Visualization and Waste Collection Route Heuristics of Smart Bins Data using Python Big Data Analytics

RICHARD B. WATSON

Ryan Watson Consulting Pty Ltd, Melbourne, Australia

PETER J. RYAN

Defence Science & Technology Group, Melbourne, Australia

This paper describes a set of waste management Application Programming Interfaces (APIs) written in the python language and using the Pandas, NumPy, Matplotlib, Basemap, Haversine and other big data analytics libraries. These access open datasets provided by the City of Wyndham, in Melbourne, Australia's western suburbs and stored on the Australian government's open data portal. These APIs read the data and process it to make it more useful to stakeholders including council administrators, waste management contractors and the general public. They provide visualization of the data in the form of plots of smart bin locations and fullnesses on maps accessed from Esri's ArcGIS API; bar charts of the frequency of fullness levels for both individual bins and all bins; and line charts of the fullness levels of a specified bin over time. The routes which can be followed by waste collection trucks are also given in terms of the legs from one bin to the next, specified in Javascript Object Notation (JSON) and also plotted on city street maps. These form heuristic solutions to the waste collection vehicle routing problem. The code used in the APIs is potentially transferable to analyses of data from other smart bin systems and other local government areas.

CCS CONCEPTS • Operations research • Data management systems • Visualization

Additional Keywords and Phrases: Smart Bins, Big Data Analytics, APIs, Python libraries

1 INTRODUCTION

A smart city is one which uses information and communication technology (ICT) to monitor and control city functions such as transport, utilities and waste management. This is facilitated by a system of sensors which monitor levels of traffic, energy usage, waste, etc and feed the data into central servers, which allow access to it and may take control action via actuators. The aim is more efficient planning and control, and enhanced liveability and sustainability. The data is termed big data as it may be much larger - terabytes, petabytes and larger - and more varied than traditional data, and may be stored on a cloud or processed in real time. Much big data is open to the public, and may be accessed over the Internet using Application Programming Interfaces (APIs). These APIs are of many kinds, some just provide access to raw data, others provide services derived from processing the data, which may include merging multiple datasets and applying various forms of Artificial Intelligence (AI) to support decision makers. An API is specified by its inputs - requests sent to a server, responses sent from the server and method of access, termed its endpoint.

Governments throughout the world are striving for greater efficiency and sustainability, and local governments seek to make their cities smarter. Australian local governments have made much of their data publicly available, but APIs are needed to get more value than can be derived from raw data. This paper describes a set of waste management APIs which access open datasets provided by the City of Wyndham, in Melbourne's western suburbs [1]. The purposes of the APIs are to enable better planning of bin placement and emptying times, to avoid unnecessary visits to empty bins, and prevent overflowing bins.

2 WASTE MANAGEMENT DECISION SUPPORT SYSTEMS

Management of a city's waste is a complex process as there are many kinds of waste which may require different disposal or recycling methods. Further, there are multiple stakeholders who play a part, including local council administrators, waste truck owners and operators, and managers of dumps and recycling facilities [2]. Decisions which must be made include how to classify waste, how to manage each kind of waste, where to locate garbage bins, when to empty the bins, what routes trucks should take, and where the waste should be taken. These decisions must take into account that the amounts and kinds of waste produced by a city's residents vary from place to place and from time to time. Household garbage bins are generally emptied in weekly collections, while council-owned smart bins may be emptied when needed, based on Internet of Things (IoT) sensor data transmitted via Wireless Sensor Networks (WSN) on their current status. The work described here only deals with council-owned smart bins.

There is a huge literature on waste management decision support systems, with many kinds of models proposed to address the different types of physical infrastructure, IoT technologies and software analytics [3, 4]. These models used can be subdivided into Operations Research (OR) and ICT [3], although these are not mutually exclusive categories. The solution methods within these categories may be further subdivided into mathematical programming and geographical information system (GIS)-based [5]. Belien et al. [6] classify solution methods into mathematical programming, metaheuristics, other heuristics and other methods including simulation.

Mathematical programming methods are described in detail by Lu et al. [7], Ramos et al. [8] and Omara et al. [9]. Recent studies using heuristic methods include the binary bat algorithm [10], genetic algorithm [11], and machine learning [12]. GIS-based solutions are described by Arribas et al. [13] and Ristic et al. [14]. Further, GIS-based software packages are available from the major GIS vendors Esri [15] and Google [16] as well as many smaller vendors [17].

It should be noted that most of the aforementioned studies evaluate the effectiveness of particular solution methods using simulation. Since the recent widespread adoption of smart bin systems, most simulations now assume that decision makers know the status of their bins in deciding how best to collect and dispose of municipal waste. However, as noted by Melare et al. [3, page 582], better decision making for waste management can be expected through the use of business intelligence tools and techniques, such as data mining and multidimensional analysis to transform raw data into meaningful and useful information. The suite of APIs developed by the authors is an exemplar for these sorts of tools. They provide various visualizations of the raw data, together with heuristic solutions to the waste collection vehicle routing problem.

3 LOCAL GOVERNMENT OPEN DATASETS IN AUSTRALIA

3.1 APIs for accessing government data

Australian government open datasets are stored and managed on the data portal: https://data.gov.au, using the platform MAGDA (Making Australian Government Data Available). The Representational State Transfer (REST) protocol and the JSON data format are used. Various APIs are provided for accessing and visualizing this data, which may be classified as Authorization, Content, Correspondence, etc. [18]. MAGDA employs software from several commercial vendors, including OpenDataSoft, Socrata and Esri. The portal has over 30,000 open datasets from federal. state and local governments. Australian local councils from all states and the Australian Capital Territory had published over 2,400 datasets as at April 2020 [2].

3.2 City of Wyndham smart bins datasets

The Wyndham City Council (WCC) has had a smart bins system since 2018 to monitor waste and recyclable garbage in the suburbs of Werribee and Point Cook in its Local Government Area (LGA). This system is similar to the Singapore-based Smartbin system described by Folianto et al. [19]. Ultrasound sensors placed in the smart bins monitor the fullness and other attributes of the bins. The data is then put online in the form of three datasets. One (the smart bins dataset) describes the bin's type, location, etc, another (the smart bins fill level dataset) adds the daily fill level, and a third (the smart bins daily fill level dataset) stores accumulated daily readings. These datasets have been analysed by one of the authors using python big data analytics, both to visualize the data in various ways, and to develop heuristic route planning algorithms for waste collection trucks. The full python code, API endpoints and instructions for using each of the APIs are documented in [20].

4 BIG DATA ANALYTICS USING PYTHON LIBRARIES

The JSON datasets used in this study are all semi-structured, ie. have internal tags and markings that identify separate data elements, that enable information grouping and hierarchies. The python language is now a mature technology with many open source libraries suitable for this sort of data [21]. The analysis process and main libraries used are depicted in Figure 1 below.

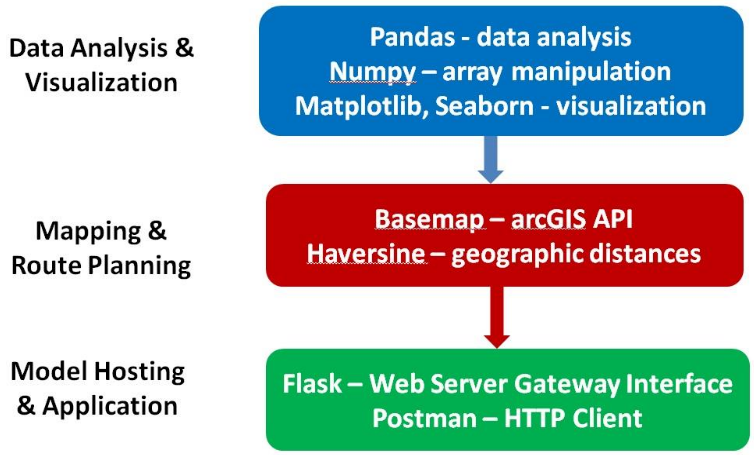


Figure 1. Python big data analytics libraries used for this study and web hosting application.

4.1 Data Analysis and Visualization

The Pandas library is used for data analysis and manipulation. It allows importing data in various file formats such as comma-separated values (CSV), JSON and Structured Query Language (SQL). It enables various data manipulations such as merging, reshaping, selecting and data cleaning.

NumPy is a scientific computing library, written in C and python, that supports large, multi-dimensional arrays with a rich collection of high-level mathematics functions. The Pandas library uses NumPy under the hood.

Matplotlib and Seaborn are plotting libraries written in python. They can be used interactively from the python shell to create pop-up windows with just a few commands.

4.2 Mapping and Route Planning

The Basemap toolkit is an extension of the Matplotlib library used for visualization of geographic data. It transforms geographical coordinates to various map projections. It can access the ArcGIS REST API [15] for plotting world street maps, which are used in the APIs described in this paper.

The Haversine or great circle distance is the angular distance between two points on the surface of a sphere [22], given their longitudes and latitudes. The formula can be coded using the NumPy library, but our APIs simply use the python Haversine library.

4.3 Model Hosting and Application

Flask is a lightweight Web Server Gateway Interface (WSGI) or web application framework used for creating web apps and APIs [23]. Its purpose is to map Hypertext Transfer Protocol (HTTP) requests to python functions. These requests include POST, GET, PUT and DELETE types, although we only used POST and GET in our APIs. Flask runs the code in the function and displays the returned response in a HTTP Client.

Postman is a simple HTTP Client which executes requests and renders responses [24]. There is a stand-alone version which runs on Microsoft Windows, and a Postman Extension to some browsers including Google Chrome, which allows POST requests to be passed to Flask.

5 APIS DEVELOPED FOR CITY OF WYNDHAM SMART BINS SYSTEM

The figures in the following subsections depict the main graphical responses produced by the APIs (Figures 2-5), and the results of further analyses of the data using Microsoft Excel® (Figures 6-8). They are meant to show the value added to the raw data by the API functionality. For example, knowing the mean fill levels of each bin can inform decisions on bin placement and frequency of emptying, and knowing the distances travelled by waste collection trucks for given collection thresholds can inform route planning.

5.1 Street Maps of Bin Fullnesses

This API reads the WCC smart bins fill level dataset and the ArcGIS API, and generates a colour-coded map showing the Werribee CBD with the bins' locations and fullness levels. This is shown in Figure 2 below.

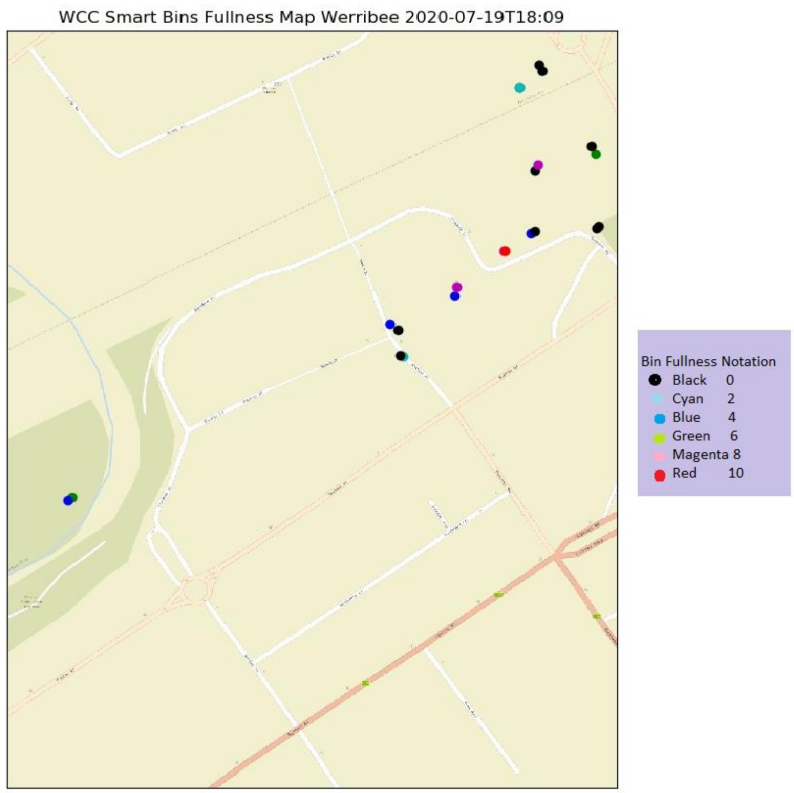


Figure 2: Street map of Werribee CBD showing smart bin fullnesses.

5.2 Bar Charts of Bin Fullnesses

This API reads the WCC smart bins fill level dataset and produces a bar chart showing the latest fullness of each bin. This is shown in Figure 3 below.

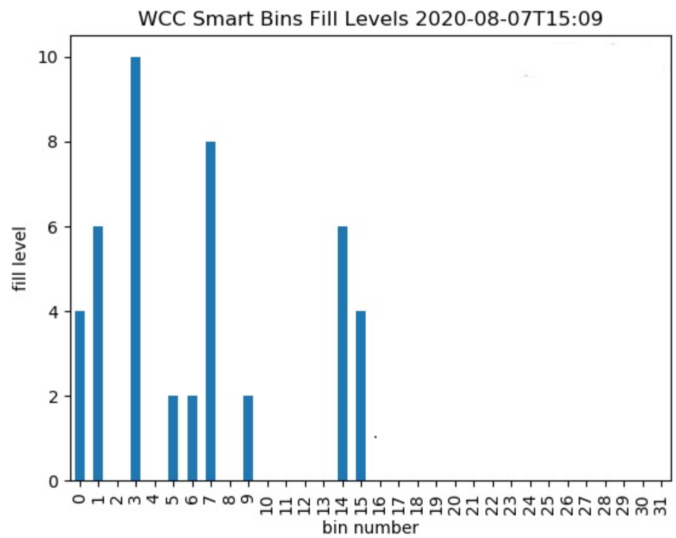
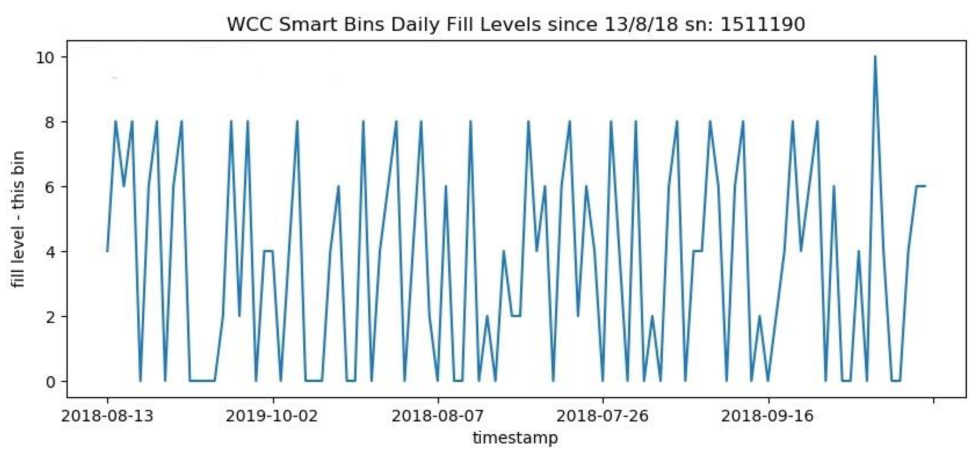


Figure 3: Bar chart of all smart bin fullnesses for a particular timestamp.

5.3 Line Charts of Fullnesses

This API reads the WCC smart bins daily fill level dataset, and inputs a JSON request with the bin serial number. The response is a line chart of that bin's fill level. As there are many days worth of data in the relevant Pandas DataFrame, it is convenient to select only 100 days of data for plotting, as shown in Figure 4 below.



Fi

Figgure 4: Line chart of a particular smart bin fullness versus timestamp.

5.4 Maps of Waste Collection Routes

This API reads the WCC smart bins fill level dataset and inputs a JSON request with the required threshold level for collection. The Pandas DataFrames are sorted by geographical coordinates and/or fullnesses, and the route determined using the Haversine python library [22]. The response is a streetmap of either the Werribee CBD, the Point Cook CBD or both, with the waste collection truck route marked for both waste and recyclables. Another API uses similar processing but outputs the route as a JSON file containing a series of pairs of bin numbers. Figure 5 below shows an example.

Figure 5: Streetmap of Werribee CBD showing waste collection vehicle routes, for a threshold level of 0.

5.5 Bar Charts of Average Fullnesses

This API reads the WCC smart bins daily fill level dataset and the response is the mean fill level of each bin since data collection began. Figure 6 below is a bar chart produced by Microsoft Excel® after sorting the data in increasing order of mean fill level.

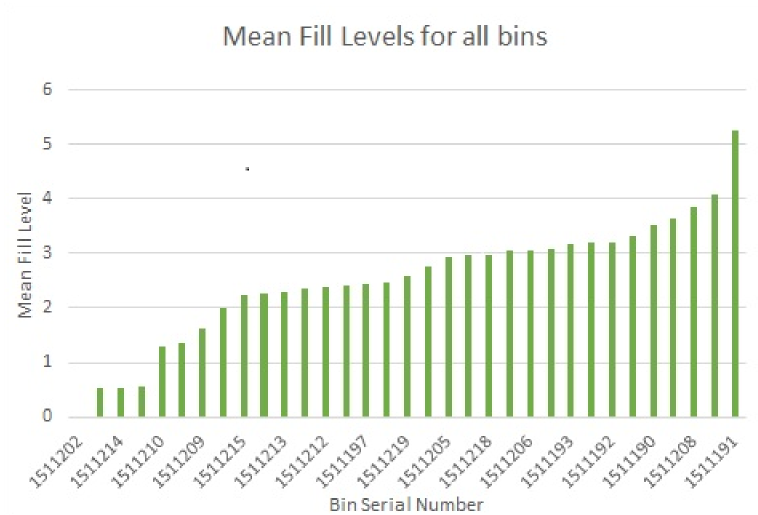


Figure 6: Bar chart of mean fill levels of all bins since data collection started.

5.6 Vehicle Routing Problem Heuristics

A special python program was written to calculate the average route distances for all days since data collection began. These are plotted for two threshold levels (2 and 4 respectively) in Figure 7 below.

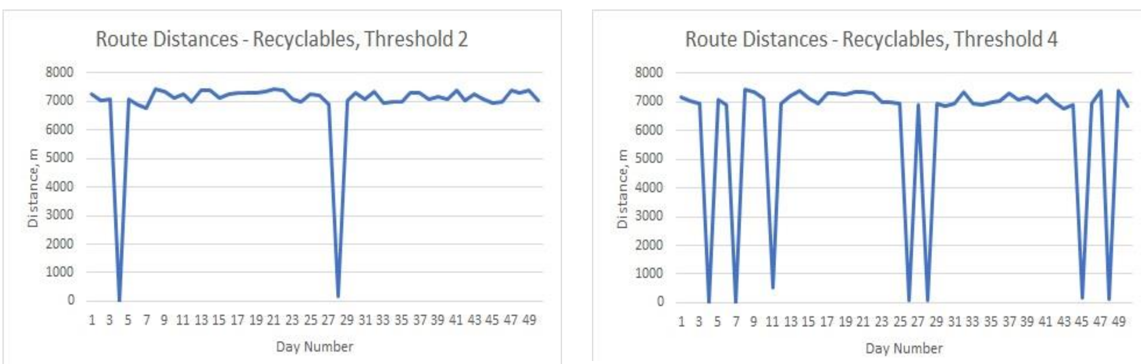


Figure 7: Line charts of average collection vehicle route distances travelled, for threshold levels of 2 and 4.

This program does not use an API, but reads both the WCC bins dataset and the WCC daily fill level dataset from their URLs. The threshold level is input via the code and the program calculates the route for each day, and the average distance travelled for both waste and recyclables. These distances are plotted against threshold in Figure 8 below (an Excel® stacked chart).

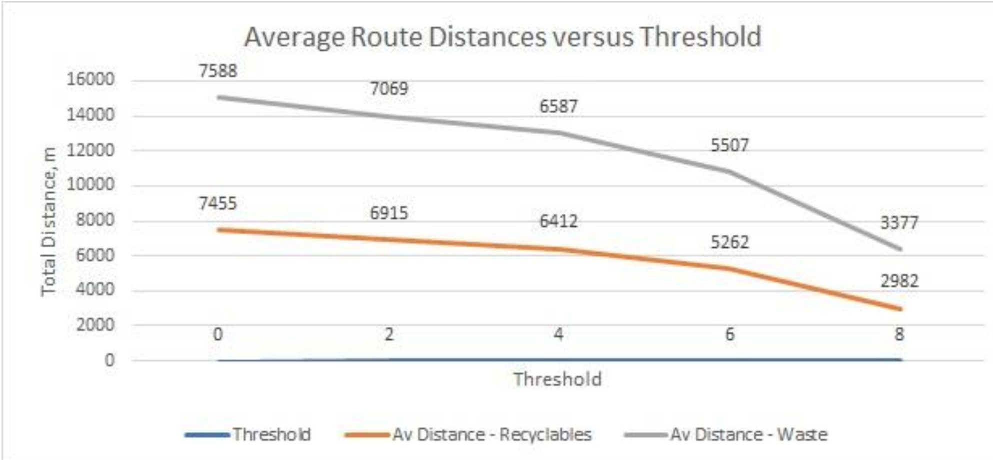


Figure 8: Plot of average collection vehicle route distances travelled, versus threshold level

6 CONCLUSIONS

This paper describes a suite of APIs developed by the authors to assist waste management decision making in an Australian local government area. The APIs convert raw IoT sensor data into charts which provide visualization of the data in the form of maps of smart bin locations and fullnesses, bar charts of the frequency of fullness levels for both individual bins and all bins; and line charts of the fullnesses of a specified bin over time. The routes which can be followed by waste collection trucks are also given in terms of the legs from one bin to the next, specified in Javascript Object Notation (JSON) and also plotted on city street maps. These form heuristic solutions to the waste collection vehicle routing problem. The code used in the APIs is potentially transferable to analyses of data from other smart bin systems and other local government areas.

The APIs described here require further testing by stakeholders and before entering the production stage will require containerization and deployment to a cloud server. The heuristic waste collection routes based on the data analysis described here could also be compared with routes based on mathematical programming and GIS models, to determine whether such more computationally intensive methods really add much value.

ACKNOWLEDGEMENT

The authors wish to acknowledge useful discussions with Paul Lowery, Spatial Information Officer, City of Wyndham.

REFERENCES

[1] Wyndham City Council. 2018. Smart Bins. Retrieved 20 September 2020 from https://www.wyndham.vic.gov.au/project/smart-bins

[2] Richard B. Watson and Peter J. Ryan. 2020. Big Data Analytics in Australian Local Government. Smart Cities 2020, 3, 657-675. https://www.mdpi.com/2624-6511/3/3/34/htm

[3] Angelina V. de S. Melare, Sahudy M. Gonzalez, Katti Faceli, and Vitor Casades. 2017. Technologies and decision support systems to aid solid-waste management: a systematic review. Waste Management 59, 567-584. https://doi.org/10.1016/j.wasman.2016.10.045

[4] Theodoros Anagnostopoulos, Arkady Zaslavsky, Kostas Kolomvatsos, Alexey Medvedev, Pouria Amirian, Jeremy Morley, and Stathes Hadjieftymiades. 2017. Challenges and Opportunities of Waste Management in IoT-enabled Smart Cities: A Survey. IEEE Transactions on Sustainable Computing 2, 275-289. https://ieeexplore.ieee.org/document/7892970

[5] Alhassan Sulemana, Emmanuel A. Donkor, Eric K. Forkuo, and Sampson Oduro-Kwarteng. 2018. Optimal Routing of Solid Waste Collection Trucks: A Review of Methods. Hindawi Journal of Engineering 2018, 4586376, 1-12. https://doi.org/10.1155/2018/4586376

[6] Jeroen Belien, Liesje De Boeck, and Jonas Van Ackere. 2011. Solid Waste Collection Problems: A Literature Review. Transportation Science 48(1), 78-102. https://dl.acm.org/doi/10.1287/trsc.1120.0448

[7] Jia-Wei Lu, Ni-Bin Chang, Li Liao, and Meng-Ying Liao. 2017. Smart and Green Urban Solid Waste Collection Systems: Advances, Challenges, and Perspectives. IEEE Systems Journal 11(4), 2804-2817. https://ieeexplore.ieee.org/document/7265018

[8] Tania R. P. Ramos, Carolina S. de Morais, and Ana P. Barbosa-Povoa. 2018. The Smart Waste Collection Routing Problem: alternative operational management approaches. Expert Systems with Applications 103, 146-158. https://doi.org/10.1016/j.eswa.2018.03.001

[9] Ahmed Omara, Damla Gulen, Burak Kantarci, and Sema F. Oktug. 2018. Trajectory-Assisted Municipal Agent Mobility: A Sensor-Driven Smart Waste Management System. Journal of Sensor and Actuator Networks 2018, 7, 1-29. https://www.mdpi.com/2224-2708/7/3/29

[10] Tamas Banyai, Peter Tamas, Bela Illes, Zivile Stankeviciute, and Agota Banyai. 2019. Optimization of Municipal Waste Collection Routing: Impact of Industry 4.0 Technologies on Environmental Awareness and Sustainability. International Journal of Environmental Research and Public Health 2019, 16, 634, 1-26. https://www.mdpi.com/1660-4601/16/4/634

[11] Sahar Idwan, Imran Mahmood, Junaid A. Zubairi, and Izzeddin Matar. 2020. Optimal Management of Solid Waste in Smart Cities using Internet of Things. Wireless Personal Communications 110, 485-501. **https://doi.org/10.1007/s11277-019-06738-8**

[12] David Rutqvist, Denis Kleyko, and Fredrik Blomstedt. 2020. An Automated Machine Learning Approach for Smart Waste Management Systems. IEEE Transactions on Industrial Informatics 16(1), 384-392. https://ieeexplore.ieee.org/document/8709695

[13] Claudia A. Arribas, Carola A. Blazquez, and Alejandro N. Lamas. 2009. Urban solid waste collection system using mathematical modelling and tools of geographic information systems. Waste Management and Research 00, 1-9. https://doi.org/10.1177%2F0734242X09353435

[14] Goran Ristic, Amelija Djordjevic, Sladjan Hristov, Predrag Umicevic, Aleksandra Petkovic, and Lidija Milosevic. 2015. Methodology for Route Optimization for Solid Waste Collection and Transportation in Urban Areas. Working and Living Environmental Protection 12(2), 187-197. http://casopisi.junis.ni.ac.rs/index.php/FUWorkLivEnvProt/article/view/777

[15] Esri. 2020. Build Powerful apps with the ArcGIS REST API. Retrieved 17 September 2020 from https://developers.arcgis.com/rest/

[16] Google. 2020. Vehicle Routing Problem. Retrieved 17 September from https://developers.google.com/optimization/routing/vrp

[17] Altexsoft. 2019. How to Solve Vehicle Routing Problems: Route Optimization Software and their APIs. Retrieved 17 September 2020 from https://www.altexsoft.com/blog/business/how-to-solve-vehicle-routing-problems-route-optimization-software-and-their-apis/

[18] Australian Government. 2020. MAGDA API Documentation. https://data.gov.au/api/v0/apidocs/index.html

[19] Fachmin Folianto, Yong Sheng Low, and Wai Leong Yeow. 2015. Smartbin: Smart Waste Management System. In Proceedings of the IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP). Singapore, 7-9 April 2015. https://ieeexplore.ieee.org/document/7106974

[20] Richard B. Watson and Peter J. Ryan. 2020. Wyndham City Council Smart Bins System, Open Data APIs – User Guide. Unpublished Report. 47 pages.

[21] Jack Chan, Ray Chung, and Jack Huang. 2019. Python API Development Fundamentals. Packt Publishing. Birmingham, UK.

[22] Wikipedia. 2020. Haversine formula. Retrieved 20 September 2020 from https://en.wikipedia.org/wiki/Haversine_formula

[23] Nicholas Hunt-Walker. 2018. An introduction to the Flask Python web app framework. Retrieved 22 September 2020 from https://opensource.com/article/18/4/flask

[24] Postman. 2020. The Collaboration Platform for API Development. Retrieved 22 September 2020 from https://www.postman.com/

Authors' background

Your Name	Title*	Research Field	Personal website
Richard Watson	PhD	Operations Research	https://sciprofiles.com/profile/RichardWatson
Peter Ryan	PhD	Operations Research	https://www.linkedin.com/in/peter-ryan-4432b3a/?originalSubdomain=au https://orcid.org/0000-0002-0365-3192

***This form helps us to understand your paper better, the form itself will not be published**