpactassess, commun, enhance, cost |
| #5 | operation on cross-boundary traffic | road, airport, region, feasibl, traffic, investing, cross, implement, land, future, improv, island, passing, highway, progress |
| #6 | local traffic in connective areas | rail, flow, benefit, tungchung, technic, passing, infrastructur, promot, demand, origin, vehicle, increase, lantau, commenc, region |
| #7 | local industry | service, termin, flow, site, highway, park, investing, time, legal, marin, environment, facil, control, problem, align |
| #8 | operation on local traffic | rail, oper, highway, local, complet, future, increase, connect, hope, facil, toll, commenc, macao, cross, road |
| #9 | local connectivity | north, opportune, viaduct, rout, railway, promot, local, time, toll, guangdong, tmclkl, tmwb, boundary, clearance, region |
| #10 | urban development | lantau, time, mainland, commenc, align, vehicle, assess, environ, preliminary, origin, progress, oper, park, problem, complet |
| #11 | finance | expenditure, cooper, viaduct, boundary, land, crossboundari, vehicle, oper, logist, enhance, rail, plan, site, preliminary, billion |
| #12 | construction | construct, island, guangdong, origin, facility, control, mainbridg, develop, facil, cost, econom, manag, hope, design, preliminari |
| #13 | design | design, assess, zhuhai, public, govern, local, estim, service, macao, hope, lantau, environment, commiss, enhance, commenc |
| #14 | regional economy | public, strategy, connect, crossboundari, govern, cost, technic, infrastructur, site, region, arrang, territory, logist, vehicle, financ |
| #15 | commercial development | shop, bridgeheadeconomi, cheklapkok, busi, advancework, eastern, facility, preliminary, infrastructur, strategy, tuenmuncheklapkok, econom, accid, commerci, hksar |

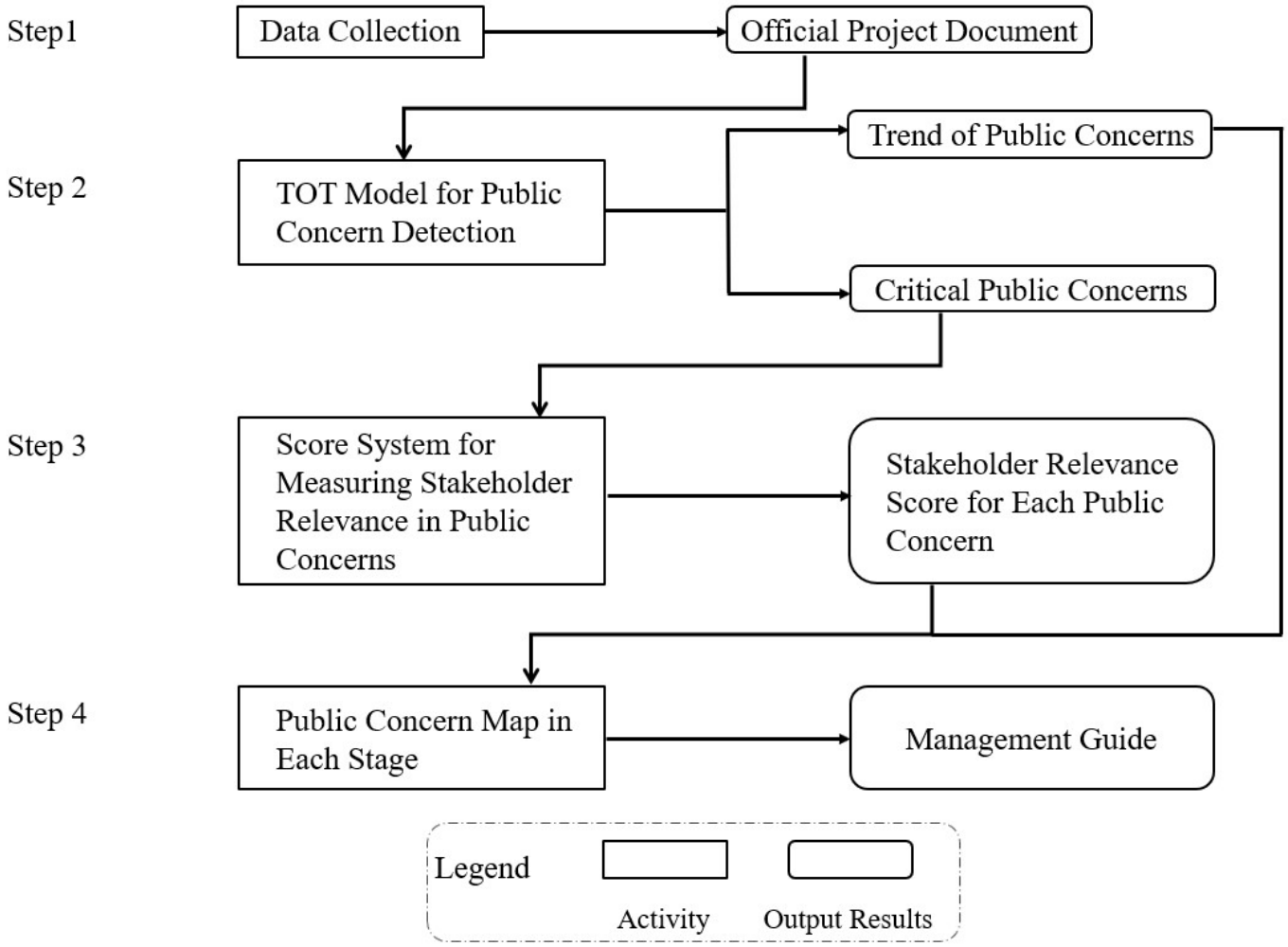



**Figure 1. The framework of DSTM**

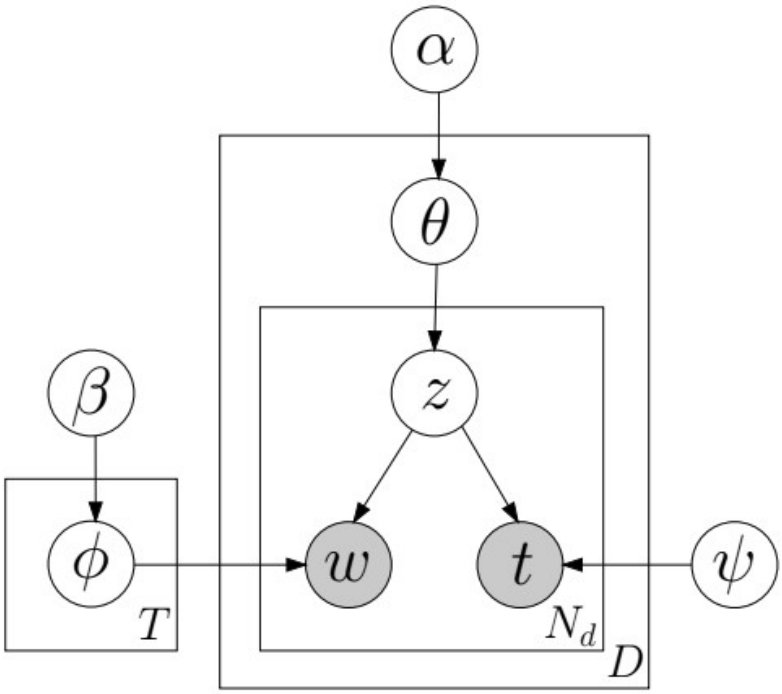



**Figure 2. Graphic representation of the TOT model for Gibbs sampling**

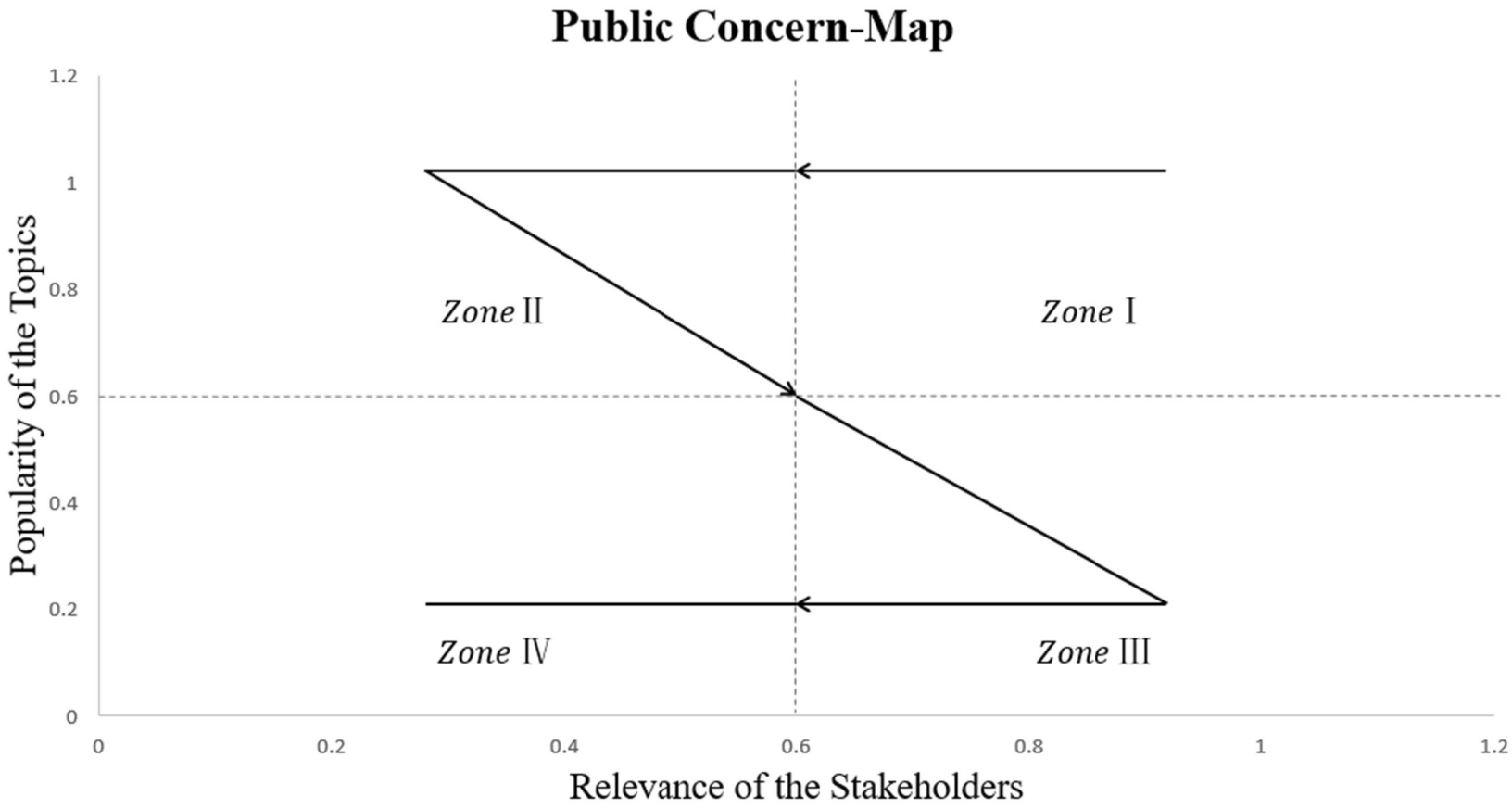



**Figure 3. "Mirror Z" approach for the management of public concerns**

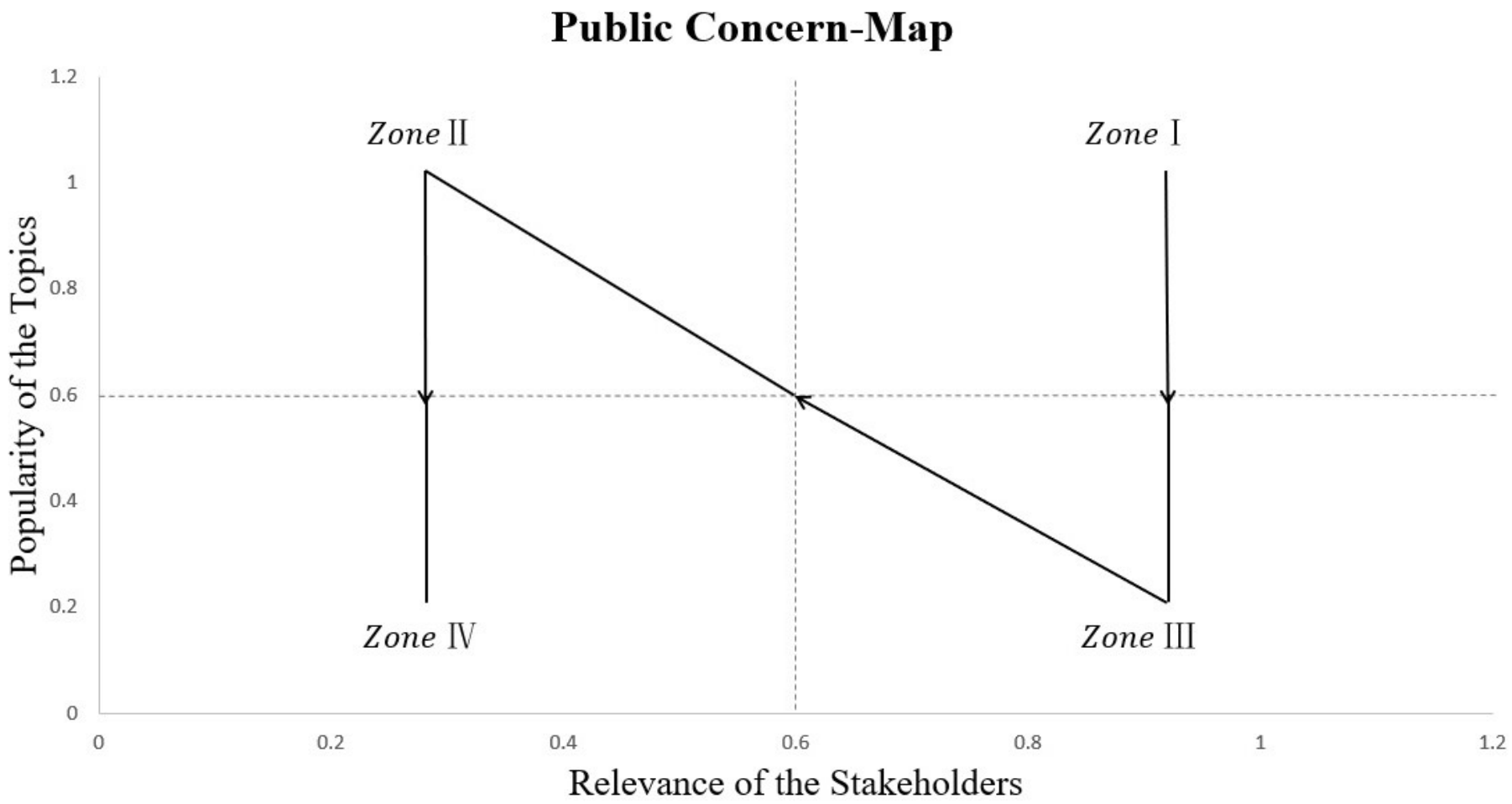



**Figure 4. "Letter N" approach for the management of public concerns**


**Figure 5. Trend of public concerns in HZMB**

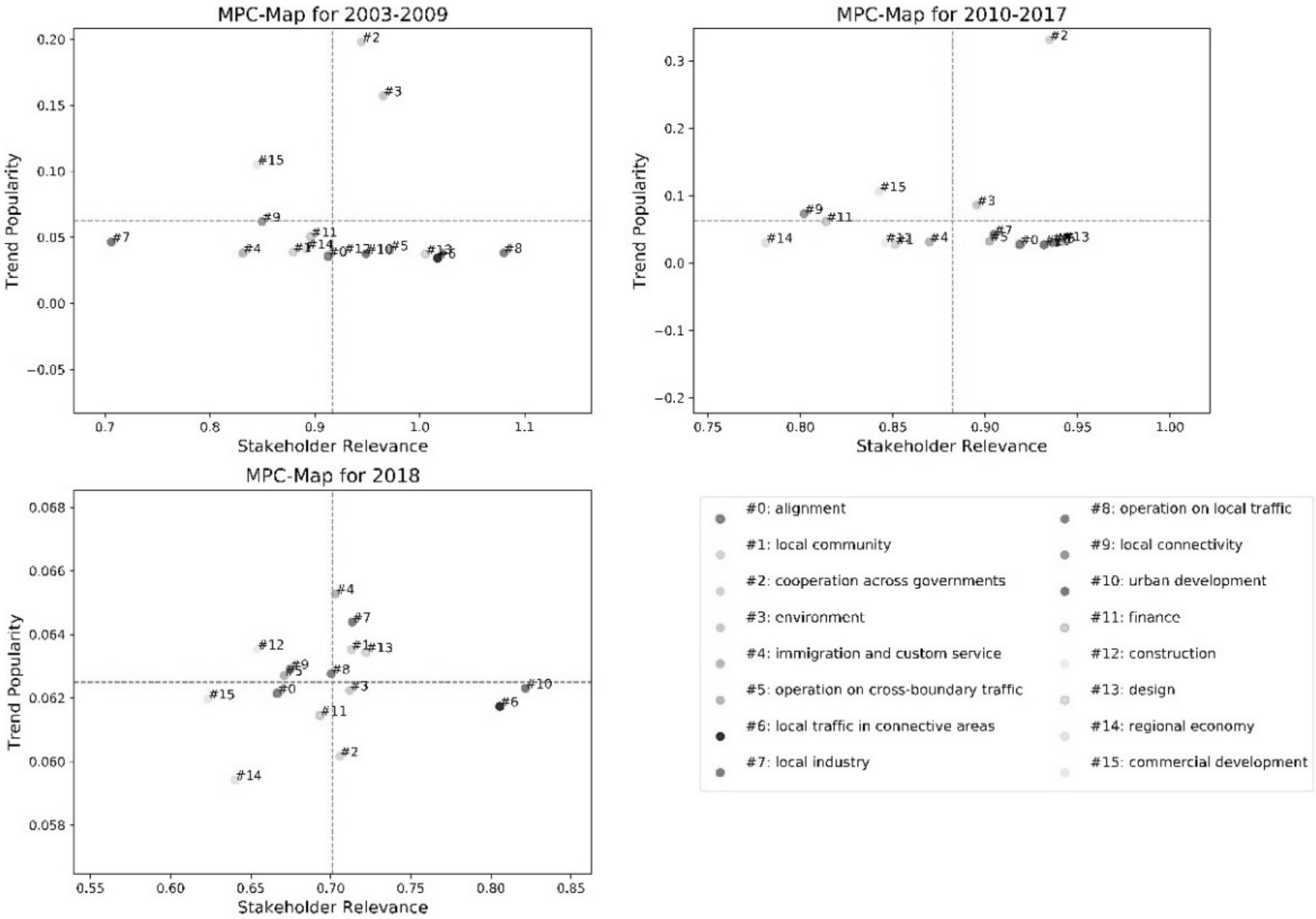



**Figure 6. Managerial map of public concerns for decision-makers**

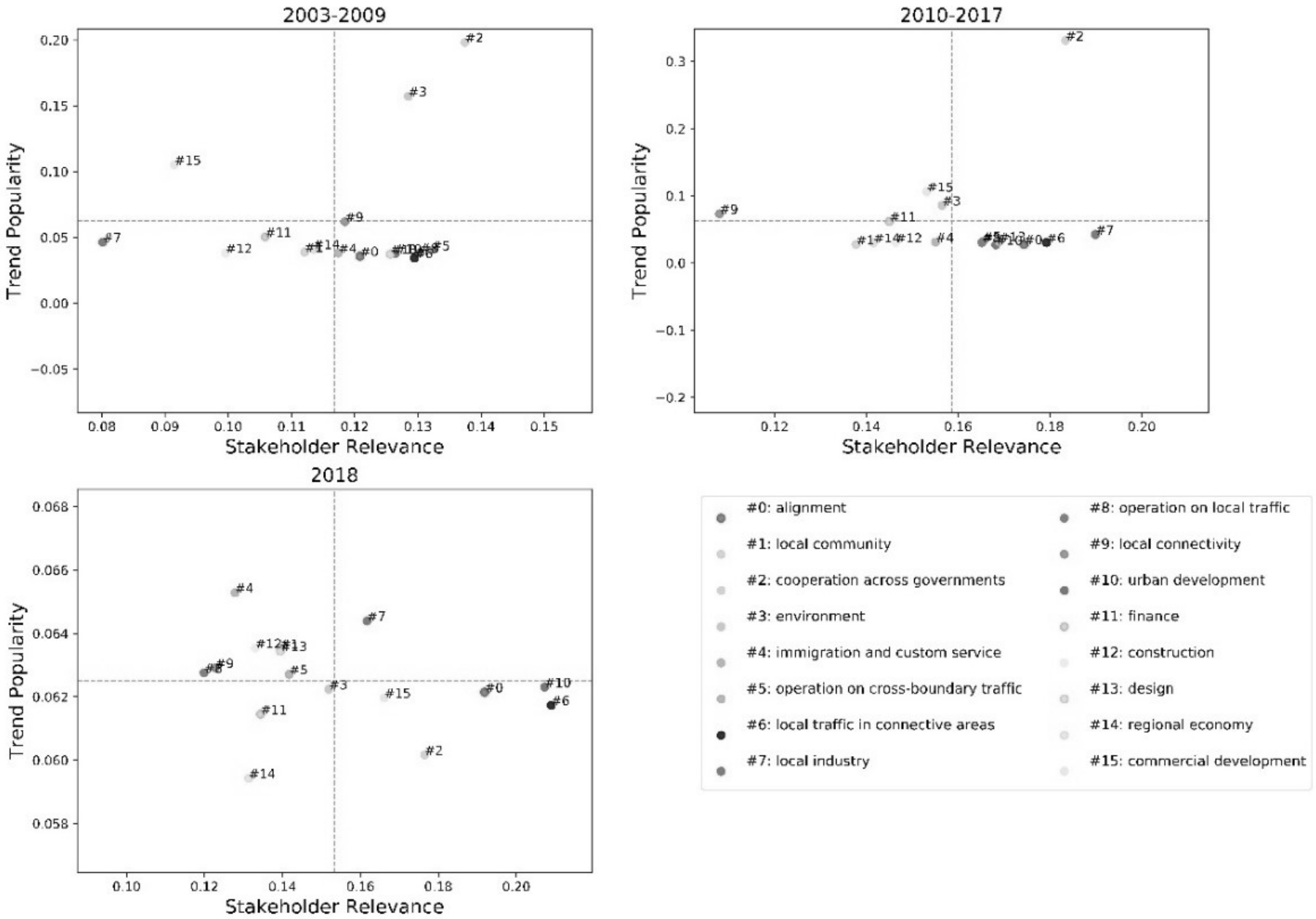



**Figure 7. Managerial map of public concerns for political groups**

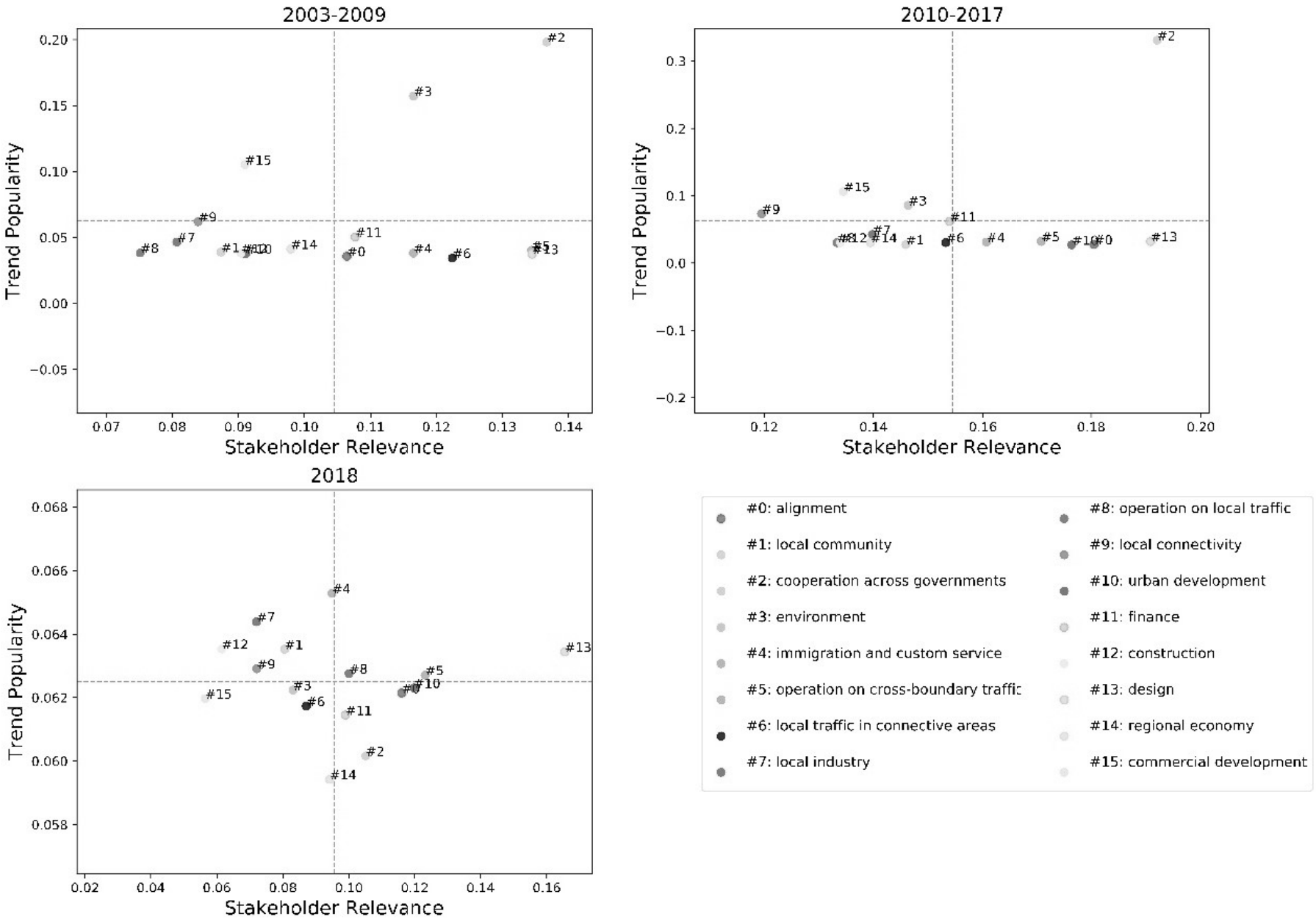



**Figure 8. Managerial map of public concerns for construction groups**

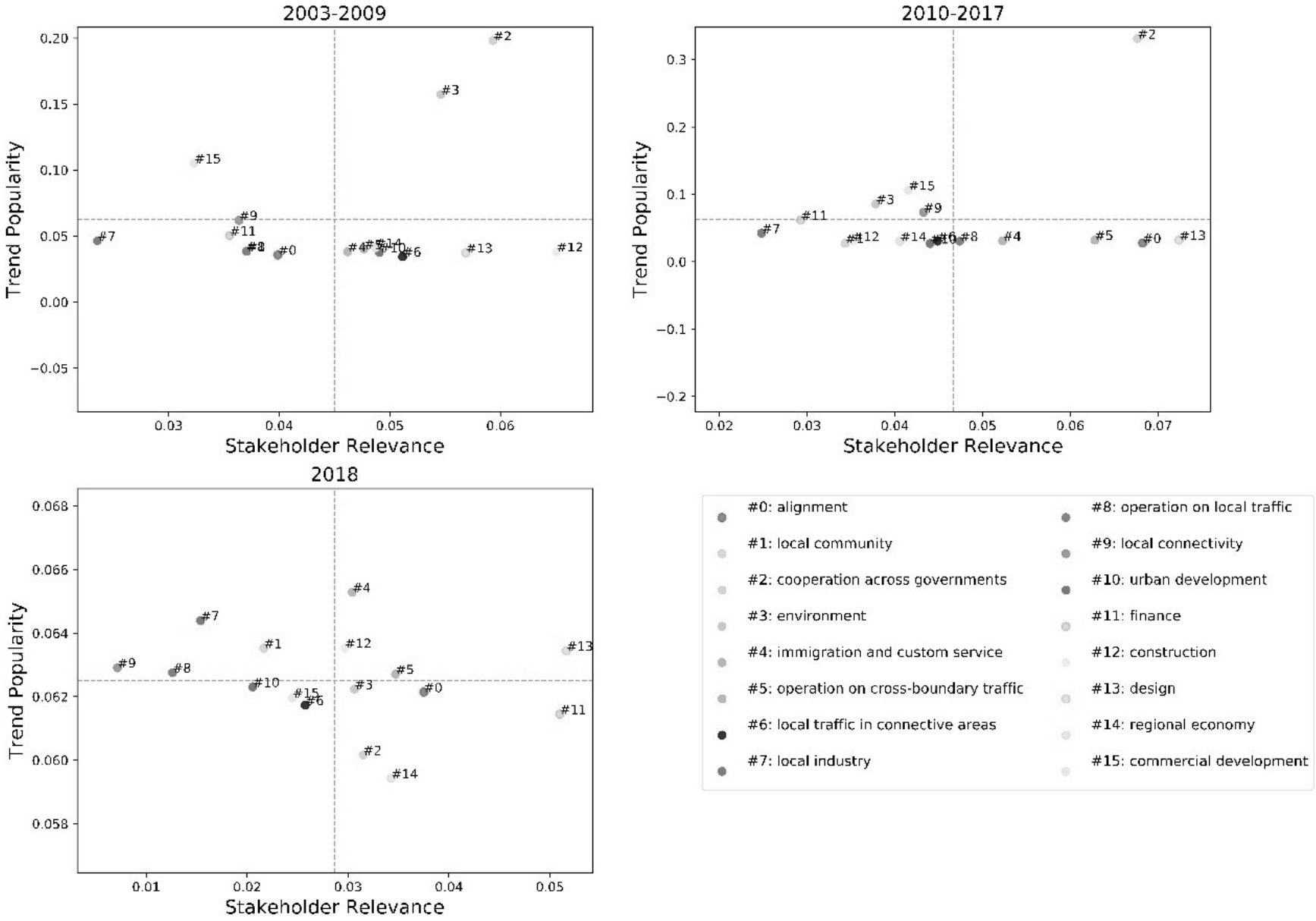



**Figure 9. Managerial map of public concerns for pressure groups**

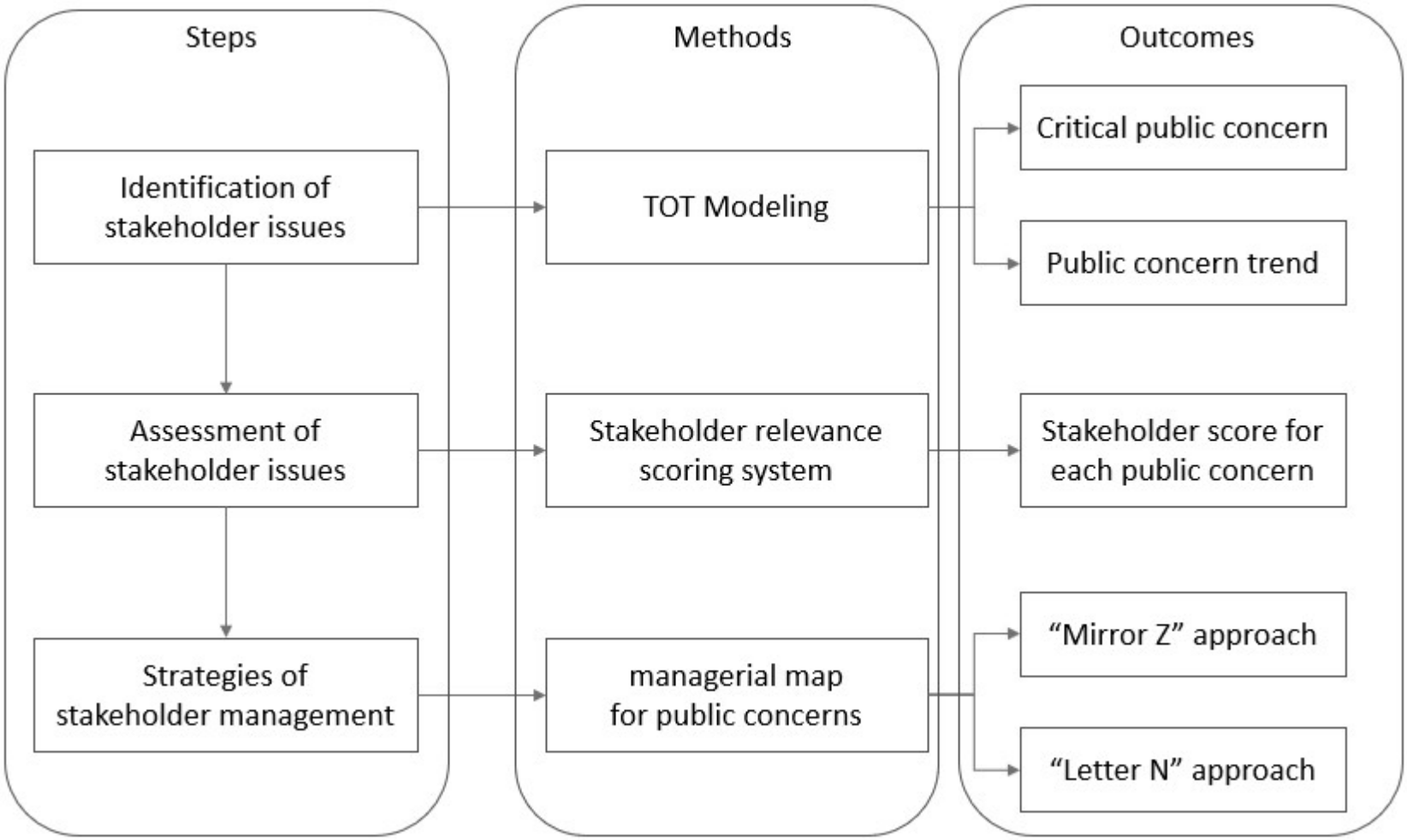



**Figure 10. The framework of stakeholder management on public concerns**

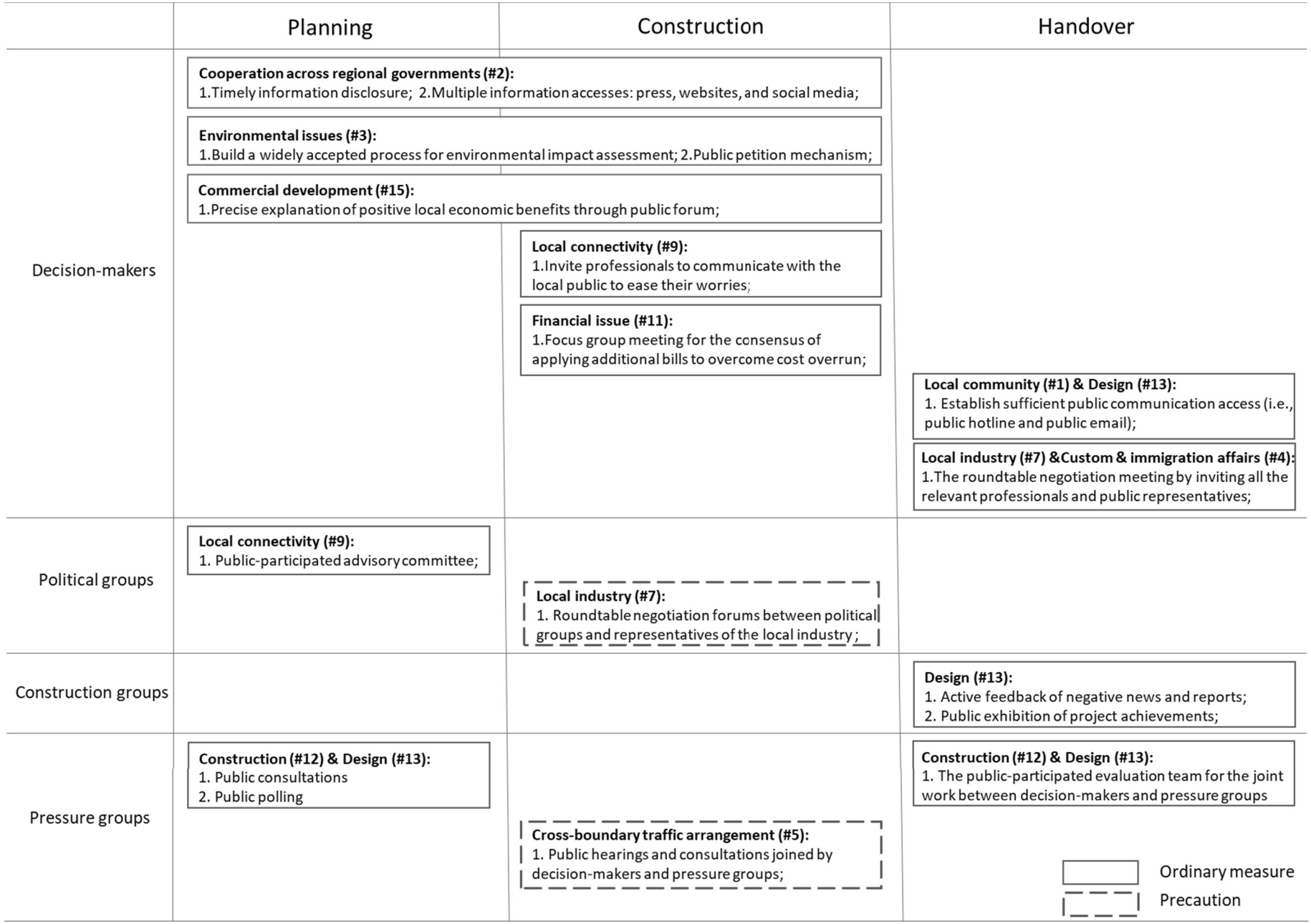
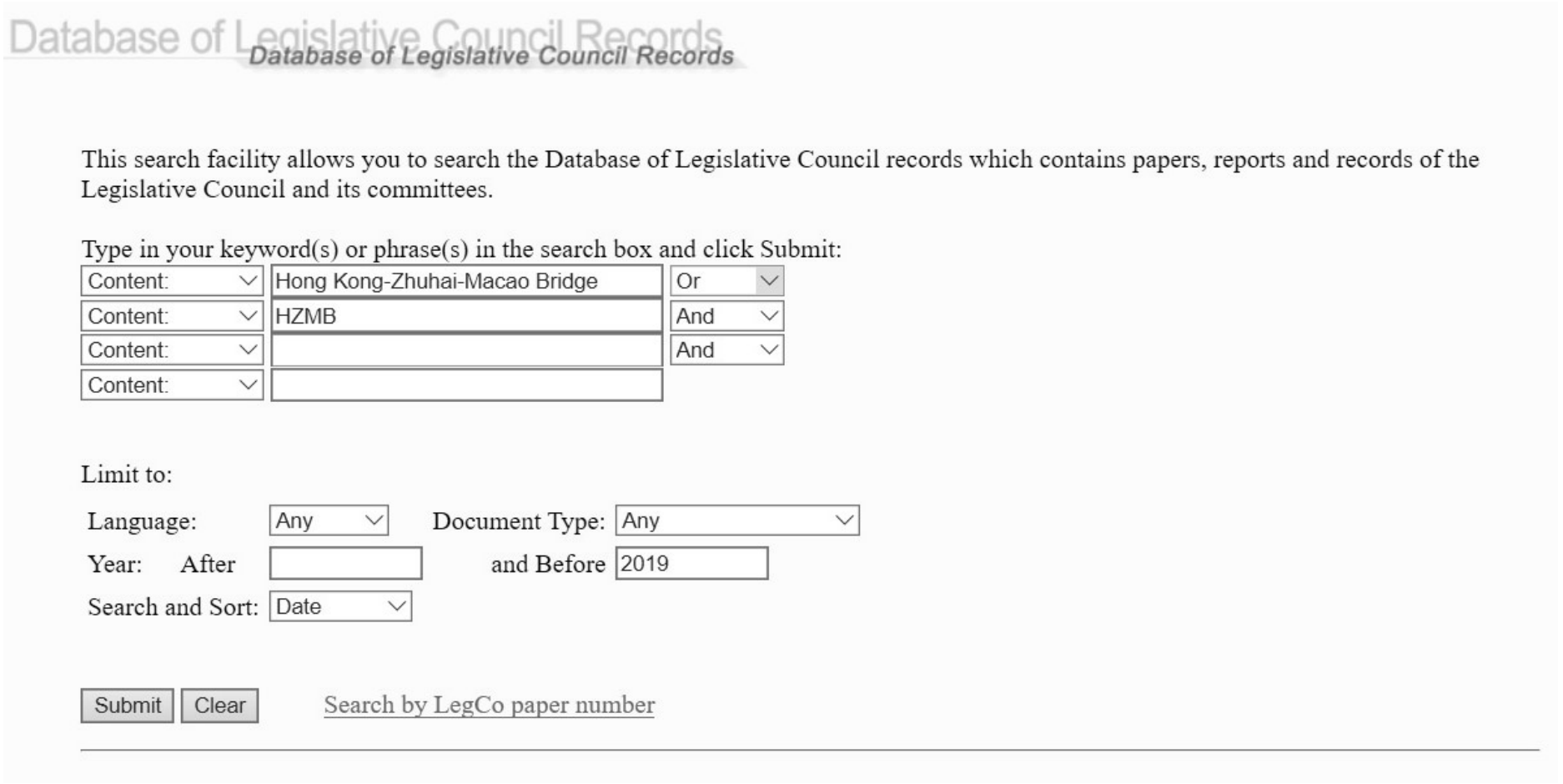
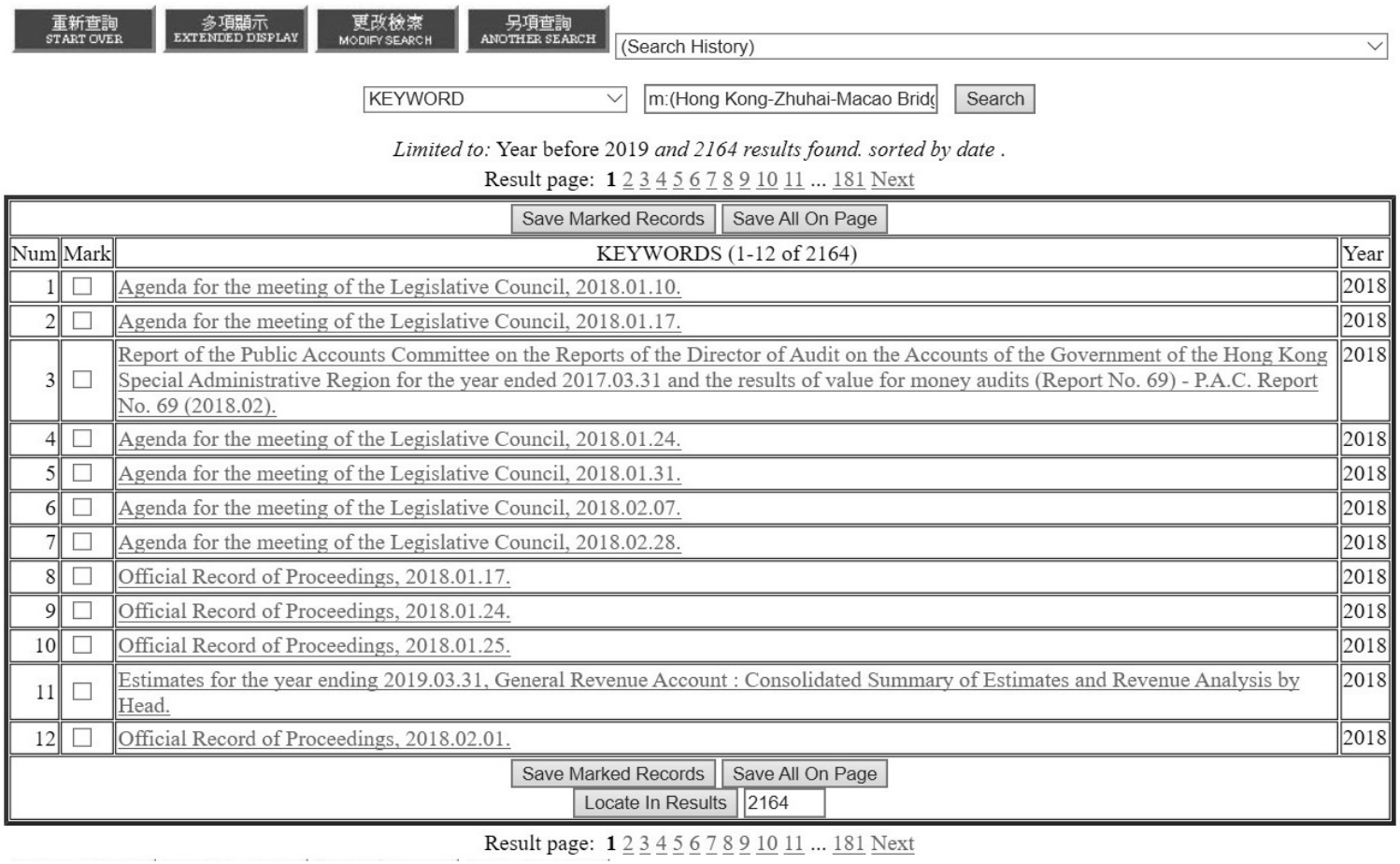



**Figure 11. Managerial implications for project stakeholders**

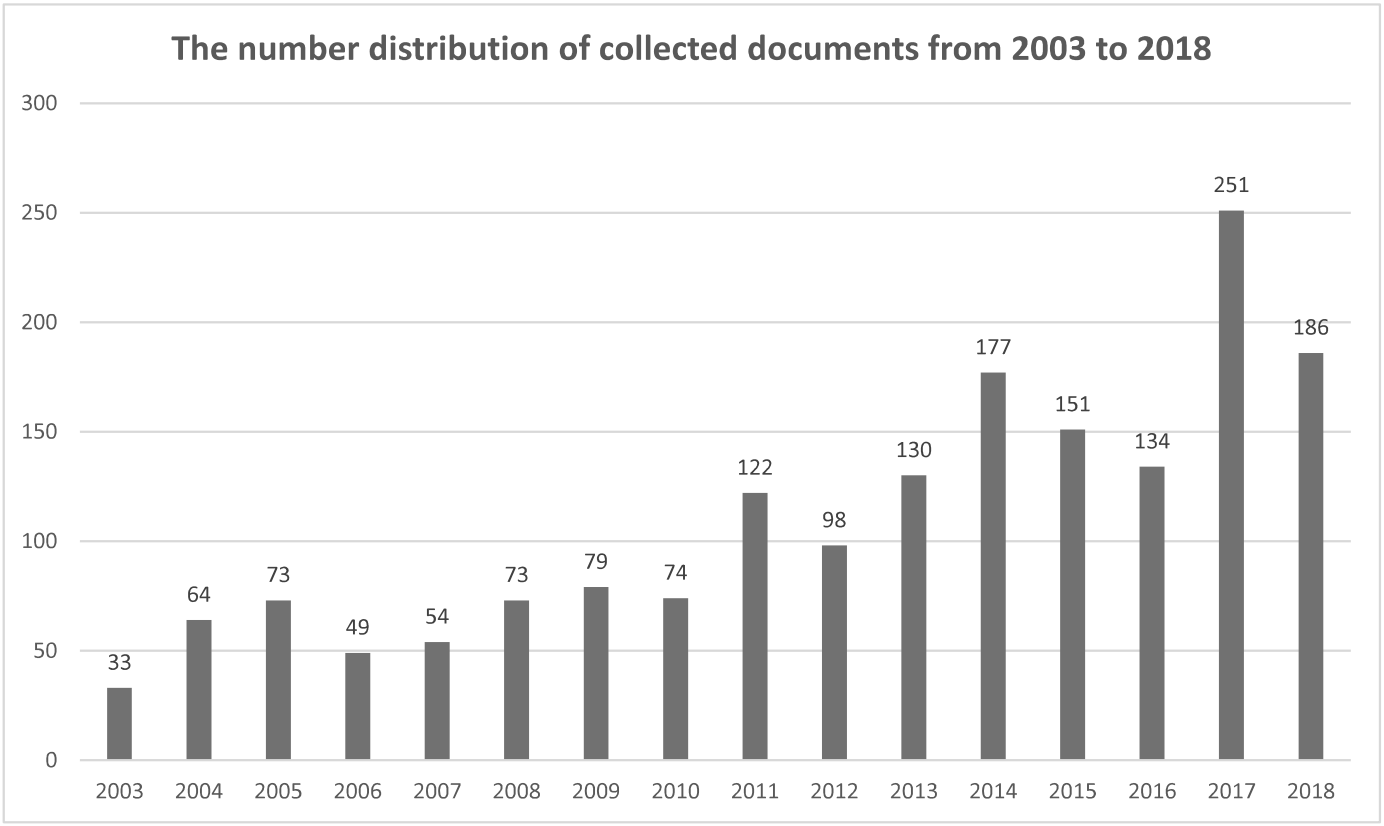



**Appendix 1. The example of a database of legislative council records**

Click here to access/download;Figure;Appendix 2.pdf




**Appendix 2. The example of raw documents related to HZMB**




**Appendix 3. The number distribution of collected documents from 2003 to 2018**