TUDelft

Delft University of Technology

Readiness-for-reuse of open government data for solving societal problems based on public participation: kindergartens in the City of Zagreb

Kević, Karlo; Miletić, Andrea; Kuveždić Divjak, Ana; Welle Donker, F.M.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

OPEN DATA USER JOURNEY: MAPPING WITH OPEN DATA FROM A USER PERSPECTIVE

Karlo Kević*[1], Andrea Miletić[2], Ana Kuveždić Divjak[3] and Frederika Welle Donker[4]

[1] Chair of Geoinformation, Institute of Cartography and Photogrammetry, University of Zagreb, Faculty of Geodesy, Croatia
[2] Chair of Geoinformation, Institute of Cartography and Photogrammetry, University of Zagreb, Faculty of Geodesy, Croatia
[3] Chair of Geoinformation, Institute of Cartography and Photogrammetry, University of Zagreb, Faculty of Geodesy, Croatia
[4] Urban Data Sciences, Department of Urbanism, Faculty of Architecture and the Built Environment, Delft University of Technology

*karlo.kevic@geof.unizg.hr

EXTENDED ABSTRACT

Over the last decade, open data has become a hot topic that researchers and governments all over the world are eager to explore. The term first appeared in a 1995 document from an American scientific agency (URL1), but its popularity increased when the U.S. president, Barrack Obama, signed the 2009 Memorandum on Transparency and Open Government, pledging to achieve an unprecedented level of openness in government (URL2). This memorandum emphasized the importance of creating a system of transparency, public participation, and collaboration through openness.

In recent years, much research has been done on the benefits arising from open government data (OGD). It is believed that OGD enables economic and social growth. According to the European Commission's 2020 report, the size of the data market in Europe is estimated to increase significantly by 2025, with possibly 2 million people employed in the open data sector. But the potential of open (government) data is going beyond just economic benefits. From a social perspective, open data enhances participation, and collaboration and empowers citizens to contribute to policies designed to their needs. Some sources state that the main motivation behind opening the data was precisely to encourage citizens to engage in solving societal problems (Huyer, 2020). This engagement is seen as a key for a successful and sustainable OGD where collaborative actions can lead to new information which could possibly solve existing societal problems (Dietrich, 2015; Purwanto et al., 2020).

For open data to reach its full potential, Pollock (2011) suggested that a traditional one-way street data model, where there is no feedback loop, and no sharing of data back to publishers and between intermediaries, should be replaced with an ecosystem approach. Charalabidis and all (2018) developed an open data life cycle model based on an open data ecosystem that brings together two main stakeholders: data providers and data users. In the inner circle are the data providers, who are responsible for creating, pre-processing, curating, storing, and publishing the data, while the outer circle consists of the data users whose role is to acquire, process, use, collaborate and provide feedback. An important part of this process is the data re-use which is responsible for unlocking benefits arising from open data.

Different open data users can have different insights derived from open data (URL3). These insights can accelerate the development of new applications that might provide solutions to

existing societal problems. For example, in some cities, the re-use of open data on public transport accessibility ended up with new applications and services facilitating the movement of disabled people (URL4). Since open data enables a better relationship with the government and enhances citizen participation in traditional policymaking (URL4), this research will investigate to what extent can citizens use open data to solve a specific societal problem and provide proposals for the improvement of the city's social policy.

With the increasing urbanization and greater population migrations to the capital of Croatia, the city of Zagreb is facing a problem of discrepancy between population growth and distribution, and kindergarten infrastructure (URL5). Despite the number of kindergartens (231) (Službeni glasnik Grada Zagreba, 2021) in Zagreb, every year a lot of children are left unrolled or, if enrolled, they might be collocated to a random institution somewhere in the city. In the future, this problem is expected to be even more prominent with additional city population growth and revisions of existing local social policies. To mitigate this problem, the city of Zagreb made a program on preschool education for the year 2022 with a strategy on how to improve the childcare service in public kindergartens (Službeni glasnik Grada Zagreba, 2021). Among others, this program foresees the expansion of existing childcare capacities either by reconstruction or construction of new facilities. However, although stating where and how they plan to do it, no studies supporting decisions in the program were found. With open data initiatives supporting public participation in decision-making, this paper aims to analyze how well an end user using open data can contribute to the policy on the improvement of kindergarten infrastructure in the city of Zagreb and make it more aligned with the population's needs. To provide an answer, the analysis is divided into two parts, Data quality and availability, and Public Engagement in Policy engagement in policy-making (Fig. 1), and focused on five specific aims:
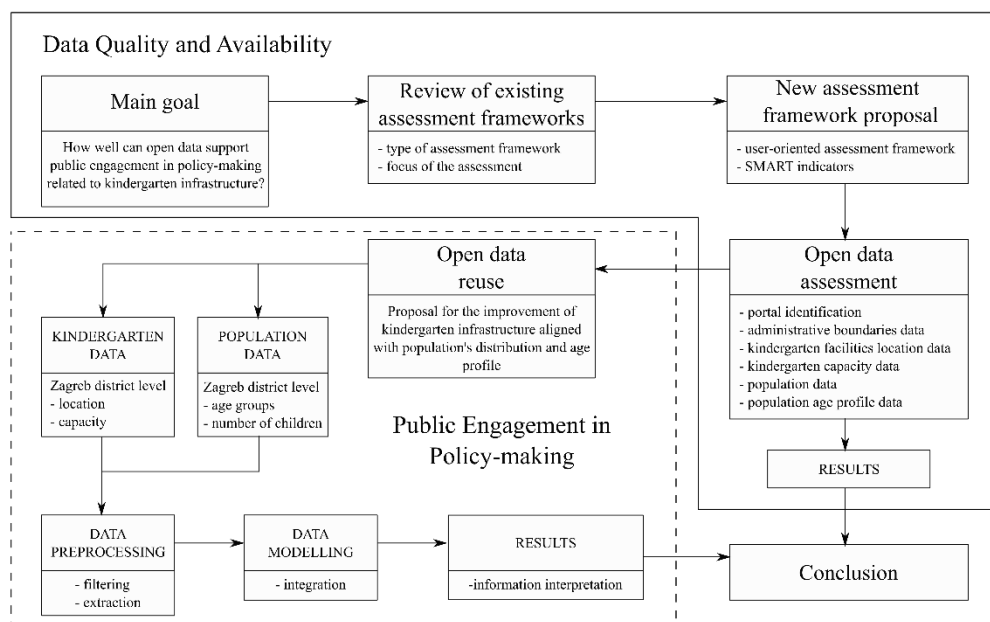

Figure 1. Data Quality and Availability Assessment

Data quality and availability
1. Identification and analysis of sources and open government data related to population and kindergarten infrastructure

2. Analysis of existing open data assessment frameworks and detection of gaps
3. Development of a new assessment framework aligned with specific user needs
4. Analysis of open government data related to administrative boundaries, population, and kindergarten infrastructure using a newly developed assessment framework

Public Engagement in Policy-making

5. Using open data, propose improvements to the policy on kindergarten infrastructure to make it more aligned with the population's distribution and age profile.

As this approach understands the re-use of open government data for solving societal problems based on public participation, the user's perspective will provide relevant and current state-of-the-art. This way, it can be assessed whether the end user, e.g., a citizen with certain data manipulation skills, can engage with data and exploit the vision of open government data initiatives. Also, providing answers to these specific aims will give information about the availability but also about suitability of open government data to solve a simple societal problem.

The methodology proposed follows the aforementioned aims. Data-quality-and-availability part of the analysis assesses the availability, findability, and level of openness of open data on administrative boundaries, kindergarten infrastructure, and statistical data about the residents of the city of Zagreb on a city district level using an assessment framework. Aligned with the aim of the analysis, the assessment framework is user-oriented meaning that its indicators are formulated to reflect on the user experience when looking for open data and based on principles of data openness. On the other hand, the public-engagement-in-policy-making part of the assessment goes deeper into the aforementioned open data and seeks if its content, resolution, and quality are appropriate to indicate the kindergarten's infrastructure alignment with the age structure of the residents in the city of Zagreb district level. This part of the analysis included data preprocessing and data modeling. Data preprocessing uses filtering to extract only data relevant to the analysis. Data modeling concerned the residents' age profile, its division into two groups: preschool children and young people, and integration with kindergarten data on the city district level. The authors consider that the division of age data into groups of preschool children and young people is relevant for two reasons. First, the number of preschool children in comparison with current kindergarten capacities will show the alignment of the service with users' needs. Second, the number of children and young people on a city district level can indicate the need for new kindergarten capacities in a specific district rather than in another district with more mature residents. Therefore, this approach is considered relevant in achieving the main goal of the analysis. In addition, with cartographic visualization being a tool of visual communication, and having in mind its role as an intermediary between data and policymakers, results are presented on a choropleth map.

All the data needed to make this analysis are government related meaning it should be available on the National Open Data Portal which serves as a central point of access to all data related to public administration (URL6).

This research however showed datasets to be scattered across different portals and agencies' websites, which makes it hard for users to find all the data needed and increases time spent on data searching. Data regarding the population by the city district was found on the official website of the Croatian Bureau of Statistics. However, data about the kindergartens and their location was found on the National Open Data Portal. There was also a problem with obtaining the administrative boundaries of the city districts. According to the new, European Open Data Directive (EU 1024/2019) administrative units belong to geospatial data and as such are considered high-value datasets and should be publicly available. In the case of the City of Zagreb's district administrative boundaries, no official data was found. Also, research showed

that there is a limited quality of data in provided datasets; there are errors in the raw data and some data is even missing. The lack of metadata is also a big issue, for some data used, there is no information if the data is up-to-date, and even the license is not clearly determined. Based on the results it can be stated that the data is still lacking some open data aspects, however, it is considered good enough for the next phase of the analysis.

After performing the necessary steps, it was stated that the current kindergarten infrastructure does not meet public needs with the number of children outnumbering kindergartens' capacities. Also, the kindergarten infrastructure does not align with the population's age profile at the city district level. For example, the kindergarten infrastructure in the city center is not exploited enough by mostly older population identified living in the area. On the contrary, newer neighborhoods with mostly young families are lacking kindergarten space.

Seen from the results, the city of Zagreb provides open data which could be used to support public engagement in decision-making. This data is however scattered across different portals and websites and is missing metadata needed for additional context. In addition, the data is of limited quality, often with no information about the last update and terms of use. When trying to provide an answer to a simple societal question of the alignment of kindergarten infrastructure with the population's age profile, the additional problem of administrative boundaries not being officially available was faced. This suggests that some of the geospatial data, which is identified by the EU Open Data Directive as high-value datasets, is still not available for re-use. The second part of the research suggests that open data looked at is sufficient (up to a certain level) to engage the public in decision-making. However, this engagement requires end users to have certain data manipulation skills which, with all the previously mentioned limitations in data, need to be more advanced. This causes layering in defining an end user. Results of the analysis based on open data also suggest a discrepancy between kindergarten infrastructure and the population's age profile which may serve as a starting point for future policy creation by the local government.

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., Ferro, E. (2018). The Multiple Life Cycles of Open Data Creation and Use. In: The World of Open Data. Public Administration and Information Technology, vol 28. Springer, Cham.

Dietrich, D. (2015). The Role of Civic Tech Communities in PSI Reuse and Open Data Policies [European Public Sector Information Platform Topic Report No. 2015/05]. Retrieved from the web page:
https://data.europa.eu/sites/default/files/report/2015_the_role_of_civic_tech_communities_in_psi_reuse_and_open_data_policies.pdf

Pollock, R. (2011): Building the (Open) Data Ecosystem, Open Knowledge Foundation, https://blog.okfn.org/2011/03/31/building-the-open-data-ecosystem/

Program javnih potreba u predškolskom odgoju i obrazovanju te skrbi o djeci rane i predškolske dobi Grada Zagreba za 2022, Službeni glasnik Grada Zagreba, 31/21

Publications Office of the European Union, Huyer, E., The economic impact of open data: opportunities for value creation in Europe, Publications Office, 2020

Purwanto, A., Zuiderwijk, A., Janssen, M. (2020): Citizen engagement with open government

data: Lessons learned from Indonesia's presidential election. Transforming Government: People, Process and Policy, vol 14 (1)., 1-30. 10.1108/TG-06-2019-0051

URL1: (https://www.paristechreview.com/2013/03/29/brief-history-open-data/)
URL2: (https://obamawhitehouse.archives.gov/the-press-office/transparency-and-open-government)
URL3: (https://data.europa.eu/en/datastories/empowering-people-open-data)
URL4: (https://data.europa.eu/en/datastories/benefits-and-value-open-data)
URL5: (https://huro.hr/alarmantno-gradu-zagrebu-nedostaje-4030-mjesta-u-vrticima/)
URL6: (https://vlada.gov.hr/highlights-15141/archives/croatian-government-launches-open-data-portal-data-gov-hr-16574/16574)