Intelligent Environments 2020
C.A. Iglesias et al. (Eds.)

doi:10.3233/AISE200041

A Soft Context-Aware Traffic Management System for Smart Cities

Davide CARNEIRO [a,b,1], António AMARAL [a] and Mariana CARVALHO [a]
[a] *CIICESI, Escola Superior de Tecnologia e Gestão, Instituto Politécnico do Porto, Felgueiras, Portugal*
[b] *Algoritmi Centre/Department of Informatics, Universidade do Minho, Portugal*

Abstract. The number of large cities is growing, as well as their inhabitants. As population density increases, so do the challenges of managing these spaces. Current cities, both large and small, have to deal with decreasing air quality, noise and air pollution, traffic congestion, and countless other issues that decrease citizens' quality of life. In this paper we propose a soft system for traffic management for smart cities. Context awareness is provided by a connection to city sensors. Traffic management is done by intelligent routing of vehicles, prioritizing not the individual's travel distance or speed, but the overall routes. Thus, routes are defined according to the city's management policies, and may allow to minimize traffic or pollution in certain areas, dynamically. That is, without explicit Human intervention. Such a system may be useful in improving the quality of life in the city for all citizens.

Keywords. Smart Cities, Traffic Management, Optimization, Graph, Shortest Path

1. Introduction

The current population growth phenomenon is being, somehow, responsible for the current way of life within urban areas. Therefore, cities are being challenged towards, properly, handling the negative externalities felt on the environment, as well as on the citizens' lifestyles and on the governance options [13]. Along the 20th century, the population living in cities increased from 220 million to nearly 2.8 billion, and according with the available forecasts in 2050, that figure will raise to about 6.9 billion, which will be close to almost 70% of the world's population [15].

Motivated partly by this social evolution in recent years, the concept of smart cities (SC) emerged. SC have been attracting an unprecedented amount of attention among different stakeholders, especially those within academia, industry and public policy making. This is mostly due to its major implications in urban planning and design, sustainability, social digitalization and cities' smart governance practices [18,1].

For that reason, it can be said that the smart city concept primary emerged as a solution to solve the problems associated with the exponential growth of urbanization [13]. Smart cities are expected to efficiently manage the growing level of urbanization,

[1]CIICESI, Escola Superior de Tecnologia e Gestão, Instituto Politécnico do Porto, Felgueiras, Portugal; E-mail: dcarneiro@estg.ipp.pt

and the amount of energy consumed, improving the citizens' economic standard, as well as raising people's awareness and capabilities to efficiently use and embed information and communication technologies (ICT) [16].

Smart cities will employ information and communication technologies towards being able to contribute for the improvement of citizens' quality of life, enhancing the overall wealth creation. They also favor the adjustment to the proper mobility solutions which could ease up the traffic management, reducing its environmental impacts, and enhancing the amount of interaction with the official authorities in charge [7,16].

Notwithstanding, the amount of complex problems that need to be handled and tackled within the smart cities' ecosystem are favoring the adoption of new technological approaches in which cities could use Big Data in transportation to guide the creation of sustainable and safer traffic systems [10]. These approaches, based in the emergence of IoT devices and the appearance of new data sources will make it possible to examine and predict traffic conditions with an utmost accuracy. This will definitely help to optimize the design of transport services management in a future automated city [9], which will play an important role in the adoption of modern Intelligent Transportation Systems (ITS), as well as in solving problems such as congestion control and peak load reduction [10].

It will also facilitate the construction of an eco-friendly environment by optimizing the vehicles' route planning [12,7], and thus the smart parking approaches that could alleviate the deadlocks in parking problems [16], as well as to create cities with low to zero carbon emissions, through the adoption and development of sustainability practices [15] and by ensuring the alignment of smart cities with the UN sustainable development goals [7].

The best practices, within the context of a smart city, usually highlight an anthropocentric approach for the development of a truly smart and sustainable city [17]. Therefore, the smart city's roadmap path should include and develop the adoption of green approaches which could support the urban development, establishing the proper conditions to ensure the attainment of a sustainable future [6].

In this paper we propose a system developed in accordance to these principles. It considers data from different sources, including sensor streaming data (e.g. environmental sensors, parking availability) and batch data (e.g. map information) and stores it in a way that facilitates the creation of services focused on facilitating traffic management. The main goal is not only for city managers to have access to live data, but also to take decisions that effectively influence traffic management in the city, in order to optimize a specific cost function.

This paper describes the proposed architecture, an implementation of a first prototype, and an example of using this architecture for developing a specific citizen-centric specific service. With the proposed service, a driver can input the desired destination, and the best parking lot will be suggested, as well as the best route there. In this case, the term *best* includes the optimization of the route according to the city's policies (i.e. not necessarily the shortest route) and the prediction of the number of free places in the parking lots, in order to reduce the negative effect of drivers searching for a parking place.

2. Problem Statement

As addressed in Section 1, cities are growing in population, and this rises significant and numerous challenges. One of them is clearly traffic management, which impacts negatively the efficiency of the transportation systems, air and life quality, among others. This paper focuses on the problem of traffic management.

Currently, traffic management in cities is done through legislation, and implemented using traffic signs, traffic lights, and similar elements. This traditional way of managing traffic is rather static in the sense that changes are slow and very rare. For instance, if the city intends to decrease traffic in a given central area, it may place traffic signs that forbid or limit the access of certain types of vehicles at given times of the day. On the one hand, this measure will take some time to implement and have associated costs. On the other hand, the measure is not context-aware in the sense that shall pollution or noise limitations be exceeded outside the defined time period, vehicles will still enter that area.

The main goal of this paper is to propose a soft traffic management system, that can complement these traditional measures. It considers a set of services for optimizing traffic at a city level. For instance, instead of each driver optimizing her/his own route through the city, a common route-finding service is used that takes into consideration not only the preferences of the driver (e.g. shortest path, quickest path) but also the preferences and policies of citizens and city managers (e.g. decreasing pollution, noise or traffic in a given area). This might result in some individual situations that are not optimized, but will result in an overall improvement towards the city.

We deem this approach *soft* traffic management in the sense that it does not really enforce decisions. Say a driver uses this route-finding algorithm to go from point A to point B. If the driver is not satisfied with it she/he can always resort to her/his own GPS application and follow a preferred route. However, as citizens gain an increased conscience about the role and impact of each individual on the society, and the importance of Social Capital [4], this kind of systems will gain increased relevance. At the same time, cities can also give back to citizens who decide to contribute to the common cause, namely by providing city services at discount prices, or through other motivators.

Finally, in an era in which driving is becoming autonomous, this kind of systems can be more easily implemented, namely through legislation that could force autonomous cars to abide by the rules and policies of each city they are driving in.

Thus, we believe that the proposed system may be of interest for implementing a wide range of citizen-centric services, that will positively impact the quality of life in cities.

3. Architecture

The proposed architecture for addressing the problem described in Section 2 is a layered one (Figure 1). The lowest layer is the *Data Ingestion* one and it includes services for connecting to two types of data sources: streaming and batch.

Streaming data sources include sensors such as environmental sensors to measure air pollution, temperature and humidity sensors, among others. Batch data sources include historic or static data sources, such as city maps, datasets of parks occupation, among others.

Once a connection to a data source is established, a *Data Transformation* stage may ensue. This may include some common data pre-processing tasks (e.g. remove missing data, filters, imputing values) or some more complex and specific tasks such as aggregations. Data transformation pipelines are thus defined *per* data source and are executed whenever new data is obtained from the respective data source.

Next there is the *Storage Layer*. We consider two types of data stores: a relational database and a graph database. The relational database are used to store objects such as historic sensor data or system meta-data (e.g. city management preferences). A MySQL database is used for this purpose.

The graph database, on the other hand, is used for storing connected data such as map information. Graph databases have been a great alternative to relational databases, especially when dealing with complex interconnected data sets. Graph databases offer high scalability, performance and flexibility advantage [11]. In the case of this work the Neo4j database is used. This is a graph database composed of nodes and relationships. Each node represents an entity and each relationship represents how two nodes are related. Neo4j also provides a web-interface and the cypher query language to conduct database operations against the data model.

Once data is stored in either of these databases, it becomes available to be used by the services that are developed and run in the *Service Layer*. A Microservice approach is followed for the development of these services in which functionalities will be implemented separately, by concern. Services can be used in two ways: by composing them to create higher-level services and through requests received from the API.

The *API layer* is the topmost layer and provides access of external applications to the services developed. An example of an application developed using existing services is described in Section 5. The web User Interface (UI), used by the city managers, is also powered by this API. The UI allows for city managers to visualize the state of the city in real time as well as to set or change city management policies (e.g. optimization functions).

4. Intelligent Path Finding Service

This section describes the implementation and functionality of one of the core services: the Path Finding Service. This service, as opposed to traditional driver routing services, does not aim to optimize the path for a specific driver. Instead, it will provide a path that takes into consideration driver preferences (e.g. shortest path, quickest path) but in the frame of city state and management policies. So, a driver may not be directed through the shortest route if this route goes through a pollution hotspot at that time. This section describes the whole process from data acquisition to its use by the driver's navigation system.

The base data that supports this service was acquired from OpenStreetMaps (OSM). OSM is an open source project that provides free geographic data. In OSM, data is represented as nodes, ways or relations. A node represents a point on the earth's surface, and it contains information about the unique identifier of the point (id number), a pair of coordinates (latitude and longitude) and tags with additional information (set of key/value pairs). A way represents linear features and boundaries and it consists of an ordered list of nodes. A relation represents additional information of both elements, nodes and ways, to explain how elements work together.

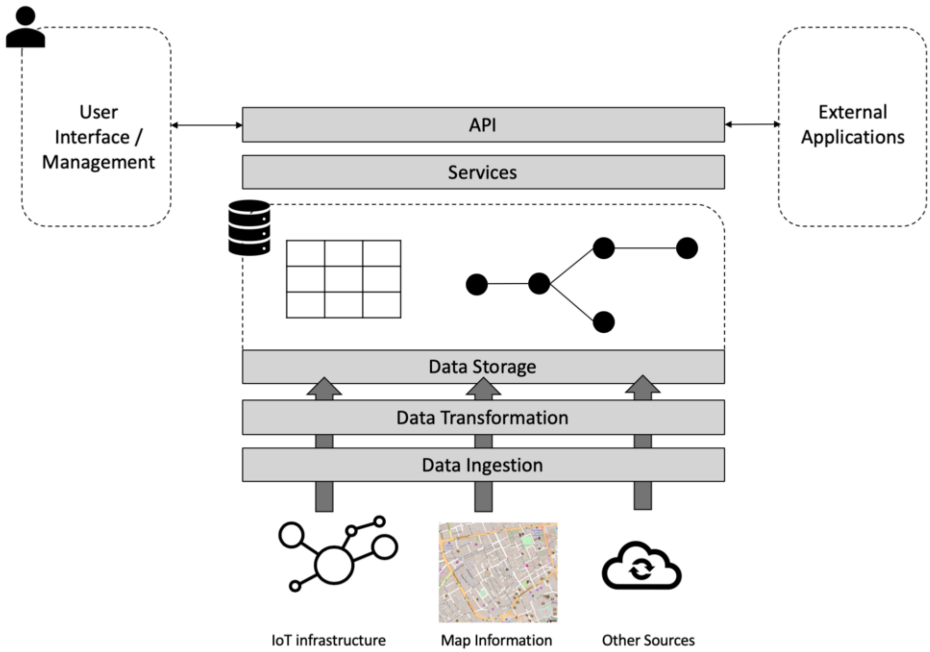

Figure 1. Overview of the main elements of the proposed architecture.

Data was extracted from OSM using the Overpass API, which is an OSM provided tool that allows to get the OSM data through an XML query, and it was stored on the graph-database.

A region which includes the whole city of Porto, in Northern Portugal, was selected to extract data. For this area all the nodes and waypoints were acquired.

Next, in the Data Transformation stage, data that was not relevant was excluded, including nodes or waypoints that referred to pedestrian or other types of paths (e.g. steps, cycle ways). New variables were also added to the data. These include the distance between each to connected nodes, calculated through the Haversine Formula [14]. After calculated, distance is normalized (min-max normalization). Three other variables were also added, which quantify three important measures: pollution level, noise level and congestion level. When the data is added these variables are set to 0. The value 0 means the absence of the phenomenon (e.g. low or no congestion) and the value 1 means a very high or critical level. Once the data is transformed it is stored into the graph database.

The Intelligent Path Finding Service is thus implemented as follows. The main goal is, as mentioned previously, to find the *best* path between two points. The notion of *best* is, however, not the traditional one (e.g. shortest, quickest). Here, the quality of a path is given by an optimization function that will take into consideration the characteristics of the path and the city's traffic management preferences.

Traffic management preferences are defined by city managers through the UI. Managers do so by creating a weights vector W that includes four values: w_d, w_n, w_c and w_p ($w_d + w_n + w_c + w_p = 1$). These variables represent, respectively, the weight of distance, noise pollution, congestion and air pollution. So, if the city manager wants to prioritize the minimization of noise, she/he may attribute weights of $0.1, 0.7, 0.1, 0.1$. On the other hand, a weights vector of $1,0,0,0$ would hold results similar to the drivers' individual navigation services.

In order to find the best path, this service uses a modified version of the Dijkstra algorithm [2]. Specifically, the concept of *distance* is now given by a weighed sum of the properties of the way (geographical distance, pollution, noise and congestion). Thus, paths will tend to reflect the city managers' preferences.

A change in the preferences of the managers and/or on the state of the city (e.g. an increase in the level of pollution or traffic in a given area) will thus lead to potential different routes. More importantly, these routes do not optimize the path of each individual driver in the city but optimize all their paths, according to the city's preferences.

5. Validation

This section describes the validation that was conducted of the architecture and of the Intelligent Path Finding Service. Validation was carried out in two ways: through the development of a prototype of the User Interface; and through the development of an external application that uses the service. Given that, at the time of the writing of this paper, we still had no access to real sensor data, this section also describes how data was generated to validate the proposed approach.

5.1. User Interface

Given the absence of sensor data, we opted to simulate these data by allowing city managers to define their own areas of pollution, congestion or noise. On the one hand, they can do so in order to simulate actual pollution/noise/congestion hotspots in the city. On the other hand, this simulation tool can also be used to analyze different scenarios and how traffic would be routed.

To insert a hotspot, a city manager has two options. Under the first one, he selects the desired type of hotspot (i.e. pollution, noise or congestion), defines its severity (a number between 0 and 1), clicks a point in the map, and then drags to create a circle with a given radius. The effect of this is that all nodes and connections inside the circle will be updated with a new value for the selected phenomenon. The intensity of the phenomenon decreases from the center to the edge of the circle, proportionally to the distance to the center.

Under the second option, the user may select a group of nodes (such as a specific road) and create a hotspot that is not circular but that applies only to the selected nodes. This second option is mostly used for signaling congestion.

Figure 2 shows three excerpts from the UI. Figure 2 (a) shows the definition of a new air pollution zone in a given region of the city. Figure 2 (b) shows how traffic is routed between two specific points if the preferences of the city favor the minimization of traveled distance ($W = 1,0,0,0$). Figure 2 (c), on the other hand, shows traffic routing if preferences favor the minimization of air pollution ($W = 0,0,0,1$).

5.2. Integration of Intelligent Path Finding Service

In order to validate the integration of the Intelligent Path Finding Service, a specific application was developed. The main goal of the application is to find the most adequate parking spot for a driver in a given area of the city.

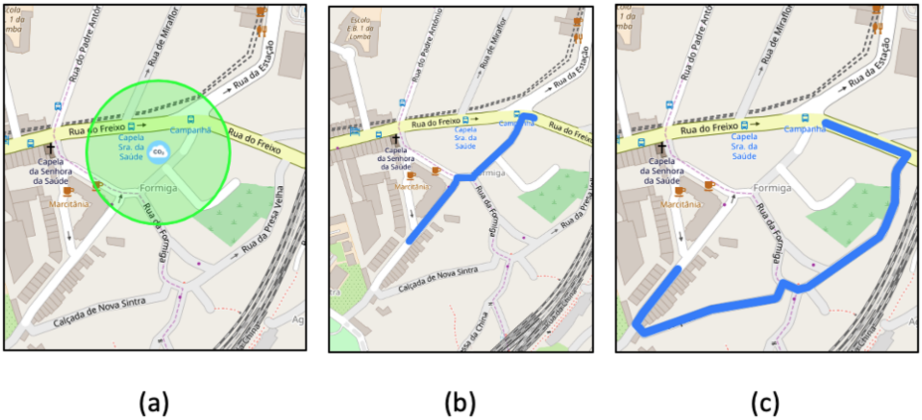

Figure 2. Details of the prototype of the interface: (a) creating a hotspot of pollution in a given region of the city; (b) traffic routing with $W = 1,0,0,0$ (minimizing distance); (c) traffic routing with $W = 0,0,0,1$ (minimizing air pollution).

To implement this application, historic data from several parking spots was used. Specifically, data from the Santa Monica dataset was used, which describes the state of parking lots every 5 minutes, 24 hours a day. Given that this dataset does not describe parking lots in the same city of the map data, parking lots were given random positions inside the map. The data was transformed in order to extract relevant features such as day, month, year, hour of the day, week day, among others. The problem was thus not treated as a time-series one. Figure 3 shows the typical fluctuation of the distribution of available parking spots during the day, for a specific parking lot. It shows how in certain hours of the day it is common for a parking lot to have very few or no free parking spots.

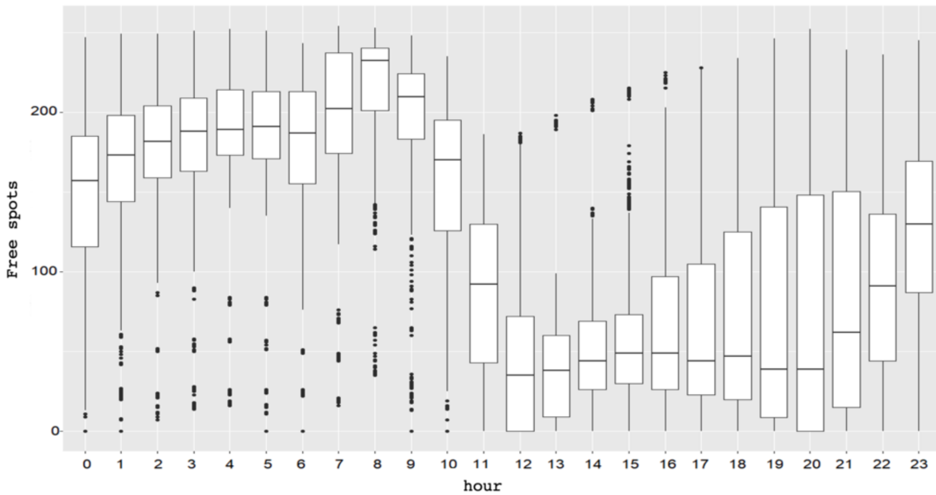

Figure 3. Typical fluctuation of free parking sports in a parking lot during the day.

Table 1. Performance of the best configuration model found for each model after parameter optimization through grid search.

Model	r^2	mae
GLM	0.56	137.27
Deep Learning	0.95	39.43
GBM	0.89	53.48

Then, several configurations of three Machine Learning models were trained to predict the number of free parking spots at a given time and parking lot, using a Grid Search scheme with cross-validation. The first model tested was a Deep Learning model [8] with different architectures of hidden layers and three different activation functions: Rectifier, Tanh and Maxout. The second model was a Generalized Linear Model (GLM) [3] and the third was a Gradient Boosting Machine (GBM) [5]. The performance metrics of the best model of each type is detailed in Table 1.

The deep learning model, given its superior performance, was thus selected to power the developed application. Thus, using this application, a driver first picks the point in the city where she/he wants to go. The application uses the Intelligent Path Finding Service to compute the best path to that point, as well as the estimated time of arrival.

Then, it predicts the number of available free parking spots in nearby parking lots and picks the one that has more expected free spots. The calculated path is then updated, to have as destination the selected parking spot. And it is this final spot that is provided to the user.

This application, which can be seen as an example of building on top of an existing service to provide a useful service for citizens, has several advantages. On the one hand, it optimizes the route of the driver in accordance with the city's rules, minimizing air pollution, noise or congestion as desired. Moreover, it selects the parking lot with the higher predicted number of free spots. This allows the driver to park more quickly, thus minimizing the negative impact of driving around searching for a free parking spot.

6. Conclusions and Limitations

In this paper we presented an innovative approach for managing traffic in a smart city. It relies on several key factors: real-time sensor data, city maps, and an engagement of citizens to abide by the systems decisions. The main limitation of the presented work is that it does not yet use real sensor data. Although a prototype of the system was implemented, it was validated using simulated data or data extracted from existing online datasets that are from a different city of the map used.

Still, we believe that the approach is worth pursuing in the sense that it might encompass advantages at several levels for a city. Its main distinguishing feature is that, instead of optimizing the behavior of each citizen in a selfish way, it optimizes for the greater good. That is, it defines individual routes while considering not only the preferences of the individual citizens but also the city preferences and policies.

Thus, as shown in Figure 2, a route may not necessarily be the shortest for a given driver if it optimizes other concerns of the city (e.g. minimizing air pollution). Moreover, the system is dynamic in the sense that routes may be different depending on the state of the city. This is especially useful to quickly react to significant changes in air pollution, for instance, something that is impossible to do with traditional methods that require the use of specific road signs, for instance.

In Future Work, we will proceed to integrate the proposed system with live data from the city's sensors and assess how traffic would be routed under different preferences and under different conditions.

Concluding, we believe that this system is particularly useful in a time in which autonomous vehicles start to reach our roads. Indeed, while a human driver may always ignore the common good and just use their own navigation system, an autonomous vehicle may be programmed to obey the management policies of the city it's driving in. All in all, this type of systems could result in optimized traffic systems that put the citizen in the center, and optimize traffic so that the quality of life in the city is improved.

Acknowledgments

This work was supported by FCT - Fundação para a Ciência e a Tecnologia, through project UIDB/04728/2020.

References

[1] BIBRI, S. E., AND KROGSTIE, J. Smart sustainable cities of the future: An extensive interdisciplinary literature review. *Sustainable cities and society 31* (2017), 183–212.

[2] DIJKSTRA, E. W., ET AL. A note on two problems in connexion with graphs. *Numerische mathematik 1*, 1 (1959), 269–271.

[3] DOBSON, A. J., AND BARNETT, A. G. *An introduction to generalized linear models*. CRC press, 2018.

[4] DUBOS, R. *Social capital: Theory and research*. Routledge, 2017.

[5] FRIEDMAN, J. H. Greedy function approximation: a gradient boosting machine. *Annals of statistics* (2001), 1189–1232.

[6] GAZZOLA, P., DEL CAMPO, A. G., AND ONYANGO, V. Going green vs going smart for sustainable development: Quo vadis? *Journal of cleaner production 214* (2019), 881–892.

[7] ISMAGILOVA, E., HUGHES, L., DWIVEDI, Y. K., AND RAMAN, K. R. Smart cities: Advances in research—an information systems perspective. *International Journal of Information Management 47* (2019), 88–100.

[8] LECUN, Y., BENGIO, Y., AND HINTON, G. Deep learning. *nature 521*, 7553 (2015), 436–444.

[9] NAGY, A. M., AND SIMON, V. Survey on traffic prediction in smart cities. *Pervasive and Mobile Computing 50* (2018), 148–163.

[10] NEILSON, A., DANIEL, B., TJANDRA, S., ET AL. Systematic review of the literature on big data in the transportation domain: concepts and applications. *Big Data Research* (2019).

[11] ROBINSON, I., WEBBER, J., AND EIFREM, E. *Graph databases*. " O'Reilly Media, Inc.", 2013.

[12] SAHARAN, S., BAWA, S., AND KUMAR, N. Dynamic pricing techniques for intelligent transportation system in smart cities: A systematic review. *Computer Communications* (2019).

[13] SILVA, B. N., KHAN, M., AND HAN, K. Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities. *Sustainable Cities and Society 38* (2018), 697–713.

[14] SINNOTT, R. W. Virtues of the haversine. *Sky Telesc. 68* (1984), 159.

[15] SODIQ, A., BALOCH, A. A., KHAN, S. A., SEZER, N., MAHMOUD, S., JAMA, M., AND ABDELAAL, A. Towards modern sustainable cities: Review of sustainability principles and trends. *Journal of Cleaner Production* (2019).

[16] ULLAH, Z., AL-TURJMAN, F., MOSTARDA, L., AND GAGLIARDI, R. Applications of artificial intelligence and machine learning in smart cities. *Computer Communications* (2020).

[17] YIGITCANLAR, T., KAMRUZZAMAN, M., FOTH, M., SABATINI-MARQUES, J., DA COSTA, E., AND IOPPOLO, G. Can cities become smart without being sustainable? a systematic review of the literature. *Sustainable cities and society 45* (2019), 348–365.

[18] Zheng, C., Yuan, J., Zhu, L., Zhang, Y., and Shao, Q. From digital to sustainable: A scientometric review of smart city literature between 1990 and 2019. *Journal of Cleaner Production* (2020), 120689.