 doi:10.20944/preprints202009.0195.v1

Article

Public participation in smart-city governance: a qualitative content analysis of public comments in urban China

Zhe Gao [1], Siqin Wang [2]* and Jiang Gu [3]

[1] The College of Urban & Environmental Sciences, Key Laboratory for Geographical Process Analysis & Simulation, Central China Normal University; gaozhe@mail.ccnu.edu.cn
[2] School of Earth and Environmental Sciences, University of Queensland; s.wang6@uq.edu.au
[3] The College of Urban & Environmental Sciences, Key Laboratory for Geographical Process Analysis & Simulation, Central China Normal University; gujiang@mail.ccnu.edu.cn
* Correspondence: s.wang6@uq.edu.au

Abstract: Public participation is crucial in the process of urban governance in smart-city initiatives to enable urban planners and policy makers to take account of the real public needs. Our study aims to develop an analytical framework using citizen-centred qualitative data to analyse urban problems and identify the areas most needed for urban governance. Taking a Chinese megacity as the study area, we first utilise a web-crawling tool to retrieve public comments from an online comment board and employ the Baidu Application Programming Interfaces and a qualitative content analysis for data reclassification. We then analyse the urban problems reflected by negative comments in terms of their statistical and spatial distribution, and the associative factors to explain their formation. Our findings show that urban problems are dominantly related to construction and housing, and most frequently appear in industry-oriented areas and newly-developed economic development zones on the urban fringe, where the reconciling of government-centered governance and private governance by real estate developers and property management companies are most needed. Areas with higher land price and a higher proportion of aged population tend to have less urban problems, while various types of civil facilities affect the prevalence of urban problems differently.

Keywords: Urban governance, public participation, public comments, web-crawling data, qualitative content analysis, urban China

1. Introduction

Public participation in the process of urban planning and governance has been unprecedentedly encouraged and promoted in the context of smart-city initiatives (Viale Pereira et al., 2017). The essential of building a cyber, digital, and wired smart city is the modern information and communications technology (ICT) infrastructures (i.e., Internet of Things and cloud computing). ICT-based urban systems provoke the new concept of e-governance comprising ICTs to support democratic processes, enhance services delivery and effective resource management for urban administration to citizens [1], and e-participation involving technology-mediated interaction among civil society, formal politics and the public administration. Most of studies on e-governance and e-participation have been developed in the Western context, largely relating to advanced technologies, frameworks, models, and methods to support decision-making in terms of urban performance, governance, and policy [2-5]. However, such studies in the Eastern context are under-explored and need to be extended to developing countries with large population base such as China for a holistic understanding of smart-city initiatives.

The increasing prevalence of mobile devices and emergence of clouding technologies generates a large volume of quantitative and qualitative information created by different agents (i.e.,

governments, organisations, and individuals) and shared through social media and online platforms. Internet-centred big data is increasingly attractive as a data source for urban analysis to provide new insights for evaluating public participation in urban governance [6]. In the current scholarship, most researchers focus on macro-level policy- and decision-making based on analysing collective-level big data produced by civil collective action in urban affairs such as transportation, emergency management, health care, or environment management [7-9]. However, citizen-centred big data at the individual level that provides temporal-spatial insights into citizen behaviour and individual opinions gain relatively scarce attention, covered merely by a few recent studies in the Chinese literature [10, 11]. There is a pressing need in the current research agenda to incorporate opinions and comments of citizens in urban governance to support decision making based on the real public needs [12].

To fulfil the aforementioned knowledge gaps, our study contributes to an emerging research agenda of e-governance facilitating the formation of smart cities, by constructing an analytical framework to retrieve qualitative big data to reflect urban problems and to examine the spatial distribution and associative factors of urban problems to guide through citizen-directed participatory practices in urban governance in a Chinese context. Taking Wuhan City, the largest capital city in Central China as our study area, we aim to address the following research questions: 1) What is the spatial distribution of urban problems? 2) What is the spatial relationship between urban problems and surrounding areas? and 3) To what extent urban problems are explained by factors related to demographic and socioeconomic profile of communities at the micro level and civil-facility provision at the macro level? To answer these questions, we commence with retrieving public comments via an online comment board by a web-crawling tool, and employ the Baidu Application Programming Interfaces and a qualitative content analysis for data reclassification. We then analyse the urban problems reflected by negative comments in terms of their statistical and spatial distribution, and the associative factors to explain their formation. By doing so, we identify the areas in most need of urban governance and provide political implications on public participation and e-governance in the smart-city initiatives, which fundamentally roots in the political-economic institutions characterized in the Chinese context.

The rest of this paper is organised in the following way. The next section reviews e-governance and e-participation in the smart-city initiatives, content analysis by using big data and the rationale of urban governance in China. Then an analytical framework for retrieving and processing data is introduced, together with the description of data sources and methodologies. This is followed by results covering the statistical and spatial distribution of urban problems identified by the negative public comments, the spatial relationship between urban problems with their surrounding areas, and the associative factors that explain urban problems. The final section presents a number of conclusions and discussion of using our analytical framework to analyse wider range of qualitative data.

2. Literature review

2.1 E-governance and e-participation in smart cities

Smart cities encompassing emerging technologies and urban sustainability have become an increasingly popular concept advocated by governments in countries world widely [13]. A smart city is defined as a public-private urban system among governments, industries, organisations, and citizens centred around sustainable economic growth and quality of life through participatory urban governance [14]. The essential of building a cyber, digital, wired, informational or knowledge-based smart city is the modern information and communications technology (ICT) infrastructures (i.e., Internet of Things and cloud computing), which not only changes in the way people live and work but also fosters innovation and knowledge creation [15]. What accompanies with a rising smart city is the e-governance, comprising ICTs to support democratic processes, enhance services delivery and effective resource management for urban administration to citizens [1]. ICT-based e-governance is

heavily in reliance on a wealth of temporal-spatial big data about the urban environment collected through ubiquitous digital devices and platforms, and contributed by various agents in urban systems (i.e., governments, industries, non-profit organisations, private sectors, and citizens). As such, e-governance inevitably involves citizen participation and collaboration among stakeholders, provoking the concept of e-participation in the implementation of smart-city initiatives. [5, 16, 17].

Public participation in urban governance original refers to a governing arrangement where one or more agencies directly engage non-state stakeholders in a collective decision-making process in order to achieve a consensus oriented and deliberative goal of collaborative governance [18]. Enabled by ICTs, e-participation or e-governance participation involves technology-mediated interaction among civil society, formal politics and the public administration. E-participation has been implemented in different countries as part of e-governance for policy-relevant issues, such as communicating information on legal issues and negotiating the benefit of different agents involved in urban planning process [19]. The popularity of digital devices (i.e., smart phones and tablets) facilities the prevalence of e-participation, through which citizens interact with people living in the same neighbourhood or having similar interests, form online communities and express their opinions on urban environment and policies. Engaging in an online environment can be especially valuable for people with disabilities, as restricted physical mobility or other life circumstances might make it more difficult for them to meet others offline. Such online public comments and options, mostly anonymous, can be easily supported and followed by other Internet users and normally characterised by highly-sentiment tendency. Studies on e-governance and e-participation are relatively more popular in the Western context and largely divided into two streams – technological and managerial [20]. The former exploits advanced technologies to enhance e-governance and e-participation of citizens [1-3], while the later aims to design conceptual frameworks, models, and methods to support decision-making in terms of urban performance, governance, and policy on the basis of experience [5, 15, 21]. However, such studies beyond Western context in non-English literature are much under-explored and need to be extended to developing countries with large population base such as China for a holistic understanding of smart-city initiatives.

2.2 Content analysis in the era of big data

Big data collection and analysis plays an indispensable role in the smart-city initiatives to enhance the reliability and efficiency of e-governance and e-participation, and in turn to strengthen the intelligence of ICT-based urban systems [22]. The increasing prevalence of mobile devices and emergence of clouding technologies, has contributed to the so-called 'data revolution' [23]. As a large volume of quantitative and qualitative information created by different agents (i.e., governments, organisations, and individuals) and shared through social media and online platforms become relatively easy to capture and analyze in real time, Internet-centred big data is increasingly attractive as a data source for urban analysis to provide new insights for evaluating public participation in urban governance [6]. Different to traditional approaches stimulating public participation in policy decisions through suggestions and indications that follow a typical top-down decision making process, ICTs, and especially social media and online platforms, are used to widen the number of participants in the public debate and to give a voice to individuals who are not usually willing to or not convenient to participate in person in public debates such as disabilities and older people. Online forums, comment boards, communicative platforms development by governments at different levels engage citizens through a mix of online initiatives which collect their points of view, opinions, comments, criticisms, and suggestions on urban environment. In the current scholarship using Internet-centred big data, most researchers focus on macro-level policy- and decision-making based on analysing collective-level big data produced by civil collective action in urban affairs such as transportation, emergency management, health care, or environment management [7-9]. However, citizen-centred big data at the individual level that provides temporal-spatial insights into citizen behaviour and individual opinion have significant potential value but gain relatively scarce attention. Although a few studies along the aforementioned direction emerged recently [10, 11], they are merely

published in the Chinese literature with limited access by international readers. Our study aims to extend this research realm with more robust research design and outcomes.

Analysing cities with citizen-centred big data brings different challenges, including approaches and tools for data retrieving and mining, and methods used to analyse such data. Citizen-centred big data to reflect public opinions are usually text-based and geotagged, retrieved from social media and online platforms by advanced technologies of web data extraction such as web crawling or harvesting. The unstructured nature of such qualitative data makes it difficult to conduct meaning research, and thus requires different analytical methodologies compared to those used to deal with traditional census or survey data. Many methods have been applied in studies of online messages and comments, for example, textual analysis, triangulation analysis, discourse analysis, narrative analysis, and content analysis [24]. Among these, content analysis is one of the most commonly used methods to deal with qualitative big data [25]. Qualitative content analysis (QCA) was first used as an analytic technique to examine a wide range of textual materials, including communications, narrative responses, survey questions, interviews, focus groups, observations, political speeches, advertisements, and printed media such as magazine articles, books, or manuals [26]. The essential of QCA is to classify written or oral materials into identified categories of similar meanings [26]. These categories need be mutually exclusive and exhaustive and no data should fall between two categories or be placed in more than one category [27]. In this case, researchers need to determine how best to categorise data if data lacks a single meaning or interpretation. QCA not only enables researchers to focus on the themes and topics of the categories to investigate the distribution, meaning and relations across categories, but also allows a great amount of textual data to be reduced into numbers and frequencies that are suitable for statistical analysis [28]. However, QCA has also been criticised as being labour-intensive and time-consuming, and inappropriate for open explorative research. Thanks for the development of advanced computing interfaces such as application programming interfaces (API) which allows the interactions between multiple software intermediaries, the reliability and efficiency of QCA has been much improved because API can be entirely customised and specific to a certain purpose of data manipulation. For example, a wide range of API have been developed to provide extension mechanisms for data conversion and categorisation, retrieving locational information, and identifying sentiment tendency [29-32]. As such, our study primarily develops an analytical framework embedding QCA and API to analyse citizen-centred qualitative big data.

2.3 E-governance and e-participation in smart cities

In the Western context, public participation in urban governance is often seen as a reflection of democracy to serve various political rationales, including improving policy decisions, realizing democratic values and stimulating policy implementation [12]. Differently, public participation in the Chinese context is largely government-led, which has been evolved with the progress of the institutional changes, fundamentally the government's deepened market reforms relating to land and housing development since the 1990s [33]. Before 1990, socialist work-unit housing was predominantly provided and governed by government and such state provision and governance of housing was ended in 1998 [34]. As alternatives, newer residential development projects including gated communities and master-planned estates have been widely proliferated as one of the dominant residential neighbourhood forms in different regions in favour of the free housing market and the private sector [35]. The growing prevalence of gated communities and master-planned estates reflects a shift from the government-initiated governance towards a neoliberal mode of urban governance, emerging with a new phenomenon of private governance within which public participation plays an increasing role [36, 37].

Private governance in the Chinese context was defined as providing public services and administration within communities through organisations that were privately held by private sectors and through joint governance by residents [38]. On one hand, through capital accumulation and implemented incentives, state and local governments draw real estate developers and property management companies into developing residential land, providing neighbourhood services and

facilities, and maintaining space, landscape, and environment within a community [39]. Residents pay property management companies for the provision of neighbourhood services and facilities, and environmental maintenance. In this sense, private governance through a fee-paying system has also been understood as a market-driven process, and such market provision of public goods and services reduces public expenditure on community governance [40]. On the other hand, gated communities have residents who elect the homeowners' association to run the community by their own rules, to secure the value of private property, to enforce neighbourhood covenants and to cope with local governance failure [37]. The use of private compacts in these gated estates recasts residents' everyday practices and reshapes residents' engagement with the public realm [41]. Changed governance and deepening in housing privatisation in China has received much attention with discussion of both advantages and disadvantages in existing housing studies [40, 42].

Many studies on liberalised housing system agree that private governance has salient goodness, including increasing public participation and awareness of community management [43, 44], retreating governments from the provision of neighbourhood services, supplementing inadequate public services and boosting property management industries in urban China. However, private governance has also stimulated public debates [45]. On one hand, although homeowners' associations have gradually emerged for managing neighbourhoods, they lack autonomy, financial independence and have poor representation of residents because residents do not have sufficient power to determine the system of neighbourhood governance but more towards to supervise property management companies. On the other hand, property management companies do not guarantee the efficiency of neighbourhood services provision because service delivery is managed in accordance with property management fees, leading to discrepancies of service and administration across communities because a wealthy community has "a supply of civil goods which apparently are better acquired there than elsewhere" in the urban area [46] (p.153). Also because of the fee-paying system, property management companies may tend to exclude the public space beyond the boundary of their own communities, which generates grey space fulfilled with nuisances, conflicts, and urban problems that may be neglected by state governments as well [40]. In this sense, the appearance of local problems in the gap between government-centralised and private governance is subject to both to the characteristics of land use and urban space at the macro level and the socioeconomic profile of communities at the micro level [47]. Thus, with the aforementioned considerations, our study selects factors relating to the demographic and socioeconomic characteristics of communities at the micro level as well as the public service configuration at the macro level to examine how such factors are associated with urban problems reflected by public comments in urban China.

3. Study context, data and Methods

3.1. Study context

Our study area is Wuhan City, the capital city of Hubei province located in the central plain of China (Figure 1). Wuhan had 11.06-million urban residential population and a total area of 8494.41 km^2 in 2018, comprising 2 national-level economic development zones (EDZ) and 13 districts, or total 186 jiedaos (sub-districts) [48]. The 2 EDZs are Wuhan EDZ and Donghu EDZ; the 13 districts include 7 districts in the main urban area (Jiang'an, Jianghan, Qiaokou, Hanyang, Wuchang, Qingshan and Hushan) and 6 new districts located in the outer city (Jiangxia, Huangpi, Xingzhou, Dongxi Hu, and Hannan). Wuhan had a GDP of RMB 1484.73 billion (Chinese Yuan) in 2018, making the city the largest economy in Central China [49]. We use Wuhan for our empirical study for the following reasons. First, as the most important political, economic, transport, educational, and cultural hub in Central China, Wuhan has been defined as the central city at the national level, and thus has profound effects on surrounding provinces. Also, Wuhan represents some common characteristics of hinterland cities in China in relation to its geographic location, industrial structures, and development trajectories. Second, Wuhan has experienced rapid urbanization and significant economic growth since the early 2000s, alongside the emergence of a series of urban problems caused by industrial transformation, urban regeneration, and land development. Third, Wuhan municipal

government has started to establish intelligent urban systems to accelerate the smart-city development, to implement the smart-city initiatives, and to widen the public participation in the process of urban planning and policy making. Such comprehensive systems for urban administration provide rich multidimensional data that are reliable in quality, quantity, and ease of access. Hence, Wuhan serves as a good test case to examine how to use citizen-centred qualitative big data to reflect urban problems and to demonstrate how public participation helps with urban government in the smart-city operation.

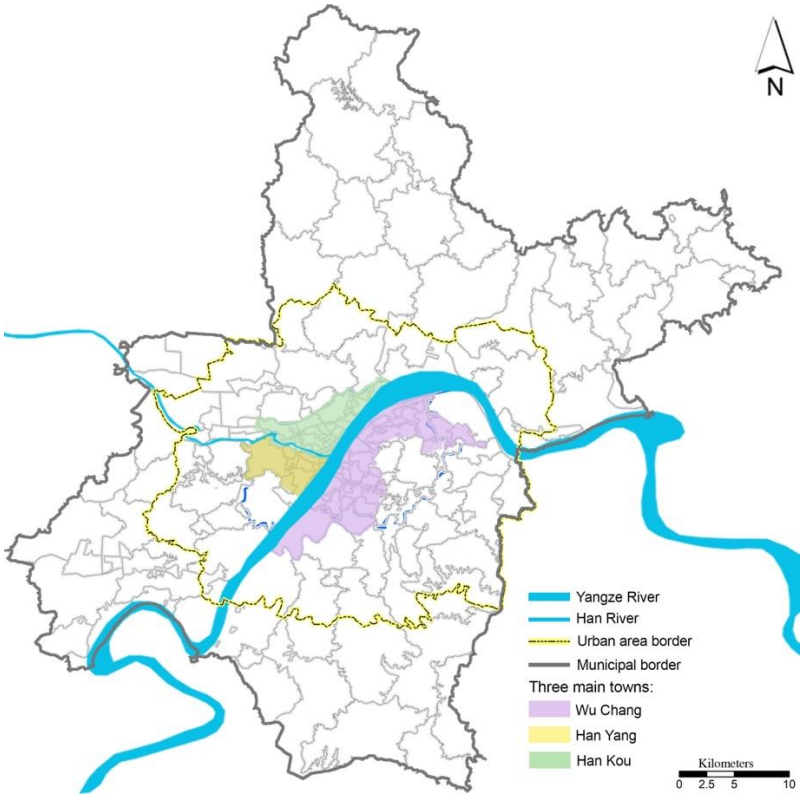

Figure 1. Study area

3.2. Data collection and manipulation

The Chinese statistical geographic hierarchy follows the structure of municipal (*shi* in Chinese), district (*qu*), sub-district (*jiedao*), and community (*xiao qu*), while the jiedao is designed as the basic building block of a city proper and the smallest census unit with the availability of digital boundaries. At the data pre-processing stage, there is a need to aggregate data at a community level to a jiedao level however the local government of Wuhan City does not provide correspondence data to convert geographic units in different levels. Thus, we construct a correspondence file for the conversion of spatial units by web crawling qualitative data from *LianJia*, one of the most popular real estate websites in China (www.lianjia.com). The correspondence file includes the name of communities in the whole municipal area of Wuhan, together with its jiedao and district where it is located, which will be used for data aggregation at the later stage. Next, we collect primary data of public comments from Wu Han Comment Board (WHCB) (http://liuyan.cjn.cn/) through web crawling. WHCB is an online board where citizens can post their comments, suggestions and complaints about urban problems and public issues. Utilising *Houyi Collector*, an open-source web crawling program (http://houyicaiji.com/), we retrieve a total number of 65,295 records of public comments posted on WHCB from Nov 16th, 2017 to Jun 26th, 2019, and match the name of jiedaos in these records with the correspondence file we previously created and exclude the unmatched ones. Through data filtering and clearing, the final records we use for the analysis are 12,164. This text-based dataset has attributes including title, description, nature, and category. We then re-classify the default category of records (originally 193 sub-categories) into 22 categories and 5 broad categories based on the nature of records

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 September 2020 doi:10.20944/preprints202009.0195.v1

(Figure 2): 1) safety and security, 2) construction and housing, 3) environment and pollution, 4) infrastructure and configuration, and 5) social service and administration. We further aggregate the dataset at the community level to the jiedao level by category, generating the number of negative and non-negative records in each category at the jiedao level. To examine the associative factors of the formation of public comments, we further collect data of land value from Wuhan land grade and benchmark land price standard [50], data of public service configurations retrieved from Point of Interest [51], and data of aged population above 65 from the 2015 population census [49].

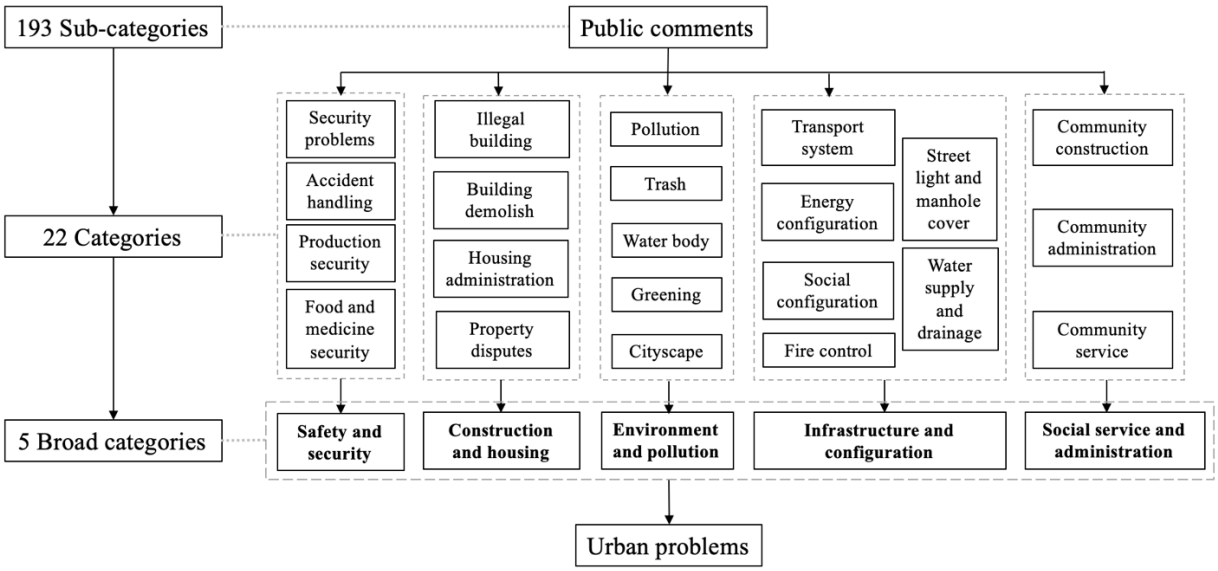

Figure 2. Data re-categorisation

3.3. Methodology

We develop an analytical framework integrating qualitative content analysis and Baidu Artificial Intelligence Open Platform (https://ai.baidu.com/), which is embedded by various application programing interfaces (API) and functionalities for natural langua ge processing. Through the API of Named Entity Recognition, we commence with extracting the name of communities from the title of records and using the correspondence file generated previously to assign these communities where public issues occur to the jiedao level. Next, we employ the API of Sentiment Tendency to analyse the sentiment tendency of comment providers based on the description of records, generating a score of sentiment and confidence. Sentiment represents the sentiment tendency of comments as negative, neutral and positive (coded as 0, 1 and 2 respectively); Confidence is a percentage indicating how much a certain public comment belongs to the sentiment tendency defined by Sentiment (80% used in our study). The above procedure distinguishes 9,488 negative records and 2676 non-negative (combining neutral and positive) records, covering 127 jiedaos. Given to the non-negative records only accounting for 22% of the total records and dominantly relating to construction and housing, our ensuing analysis and visualisation focus on the negative comments to reflect urban problems as they involve more diversities and variations across category and space.

We then conduct a descriptive analysis to summarise the number of urban problems and map out their spatial distribution at the jiedao level and further examine the spatial relationship between problem-concentrated areas with their surrounding areas by a local index of spatial autocorrelation (LISA). LISA, in particular the local Moran's I, was utilised to capture the spatial distribution of socio-economic characteristics. LISA has been used widely in the literature to indicate spatial clustering, commonly referred to as hotspot analysis [52]. In GeoDa software, the LISA cluster map is generated to illustrate the distribution of the local Moran's I statistic across observations [53]. LISA values are colour coded by type of spatial autocorrelation: the high-high and low-low clusters (positive local spatial autocorrelation) are typically referred to as spatial clusters; while the high-low and low-high clusters (negative local spatial autocorrelation) are spatial outliers. The global Moran's I statistic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 September 2020 doi:10.20944/preprints202009.0195.v1

indicates the mean of the local Moran's I statistics, working as an indicator of the spatial autocorrelation across the whole region. In order to obtain more robust results, the number of permutations was set to 999 along with a threshold significance level of p = 0.05.

Finally, we employ a pairwise Pearson correlation analysis and then a spatial lag regression (SLR) analysis to examine the association between urban problems and factors relating to the demographic and socioeconomic profile of areas at the micro level and public service provision at the macro level. The bivariate Pearson correlation serves as a pre-test of the relationship among above variables to reduce the effect of variables' collinearity on calculation. A Pearson correlation coefficient ranges between -1 and 1; 1 means total positive correlation, 0 means no correlation, and −1 mean total negative correlation [54]. SLR is a linear spatial autoregressive method with the advantage of capturing spatial dependency in regression analysis, avoiding statistical problems such as unstable parameters and unreliable significance test, as well as providing information on spatial relationships among the independent variables [55]. In GeoDa, SLR sets up the first-order rook's move contiguity (that is, adjacent edges) and uses the diagnostics from GeoDa to determine the most appropriate weights matrix [53]. The spatial lag model can be specified by [56]:

$$Y = \alpha W_y + \beta X + \varphi \quad (1)$$

$$\varphi = (1 - W_y)^{-1}\varepsilon \quad (2)$$

where Y is the spatially lagged dependent variable (the number of negative comments normalised by the number of population at the jiedao level in our study); α is the spatial autoregressive structure of the disturbance φ; W_y represents the spatial weight matrix of the dependent variable Y; X is a matrix containing a set of independent variables (the average land price, the number of aged population and public facilities by type at the jiedao level in our study); β is the coefficient of X; and ε =the normally distributed error. The overall analytical workflow is presented in Figure 3.

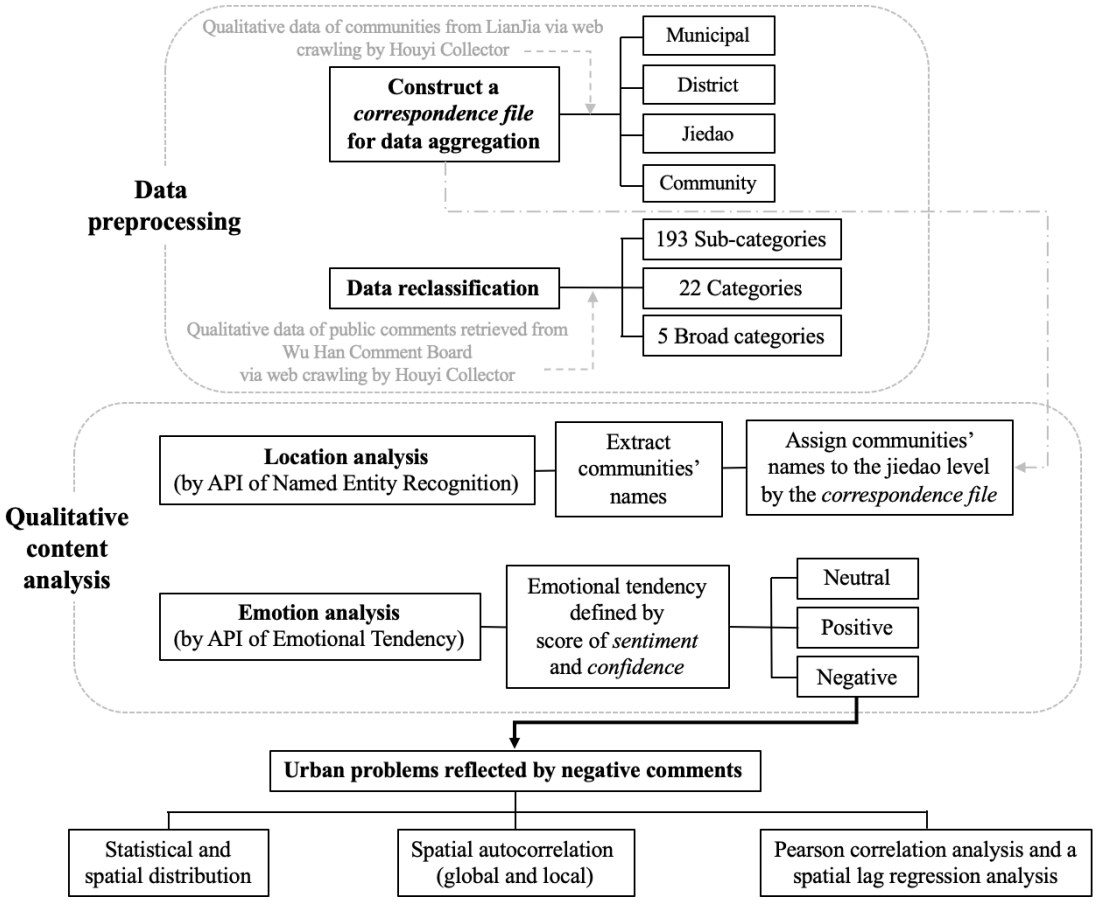

Figure 3. Analytical workflow

4. Results

4.1. Statistical and spatial distribution of urban problems

There are 124 jiedaos (97.6% of the jiedaos with public comments) observed with urban problems reflected by negative comments. It is interesting to notice that the top ten jiedaos with the highest number of urban problems appear in industry-oriented and newly-developed areas beyond the 3rd transport circle (Figure 4), spreading out to the remote region on the urban fringe. The top 10 jiedaos include several economic development zones (EDZs) (e.g. Pan Long Cheng, Yang Luo, Miao Shan) and industrial parks (e.g. Wuhan Iron and Steel Company (WISC) and Yaojiashan). Such areas have relatively large in- and out-flow of migrants and temporary residents, which thus are observed to have more problems in a pressing need for urban governance compared to the well-developed areas in and around the inner city. We further analyse the proportion of urban problems by category in Figure 5 with X axis showing 124 jiedaos ordered by a descending number of urban problems. In most of jiedaos, construction and housing accounts for the dominant proportion of urban problems, followed by infrastructure and configuration, and social service and administration. Safety and security issues tend to appear in a high proportion in jiedaos with a relatively smaller number of urban problems (towards the right of Figure 5).

We further map the spatial distribution of urban problems by category (Figure 6). Overall, the red-spots of urban problems in each category all tend to sporadically distributed outside of the 3rd transport circle, in the newly developed region and remote districts including Qingshan, Caidian, Jiangxia, Hanyang, Xingzhou Districts. When urban problems are broken down into each category, the most obvious prevalence of safety and security issues, and environmental pollution issues is observed in old industrial jiedaos in Qingshan District (e.g. Changqian, worker villages, WISC industrial park) and some of EDZs (e.g. Yangluo, Yaojiashan, and Panlongcheng). Furthermore, the spatial pattern of infrastructure and configuration is similar to that of social service and administration and red-spots of both categories are distributed outside of the 3rd transport circle mostly in newly-developed EDZs (e.g., Yangluo, Panlongcheng, Yaojiashan, and Miaoshan) as well as in the WISC industrial park. However, different to the aforementioned 4 categories of urban problem distributed more towards remote districts on the urban fringe, construction and housing issues spread out in the inner-city districts with a large number of historical jiedaos (e.g., Jiang'an, Jianghan, and Qingshan District), where urban regeneration and gentrification occurred.

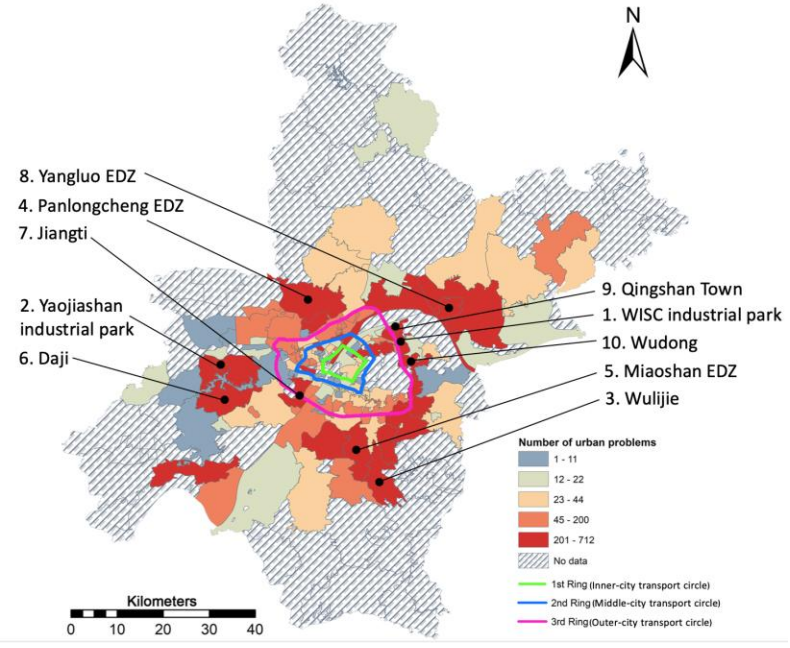

Figure 4. Spatial distribution of total urban problems and the top ten jiedaos with the high prevalence of urban problems

(Note: The number of urban problems is normalised by the number of population at the jiedao level)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 September 2020 doi:10.20944/preprints202009.0195.v1

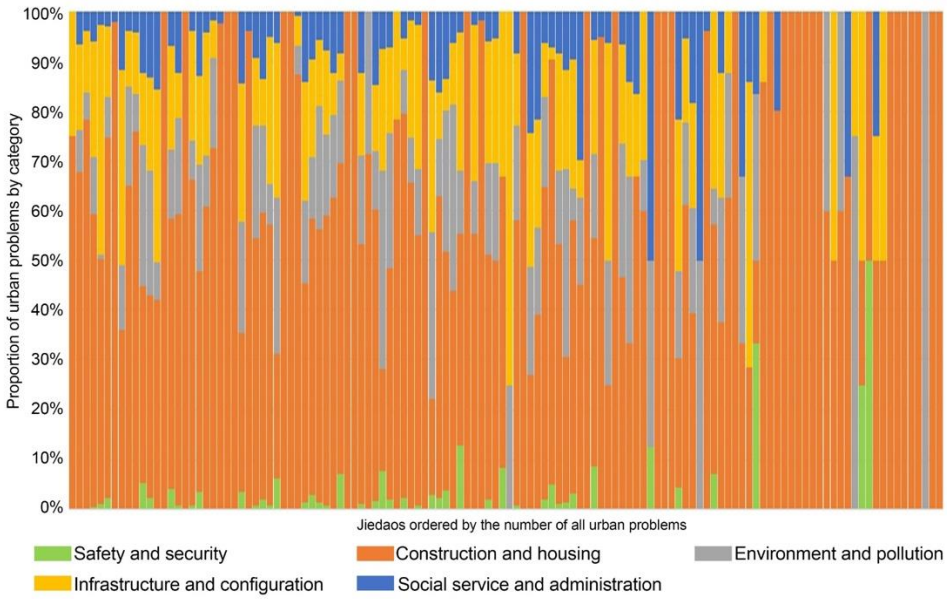

Figure 5. Proportion of urban problems by category at the jiedao level

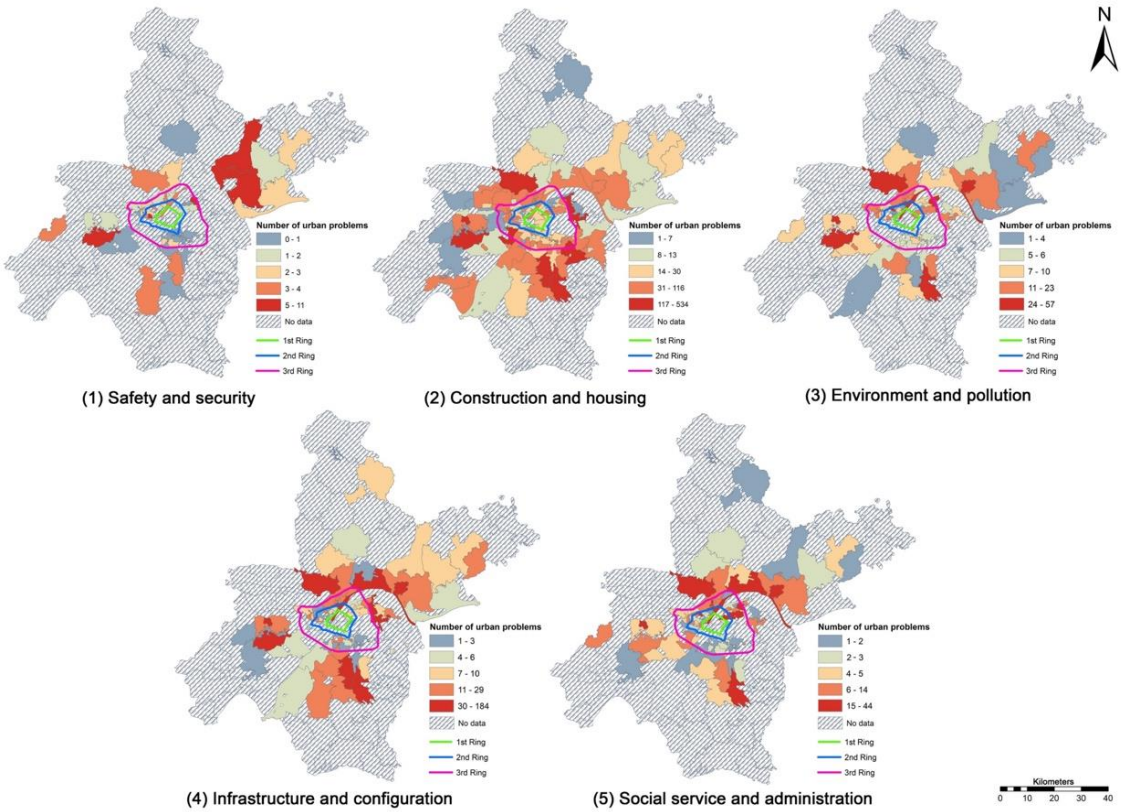

Figure 6. Spatial distribution of urban problem by category

(Note: The number of urban problems is normalised by the number of population at the jiedao level)

4.2. Spatial relationship of urban problems

The result of LISA shows the spatial relationship of urban problems with their surrounding areas, where our main interests focus on the clusters of urban problems (high-high red-spots in Figure 7) indicating the concentrations of adjacent jiedaos with high prevalence of urban problems. Globally, the magnitude of global Moran's I for all types of urban problems are small, indicating that the distribution of urban problems is dispersed in the whole study area. However, several local clusters are observed around three EDZs (e.g., Miaoshan, Yangluo, Yaojiashan), indicating a local

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 September 2020 doi:10.20944/preprints202009.0195.v1

concentration of urban problems across each EDZ and its surrounding areas. When urban problems are divided to each category, with the exception of safety and security which is most obviously clustered in Wangji in Xingzhou District, all other categories of urban problems are highly concentrated in several EDZs. For instance, Yaojiashan EDZ is a high concentration of construction and housing, environment and pollution, and social service and administration; Miaoshan EDZ is a high concentration of construction and housing, and infrastructure and configuration; WISC industrial park is a high cluster of construction and housing, environment and pollution, and infrastructure and configuration. In sum, the patterns of these local clusters indicate that urban problems are highly prevalent in old industrial areas and newly-developed EDZs, where urban problems may spill over and have chain effects to surrounding areas to form local clusters in most need for urban governance endorsed by governments or related sectors.

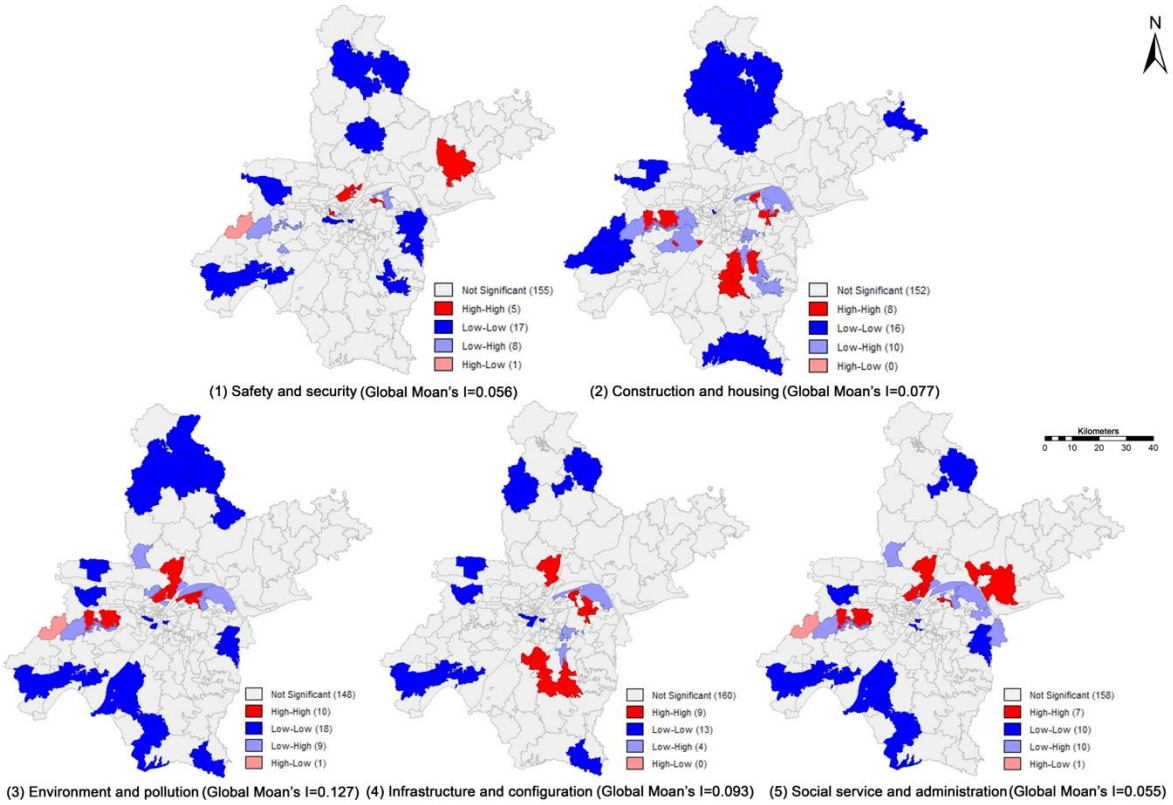

Figure 7. Local clusters of urban problems by category

4.3. Associative factors of urban problems to guide urban governance

We first conduct a pairwise correlation analysis among all measures in the later regression model to test data collinearity, including 8 independents variables (civil facilities contain 5 categories) and 6 dependent variables representing all urban problems and each category of urban problems (Table 1). The result shows that both land price and age above 65 are significantly ($p<0.01$) correlated to urban problems as a whole, Category 2 and Category 4. In addition, five types of civil facilities are also significantly ($p<0.01$) correlated to urban problems as a whole, Category 2 and Category 4. The extent that land price, age above 65, and civil facilities affect urban problems needs further exploration by the ensuring SLR model.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 September 2020 doi:10.20944/preprints202009.0195.v1

Table 1. Coefficients of the pairwise Pearson correlation between urban problems and selected associative factors

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Land price	1													
2	Age above 65	0.195	1												
3	Civil facilities	0.163	-0.159	1											
4	*Traffic*	-0.002	-0.143	**0.936**	1										
5	*Commercial*	0.173	-0.162	**0.999**	**0.924**	1									
6	*Public*	0.098	-0.047	**0.925**	**0.950**	**0.913**	1								
7	*Educational*	0.043	**-0.259**	**0.583**	**0.608**	**0.567**	**0.603**	1							
8	*Medical*	**0.307**	-0.022	**0.910**	**0.834**	**0.907**	**0.848**	**0.514**	1						
9	Urban problems	**-0.298**	**-0.323**	**0.372**	**0.478**	**0.361**	**0.416**	**0.258**	**0.279**	1					
10	*Category1*	-0.017	-0.024	-0.036	-0.046	-0.037	0.022	-0.026	0.006	**0.232**	1				
11	*Category2*	**-0.283**	**-0.320**	**0.392**	**0.509**	**0.380**	**0.439**	**0.280**	**0.302**	**0.965**	0.120	1			
12	*Category3*	-0.143	-0.171	-0.038	-0.030	-0.037	-0.026	-0.096	-0.040	**0.586**	**0.512**	**0.440**	1		
13	*Category4*	**-0.288**	**-0.255**	**0.411**	**0.497**	**0.402**	**0.443**	**0.298**	**0.300**	**0.852**	**0.239**	**0.717**	**0.488**	1	
14	*Category5*	-0.163	-0.218	-0.054	-0.030	-0.054	-0.035	-0.097	-0.067	**0.664**	**0.360**	**0.515**	**0.845**	**0.610**	1

Note: Bold: $p<0.01$

Category 1: Safety and security

Category 2: Construction and housing

Category 3: Environment and pollution

Category 4: Infrastructure and configuration

Category 5: Social service and administration

The SLR result shows that land price, age above 65 and civil facilities are significantly ($p<0.05$) associated with urban problems as a whole (the first column in Table 2). More specifically, jiedaos with higher land price tend to have less urban problems, indicating socioeconomically advanced communities in such jiedaos may have better urban governance and administration thus are observed with less problems. Furthermore, Jiedaos with a higher proportion of aged population tend to have less urban problems, possibly due to that aged population have less access to Internet and thus less chances to lodge complaints and report problems. However, it is surprising to observe that jiedaos with more provisions of civil facilities tend to have more urban problems possible due to some facilities that bring nuisance, noise and disturbance to local communities. For example, traffic facilities (e.g., train and bus stations, and transit hubs) are observed to be positively ($\beta = 0.687$, $p<0.05$) relevant to the prevalence of urban problems.

When urban problems are broken down into each category, both land price and age above 65 are negatively (significant at $p<0.05$) associated with each category of urban problems except safety and security (Category 1), while civil facilities are only positively (significant at $p<0.01$) associated with Category 2 and 4. In other words, urban problems related to construction and housing, and infrastructure and configuration. Meanwhile, it is noteworthy that different types of facilities are observed to associate with different categories of urban problems. For example, less commercial facilities (e.g., supermarkets and shopping centers) are associated with more problems in construction and housing, and safety and security; less educational facilities (e.g., schools) are associated with more problems in social service and administration, while more traffic facilities are consistently associated with more problems in most categories, in particular, in infrastructure and configuration ($\beta = 1.061$), and construction and housing ($\beta = 0.827$). These observations indicate that different types of civil facilities have various impacts on urban problems; jiedaos with deficits in the provision of or lacking of civil facilities may provoke the need of urban governance in infrastructure and configuration, and social service and administration, whereas some civil facilities especially

traffic facilities that increase human mobility and bring in annoyance, nuisance, and disruptions to local residents would impose adverse effects on local neighbourhoods.

Table 2. SLR coefficients of urban problems by category

	All	Category1	Category2	Category3	Category4	Category5
Land price	-0.196**	-0.137	-0.154*	-0.36***	-0.271**	-0.273**
Age above 65	-0.287***	-0.189	-0.29***	-0.437***	-0.187*	-0.341***
Civil facilities	0.393***	0.015	0.407***	0.100	0.461***	0.114
Traffic	0.687**	-0.033	0.827***	0.411***	1.061**	0.471***
Commercial	-0.508*	-0.461**	-0.649**	-0.453***	-0.281	-0.225
Public	0.234	0.721***	0.218	0.515***	-0.238	0.257**
Educational	-0.122	-0.237	-0.146	-0.311***	0.007	-0.279**
Medical	0.085	0.145	0.135	0.129	-0.086	-0.011
R-squared	0.386	0.292	0.421	0.395	0.439	0.342
Number	124	44	99	73	76	76

Note: *:p<0.1;**:p<0.05; ***:p<0.01
Category 1: Safety and security
Category 2: Construction and housing
Category 3: Environment and pollution
Category 4: Infrastructure and configuration
Category 5: Social service and administration

5. Discussion and conclusion

Our study proposes an analytical framework to retrieve citizen-centred data from an online comment board and through a qualitative content analysis of public comments to reflect urban problems, on the basis of which to examine the spatial distribution of urban problems, and their relation with surrounding areas and associative factors in order to ultimately imply how public participation helps with urban governance. The methodological workflow in our study integrates data web-crawling and the usage of APIs for location and sentiment analysis, which can be applied to retrieve non-English qualitative big data in other contexts. By analysing urban problems reflected by negative public comments, our findings reveal that urban problems most frequently appear in industry-oriented and newly-developed areas beyond the 3rd transport circle, spreading out to the remote region on the urban fringe. In most of jiedaos, construction and housing is the dominant type of urban problems, most obviously observed in the inner-city districts with a large number of historical jiedaos where urban regeneration and gentrification occurred. The local clusters of urban problems by category, in exception of safety and security, are obviously observed in old industrial areas and newly-developed EDZs, and such problems spill over to adjacent areas to form a larger clustering spot in the most need for urban governance given by governments or related sectors. Moreover, jiedaos with higher land prices and a higher proportion of aged population tend to have less urban problems, while various types of civil facilities affect the prevalence of urban problems differently.

The spatial distribution of urban problems and the explanation of their formations lie in the institutional and political-economic mechanisms, characterized in the Chinese context. Prior to the China's reform and opening up in 1978, Chinese cities had a slow progression of socioeconomic development with weak capability of urban governance. With the acceleration of neo-liberalism, and the implementation of land system and housing reform since 1978, the reconstruction of traditional work-unit housing (*danweifang* in Chinese) and the privatisation of housing have stimulated large investment in real estate and infrastructure construction [45]. Consequently, urban governance has been transformed from under the full supervision of governments to under the dispersed supervision of multiple parties (e.g., real estate developers and private property management companies)

together with the control of governments [42]. Private governance endorsed by real estate developers and property management companies plays an increasingly important role in providing service and administration in post-reform commercially built commodity housing (*shangpingfang* in Chinese), mostly gated communities which are privately managed and designed according to the tastes of the growing urban middle class [45]. Such gated communities with high land values and housing prices are usually well located around city centers with easy access to civil facilities and cityscape, or newly developed on the urban fringe with high-class neighbourhood environment, good service provision and administration, and well-designed public space (e.g., clear streets, sufficient green space and parking lots, and visually-present landscape). Residents living in high-class communities pay high body corporate and administrative fee for good life quality and well-organised community service provided by real estate developers or property management companies who take over the managerial role from local governments and provide governance at the community level. Conversely, communities with low land value and housing prices are usually old communities built in the early stage of China's land system reform or traditional work-unit housing mostly located in and around the inner city. Residents living in these low-valued neighbourhoods have limited access to service provided by real estate developers or property management companies, concurrently lacking of sufficient governance and public service given by local governments, and thus associating with a high prevalence of urban problems.

The dilemma of government-centred and private governance is accompanied by the deepening process of industrial transformation and upgrading, and such dilemmatic phenomena have been widely observed in old industrial areas in Chinese cities [57]. The particular example is WISC industrial park with the highest number of urban problem in construction and housing, followed by environmental pollution, infrastructure and configuration. As a large state-owned company, WISC provided both working and living space as well as public service and administration to employees in the old era before the economic resolution of free market in China in 1978 [58]. After the economic resolution and industrial transformation since 1978, WISC changed its company structure, no longer playing a managerial role in providing social service and administration in the jiedao, where providers of urban governance were replaced by government and private sectors [59]. If such replacements were not initiated efficiently, deficits of governance would appear, leading to increasing urban problems and complaints from local residents. This governance-deficit phenomenon is also observed in another 3 EDZs as the red-spots of urban problem including Miaoshan, Yangluo, Yaojiashan EDZ. These zones have been transformed from old industrial districts to more mixed land-use areas for residence and recreation in the past two decades of industrial transformation, but lacking of the efficient provision of public service and amenities.

The provision of civil facilities needs to be better assessed by type, catering to the need of local residents, and different types of civil facilities and their supply-demand balance/imbalance would trigger urban problems differently. This finding surprisingly runs counter to the empirical observation in existing literature that the better provision of public facilities results in less urban problems [60]. The possible explanation is that the diversity of public facilities brings convenience but simultaneously cause some issues to local residents. In reality, citizens seem to have conflicting attitudes on the large-scale civil facilities. In particular, commercial and educational facilities are expected to build close to communities by residents due to the convenience and easy access while lacking of such facilities, as seen in our findings, would increase the voice for better urban governance in infrastructure and configuration, and social service and administration. Whereas, the large-scale commercial and entertainment facilities as well as traffic facilities also bring in annoyance, nuisance, and disruptions to local residents and impose adverse effects on local neighbourhoods due to the increasing human mobility, noise, parking shortage and traffic congestion. Moreover, power stations, garbage dump grounds, waste transit sites and stations, and other aging facilities have been also mentioned in public comments as facilities 'kept away from my backyard', which have adverse impacts on the health, visibility, and environment of surrounding areas and indirectly decrease property prices.

Our findings based on the *actual* voice from citizens reflect their real problems and needs, and thus provide evidences for political implications that can be considered to improve the efficiency of future public participation and urban governance. First, state and municipal governments are advised to establish policies, such as national planning codes, to standardise the provision of public goods and service at the community level and to further guide through private governance. Second, real estate developers and property management companies are encouraged to participate into the process of policy making and to nominate state agencies to regulate private provision of community service and administration, in terms of detailed responsibilities and obligations as well as the spatial boundaries of administration. The main purpose of introducing private governance should be responding to residential demands, rather than stimulating the housing market and generating local revenues from residential land leasing. Third, online comment boards and forums are effective platforms where citizens can express their demand before decision making, and supervise and give feedbacks after decision making. The development of such online tools can be extended to interactive maps and cloud sharing empowered by advanced technologies such as machine learning and artificial intelligence, in order to enable and broaden public participation throughout the whole process of decision making. Finally, public participation needs be promoted through multiple approaches, from online platforms to offline surveys to involve aged population who may have limited access to Internet and be somehow neglected in current theme of public participation [11, 61]. Accounting for the demand of aged groups and aging communities needs to be prioritised in urban governance for a better social justice and equity.

There are a number of limitations in this study, where future research could expand. First, the citizen-centred data retrieved from the online comment board lacks of demographic and socioeconomic information of comment providers, which can be part of associative factors to explain urban problems. Second, the collection of citizen-centred data as a stratified sampling approach inevitably brings in some bias because the data is limited to comment providers within a certain age group. Such bias can be calibrated by involving multi-sourced data such as census data and survey data in future work. Third, the methodological approach proposed in our study integrates web-scrawling tools and the usage of APIs, which can be applied to non-Chinese qualitative data retrieving and analysis. The accuracy and application of this approach can be improved by advanced technologies such as deep learning and AI optimisation. Therefore, our study, as an initial attempt of retrieving non-English qualitative data to analyse urban problems, explores a promising research direction of integrating citizen-centred big data into urban governance and public participation, and calls for future research in different linguistic and geographic contexts.

Author Contributions: Conceptualization, Zhe Gao and Jiang Gu; Data curation, Zhe Gao; Formal analysis, Siqin Wang; Funding acquisition, Zhe Gao; Investigation, Zhe Gao; Methodology, Zhe Gao and Siqin Wang; Project administration, Zhe Gao; Resources, Zhe Gao; Software, Zhe Gao and Siqin Wang; Supervision, Zhe Gao; Validation, Zhe Gao, Siqin Wang and Jiang Gu; Visualization, Siqin Wang; Writing – original draft, Zhe Gao and Siqin Wang; Writing – review & editing, Jiang Gu. All authors have read and agreed to the published version of the manuscript.

Acknowledgement: We would thank the efforts from Yuxiao Li, Wenjun Jiang and Tingting Qiang, the undergraduates from the College of Urban & Environmental Sciences, Central China Normal University, for data collection and preparation.

Funding: This research was funded by Hubei Provincial Natural Science Foundation of China (2020CFB350) and Fundamental Research Funds for the Central Universities (CCNU20QN031).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jin, J., Gubbi, J., Marusic, S., & Palaniswami, M. An information framework for creating a smart city through internet of things. *IEEE Internet of Things journal* 2014, *1*(2), 112-121.

2. Perera, C., Zaslavsky, A., Christen, P., & Georgakopoulos, D. Sensing as a service model for smart cities supported by internet of things. *Transactions on emerging telecommunications technologies* 2014, *25*(1), 81-93.
3. Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. Internet of things for smart cities. *IEEE Internet of Things journal* 2014, *1*(1), 22-32.
4. Angelidou, M. Smart city policies: A spatial approach. *Cities* 2014, *41*, S3-S11.
5. Lombardi, P., Giordano, S., Farouh, H., & Yousef, W. Modelling the smart city performance. *Innovation: The Europe* 2012, *25*(2), 137-149
6. Poorthuis, A., Zook, M., Taylor Shelton, M. G., & Stephens, M. Using Geotagged Digital Social Datain Geographic Research. In *Key methods in geography*, 3rd nd; Nicholas Clifford, Meghan Cope, Thomas Gillespie, Shaun French; SAGE: London, 2016, *16*, 248-269.
7. Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Affairs* 2014, *33*(7), 1123-1131.
8. Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. Big data and its technical challenges. *Communications of the ACM* 2014, *57*(7), 86-94.
9. Song, M. L., Fisher, R., Wang, J. L., & Cui, L. B. Environmental performance evaluation with big data: Theories and methods. *Annals of Operations Research* 2018, *270*(1-2), 459-472.
10. Li, S. Research on Public Participation and Government Response Based on Big Data Mining on the Internet Political Platform. Huazhong University of Science and Technology, 2019. [In Chinese]
11. Zhang, Y. Grade Structure of Public Planning Values: Based on Content Analysis of Public Messages in the Field of Urban Planning. *City Planning Review* 2019, *43*(08), 108-116. [In Chinese]
12. Pløger, J. Public participation and the art of governance. *Environment and Planning B: Planning and Design* 2001, *28*(2), 219–241.
13. Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., ... & Portugali, Y. Smart cities of the future. *The European Physical Journal Special Topics* 2012, *214*(1), 481-518.
14. Albino, V., Berardi, U., & Dangelico, R. M. Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of urban technology* 2015, *22*(1), 3-21.
15. Angelidou, M., Psaltoglou, A., Komninos, N., Kakderi, C., Tsarchopoulos, P., & Panori, A. Enhancing sustainable urban development through smart city applications. *Journal of Science and Technology Policy Management* 2018, *9*(2), 146-169.
16. Baccarne, B., Mechant, P., & Schuurman, D. Empowered cities? An analysis of the structure and generated value of the smart city Ghent. In *Smart city*; Springer, Cham, 2014 (pp. 157-182).
17. Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., ... & Scholl, H. J. Understanding smart cities: An integrative framework. In *2012 45th Hawaii international conference on system sciences*, IEEE, USA, 2012 (pp. 2289-2297).
18. Ansell, C., & Gash, A. Collaborative governance in theory and practice. *Journal of public administration research and theory* 2008, *18*(4), 543-571.

19. Misuraca, G., Alfano, G., & Viscusi, G. Interoperability challenges for ICT-enabled governance: towards a pan-European conceptual framework. *Journal of theoretical and applied electronic commerce research* 2011, *6*(1), 95-111.
20. Ju, J., Liu, L., & Feng, Y. Citizen-centered big data analysis-driven governance intelligence framework for smart cities. *Telecommunications Policy* 2018, *42*(10), 881-896.
21. Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Heredero, C. Information reuse in smart cities ecosystems. *El profesional de la información* 2015, *24*(6), 838-844.
22. Al Nuaimi, E., Al Neyadi, H., Mohamed, N., & Al-Jaroodi, J. Applications of big data to smart cities. *Journal of Internet Services and Applications* 2015, *6*(1), 25.
23. Kitchin, R. Big Data, new epistemologies and paradigm shifts. *Big data & society* 2014, *1*(1), 2053951714528481.
24. Maloney-Krichmar, D., & Preece, J. A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Transactions on Computer-Human Interaction (TOCHI)* 2005, *12*(2), 201-232.
25. Mayring, P. Qualitative content analysis. *A companion to qualitative research* 2004, *1*(2004), 159-176.
26. Elo, S., & Kyngäs, H. The qualitative content analysis process. *Journal of advanced nursing* 2008, *62*(1), 107-115.
27. Crowley, B. P., & Delfico, J. F. Content analysis: A methodology for structuring and analyzing written material. *United States General Accounting Office (GAO), Program Evaluation and Methodology Division* 1996.
28. Schreier, M. *Qualitative content analysis in practice*. Sage publications, Londaon, 2012.
29. Liu, Y., Liu, D., & Chen, Y. Research on Sentiment Tendency and Evolution of Public Opinions in Social Networks of Smart City. *Complexity* 2020.
30. Wang, X., Wei, F., Liu, X., Zhou, M., & Zhang, M. Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach. In *Proceedings of the 20th ACM international conference on Information and knowledge management* 2011, (pp. 1031-1040).
31. Nodarakis, N., Sioutas, S., Tsakalidis, A. K., & Tzimas, G. Large Scale Sentiment Analysis on Twitter with Spark. In *EDBT/ICDT Workshops* 2016 (pp. 1-8).
32. Kanavos, A., Nodarakis, N., Sioutas, S., Tsakalidis, A., Tsolis, D., & Tzimas, G. Large scale implementations for twitter sentiment classification. *Algorithms* 2017, *10*(1), 33.
33. Enserink, B., & Koppenjan, J. Public participation in China: sustainable urbanization and governance. *Management of Environmental Quality: An International Journal* 2007, *18*(4), 459-474.
34. Wu, F. Housing privatization and the return of the state: changing governance in China. *Urban geography* 2018, *39*(8), 1177-1194.
35. Cheshire, C., Antin, J., Cook, K. S., & Churchill, E. General and familiar trust in websites. *Knowledge, Technology & Policy* 2010, *23*(3-4), 311-331.
36. Chen, S. C., & Webster, C. J. Homeowners associations, collective action and the costs of private governance. *Housing studies* 2005, *20*(2), 205-220.
37. Csefalvay, Z. Gated communities for security or prestige? a public choice approach and the case of Budapest. *International journal of urban and regional research* 2011, *35*(4), 735-752.

38. Lu, T., Zhang, F., & Wu, F. Place attachment in gated neighbourhoods in China: Evidence from Wenzhou. *Geoforum* 2018, *92*, 144-151.
39. Shen, J., & Wu, F. The suburb as a space of capital accumulation: The development of new towns in Shanghai, China. *Antipode* 2017, *49*(3), 761-780.
40. Wu, F. Rediscovering the 'gate' under market transition: from work-unit compounds to commodity housing enclaves. *Housing studies* 2005, *20*(2), 235-254.
41. McGuirk, P., Dowling, R., Gibson, C. R., Iveson, K., & Mee, K. Urban vitality, culture and the public realm. University of Wollongong, 2007
42. Lu, T., Zhang, F., & Wu, F. Place attachment in gated neighbourhoods in China: Evidence from Wenzhou. *Geoforum* 2018, *92*, 144-151.
43. Atkinson, R., & Blandy, S. Introduction: International perspectives on the new enclavism and the rise of gated communities. *Housing studies* 2005, *20*(2), 177-186.
44. Goodman, R., & Douglas, K. Life in a master planned estate—community and lifestyle or conflict and liability? *Urban Policy and Research* 2010, *28*(4), 451-469.
45. Breitung, W. Enclave urbanism in China: Attitudes towards gated communities in Guangzhou. *Urban geography* 2012, *33*(2), 278-294.
46. Webster, C. Gated cities of tomorrow. *The town planning review* 2001, *72*(2), 149-170.
47. He, S. Homeowner associations and neighborhood governance in Guangzhou, China. *Eurasian Geography and Economics* 2015, *56*(3), 260-284.
48. Wuhan Bureau of Statistics. Wuhan Statistical Yearbook. Available online: http://www.wuhandata.gov.cn/whData/directory.html?typeId=2&id=937568b8-1a69-4194-a823-edc8cc0f4a2d (accessed on 7 September 2020)
49. Wuhan Bureau of Statistics. Available online: http://www.stats.gov.cn/tjsj/zxfb/201104/t20110428_12705.html (accessed on 7 September 2020)
50. Wuhan Natural Resources and Planning Bureau, 2018. Wuhan land grade and benchmark land price standard. Available online: http://zrzyhgh.wuhan.gov.cn/bsfw_18/tdsc/jzdj/202004/t20200420_1032923.shtml
51. Amap, 2018. Map POI (Point of Interest) data. Available at: https://opendata.pku.edu.cn/dataset.xhtml?persistentId=doi:10.18170/DVN/WSXCNM (accessed on 7 September 2020)
52. Anselin, L. Local indicators of spatial association—LISA. *Geographical analysis* 1995, *27*(2), 93-115.
53. Anselin, L., Syabri, I., & Kho, Y. GeoDa: an introduction to spatial data analysis. *Geographical Analysis* 2006, *38*, 5–22.
54. Benesty, J., Chen, J., Huang, Y., & Cohen, I. Pearson correlation coefficient. In *Noise reduction in speech processing*, Springer, Berlin, Heidelberg, 2009; pp. 1-4.
55. Ward, M. D., & Gleditsch, K. S. *Spatial regression models*, Sage Publications, London, 2018
56. Anselin, L. *Spatial econometrics: methods and models*, Springer Science & Business Media, 2013
57. Zhang, X., Hu, J., Skitmore, M., & Leung, B. Y. Inner-city urban redevelopment in China metropolises and the emergence of gentrification: Case of Yuexiu, Guangzhou. *Journal of Urban Planning and Development* 2014, *140*(4), 05014004.

 doi:10.20944/preprints202009.0195.v1

58. Guthrie, D. *China and globalization: The social, economic, and political transformation of Chinese society*. Routledge. 2012
59. Gao Z., Yin N., Tong X., Li D., Gu J. Shrinkage under Urban Growth: A Case Study of Wuhan City. *Tropical Geography* 2019, *39*(1), 31-38 [In Chinese]
60. Chang, H. S., & Liao, C. H. Exploring an integrated method for measuring the relative spatial equity in public facilities in the context of urban parks. *Cities* 2011, *28*(5), 361-371.
61. Yang, X. and Long, Y. Opportunities and Limitation of Big Data Application to human and economic geography: the state of the art. *Progress in Geography* 2015, *34*(4), 410-417. [In Chinese]