Journal of Global Information Management
Volume 30 • Issue 6

Citizen Participation in the Co-Production of Urban Natural Resource Assets:
Analysis Based on Social Media Big Data

Lei Li, College of Management and Economics, Tianjin University, Tianjin, China & Qingdao Institute for Marine Technology of Tianjin University, Qingdao, China

Shaojun Ma, Tianjin University, China*

Runqi Wang, Tianjin University, China

Yiping Wang, Tianjin Sino-German University of Applied Sciences, China

Yilin Zheng, Tianjin University, China

ABSTRACT

Abundant natural resources are the basis of urbanisation and industrialisation. Citizens are the key factor in promoting a sustainable supply of natural resources and the high-quality development of urban areas. This study focuses on the co-production behaviours of citizens regarding urban natural resource assets in the age of big data and uses the latent Dirichlet allocation algorithm and the stepwise regression analysis method to evaluate citizens' experiences and feelings related to the urban capitalisation of natural resources. Results show that, firstly, the machine learning algorithm based on natural language processing can effectively identify and deal with the demands of urban natural resource assets. Secondly, in the experience of urban natural resources, citizens pay more attention to the combination of history, culture, infrastructure, and natural landscape. Unique natural resources can enhance citizens' sense of participation. Finally, the scenery, entertainment, and quality and value of urban natural resources are the influencing factors of citizen satisfaction.

KEYWORDS

Citizen Participation, Co-Production, LDA, Natural Resource Assets, Social Big Data, Stepwise Regression, Urban

1. INTRODUCTION

1.1. Capitalisation and Economic Value of Urban Natural Resources

Urban natural resources are an important factor in determining regional economic security and social development. Due to variations in natural and economic geographies in different regions, the factor endowment of urban natural resources also tend to vary. Generally, urban natural resources are non-excludable, competitive and decaying. They are also called common-pool resources (CPR) in urban development and include land, forest, pasture, water, wetland, mineral resources, and so on. One of the most difficult problems faced by the government or public sector managers is how to achieve the optimal allocation as well as the long-term and healthy development of limited CPR (Hardin, 2009; Grainger & Costello, 2014). In modern urban construction, natural resources may involve various property rights of different stakeholders, whilst citizen behaviours driven by rationality may cause

DOI: 10.4018/JGIM.291514 *Corresponding Author

the so-called 'tragedy of the commons' phenomenon of resource crowding or resource degradation to CPR. In relation to this, capitalisation has become one of the common management methods of urban natural resources (Arnason, 2012; Grafton et al., 2000). To date, most Chinese cities have successively established and improved natural resource asset operation systems in their respective administrative regions. In 2019, the Chinese government also proposed the 'Guiding Opinions on the Overall Promotion of the Reform of the Property Right System of Natural Resource Assets' and the 'Notice on Organising and Carrying Out the Pilot Work of the Inventory of All Public-Owned Natural Resource Assets'.

Unlike ecological products, urban natural resource assets have ecological and economic benefits. In terms of economic benefits, the exploitation, utilisation and protection of urban natural resource assets often require a labour force or more human resources (Song et al., 2021). These are needed to increase the outputs of natural resource products and create more economic benefits for various industries, such as the eco-tourism industry (Pueyo-Ros et al,. 2018). Eco-tourism is based on the concept of the sustainable development of natural resources (Okan et al., 2016; Weaver and Lawton, 2007). This is a tourism mode that operates on the principle of protecting nature and the ecological environments where scenic sites can be found, including forest parks, nature conservation areas, geological (or wetland) parks and other natural resource assets (Sun et al., 2020). According to a report by the Chinese Government Network (CGN), in 2019, China's eco-tourism industry received over 3.3 billion tourists, generating revenues of over 850 billion yuan. Thus, in recent years, the characteristic eco-tourism industry based on urban natural resource assets has become a new driver of urban economic development.

1.2. Co-production: A New Trend of Data-Driven Urban Governance

Co-production is an important concept first introduced by Ostrom's urban governance studies in the 1970s (Ostrom, 1972; Brudney & Jeffrey, 1983). This was originally designed to solve the 'free-riding' phenomenon of urban natural resources (e.g. water and grassland). Since the 1990s, citizen participation has been reintroduced in the study of co-production in which citizens, previously considered 'customers', have been transformed into 'partners' (Levine & Fisher, 1984; Nabatchi et al., 2017). Ostrom (1996) also discussed the mode of citizen participation in the cooperative supply or collaborative governance of CPR. Unlike political democratic participation (e.g. party elections or voting in European and American countries), citizen participation in co-production requires the government to pay more attention to the voice of citizens in urban operation, which is reflected in the value co-creation of urban living environments, quality of life, citizen welfare and social stability (Lember et al., 2019; West, 2004).

Within the data-driven context, governments around the world have proposed various policies or plans involving citizen participation in digital government, such as the data opening and information response policies of local governments in China, the 'Government with You' e-government strategy in Singapore and the Big Society programme in the UK. At present, there are now more practical platforms and channels that allow citizens to participate in co-production. Taking China as an example, the most common example of co-production is the country's health code policy on COVID-19 in 2020. China has designed an electronic e-permit (citizen's health code) through information and communication technology (ICT). In response, citizens have perfected their travel information through the application to ensure that their locations and environments are both safe. This kind of co-production practice in the digital era has effectively helped China in its efforts to control the spread of the virus. As far as natural resource assets related to regional development and economic security are concerned, the above measures cannot fully demonstrate the citizens' ability to directly participate in the governance of urban natural resource assets. Nevertheless, they have successfully introduced a new mechanism through which citizens can indirectly participate in the allocation of urban natural resource assets from the bottom up in the long run.

1.3. Research Questions

According to Guo et al. (2017), in the era of big data, virtual networks have generated more visible and referential user-generated content (UGC). Online comments, diaries and ratings based on urban natural resource assets (ecotourism resources) have become new forms of citizens' participation in the co-production of social media data resources (Liu & Park, 2015; Ye et al., 2011). The generated information can help in analysing citizens' online preferences and public opinions from the demand side (Li et al., 2020), accurately identifying citizens' actual demands for urban natural resources and improving the service quality of urban natural resource assets (Mondal & Samaddar,2021). Therefore, the purpose of this study is to answer the following research questions (RQs):

RQ1. Can big data analysis and algorithms effectively identify and process citizens' needs related to urban natural resource assets?
RQ2. From the large sample text-based data and topic modelling method, what are citizens concerned about in relation to urban natural resource assets?
RQ3. Based on the numerical data of a large sample and stepwise regression analysis, what are the factors influencing citizens' satisfaction with urban natural resource assets?

The remainder of this study is arranged into sections. Section 2 compares the literature review. Section 3 proposes the analytical framework and data collection. This study takes natural resource assets in Tianjin as the research object. Section 4 describes research methods of the study, which uses a topic algorithm and regression model in the analysis. Section 5 presents the results of the case study. Finally, Section 6 and Section 7 present the conclusions, significance and future research of the study.

2. LITERATURE REVIEW

2.1. Government Performance and Capitalisation of Urban Natural Resources

Urban natural resource assets are an important part of government governance. It has also become one of the indicators to measure government performance. This paper will elaborate on the relationship between them. More and more studies have examined government reform and organisational performance in modern city operations (Huang et al., 2020; Moynihan, 2008). In these studies, public management scholars believe that the cooperation network composed of collective efforts in different fields can promote the performance of government departments (Provan et al., 2007; Provan & Kenis, 2008; Yi et al., 2017), thereby showing that policy implementation is more effective and public service supply is more efficient (Lubell & Fulton, 2008). Performance appraisal under the target responsibility system is more common in Chinese governments at all levels (Cui et al., 2021). However, some studies believe that the impact of the property rights system on urban natural resources, as implemented by the government, may reveal uncertainties (Lisco, 2013). Under the guidance of incentive mechanisms, such as government performance evaluation, the establishment of the property right system of natural resources, such as river, forest and farmland chief systems maintained by limited franchises, can promote the efficient utilisation and protection of urban natural resources (Li et al., 2017; Zhu et al., 2017). In this way, authorities can develop urban natural resources in a planned way, promote the recycling of idle resources and serve more public interests in the long run (Bohn & Deacon, 2000; Barbier & Burgess, 2001).

However, given that urban natural resources have different property structures and characteristics (Grainger & Costello, 2014), under capitalisation, the stakeholders have the right to develop urban natural resources. Meanwhile, government behaviour under short-sighted logic tends to ignore the potential of urban natural resources. Furthermore, unsecure property rights cannot facilitate overall planning and suitable development (Godoy et al., 2001), which in turn, can lead to various problems, such as damaged mountains and polluted air, rivers and lakes. Generally, the evaluation of resource

value attributes (economic and ecological values) and the driving force of urban renewal determine the manner and degree of the government's development of urban natural resources. These can also accelerate the transformation of urban natural resources from assets to capital.

2.2. Citizens' Co-production Behaviours in the Digital Age

The protection of urban natural resource assets is inseparable from citizens' co-production. Co-production, an important theory in the field of public administration, asserts that citizens participate in the cooperative supply of urban public services (Ostrom, 1996; Parks et al., 1981). In recent years, the concept of co-production has gradually changed from citizens to communities and non-profit organisations (Nabatchi et al., 2017; Brudney, 2020). Co-production not only improves citizens' participation in urban governance but also meets citizens' differentiated needs for urban development (Anand & Gaur, 2019; Voorberg et al., 2015; Zhang et al., 2020).

The concept of using data-based technology empowerment to build a more intelligent and low-carbon city is highly valued in many countries throughout the world. The Internet of things (IoT) and ICT are new infrastructures to realise urban data governance (Clifton, 2019; Meijer, 2015; Chen et al., 2020). Furthermore, under the influence of this environment, co-production theory pays more attention to new forms of citizens' participation in co-governance to promote urban digital transformation (Paletti, 2016). However, an increasing number of scholars have proposed different views on how ICT can change the behaviours of citizens and governments. This section discusses this topic from two perspectives: value co-creation and value co-destruction.

On the one hand, symmetric information and trust are antecedents of effective communication between citizens and their government (Jennings & Hall, 2012). As ICT reduces the entry threshold of citizen participation (Lember et al., 2019), e-government or social media platforms supported by digital technology create new platforms for the interaction of citizen information resources (Linders, 2012). Information type, content, channel and other quality factors can affect the outcomes of citizens' participation in co-production (McColl Kennedy et al., 2012). One study argued that ICT and big data analytics (BDA) bring about the diversification of citizen co-production forms and practices, thus encouraging citizens to participate in resource sharing, technology co-construction and decision-making (Clifton, 2019; Li et al., 2020).

On the other hand, Osborne et al. (2016) argued that existing co-production policies and public sector practices do not fully recognise the existence of value co-destruction. The public value created by citizen participation is only reflected in the supply of potentially meaningful behaviour data (Athey, 2017; Cardullo & Kitchin, 2018). This kind of co-production may also expose it to unknown risks (Plé & Cáceres, 2010). However, the data cyberspace constructed by ICT can always supervise and regulate citizen behaviours (e.g. compliance with restrictive policies). Therefore, citizens with negative attitudes towards ICT tend to gradually reduce their participation in digital co-production (Gao, 2018; Linders, 2012; King & Cotterill, 2007).

2.3. Summary of the Literature Review

Sections 2.1 to 2.2 reviewed and summarised current research on urban natural resources, government behaviours and citizens' co-production enabled by big data. The review reveals that a positive correlation does not always exist between the development of urban natural resources and the government's measures to establish the property rights system of assets. Few studies have analysed urban natural resource assets from the perspective of big data. Thus, the current study begins the discussion by examining urban tourism resources and combining citizens' subjective experiences and evaluations of natural resource assets. Using big data on residents' behaviours from social media, this study attempts to analyse the focus and influencing factors of the protective development of such assets.

3. ANALYTICAL FRAMEWORK AND DATA COLLECTION

3.1. Analytical Framework

For the case study, we selected Tianjin (116°E to 118°E,38°N to 40°N), which is one of the four municipalities directly under the central government in China. Tianjin is also the largest open city in the North, with the reputation of 'international consumption center city'. Tianjin's history is based on a water transport culture and a high level of urbanization. As of December 2019, there are two National 5A scenic sites and 32 National 4A sites in Tianjin, which are rich in natural and cultural tourism resources. Tianjin is the appropriate case for this research topic.

Ctrip(www.ctrip.com) is an online service enterprise integrating tourism, ticketing and hotel reservation in China. Ctrip is also one of the largest tourism service platforms in China. In 2016, it launched a B2C and C2C tourism experience sharing service model on the Internet platform. It can reflect the needs of citizens through their evaluation in the process of urban natural resource assets, and improve the construction of the service platform through citizen participation. Therefore, it provides convenience for the study of citizen participation in digital co-production.

The study used Python to develop a web crawler programme to regularly collect big data (including text-based and numerical big data) from urban natural resource assets in Ctrip. This study started from two aspects and established the analysis framework of this problem, as shown in Figure 1.

Figure 1. Research framework

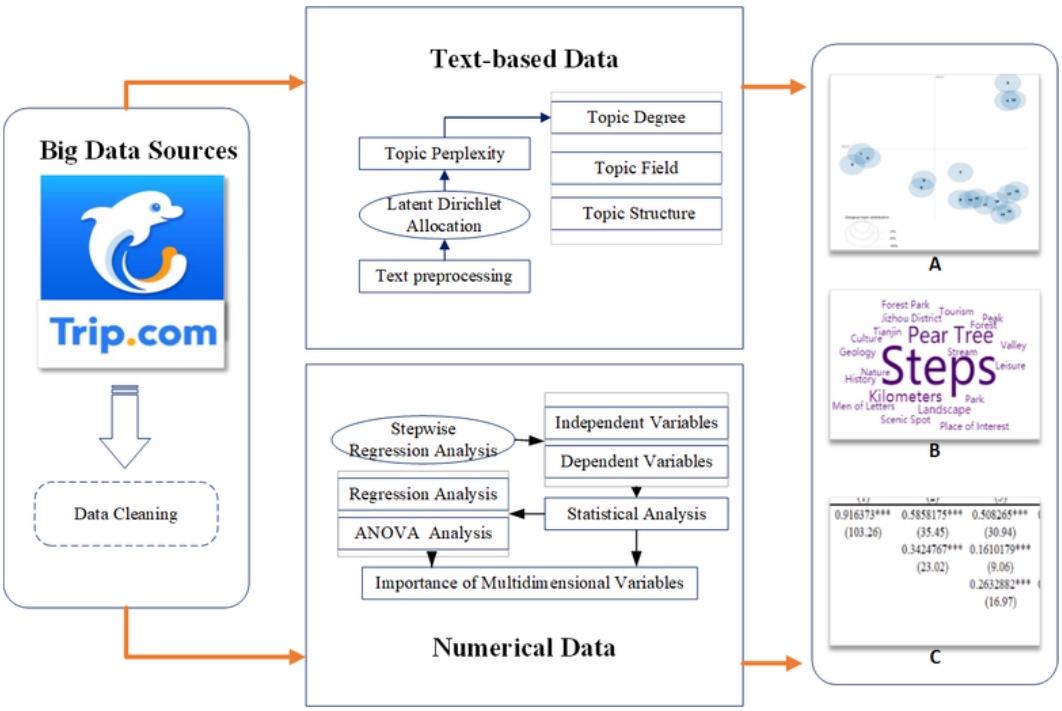


Table 1. Characteristics of rating big data

Variable	M	SD	Min	Max	Std	VIF
Rating (R)	4.439	0.928	1	5	0.928	N/A
Scenery (S)	4.440	0.897	1	5	0.897	4.51
Entertainment (E)	4.280	0.995	1	5	0.897	6.45
Quality & value (Q&V)	4.274	1.030	1	5	1.030	5.28
Level (L)	3.830	1.339	1	5	1.339	1.64
Mountain (M)	0.626	0.484	0	1	0.484	4.60
Beach (B)	0.001	0.037	0	1	0.037	1.03
Forest Park (F)	0.261	0.439	0	1	0.439	3.46
Lake (La)	0.138	0.345	0	1	0.345	3.61
Geopark (G)	0.078	0.268	0	1	0.268	3.00

3.2. Data Collection

A total of 6930 text-based big data were obtained. The R studio open source module Jieba was used to realise text semantic segmentation and part of speech tagging (POST). Then, text2vec was used to create tables of vocabulary, word frequency and document frequency (Hou et al., 2019; Sun et al., 2020), which facilitated the pre-processing of text big data.

In addition, on the basis of big data collection, this study deleted the missing values in the text-based data and combined these with the characteristic data of urban natural resource assets (e.g. type and grade). These resulted in 2967 numerical datasets that can realise regression analysis, in which the dependent variable is the overall rating of public participation and the independent variables are scenery (S), entertainment (E), quality and value (Q&V) and level (L). The control variables are the types of scenic sites, including mountain (M), beach (B), forest park (F), lake (La) and geopark (G). The data characteristics are shown in Table 1. The variance inflation factors (VIFs) of the variables are all less than 10; thus, there is no collinearity between the variables.

4. METHODOLOGY

4.1. LDA

LDA is a commonly used and efficient natural language processing method (Blei et al., 2003; Vu et al.,2019). It is also considered an unsupervised machine learning (ML) algorithm (Guo et al., 2017). As a topic modelling method for extracting big data text, it can evaluate the generated probability distribution model based on the perplexity of the corpus (Bastani et al.,2019; Hoffman et al.,2013) and estimate the number of optimal topics and clustering results through multinomial distribution (Blei, 2012). In this way, researchers can better understand the value of a large number of text-based

data resources. In the current study, LDA is used to identify citizens' feelings and needs for urban natural resource asset. In relation to this, the following hypotheses are presented:

Hypothesis 1. There are D documents in the corpus, and there are V non-repetitive words in the corpus ω. The topic of each document d_i is subject to the probability distribution composed of T topics. The probability distribution of d_i is vector θ_T with length T, which is called 'document–subject' distribution, where $P(t_t \mid d_d)$ means that document d belongs to the probability of subject d, and $P(\omega_v \mid T_t)$ is the probability of the change subject t appears the word ω.

Hypothesis 2. The probability distribution matrix of 'word–subject' is given as $\varnothing$, $\varnothing \sim Dir(\beta)$, where $Dir(\beta)$ indicates that the parameter is β's Dirichlet distribution. This is expressed as follows:

$$\varnothing = \begin{pmatrix} \varnothing_1 & \dots & \varnothing_T \end{pmatrix} = \begin{pmatrix} P(\omega_1 \mid t_1) & \dots & P(\omega_1 \mid t_T) \\ \vdots & \ddots & \vdots \\ P(\omega_V \mid t_1) & \dots & P(\omega_1 \mid t_T) \end{pmatrix}.$$

Hypothesis 3. Given that the probability distribution of 'word–subject' is θ_d,we have $\theta_d \sim Dir(\alpha)$, where $Dir(\alpha)$ indicates that the parameter is α. Its Dirichlet distribution is expressed as follows:

$$\theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_D \end{pmatrix} = \begin{pmatrix} P(t_1 \mid d_1) & \dots & P(t_T \mid d_1) \\ \vdots & \ddots & \vdots \\ P(t_1 \mid d_D) & \dots & P(t_T \mid d_D) \end{pmatrix}.$$

Hypothesis 4. Document d_i has N_d words, each of which is expressed as ω_d. A topic is expressed as $Z_{d,n}$, where $Z_{d,n} \sim multinomial(\theta)$. Then, we choose words $\omega_{d,n}$ from the conditional probability $p(\omega_{d,n} \mid Z_{d,n}, \varnothing)$.

Hypothesis 5. In the LDA model, $Dir(\alpha)$ and $Dir(\beta)$) belong to the conjugate distribution. Document D updates $Z_{d,n}$ and $\varnothing$ continuously. The maximum likelihood function model is improved and optimised. The function is expressed as follows:

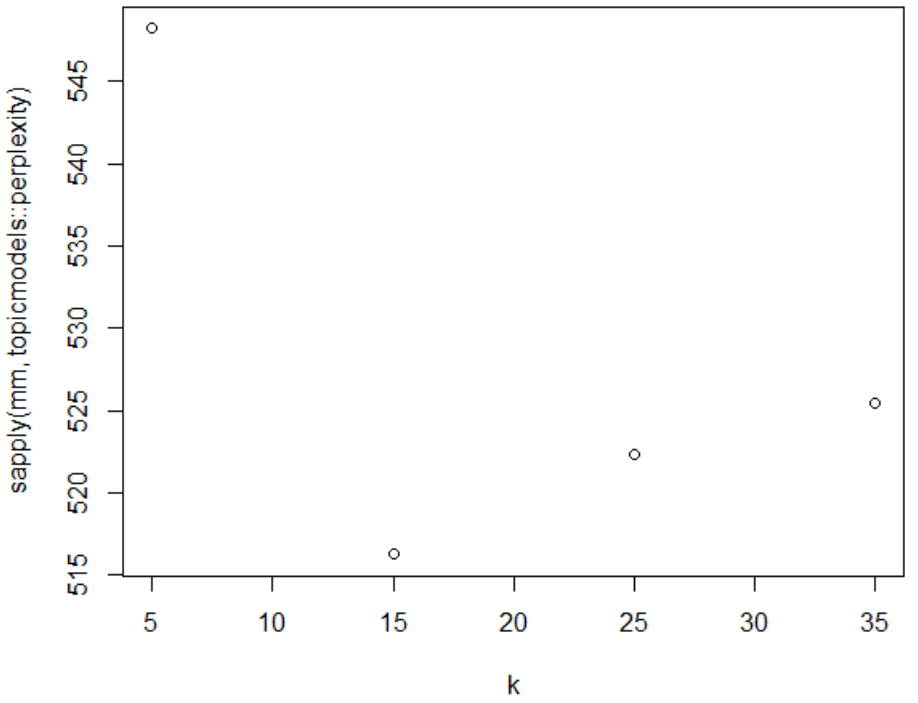
$$\prod_{d=1}^{D} P(\theta_d) \{ \prod_{n=1}^{N_d} \sum_{Z_{d,n}} P(Z_{d,n} \mid \theta_d) P(\omega_{d,n} \mid Z_{d,n}, \varnothing) \}.$$

4.2. Stepwise Regression Analysis

This study also analyses the numerical big data of citizens' experiences of urban natural resource assets and gradually introduces different dimensions of explanatory variables into the model to reduce the explanatory variables that may have multicollinearity. This work also refines the variable set that significantly affects the satisfaction of urban natural resource assets from the perspective of co-production. The procedure is described below.

Step 1. For p, we conduct the regression of independent variables $X_1, X_2, X_3 \dots X_P$ with the dependent variable Y to establish a univariate regression model $Y = \beta_0 + \beta_i X_i + \varepsilon$, where

Figure 2. perplexity degree of text-based big data



$i = 1,2,3 \ldots p$. The explanatory variable X_i is calculated, along with the values of the F test of the corresponding regression coefficient, which are respectively denoted as $F_1^{(1)}, F_2^{(1)} \ldots F_P^{(1)}$. Taking the largest value $F_{i1}^{(1)}$, i.e. $F_1^{(1)} = \max\left\{F_1^{(1)}, \ldots, F_P^{(1)}\right\}$., for a given significance level τ, we note the corresponding critical value F^1, where $F_{i1}^{(1)} \geq F^1$. Then, the variable X_{i1} introduces the regression model, and record I_1 is the selected variable index set.

Step 2. In establishing dependent variable Y and independent variables $\left\{X_{i1}, X_1\right\}, \ldots, \left\{X_{i1}, X_{i1-1}\right\} \ldots \left\{X_{i1}, X_p\right\}$, there are $p-1$ binary regression models. We calculate the regression coefficients of the variables and the statistic value of the F test $F_k^{(2)}\left(k \neq I_1\right)$. Then, we choose the largest value $F_{i2}^{(2)}$, i.e. $F_{i2}^{(2)} = \max\left\{F_1^{(2)}, \ldots, F_P^{(2)}\right\}$ for a given significance level τ. Note the corresponding critical value F^1, where $F_{i1}^{(2)} \geq F^2$. Then, the variable X_{i2} is introduced into the regression model. Otherwise, we terminate the variable introduction process.

Step 3. Consider a dependent variable to variable subset $\left\{X_{i1}, X_{i2}, X_P\right\}$. Repeat Step 2 for regression.

5. RESULTS

5.1. Topic Model for the Evaluation of Urban Natural Resource Assets

5.1.1. Topic Perplexity

As an ML-based text content analysis technology, the LDA topic model uses an unsupervised ML algorithm to determine the existence and distribution of different topics in a large sample text. Therefore, it is also considered an automatic text mining and analysis technology. As we all know, in the text big data of natural resource assets, parameter selection has important implications for the result analysis. In information science, the degree of perplexity is often used to indicate either the superiority or inferiority of the prediction results of probability models. In this section, we simulated and trained the big data of urban natural resource assets (ecotourism resources) comments and then obtained the perplexity degree diagram of urban natural asset evaluation simulated by LDA. As shown in Figure 2, when $K\left[15,20\right]$, the LDA topic model had the lowest perplexity and the best performance for the corpus. To further obtain the optimal parameters, according to the principle of interval selection, this study selected K to calculate 15, 16, 18 and 20, respectively, after which we obtained the distance between topic maps from Figure 3 to Figure 6. This process uses multidimensional scale analysis, extracts principal components as dimensions and distributes topics into two dimensions. The circle indicates the distribution of the corresponding topics, the size of the circle indicates the frequency (degree) of the topic and the distance between the centres of the blue circles reflects the proximity between topics.

Figure 3 shows the distribution of 15 topics in the data related to comments about urban natural resource assets. Comparing Figure 4 to Figure 6, when $K=15$, there is little difference in topic degree amongst the 15 categories (between 6% and 7% only), and the simulation effect is the best. This study will analyse the results (k = 15) later.

5.1.2. Topic Analysis

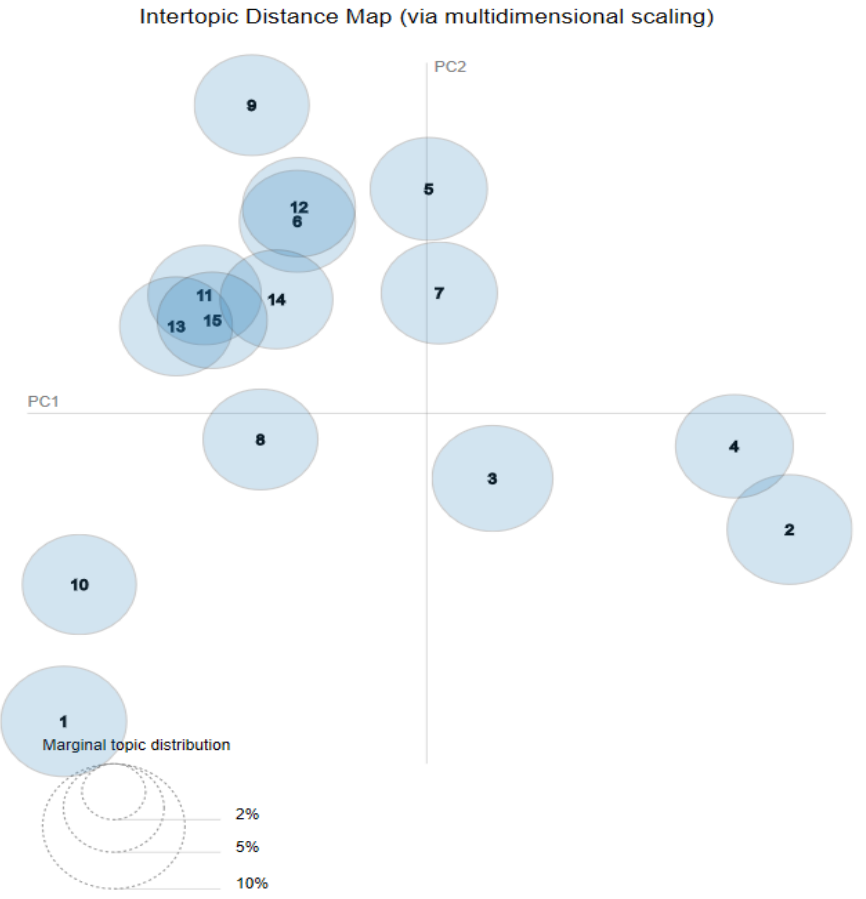
After combining the results of the optimal selection of the perplexity degree, this study extracts 15 topic probabilities of comments on urban natural tourist sites and obtains a keyword table corresponding to each topic. In this way, we are able to capture the structures of keywords in different topics and mine the potential needs of citizens amongst the identified keywords. Figure 7 shows the topic keywords in descending order of degree value. The word cloud visualisation method was used to display the keywords. As can be seen, the size of each word reflects the degree of correlation between the keywords and their topics. For example, topic 1 is the clustering word group with the largest degree value of 7.8%, and the step is the keyword with the closest relationship to topic 1.

On the whole, this study finds that, in addition to the scenery of urban natural tourist sites, citizens' co-production of urban natural resource assets is reflected in the entertainment and Q&V of scenic sites. In the evaluation of Q&V, there are some keywords, such as 'ticket price' (topic 6), 'discount' (topic 5), 'choice' (topic 4), 'come again' (topic 14), 'next time' (topic 14) and 'recommend' (topic 15). 'the overall feeling is that there is no equivalence between scenic spots and ticket prices. The price of the cable car is a little expensive, and the browsing content needs to be further developed'. 'the price is budget- friendly, the scenery is good and interesting, and it is worth recommending'.

In terms of entertainment, we found certain keywords, such as 'comfortable' (topic 9), 'not too bad' (topic 12), 'enthusiasm' (topic 9), 'joyful' (topic 14), 'relax' (topic 13) and other satisfactory experiences. We also found 'tired' (topic 14), 'too expensive' (topic 6), 'too crowded' (topic 5) and other negative feelings. For example, 'there are green grass and streams in the mountains, which are very comfortable and good places to swim on weekends', 'the elderly should take a cable car; it's too tired'.

In addition, the heat of the highest sites in Tianjin is mainly distributed in Jizhou District (topic 1). Skylon Tower (topic 10) and karst caves (topic 12) are also representative eco-tourism sites consistent

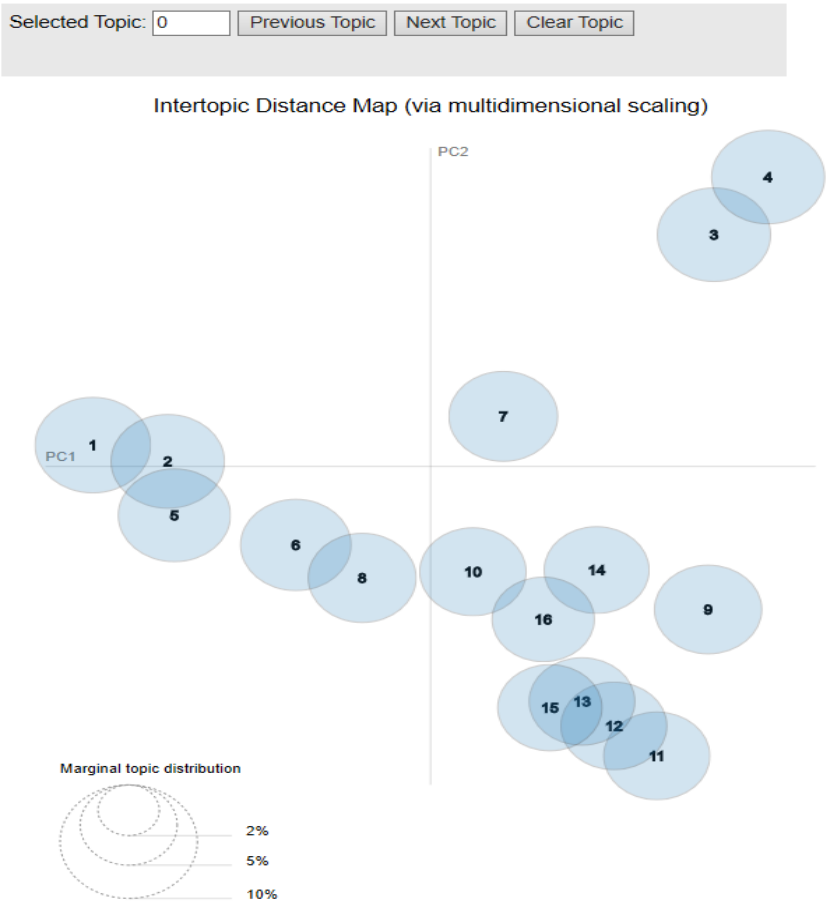
Figure 3. Topic degree (k = 15)



with the city brand and image publicised by Tianjin. As citizens often travel with their families on weekends, they attach great importance to the feelings of children, as reflected in keywords, such as 'youngster' (topic 4), 'husband' (topic 13), 'son' (topic 14), 'kid' (topic 14) and other keywords. For example, 'the mountain path is too cool. It is a good place for summer vacation. I saw a snake and many squirrels on the way, which frightened my son', 'my son likes it very much, and there are many outdoor swimming pools.'

Interestingly, citizens pay more attention to the combination of urban natural resource assets and historical culture, as reflected in keywords, such as 'culture' (topic 1), 'history' (topic 1), 'Qianlong' (topic 8) and 'Gangnam' (topic 8). For example, 'Emperor Qianlong of the Qing Dynasty visited Panshan mountain 32 times, leaving 1702 poems singing about Panshan'. As for the features and seasonal characteristics of different natural scenic sites in cities, citizens also recommend appropriate travel time intervals, as shown in several keywords, including 'summer' (topic 10), 'autumn' (topic 11) and 'winter' (topic 15) among others. As for citizens' impressions of urban natural resource assets, they generally pay attention to the characteristic natural landscapes, as shown in certain keywords, such as 'wonderful peak' (topic 8), 'strange stone' (topic 8), 'karst caves' (topic 12), 'thin strip of sky' (topic 11), 'stalactite' (topic 12) and 'ice falls' (topic 15), and the infrastructure of ecological tourism sites, as shown in several keywords, such as 'boardwalk' (topic 12), 'restaurant' (topic 10), 'scenic zone' (topic 10) and 'hotel' (topic 9). For example, 'the Karst Cave Scenic Spot has good

Figure 4. Topic degree (k = 16)



scenery and rich entertainment items, including natural landscape karst cave, glass plank road, dry sliding' and 'stalactites with different shapes are beautiful under the light'.

5.2. Factors Affecting the Evaluation of Urban Natural Resource Assets

In the formula shown in Section 3.2, we find that the rating data (scenery, entertainment and Q&V) serve as the most direct indicators reflecting citizens' perceptions of urban natural resource assets. The levels they indicate also reveal citizens' long-term comprehensive evaluation of urban natural resource assets, which refer to the four independent variables highlighted in this study. In addition, when urban natural resource assets belong to any of these categories (mountains, beaches, forest parks, lakes and geoparks), the value of the variable is 1; otherwise, the value of the variable is 0.

Figure 5. Topic degree (k=18)

Table 2 shows the results of the stepwise regression analysis. The positive and negative values of the coefficients in columns (1), (2), (3) and (4) indicate the degree of influence of the relevant models, while * reflects the significance level of related variables in different models. Meanwhile, the coefficient of R-squared (R) and the adjusted coefficient of R-squared (R2) reflect the regression effects of the different models. The R2 values of models (1)–(3) range from 0.7839–0.8329, indicating that the effect of the regression model is closer to the real situation as the number of independent variables increases. Model (4) is based on model (3) by adding a level variable. The value of R2 shows little change (from 0.8329 to 0.8328), indicating that model (3) is the best model result in this research topic.

In the three independent variables of model (3), the coefficients of scenery, taste and cost performance are all positive and have statistical significance ($\beta_1 = 0.508265$, $p < 0.01$; $\beta_2 = 0.1610179$, $p < 0.01$; $\beta_3 = 0.2632882$, $p < 0.01$, respectively), compared with all variables in the model (4). Moreover, even though the correlation coefficients of the above three variables are not significantly different from those of model (3), the level is not statistically significant, indicating that the level of scenic site does not directly affect the subjective feelings of citizens. In addition, from the control variables, we find that the geopark has a high significance level, indicating that tourists visiting Tianjin have a high evaluation of its geoparks and that the representative scenic site in the city is the Tianjin Jizhou National Geopark.

Figure 6. Topic degree (k = 20)

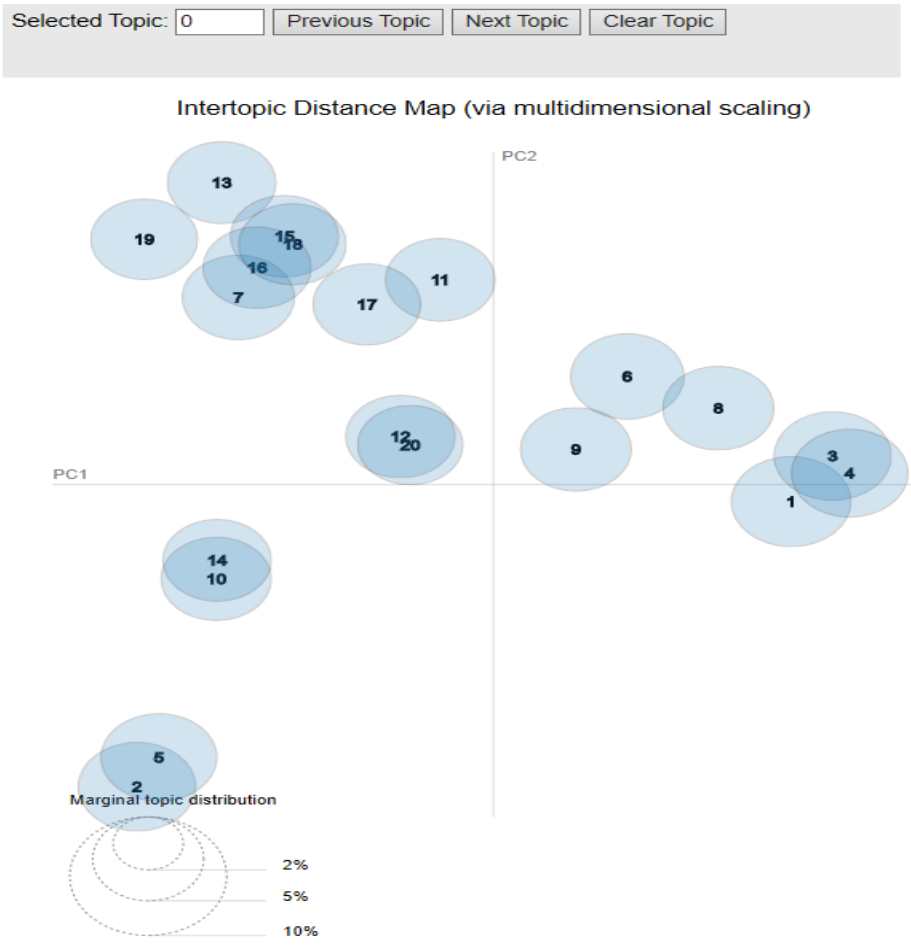


Figure 7. Word clouds of different topics (k = 15)

Table 2. Model evaluation results

	(1)	(2)	(3)	(4)
Scenery (S)	0.916373*** (103.26)	0.5858175*** (35.45)	0.508265*** (30.94)	0.5067746*** (30.72)
Entertainment (E)		0.3424767*** (23.02)	0.1610179*** (9.06)	0.16135*** (9.07)
quality & value (Q&V)			0.2632882*** (16.97)	0.2641521*** (17.00)
Level (L)				0.0064271 (0.96)
Mountain (M)	−0.0124974 (−0.39)	−0.018421 (−0.63)	−0.0190293 (−0.68)	−0.0062553 (−0.20)
Beach (B)	0.0045636 (0.02)	0.0221447 (0.11)	−0.0461952 (−0.24)	−0.0312141 (−0.16)
Forest Park (FP)	0.0092869 (0.34)	−0.0060719 (−0.24)	−0.0234162 (−0.99)	−0.0065224 (−0.22)
Lake (La)	0.0366052 (0.92)	0.0037036 (0.10)	−0.0216816 (−0.62)	−0.0064815 (−0.17)
Geopark (G)	0.2579085*** (5.86)	0.2147518*** (5.30)	0.2068012*** (5.34)	0.2290314*** (5.08)
_Cons	0.350851*** (6.74)	0.3682239*** (7.68)	0.3730227*** (8.15)	0.3336667*** (5.44)
N	2967	2967	2967	2967
R	0.7843	0.8171	0.8333	0.8334
R^2	0.7839	0.8166	0.8329	0.8328

Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.10

6. CONCLUSION AND DISCUSSION

Text has always been an important research object that enables social scientists to analyse and study certain problems (Lasswell et al., 1952). The reason why a large number of comment texts are called 'data' is that they possess valuable information and resources that need to be mined. The current study uses an ML algorithm and a statistical analysis model as analysis tools to explore the research problems regarding citizens' feelings and experiences of sharing urban natural resources through social network platforms in the era of big data. Results show that citizens' co-production behaviours play an indispensable role in the development of urban natural resource assets.

6.1. Conclusion

This study demonstrates the application of LDA topic modelling and stepwise regression methods from the perspectives of co-production and value co-creation theory to reveal citizens' preferences from the demand-side text-based and numerical data with regard to urban natural resource assets. The findings can help urban planners and managers identify and explore citizens' urban eco-tourism needs. At the same time, the results can promote the protective development of urban natural resources according to public demand. In this way, authorities can reduce the losses and damages sustained by the internal ecologies and natural landscapes of urban spaces and help realise the green and sustainable development of the city. In line with this, the following conclusions were obtained:

Firstly, the LDA topic model can identify the needs of citizens related to the natural resource assets of cities via BDA and the algorithm of natural processing. Through the comparison and selection of puzzle parameters, we find that when the number of subjects is 15, the semantic of text data is the clearest and the degree of the topic and the coverage between topics are more accurate.

Secondly, in the process of participating in and experiencing urban natural resource assets, citizens can more truly reflect their feelings whilst visiting the urban natural resource assets, as indicted in the emergence of multi-level emotional keywords (e.g. 'interesting', 'not too bad', 'too expensive', etc.). At the same time, cultural history, the surrounding infrastructures and unique landscape are the reference points that citizens focus on while exploring the surrounding tourism sites. Thus, modern government proposals should include the construction of ecological cities, combined with the three aspects mentioned above.

Finally, this study summarises the needs of LDA topic clustering into four aspects (scenery, entertainment, Q&V and level) and then collects the experiences of citizens related to urban natural resource assets. The conclusion reveals that scenery, entertainment and Q&V are all statistically significant. In addition, we find that geological scenic sites are closely related to the overall scores of citizens (also significant) and that they have common features with the results of LDA topic clustering, such as 'wonderful peak' (topic 8), 'strange stone' (topic 8), 'karst caves' (topic 12) and other keywords, which we generated in the LDA visualisation keywords.

6.2. Discussion

Combined with many conclusions of the above research, this study extends some management enlightenment. Firstly, urban natural resource assets need to reflect the citizens' dominant position (Levine, 1984), enhancing citizens' sense of acquisition and happiness in urban capitalized natural resources, which is helpful to improve the administrative efficiency of the government in protecting natural resource assets, and achieve the goal of co-construction and sharing of natural resources in the whole society; Secondly, with the enhancement of citizens' awareness of participation (lember et al., 2019), the government needs to further establish citizens' environmental awareness, pay attention to the value of natural resource assets, promote the co-protection and co-governance of natural asset system, especially the restoration of ecological environment, build wetland park system and promote the construction of multi-type and multi-functional wetland parks; Finally, managers and planners should give full play to the supporting role of science and technology and pay attention to citizens' behavioral preferences in urban natural resources (Athey, 2017; Li et al., 2020), build a technology platform based on citizen experience to promote the dynamic monitoring, supervision, evaluation and decision-making of natural resource assets.

7. SIGNIFICANCE AND FUTURE RESEARCH

7.1. Significance

This study contributes to theoretical research on urban natural resource assets from many aspects. Firstly, this study analysed the significance of citizen co-production in urban eco-tourism from the perspective of social media data. In addition, the sample data were divided into text data and scoring data. Finally, the advantages of the new ML algorithm and traditional regression analysis were combined to expand and enrich the theory of urban natural resource assets, which uses the LDA topic model.

This study also contributes to practical research on urban operation and natural resource asset protection. On the one hand, the findings can help the government assess the importance and heterogeneity of the potential dimensions of citizens' needs from user-generated data, which in turn, can help identify the inherent needs of the consumer market. On the other hand, this study also suggests

that relevant government departments should improve the infrastructure and brand construction of urban natural scenic sites to realise the goal of ecological and environmental protection.

7.2. Future Research

In this study, we found some interesting research results and, to some extent, addressed the research problems. However, it also has its share of limitations. For example, the optimal perplexity value obtained in this study is relative. Furthermore, the study has nesting defects between topics in data processing; thus, we will consider a more scientific LDA topic model for clustering in the future. Finally, given that public evaluation of urban natural resource assets is subjective and arbitrary, in the future, we will consider identifying the factors influencing citizens' satisfaction with urban natural resource assets (ecotourism resources) from a qualitative perspective.

ACKNOWLEDGMENT

This study is partially supported by the National Natural Science Foundation of China (No.71874120;No.72174139), the Philosophical and Social Science Planning Project of Tianjin (No. TJGL16-016), and the Postgraduate Research and Innovation Project of Tianjin (No.2019YJSB186).

REFERENCES

Anand, I. M., & Gaur, S. S. (2019). Consequences of consumers' emotional responses to government's green initiatives: Insights from a scenario-based experimental study. *Management of Environmental Quality*, *30*(1), 243–259. doi:10.1108/MEQ-02-2018-0045

Arnason, R. (2012). Property rights in fisheries: How much can individual transferable quotas accomplish? *Review of Environmental Economics and Policy*, *6*(2), 217–236. doi:10.1093/reep/res011

Athey, S. (2017). Beyond prediction: Using big data for policy problems. *Science*, *355*(6324), 483–485. doi:10.1126/science.aal4321 PMID:28154050

Barbier, E. B., & Burgess, J. C. (2001). Tropical deforestation, tenure insecurity, and unsustainability. *Forest Science*, *47*(4), 497–509.

Bastani, K., Namavari, H., & Shaffer, J. (2019). Latent Dirichlet allocation (LDA) for topic modeling of the CFPB consumer complaints. *Expert Systems with Applications*, *127*, 256–271. doi:10.1016/j.eswa.2019.03.001

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. doi:10.1145/2133806.2133826

Blei, D. M., Andrew, Y. N. G., & Michael, I. J. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Bohn, H., & Deacon, R. T. (2000). Ownership risk, investment, and the use of natural resources. *The American Economic Review*, *90*(3), 526–549. doi:10.1257/aer.90.3.526

Brudney, J. L. (2020). Rethinking coproduction: Amplifying involvement and effectiveness. *Journal of Chinese Governance*, *5*(1), 8–27. doi:10.1080/23812346.2019.1666542

Brudney, J. L., & England, R. E. (1983). Toward a definition of the co-production concept. *Public Administration Review*, *43*(1), 59–65. doi:10.2307/975300

Cardullo, P., & Kitchin, R. (2018). Being a 'citizen' in the smart city: Up and down the scaffold of smart citizen participation in Dublin, Ireland. *GeoJournal*, *84*(1), 1–13. doi:10.1007/s10708-018-9845-8

Chen, J., Gao, M., Mangla, S. K., Song, M., & Wen, J. (2020). Effects of technological changes on China's carbon emissions. *Technological Forecasting and Social Change*, *153*, 119938. doi:10.1016/j.techfore.2020.119938

Clifton, J. (2019). ICT-enabled co-production of public services: Barriers and enablers. A systematic review. *Information Polity*, 1–24. doi:10.3233/IP-190122

Cui, L., Duan, H., Mo, J., & Song, M. (2021). Ecological compensation in air pollution governance: China's efforts, challenges, and potential solutions. *International Review of Financial Analysis*, *74*, 101701. doi:10.1016/j.irfa.2021.101701

Gao, X. (2018). Networked co-production of 311 services: Investigating the use of Twitter in five US cities. *International Journal of Public Administration*, *41*(9), 712–724. doi:10.1080/01900692.2017.1298126

Godoy, R., Kirby, K., & Wilkie, D. (2001). Tenure security, private time preference, and use of natural resources amongst lowland Bolivian Amerindians. *Ecological Economics*, *38*(1), 105–118. doi:10.1016/S0921-8009(01)00144-6

Grafton, Q., Squires, D., & Fox, K. J. (2000). Private property and economic efficiency: A study of a common-pool resource. *The Journal of Law & Economics*, *43*(2), 679–714. doi:10.1086/467469

Grainger, C. A., & Costello, C. J. (2014). Capitalizing property rights insecurity in natural resource assets. *Journal of Environmental Economics and Management*, *67*(2), 224–240. doi:10.1016/j.jeem.2013.12.005

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management*, *59*, 467–483. doi:10.1016/j.tourman.2016.09.009

Hardin, G. (2009). The tragedy of the commons. *Journal of Natural Resources Policy Research*, *1*(3), 243–253. doi:10.1080/19390450903037302

Hoffman, M. D., Blei, D. M., Wang, C., & Paisley, J. (2013). Stochastic variational inference. *Journal of Machine Learning Research*, *14*(5), 1303–1347.

Hou, Z., Cui, F., Meng, Y., Lian, T., & Yu, C. (2019). Opinion mining from online travel reviews: A comparative analysis of Chinese major OTAs using semantic association analysis. *Tourism Management*, *74*, 276–289. doi:10.1016/j.tourman.2019.03.009

Huang, C., Chen, W., & Yi, H. (2020). Collaborative networks and environmental governance performance: A social influence model. *Public Management Review*, ●●●, 1–22. doi:10.1080/14719037.2020.1795229

Jennings, E. T. Jr, & Hall, J. L. (2012). Evidence-based practice and the use of information in state agency decision making. *Journal of Public Administration: Research and Theory*, *22*(2), 245–266. doi:10.1093/jopart/mur040

King, S., & Cotterill, S. (2007). Transformational government? The role of information technology in delivering citizen-centric local public services. *Local Government Studies*, *33*(3), 333–354. doi:10.1080/03003930701289430

Lasswell, H. D., Lerner, D., & Pool, I. (1952). *The comparative study of symbols: An introduction*. Stanford University Press. doi:10.1037/14763-000

Lember, V., Brandsen, T., & Tõnurist, P. (2019). The potential impacts of digital technologies on co-production and co-creation. *Public Management Review*, *21*(11), 1665–1686. doi:10.1080/14719037.2019.1619807

Levine, C. H., & Fisher, G. (1984). Citizenship and service delivery: The promise of coproduction. *Public Administration Review*, *44*, 178–189. doi:10.2307/975559

Li, L., Hao, T., & Chi, T. (2017). Evaluation on China's forestry resources efficiency based on big data. *Journal of Cleaner Production*, *142*, 513–523. doi:10.1016/j.jclepro.2016.02.078

Li, L., Ma, S., Han, X., Zheng, C., & Wang, D. (2020). Data-driven online service supply chain: A demand-side and supply-side perspective. *Journal of Enterprise Information Management*, *34*(1), 365–381. Advance online publication. doi:10.1108/JEIM-11-2019-0352

Linders, D. (2012). Frome-government to we-government: Defining a typology for citizen coproduction in the age of social media. *Government Information Quarterly*, *29*(4), 446–454. doi:10.1016/j.giq.2012.06.003

Liu, Z. W., & Park, S. W. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management, 47*, 140-151.

Lubell, M., & Fulton, A. (2008). Local policy networks and agricultural watershed management. *Journal of Public Administration: Research and Theory*, *18*(4), 673–696. doi:10.1093/jopart/mum031

McColl-Kennedy, J., Vargo, S., Dagger, T. S., Sweeney, J. C., & van Kasteren, Y. (2012). Health care customer value cocreation practice styles. *Journal of Service Research*, *15*(4), 370–389. doi:10.1177/1094670512442806

Meijer, A. (2015). E-governance innovation: Barriers and strategies. *Government Information Quarterly*, *32*(2), 198–206. doi:10.1016/j.giq.2015.01.001

Mondal, S., & Samaddar, K. (2021). Issues and challenges in implementing sharing economy in tourism: A triangulation study. *Management of Environmental Quality*, *32*(1), 64–81. doi:10.1108/MEQ-03-2020-0054

Moynihan, D. P. (2008). *The dynamics of performance management: Constructing information and reform*. Georgetown University Press.

Nabatchi, T., Sancino, A., & Sicilia, M. (2017). Varieties of participation in public services: The who, when, and what of coproduction. *Public Administration Review*, *77*(5), 766–776. doi:10.1111/puar.12765

Okan, T., Köse, N., Arifoğlu, E., & Köse, C. (2016). Assessing ecotourism potential of traditional wooden architecture in rural areas: The case of Papart Valley. *Sustainability*, *8*(10), 974. doi:10.3390/su8100974

Osborne, S. P., Radnor, Z., & Strokosch, K. (2016). Co-production and the co-creation of value in public services: A suitable case for treatment? *Public Management Review*, *18*(5), 639–653. doi:10.1080/14719037.2015.1111927

Ostrom, E. (1972). Metropolitan reform: Propositions derived from two traditions. *Social Science Quarterly*, *53*(3), 474–493.

Ostrom, E. (1996). Crossing the great divide: Coproduction, synergy, and development. *World Development*, *24*(6), 1073–1087. doi:10.1016/0305-750X(96)00023-X

Paletti, A. (2016). Co-production through ICT in the public sector: When citizens reframe the production of public services. In *Digitally supported innovation*. Springer. doi:10.1007/978-3-319-40265-9_10

Parks, R. B., Baker, P. C., Kiser, L., Oakerson, R., Ostrom, E., Ostrom, V., Percy, S. L., Vandivort, M. B., Whitaker, G. P., & Wilson, R. (1981). Consumers as co-producers of public services: Some economic and institutional considerations. *Policy Studies Journal: the Journal of the Policy Studies Organization*, *9*(7), 1001–1011. doi:10.1111/j.1541-0072.1981.tb01208.x

Plé, L., & Cáceres, R. C. (2010). Not always co-creation: Introducing interactional codestruction of value in service-dominant logic. *Journal of Services Marketing*, *24*(6), 430–437. doi:10.1108/08876041011072546

Provan, K. G., Fish, A., & Sydow, J. (2007). Interorganisational networks at the network level: A review of the empirical literature on whole networks. *Journal of Management*, *33*(3), 479–516. doi:10.1177/0149206307302554

Provan, K. G., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. *Journal of Public Administration: Research and Theory*, *18*(2), 229–252. doi:10.1093/jopart/mum015

Pueyo-Ros, J., Ribas, A., & Fraguell, R. M. (2018). Uses and preferences of visitors to coastal wetlands in tourism destinations (Costa Brava, Spain). *Wetlands*, *38*(6), 1183–1197. doi:10.1007/s13157-017-0954-9

Song, M., Xie, Q., Wang, S., & Zhou, L. (2021). Intensity of environmental regulation and environmentally biased technology in the employment market. *Omega*, *100*, 102201. doi:10.1016/j.omega.2020.102201

Sun, B., Ao, C., Wang, J., Mao, B., & Xu, L. (2020). Listen to the voices from tourists: Evaluation of wetland ecotourism satisfaction using an online reviews mining approach. *Wetlands*, *40*(5), 1–15. doi:10.1007/s13157-020-01274-9

Voorberg, W. H., Bekkers, V. J. J. M., & Tummers, L. G. (2015). A systematic review of co-creation and co-production: Embarking on the social innovation journey. *Public Management Review*, *17*(9), 1333–1357. doi:10.1080/14719037.2014.930505

Vu, H. Q., Li, G., & Law, R. (2019). Discovering implicit activity preferences in travel itineraries by topic modeling. *Tourism Management*, *75*, 435–446. doi:10.1016/j.tourman.2019.06.011

Weaver, D. B., & Lawton, L. J. (2007). Twenty years on: The state of contemporary ecotourism research. *Tourism Management*, *28*(5), 1168–1179. doi:10.1016/j.tourman.2007.03.004

West, D. M. (2004). E-government and the transformation of service delivery and citizen attitudes. *Public Administration Review*, *64*(1), 15–27. doi:10.1111/j.1540-6210.2004.00343.x

Ye, Q., Law, R., Li, S., & Li, Y. (2011). Feature extraction of travel destinations from online Chinese-language customer reviews. *International Journal of Services Technology and Management*, *15*(1–2), 106–118. doi:10.1504/IJSTM.2011.038665

Yi, H., Suo, L., Shen, R., Zhang, J., Ramaswami, A., & Feiock, R. C. (2017). Regional governance and institutional collective action for environmental sustainability. *Public Administration Review*, *78*(4), 556–566. doi:10.1111/puar.12799

Zhang, Y. L., Liu, X. S., & Vedlitz, A. (2020). How social capital shapes citizen willingness to co-invest in public service: The case of flood control. *Public Administration*, *98*(3), 1–17. doi:10.1111/padm.12646

Zhu, Q., Wu, J., Li, X., & Xiong, B. (2017). China's regional natural resource allocation and utilisation: A DEA-based approach in a big data environment. *Journal of Cleaner Production*, *142*, 809–818. doi:10.1016/j.jclepro.2016.02.100

Lei Li is a professor in College of Management and Economics at Tianjin University. His research interests include service supply chain, urban sustainable development, and performance management. He has published in Technological Forecasting and Social Change, Sustainability, Ecological Indicators, Science of the Total Environment, Mathematical and Computer Modelling, and Transport Policy.

Shaojun Ma is a Ph.D. student in College of Management and Economics at Tianjin University. His research interests include data mining, policy analysis and management, co-production, resources allocation in urban agglomeration. He has published in Journal of Enterprise Information Management, Journal of Public Administration (in Chinese), and International Journal of Innovation, Management and Technology.

Runqi Wang is an undergraduate student in College of Management and Economics at Tianjin University. Her research interests include quantitative methods, data analysis and data security.

Yiping Wang is an undergraduate student in Software & Communication School, Tianjin Sino-German University of Applied Sciences. His research interests include machine learning, data mining.

Yilin Zheng is a Ph.D. student in College of Management and Economics at Tianjin University. Her research interests include urban human settlements, environmental collaborative governance, ecological benefits of performance. She has published in Science of the Total Environment.