MANAGEMENT
2025, Vol. 29, No. 1
www.management-poland.com
DOI: 10.58691/man/207993

MARCIN ZDRADZISZ
MARZENA NOWAKOWSKA

Quantitative and qualitative aspects of participatory budgeting – the Kielce case study

Abstract

Research background and purpose: Participatory budgeting (PB) plays an important role in a city live. It strengthens democracy, improves transparency of public expenditures, helps allocate resources more efficiently, and increases civic responsibility. In order to assess how such a mechanism contributes to public involvement and which initiatives are prioritized by the community engaged in the participatory budgeting process, a pilot study was undertaken for a selected Polish voivodeship city - Kielce. Attention was paid to the category of winning tasks, considering the classification by project type and its narrative.

Design/methodology/approach: The exploratory analysis was carried out on real unit data. Data on projects submitted by residents were collected from websites using web scraping tools. The article relies on the case-study method as well as on the methods of quantitative and qualitative analysis. The narrative data were explored with the assistance of machine learning technology – Natural Language Processing (NLP) and the results were illustrated in the form of a Word Cloud.

Findings: The conclusions of the picture of participatory budgeting in the chosen city are decidedly positive. It is a tool that promotes democratization of decision-making processes and involvement of citizens in public life. It is significant that the local community opts for soft projects, particularly green ones.

Value added and limitations: This study investigates the changes in citizen engagement in PB using a time- and project-type-oriented quantitative approach. Furthermore, a machine learning-based qualitative approach in terms of NLP was employed to capture the major PB issues and topics, which is rather absent in this research field. As a pilot study, this research is limited to a single city and begins to explore text processing algorithms in such analysis. The promising results suggest potential for research development in the future.

Keywords: *participatory budgeting, civil society, case-study, qualitative and quantitative analysis, NLP*

JEL Classification: H72

Received: 2024-10-30; **Revised:** 2025-03-05; **Accepted:** 2025-07-07

Marcin Zdradzisz ✉
Faculty of Management and Computer Modelling, Kielce University of Technology, al. Tysiąclecia Państwa Polskiego 7, 25-314, Kielce, Poland; email: mzdradzisz@tu.kielce.pl
ORCID: 0000-0002-8929-2956

Marzena Nowakowska
Faculty of Management and Computer Modelling, Kielce University of Technology, Poland
ORCID: 0000-0002-6934-523X

1. Introduction

Participatory budgeting (PB) is defined as a process which allows citizens to participate in deciding on the allocation of all or a part of public funds. It is a manner in which citizens may become involved in one of the key public policy processes, which is the budget procedure. It is an innovative approach that enables residents to actively participate in shaping public expenditures.

Participatory budgeting as a form of direct democracy at a local level has gained international recognition for its ability to engage citizens in decision-making processes regarding public spending. Its roots date back to the late 1980s, when it was first implemented in Porto Alegre, Brazil (de Sousa Santos, 1998). That pioneering model allowed city residents to influence the allocation of part of the municipal budget, which was a revolutionary step towards increasing transparency and accountability in the management of public funds. In Porto Alegre, participatory budgeting contributed to significant improvements in infrastructure and public services, as well as to increased social and political involvement of residents.

In order to recognize the PB literature, two global bibliographic platforms, Web of Science – WoS and Scopus, were searched on 2025/02/16 for publications related to participatory budgeting. The search criteria specified publications from 2000 to 2024 and used the following keywords in the title, abstract, or keywords: "*participat* budget*". In Scopus and WoS, there were over 1,200 and 950 documents, respectively. In both databases, during the first five years of the assumed period, academic production on PB did not exceed forty and twenty, while in the last five years, it reached nearly six hundred and slightly more than four hundred. The number of research issues is also significant, with a wide range of topics in both Poland and other countries (Kociuba & Rabczewska, 2019; Bartocci et al., 2023). Many PB publications address single locations, creating case studies in a variety of contexts. For example, Falanga (2024) focuses on the Lisbon (Portugal) edition of the 2021 PB, considering the engagement of the youth in environmental sustainability through participatory budgeting, while Tatarowska and Furmankiewicz (2018) analyze types and spatial distribution of projects for selected implementations in 2014 in Wrocław (Poland). Participatory budgeting is also studied from a financial point of view, at country level (Bogo & Falanga, 2024) or at a city level (Dubicki, 2021), but also taking into account sociological issues (Šabović et al., 2021), or determinants influencing the sustainability of PB (Svidroňová, 2024). Technical aspects are considered as well; for example, Faliszewski et al. (2023) discuss how various selection rules in participatory budgeting work in practice. In an in-depth technical analysis, they showed which voting rules are fair and indicated issues that should be considered when designing elections, particularly with regard to their overall usability for voters'. The analyses of total research outputs in the field of PB can be found in

works by Bartocci et al. (2023), Milosavljević et al. (2023), or Klimovský et al. (2024), the authors acknowledge that PBs are research subject from variety of science disciplines, thus becoming an important and interdisciplinary issue.

The investigation can focus not only on numerical data related to participatory budgeting projects but also on identifying the general priorities and preferences of residents regarding the improvement of urban quality of life. The latter aspect is directly linked to the analysis of textual data contained within these projects. Therefore, the search string criterion was refined by modifying the condition applied to the title, abstract, or keywords as follows: "*participat* budget*" AND ("NLP" OR "natural language processing"). A total of four publications were identified in both bibliographic databases. The article by Segura-Tinoco et al. (2022) presents and evaluates a new chatbot that helps users explore citizen-generated content on the Decide Madrid forum-based e-participatory budgeting platform, using Natural Language Processing and artificial intelligence. Davies et al. (2021a, 2021b) identify areas where IT can support selected processes related to participatory budgeting in Scotland. They discuss and evaluate the capabilities of the Consul digital platform, which incorporates NLP to enhance citizen engagement and its use by public decision-makers. The works examine how NLP technology helps address the challenges of scaling PB across 32 Scottish local authorities and how it can enhance the effectiveness of citizen participation and policy development. Peláez et al. (2016) discuss a framework for decision-making in e-government systems, leveraging fuzzy logic, mobile technologies, and linguistic information. The authors analyze how combining these elements can contribute to obtaining a representative final opinion decision-making, leading to better engagement and more efficient government services with regard to participatory budgeting.

Although PB is a fairly common research subject, particularly in social sciences, the authors have not found a publication that provides a picture of participatory budgeting throughout its life, taking into account both quantitative and qualitative contexts, with the use of Natural Language Processing in particular. This work undertakes such a task. The study is of a pilot nature and aims to present an analytical perspective on participatory budgeting (PB) in Kielce, one of the 18 voivodeship capitals in Poland, during the years 2013-2023, based on selected characteristics of citizen projects. The research questions are therefore:

1. How has the participatory budgeting implementation in Kielce evolved over time concerning the numerical characteristics of various project types?
2. What are the priorities and preferences of residents regarding urban development to improve live quality in the city?
3. How can machine learning, specifically Natural Language Processing, be used in analyzing PB project narratives?

The article is structured as follows. First, information about Kielce participatory budgeting (KPB) is given. Next, the PB data selected for the study and research methodology are explained. Then, results are presented and discussed. The last section provides concluding remarks along with the main research limitations.

The analysis of projects submitted by residents provides insights into their preferences, which can help optimize public spending and more efficient distribution of funds as well as assess the social value of projects financed through participatory budgeting. The study highlights both the PB overall picture and the potential of using modern data analysis methods in public management, which may contribute to better understanding and further development of participatory budgeting concept. The proposed approach provides a useful research methodology that can inspire readers, offering a framework for comparative analysis of results obtained in similar studies and contexts.

2. Kielce Participatory Budgeting

In Poland, the first participatory budget was introduced in Sopot in 2011 (Kempa & Kozłowski, 2020), which was an important step towards strengthening local democracy and increasing the influence of residents on the development of their communities. Since then, the participatory budget has become a popular instrument in many Polish cities, enabling the implementation of projects that directly respond to the needs of local communities (Soboń, 2017). These initiatives are undertaken in accordance with PB regulations, which are set independently by local authorities. The city of Kielce, along with Poznań, Łódź, Wrocław, and Gorzów Wielkopolski, was among the voivodeship capitals that first introduced PB – in 2013.

Kielce is a city located in central Poland, the capital of the Świętokrzyskie Voivodeship, serving as its administrative and economic center. Among the 16 voivodeship capitals, it ranks roughly in the middle in terms of demographics (200,000 residents) and budget (381 million EUR in 2024). The city is situated in the heart of the Świętokrzyskie Mountains, characterized by a mix of urban and natural landscapes. Kielce's economy is diverse, with key sectors including commerce, services, and regional industry. It is also known for hosting trade fairs and industrial exhibitions, giving it a distinct economic niche within the country. Its development strikes a balance between urbanization and the preservation of green spaces. Cultural and educational institutions play a significant role in regional development and the shaping of local identity.

In Kielce, the history of the PB idea in question dates back to 2013, when projects selected by Kielce residents were implemented for the first time. Since then, the Kielce Participatory Budgeting (KPB) has become an important element of local democracy, enabling residents to decide on projects that have a real impact on the quality of their lives and surroundings. The regulations of KPB were gradually updated, taking into

consideration various aspects of the functioning of the local community as well as changes in legislation. Initially, until 2018, there were only investment project categories, which were divided into two types, depending on the proposed costs: *Large* – over PLN 150000 (EUR 34400-35900[1]) and *Small*, introduced in 2014, – not more than PLN 150000. In 2019, the *Non-investment* projects category launched, and in 2021 the *Green* category. Also in 2021, the limits for investment projects were changed: *Large* – over PLN 200000 (EUR 32000-33000), *Small* – not more than PLN 200000. Investment projects, both *Large* and *Small*, are initiatives that lead to the creation of new or the modernization of existing urban infrastructure. These may include playgrounds, parks, roads, sidewalks, lighting, and sports or cultural facilities. Their aim is to improve the quality of life for residents by creating functional and aesthetically pleasing public spaces. *Non-investment* projects are endeavors that do not require significant financial expenditure on infrastructure but aim to promote the social, cultural, or educational development of residents. These may include workshops, training sessions, cultural, sports, or educational events, social programs, charitable actions, or efforts aimed at integrating the local community. Their goal is to activate residents, develop their skills and interests, build a sense of community, and strengthen social bonds. *Green* projects focus on environmental protection and the creation of green spaces in the city. They include activities such as planting trees and shrubs, establishing flower meadows, pocket parks, creating green roofs, and rainwater systems. Their goal is to improve air quality, increase biodiversity, protect water resources, and create inviting spaces for recreation and relaxation in nature.

The KPB procedure begins when city residents apply to the City Hall by filling out an electronic form on the website www.bo.kielce.pl, which is temporarily available before the project submission deadline. The application is processed according to the schedule presented in Figure 1.

Figure 1. **Schedule for processing KPB projects**

Source: prepared on the basis of www.bo.kielce.pl

1 According to the information from the National Bank of Poland, there was the following yearly average exchange rates for converting EUR into PLN ((Narodowy Bank Polski, b.d.): 2013 – 4.1976; 2014 – 4.1819; 2015 – 4.1746; 2016 – 4.3574; 2017 – 4.2625; m2018 – 4.2617; 2019 – 4.2988; 2020 – 4.4459; 2021 – 4.5670; 2022 – 4.6872; 2023 – 4.5437.

The procedure begins with the submission of a project by a resident, with the support of at least 15 people being required. Submitted applications are published on the website – projects' showcase. Projects undergo formal verification process, conducted by the City Hall department responsible for coordinating PB. In the case of a negative assessment, the applicant has the right to appeal. Subsequently, projects positively verified in terms of formalities are subjected to substantive and legal verifications by relevant City Hall units. Accepted projects create the list of those eligible for voting and are published on the website. Eligible to vote is any resident of Kielce; in the case of a person under 13 years of age, the consent of a legal guardian is required. KPB projects are promoted by their authors and supporters in various ways to reach as wide an audience as possible and advocate for the initiatives. Often, the authors and supporters of these initiatives begin promotional campaigns early, during the project preparation phase. The intensity of informational and encouraging activities increases right after the list of projects approved for voting is published. Promotion is carried out in various ways, such as meetings with residents, posting content on social media, posters and leaflets, and direct campaigning on the streets and at the workplaces of authors and supporters. Voting takes place electronically via the website. A maximum of 4 projects can be supported per vote as follows:

- one *Large* investment project,
- one *Small* investment project,
- one *Non-investment* project,
- one *Green* project.

The voting results are announced on the website within 14 days after the end of voting. The announcement contains three pieces of information:

- the total number of people who took part in the vote, specifying the number of valid and invalid votes,
- the number of votes received by each project,
- projects targeted for implementation.

Projects that receive the highest number of votes consecutively and whose total costs are within the pool of funds allocated to each category are subject to implementation.

Over the years, thanks to participatory budgeting, a wide variety of initiatives have been carried out in Kielce, from infrastructure projects, such as the construction of playgrounds or the modernization of public spaces, to projects of a social, cultural and educational nature. The value of the projects that were successfully realized by 2023 exceeded EUR 16.4 million, which proves the scale of the involvement and influence of people of Kielce on the development of their city.

3. Data and Methodology

3.1. The data

The data for the analysis cover the period 2013-2023. The data were downloaded from the Kielce City Hall website www.bo.kielce.eu, and were also obtained through individual contacts with the local authority. A proprietary program in the Python programming language was developed to retrieve data from the webpage, which used web scraping library tools. Once downloaded, the page content was preprocessed and reformatted. The required data were then combined and saved in Excel format, creating a structure necessary for the analysis. The list of fields describing the data initially consisted of 8 fields in 2013, gradually increasing to 36 in 2019, and reaching 53 positions in 2023. This structure was unified (mapped to a common pattern), followed by data preparation carried out by selecting fields that define the features (variables) used in the analysis directly (inherited from the data set) or indirectly (used in the calculation formulas). Their brief characteristics are provided below:

- *Year* – inherited, specifies the year in which the project was uploaded and processed.
- *Name* – inherited, title of the project submitted to KPB,
- *Type* – calculated, classifies the project into one of four categories, with this classification appearing in different years (see the "Kielce participatory budgeting" section):
  - *Large*,
  - *Small*,
  - *Non-investment*,
  - *Green*.
- *Cost* – calculated, determines the cost of implementing the investment in EUR.
- *VtngAccp* – inherited, informs about the correctness of the project and its further processing; has the *Yes* value when the project was approved for voting (met the requirements), otherwise the variable has the *No* value,
- *RlztAccp* – inherited, has the *Yes* value when the project was approved for implementation, otherwise the variable has the *No* value. A project not approved for voting cannot be executed (implemented).

3.2. The methodology

Two distinct approaches to research and data evaluation were applied in this work: quantitative analysis and qualitative analysis. Each focuses on different aspects of information, and when used together, they complement each other and can provide a comprehensive understanding of the subject.

Quantitative analysis used to obtain the quantitative characteristics of PB projects involves a combination of basic counting, percentage calculation, aggregation, and statistical visualization techniques. The changes in nominal values and structure indicators for the total numbers and values of implemented projects by project type in consecutive years of the analysis period are presented. Additionally, for each project type, box plots are presented for consecutive years to illustrate the distribution of total project costs, enabling comparisons and the identification of fund allocation during the specified period. Thus, a well-rounded picture of the participatory budgeting process over time was obtained, allowing stakeholders to assess project distribution, costs, and tendencies.

In cases where qualitative analysis is used, textual data is processed using Natural Language Processing techniques, including the Bag-of-Words model and Word Cloud method. NLP is a machine learning technology designed for analyzing text data (text mining). It allows computers to process ("understand") human language, written or spoken alike, by performing a dedicated type of linguistic investigation on narrative or descriptive data. It uses various methods to organize unstructured text data to prepare them for computer algorithms. A corpus is an essential resource for NLP. It is a substantial, organized collection of data that often includes a wide range of documents, texts, or voices in one or more specific languages. NLP was used to analyze descriptive information on participatory budgeting projects.

Natural Language Processing software requires a consistent knowledge base such as a detailed thesaurus, a lexicon of words, a data set for linguistic and grammatical rules, or ontology, in order to work. In the KPB analysis, two such elements support NLP: the set of unnecessary words and the thesaurus.

There are certain common words that typically carry little information in the analyzed text, either because they occur too often in various contexts or because they belong to a category called function words. These uninformative, unnecessary words, which do not add substance, are called stop words and their collection is called a stop list. Examples are shown in Table 1. Computer programs should be told not to include stop list items in NLP results. Using a thesaurus improves text analysis. A thesaurus is a resource containing synonyms and groups of words assigned to them based on similarity. A synonym is a word or several-word term equivalent to another or close enough to replace the latter in an appropriate context. If a thesaurus is provided to NLP algorithms, it usually has a two-column structure, as shown in Table 1.

Table 1. **Examples of dictionary sources for NLP algorithms**

Stop list	Thesaurus	
	synonym	word (to be replaced)
able	animals	animal
conscious	animals	animals
early	animals	cat
iconic	animals	cats
let	animals	dog
near	animals	dogs
without	animals	kocice (Polish word for catkin)

Source: authors' elaboration

Bag of Words (BoW) is a fundamental Natural Language Processing technique that allows text data to be represented in a numerical format. In the Bag-of-Words model, an analyzed text (a document or a collection of documents) is represented by the frequency of each word. Thus, the representation of text disregards word order but captures multiplicity. Bag-of-Words model is a simplified way of representing text data. The interpretation focuses on which words appear most often in the text under consideration. Higher frequency suggests greater importance of that word to the document's content. A graphical representation of word occurrence providing better visibility to words that appear more frequently in a source text is the Word Cloud (or the Cloud of Words – CoW). The larger the word in the visualization, the more often it appears in the text. This visual representation helps highlight key themes or concepts within the text, making it easy to spot frequently discussed ideas or terms at a glance. Words with a higher visual prominence suggest topics that are central or of greater importance to the subject matter. Word clouds usually look better with top 20–100 words or phrases and enable to correctly portray the content.

Descriptive data on participatory budgets are included in the *Name* field presenting a general concept of proposed investments. The set of names was split into four groups, representing the corpus of titles of winning projects separately for each type: *Large*, *Small*, *Non-investment*, and *Green*. In the study, the corpus structure is simple (one column) as serving a pilot study on participatory budgeting in the NLP aspect. MS Excel, specialized Python tools, and the MS Access database management system were used to analyze the data and to obtain four final Bags of Words and Clouds of Words for the four KPB project types.

It is more effective to apply NLP algorithms to English than to Polish. The latter is characterized by fusional morphology, which makes it difficult to analyze texts written in that language. English is considered a weakly inflected language that has moved to

a more analytical structure. Therefore, NLP tools, including those in Python, are better developed for it. This fact was considered by translating the project names into English, and then performing text mining on the resulting translation. The workflow of the research is described by the points listed below. The IT tools used are also provided at the end of each point.

A. Data restructuring. Preparation of the scrapped data to fit them to the structure as follows: *Name*, *Type*, and *RlztAccp* = *True*. MS Excel.
B. Preprocessing Polish names. (1) Exclusion of texts in parentheses from the titles. (2) Conversion of Polish common abbreviations into full words (refers to: *street*, *square*, *avenue*). Python.
C. Data splitting. Dividing the dataset into four groups by project type: *Large*, *Small*, *Non-investment*, *Green*. Creating four corpora. MS Excel.
D. For each corpus:
    D.1. Translation to English. Automatic translation of individual titles into English – applying Machine Translation (MT) in an iterative cycle (title by title) resulting in the *Name_EN* field. Python.
    D.2. Translation check. Verification of 10% of randomly chosen English titles in each corpus on the correctness of translation. Inappropriate translation was corrected. Python and manual work in MS Excel.
    D.3. First text processing. Preparation of a preliminary BoW from the *Name_EN* field. Elimination of irrelevant elements such as punctuation marks and colloquial English words from the BoW, based on the standard stop list provided by Python's *nltk* library. Python.
    D.4. Creation of a non-standard stop list. The list consists of present in the corpus: words containing numbers (usually dimensions given with a unit of measurement or separate numbers), proper names, other irrelevant terms not present in the standard stop list (e.g., *almost*, *also*, *near*, *next*, *left*), special characters, and strings of black characters meaning nothing (e.g., *oi*, *rev*, *xxi*), the city name *Kielce*, and the word *city*. Manual work in MS Excel.
    D.5. Second text processing. Modification of the preliminary BoW by elimination of non-standard stop words, based on the non-standard stop list. Python.
    D.6. Synonyms elaboration. Preparation of synonyms for words from the BoW obtained in step D.5. MS Access and manual work.
E. Creation of thesauruses. Merging synonyms obtained for each corpus and unifying them. After unification, dividing the set of synonyms into subsets for projects of each type, delivering project type-oriented thesaurus. MS Access.
F. For each project type-oriented BoW:
    F.1. Applying the respective thesaurus, resulting in final BoW. Python.
    F.2. Creation of CoW from the final BoW. Python.

4. Results and Discussion

4.1. The analysis of quantitative KPB data

To provide an overview of the participatory budgeting process in Kielce, the following analysis presents key quantitative information on the projects submitted by city residents over the 2013-2023 period.

Figure 2 shows basic information on PB projects submitted by Kielce city residents in subsequent years in the 2013-2023 period. The centers of the bubbles indicate the total number of projects (see the vertical axis of the graph), and their radiuses illustrate the percentages of projects that passed formal and substantive verification.

The number of applications submitted has fluctuated over the years. In the first half of the period, the initial activity of applicants is followed by a decline; starting with a value of 67 in 2013, it reached a peak of 147 in 2015, and then, in 2017, fell to the lowest level of 49 in the entire period under study in 2017. From 2018 to 2021, the number of applications increased by about 50 each year. It seems that the reason for this optimistic change was the emergence of the possibility of applying for *Non-investment* projects. The peak value (227) occurred in 2021 when the next project type, *Green*, appeared, to take on a level of about 180 from 2022. The number of projects submitted in the second half of the period was on average higher than in the first half.

The percentage of rejected applications was zero in 2013, when the possibility of submitting projects under PB was introduced for the first time, and very low (only 5%) in 2019, when the possibility of submitting applications for *Non-investment* projects was introduced for the first time. In this case, the novelty effect could have implied a high level of people's involvement and their care to ensure that applications submitted for projects of new types were correct in terms of form and content. In the 2014-2016 period, the percentage of accepted projects was on average lower (it oscillated around 60%) than in all other years (the statistic was not lower than 72%).

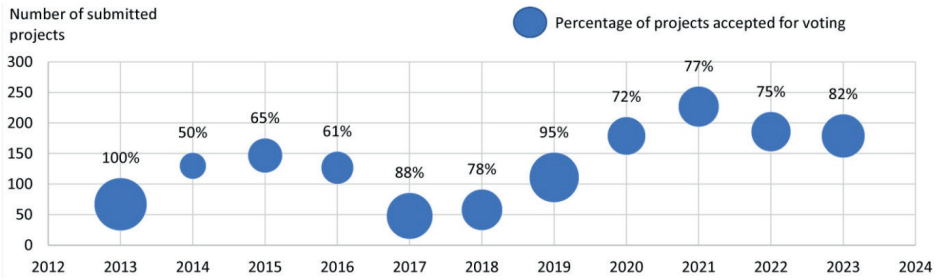


Figure 2. **Basic information on KPB projects**

Source: authors' elaboration

Figure 3 shows the number of projects accepted for voting (i.e., after positive formal and substantive verification) and projects intended for implementation (i.e., those that received consecutively the highest number of votes, and their total cost is within the amount of financial resources). During the 2014-2023 period, the variation profiles of the total number of applications submitted (Figure 2) and the number of applications voted on (Figure 3) are very similar. In 2013, there was only a *Large* investment project type, and all projects, three in number, received funding.

In almost every year (with the exception of 2017 and 2019), the number of projects designated for implementation was from nearly 3 to over 5 times smaller than those voted on. This disproportion indicates that the city's development needs, as perceived by residents, are relatively large, but their implementation would exceed the financial capacity established for PBs.

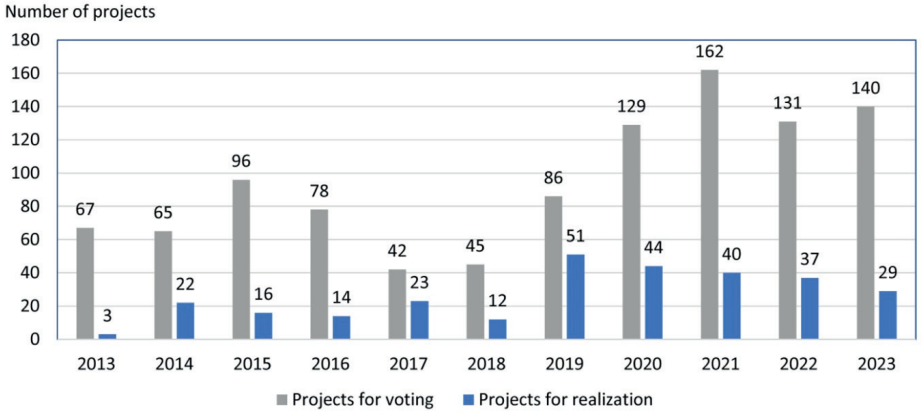


Figure 3. **Number of KPB projects voted and implemented**

Source: authors' elaboration

In Figure 4, the percentage of projects of each type in the total number of projects to be implemented (bars) and their total costs (lines) in consecutive years is illustrated. When there were only two types of projects: investment *Large* and investment *Small*, the latter were significantly more numerous; their percentage ranged from 83% up to 95%. Unlike small ones, large-scale ventures require significant expenditures, which in most cases explains the difference. *Non-investment* projects are noticeably popular, which means that residents appreciate this type of urban development. At the same time, the percentage of investment projects of both types is decreasing. This could be due to the fact that some of the needs of this type were carried out in previous years,

the pool of funds intended for them diminished, or the unit cost of the project was too high. The percentage increase in the number of *Green* applications is small but gradual, likely indicating the beginning of an increasing trend implied by climate change, the perception of the problem by politicians and policymakers, and growing support for green initiatives by various communities.

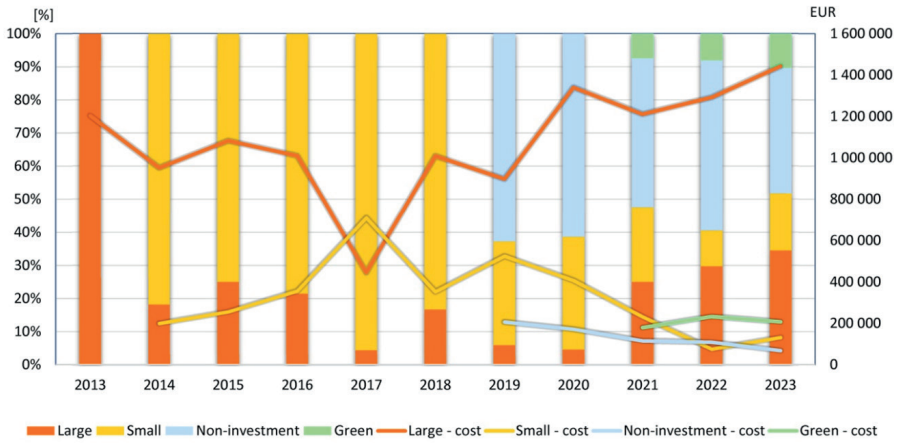


Figure 4. **Overall picture of KPB projects accepted for realization**

Source: authors' elaboration

Figure 5 illustrates the basic statistical measures of project costs by type. The horizontal axis tick labels indicate the year and, next to it, the number of projects to be executed.

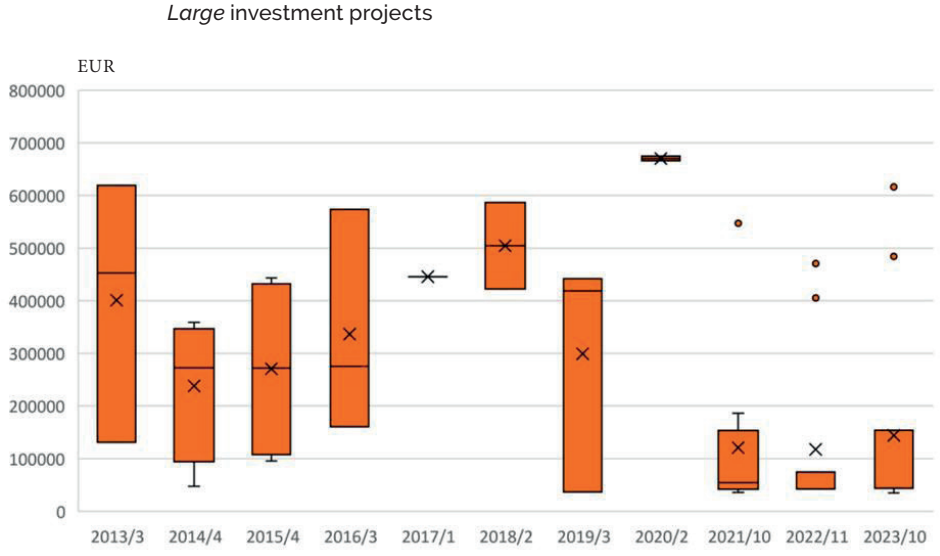
Until 2020, the number of *Large* investment projects implemented was small – no more than 4 per year. The amounts allocated to individual projects during this period ranged from around EUR 40000 (2019) to almost EUR 700000 (2020). The last three years are characterized by an increase in the number of ongoing projects; at least 10. Three quarters of them did not exceed EUR 160000. Five high-cost projects appeared in this group (see also Figure 4), singled out as outliers.

Unlike *Large* investment projects, by 2020 the number of *Small* projects underway was relatively significant; at least 10 per year. The amounts allocated to individual projects in this period ranged from approximately EUR 2500 (2014) to just over EUR 35000 (2018). In most cases, the average (cross (×) mark on the graphs) is smaller

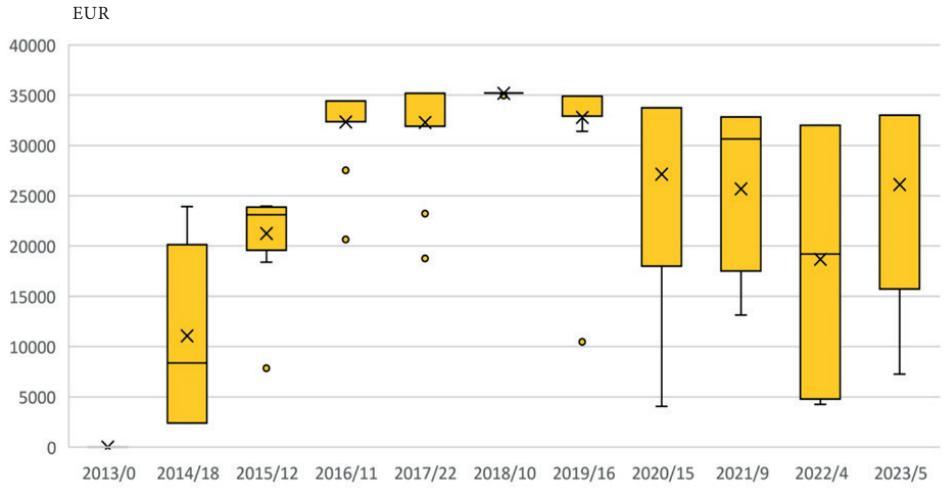
than the median (dash (–) mark on the graphs), which means that there were relatively many high-cost projects and few low-cost ones. There are six projects that were clear outliers with lower costs than the group of other projects in the respective years. The last three years are characterized by a decrease in the number of projects implemented; 9, 4 and 5 respectively.

The costs of half of the *Non-investment* projects are close to the allowable upper limit, i.e., close to EUR 7000, and heavily clustered around the central tendency (small box plots). The remaining half of the cost is characterized by a wide dispersion of values, ranging from just over EUR 1000 to around EUR 6500. A significant proportion of these projects can improve the quality of life of residents with relatively little financial outlay.

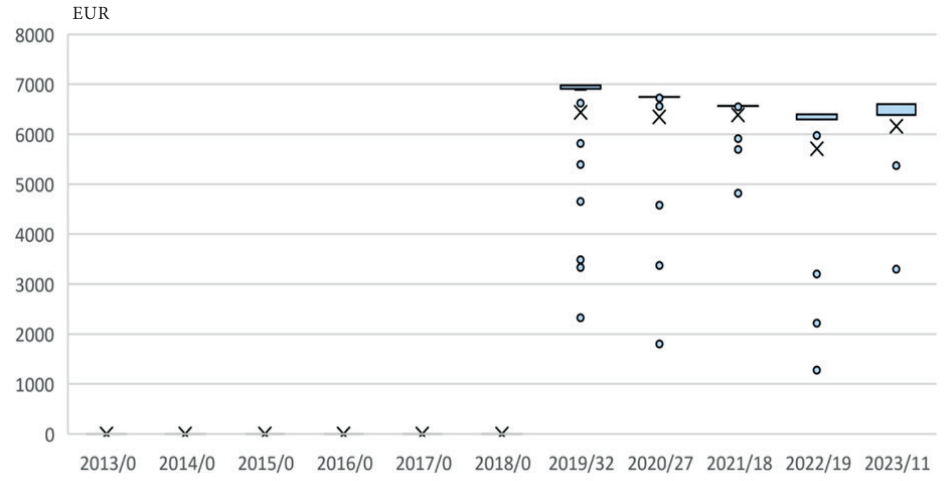
Three *Green* projects per year were implemented in the period 2021-2023. The cost of each exceeded EUR 17500. Impressive sums of EUR 141000 and EUR 126000 were allocated to two projects: both involved planting a total of nearly 1100 trees in designated areas of the city. It is possible that *Green* projects have affected the picture of both *Large* and *Small* investment projects (see Figure 4).



Small investment projects



Non-investment projects



Green projects

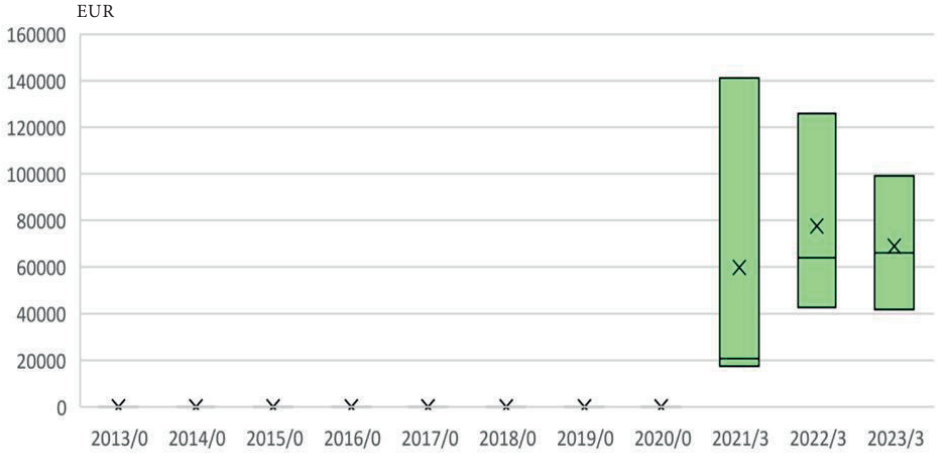


Figure 5. **KPB projects costs by type and year**

Source: authors' elaboration

4.2. The analysis of qualitative KPB data

Text data found in the English titles of projects were initially prepared to create the Bag of Words and Word Cloud for each project type. Table 2 provides a summary of this preprocessing. Original BoW is the set of words extracted from the English project titles, after applying standard stop list – the result of step D.3. Non-standard stop list is the set of words extracted manually from the original BoW (step D.4).

Table 2. **Statistics of Bag of Words of participatory budgeting by project type**

Project type	Number of projects	Original BoW size	Non-standard stop list size	Number of applied synonyms	Final BoW size	Number of one-occurrence words	Maximum word frequency in final BoW	Minimum word frequency in CoW	CoW size
Large	53	223	55	69	141	89	32 (sports)	**2**/3/4/5	**52**/31/21/16
Small	122	484	161	137	247	123	41 (street)	2/3/**4**/5	124/71/**51**/36
Non-investment	107	305	54	114	191	107	20 (entertainment)	2/**3**/4	84/**59**/40
Green	9	47	14	17	25	16	7 (greenery)	**1**/2	**25**/8

Source: authors' elaboration

The number of applied synonyms refers to those synonyms that were matched to words belonging to projects of a specific type (step E). Final BoW is the set of words of specific type projects obtained by deleting non-standard stop words (step D.5) and applying synonyms (step F.1) – this is the Bag of Words subject to the word cloud creation (step F).

The size of final bags of words varies from 25 for *Green* projects to *247* for *Small* projects. In each case, 50% or more are words that occurred once. In order to limit the CoW size, the word frequency method was applied. Different frequency values were tested, resulting in different word cloud sizes. This is presented by the numbers separated by a slash in the two last columns of the table. In the end, it was decided that the number of words in each word cloud should be as close as possible to the value of 50, which implicates the minimum word frequency – in the table, bold font indicates the choice. Figure 6 pictures the importance of words in the final word clouds for projects of each type. The Word Clouds illustrate citizens' preferences for improving their urban surroundings, enhance sustainability and enrich both social and community life in the city.

In the *Large* project corpus, the most important word is *sports*, which appears 32 times. As a result of the adaptation of synonyms in the NLP processing, it contains, in addition to the original *sports* term (16 occurrences), eleven other words, such as: *volleyball*, *sprint*, *skatepark*, *handball*, *gym*, or *football*. Next, in order of importance are *construction* (21 occurrences), *school* (20 occurrences), *complex* (16 occurrences), and

street (14 occurrences). Two general concepts emerge from the whole CoW picture of the *Large* projects. Both are connected with big, complex investments, relatively expensive (exceeding the cost of PLN 150000; in EUR 32000-36000, depending on the exchange rate). One of them is the construction of sports facilities, focused mainly on the needs of young people (additional words supporting the concept: *kindergarten*, *children*, *primary*, *education*); these can be sports fields at schools. The second one advocates the construction or modernization of communication routes in the city, both for pedestrian and vehicular traffic (additional words supporting the concept: *track*, *lane*, *avenue*, *roads*).

In the *Small* project corpus, the most important word is *street*, which appears 41 times. Next, in order of importance are: *construction* (38 occurrences), *school* and *sports* (32 occurrences each), *playground* (30 occurrences), *recreation* (20 occurrences). The model represents smaller construction investments, including modernization, not exceeding the cost of PLN 150,000, focusing mainly on streets. Like the *Large* ones, the *Small* participatory projects include sports facilities. These are proposals for investments dedicated to the city's youngest residents and their parents (words: *children*, *playground*, *recreation*, *kindergarten*), not significantly noted in the *Large* projects.

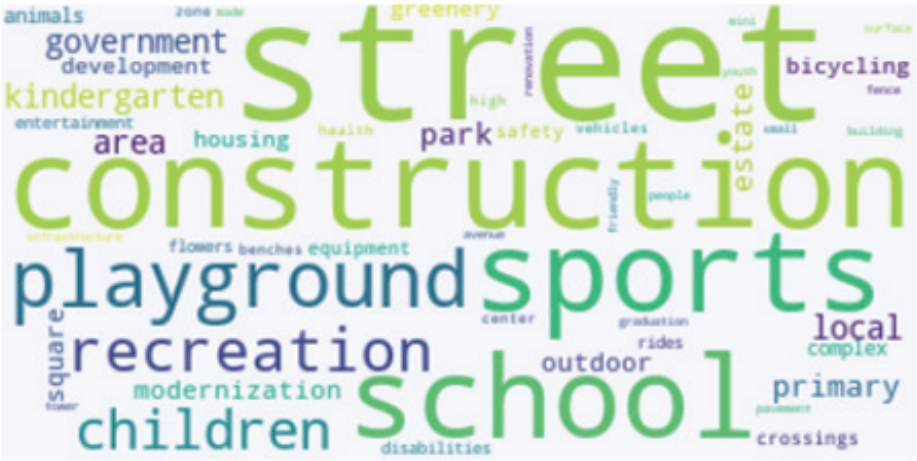
In the corpus of *Non-investment* projects, the most important words are *entertainment* and, almost as important, *festival*, which appear 20 and 19 times, respectively. Next in order of importance are *animals* (a word covering 7 terms), *sports* (4 terms) and *trees* (3 terms), each of which appears 15 times. *Non-investment* projects focus on events related to mass entertainment and family integration (*entertainment*, *festival*, *picnic*, *cinema*, *family*, and *children*). Animal-related projects are often undertaken to reduce the animal population (*castration*); cats are their majority object, most likely left unattended.

In the *Green* project corpus, the most important word is *greenery*, with 7 occurrences. The term, along with the words *trees*, *shrubs*, and *flowers* on the one hand and *street* and *concrete* on the other, indicates the need to green the city and mitigate the effects of concrete overuse leading to a lack of green spaces and natural environments.

It is worth mentioning that the *bicycling* aspect is present in all but *Green* projects. This shows that the use of a bicycle for sports, recreation, or transportation is becoming increasingly important to residents, which can be linked to an ecological approach and a desire to improve the quality of local life.

Large, minimum word frequency = 2

Small, minimum word frequency = 4



Non-investment, minimum word frequency = 3

Green, minimum word frequency = 1

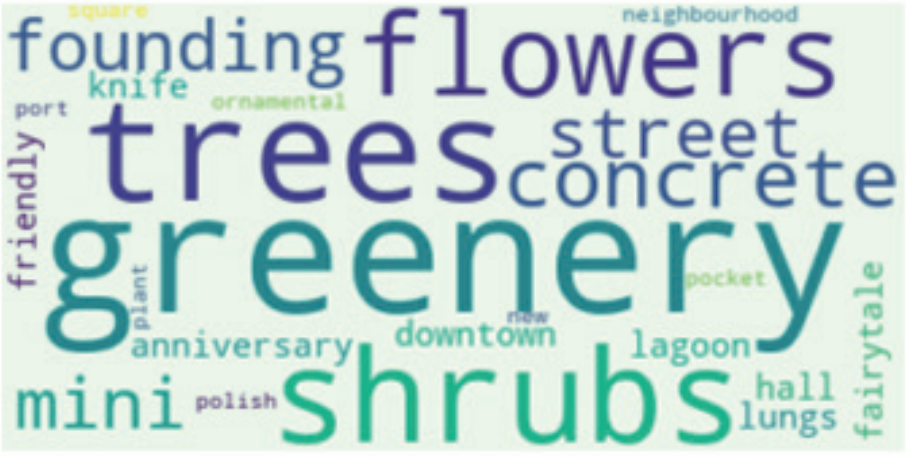


Figure 6. **Clouds of words by participatory budget type**

Source: authors' elaboration

5. Conclusions

The history of participatory budgeting in Kielce is also the history of the city's local civic participation. As in other cities, it is proof that residents can have a real impact on shaping the space and life in their surroundings, and their involvement translates into specific, positive changes. Each city defines its own participatory budget regulations, which typically outline the types of civic projects eligible for funding. At the time of writing this article, the following types of projects are distinguished in Kielce: *Large* and *Small* investment projects, *Non-investment* projects, and *Green* projects. Participatory budgeting is a tool that allows the residents of Kielce not only to implement their own ideas, but also gives them a sense of shared responsibility for the development of the local community. The paper discusses issues that allow creating a picture of participatory budgeting, which may contribute to its better understanding, further development, and improvement. Local authorities receive a barometer of democratic involvement and community preferences in the context of urban development, which can become an inspiration when planning the city budget.

The following depiction of participatory budgeting in Kielce can be presented from an analytical perspective employed in the study.

The number of submitted KPB projects (including erroneous ones) varies over the period under consideration (2013-2023); it fluctuates between 0.2 and 0.8 per 1000 inhabitants. In the second half of the period under review, the percentage of incorrect applications remains relatively stable, at about 20%-30% and is generally smaller than in the preceding years. It seems that the introduction of new project types led to higher involvement and more careful submissions. Surprisingly few projects were forwarded for implementation: leaving aside the first year (2013), there were on average four times fewer projects than all those accepted for voting. Such a large disproportion may indicate that the city's financial offer for PB is too modest, and the development needs exceed the available funding. Over time, the number of *Large* investment projects decreases, but their costs increase, while involvement in *Small* projects, initially considerable, drops from the middle of the period. *Non-investment* projects gain popularity due to their lower costs and potential to improve residents' quality of life. *Green* projects gradually increase, reflecting support for environmental initiatives.

Kielce residents' priorities for urban development focus on improving both the physical infrastructure and social aspects of the city. For *Large* investment projects, there is a strong emphasis on sports facilities and communication routes, particularly for young people and pedestrians. *Small* investment projects prioritize expenditure on the construction or reconstruction of streets, playgrounds, and recreational spaces, with a focus on enhancing life for families and children. Proposals for soft projects – *Non-investment* and *Green* ones – are playing an increasing role. *Non-investment* projects, integrating local communities, are centered around entertainment, festivals,

and family-oriented events, while *Green* projects, emphasizing the importance of connection with nature, highlight the need for more greenery to mitigate overuse or excessive spread of concrete in urban environments. The growing interest in bicycling reflects a pursuit of sustainable urban transportation and apparently a need for a deeper connection with the urban natural environment and an improvement in physical well-being.

Machine learning, specifically Natural Language Processing, can be used to analyze participatory budgeting project narratives by extracting and processing textual data, such as project titles. NLP should apply synonym matching to broaden the analysis, clustering similar terms for a more comprehensive understanding. Techniques like Bag of Words, applied to PB text corpora by project type, help identify frequently occurring terms in each, enabling the filtering of less significant words. Word Clouds provide an efficient way to identify and highlight the most important terms, making it easier to grasp key themes in *Large*, *Small*, *Non-investment*, and *Green* projects.

The authors are aware of some limitations of the study, mainly due to its pilot nature. The results come from quantitative and qualitative analyses of cases located in a single city, so their interpretation requires an awareness of the environment of local democracy. Consequently, any generalizations beyond the studied context should be treated with caution. Nevertheless, the research has set in motion areas for future investigations, as outlined below:

1. The Bag-of-Words model ignores the order of words in a text document and their meanings, which means that it loses the context and semantic information. Thus, considering NLP with n-gram approach, aiming to picture the context, seems reasonable. Since the representation of a bag of bigrams has much greater impact than a Bag of Words and in many cases proves to be more successful (Dang and Luk, 2014), conducting such an analysis for participatory projects looks promising, especially when it comes to analyzing not just the titles themselves, but entire project descriptions.
2. A promising approach is to investigate weighting terms according to assumed criteria to illustrate its importance in a specified context. This can be developed along with the sentiment analysis (considering the basis of assessing the "feeling" about proposed investments).
3. Synonyms are not recognized unless a proper thesaurus is developed and implemented. The same applies to the content of the stop list. Therefore, it would be a good idea to elaborate the PB-oriented thesaurus and the PB-oriented non-standard stop list so that they would be employed in a variety of text analyses in that area.
4. Expanding the analysis by considering the number of votes, voters' gender and age structure, geographic areas of initiatives, projects that passed the verification, but did not fit into the pool of tasks for implementation can be helpful in establishing

a barometer for the city authorities for investment works, not only within the participatory budgeting.
5. So far recognized and then developed methods of PB analyses, especially in NLP context, can be applied to other Polish cities of similar size to Kielce and to perform comparative analyses of the results obtained.

Authors' contribution

M.Z.: research methods applied, data collection, analysis and interpretation of results, draft manuscript preparation. **M.N.:** article conception, theoretical content of the article, research methods applied, conducting the research, analysis and interpretation of results, draft manuscript preparation

Acknowledgements

This research was supported by: Kielce University of Technology, Faculty of Management and Computer Modelling, Kielce, Poland, grant number 04.0.06.00/1.02.001 SUBB. ZKIS.25.001.

References

Bartocci, L., Grossi, G., Mauro, S.G., & Ebdon, C. (2023). The journey of participatory budgeting: a systematic literature review and future research directions. *International Review of Administrative Sciences*, *89*(3), 757–774. https://doi.org/10.1177/00208523221078938

Bogo, R.S., & Falanga, R. (2024). The cost of participation: An analysis of the financial dimensions of participatory budgets in Portugal. *European Policy Analysis*, *10*, 279–299. https://doi.org/10.1002/epa2.1200

Dang, E.K.F., Luk, R.W.P., & Allan, J. (2014). Beyond Bag-of-Words: Bigram-Enhanced Context-Dependent Term Weights. *Journal of the Association for Information Science and Technology*, *65*(6), 1134–1148. https://doi.org/10.1002/asi.23024.

Davies, J., Arana-Catania, M., Procter, R., van Lier, F.-A., & He, Y. (2021a). A mixed-methods ethnographic approach to participatory budgeting in Scotland. *Proceedings of the Conference on Information Technology for Social Good*, 317–320. https://doi.org/10.1145/3462203.3475891

Davies, J., Arana-Catania, M., Procter, R., van Lier, F.A., & He, Y. (2021b). Evaluating the application of NLP tools in mainstream participatory budgeting processes in Scotland. *Proceedings of the 14th International Conference on Theory and Practice of Electronic Governance*, 362–366. https://doi.org/10.1145/3494193.3494242

Dubicki, P. (2021). Budżet Obywatelski jako element partycypacji społecznej [The civic budget as an element of social participation].*Studia z Polityki Publicznej*, *8*(1(29)), 91–105. https://doi.org/10.33119/KSzPP/2021.1.5

Falanga, R. (2024). Youth participation in environmental sustainability: insights from the Lisbon participatory budget. *Journal of Public Budgeting, Accounting & Financial Management*, *36*(1), 20-39. https://doi.org/10.1108/JPBAFM-12-2021-0176

Faliszewski, P., Flis, J., Peters, D., Pierczyński, G., Skowron, P., Stolicki, D., Szufa, S., & Talmon, N. (2023). Participatory Budgeting: Data, Tools, and Analysis. *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence* (IJCAI-23), 2667–2674.

Kempa, J., & Kozłowski, A.R. (2020). Participatory Budget as a Tool Supporting the Development of Civil Society in Poland. *NISPAcee Journal of Public Administration and Policy*, *13*(1), 61–79. https://doi.org/10.2478/nispa-2020-0003

Klimovský, D., Secinaro, S., Baláž, B., & Brescia, V. (2024). Participatory Budgeting as a Democratic and Managerial Innovation: Recent Trends and Avenues for Further Research. *Central European Journal of Public Policy*, *18*(1), 2667–2674. https://doi.org/10.2478/cejpp-2024-0004

Kociuba, D., & Rabczewska, K. (2019). Rola budżetów partycypacyjnych w zagospodarowaniu przestrzeni publicznych polskich miast – studium przypadku Lublina [The role of participatory budgets on the development of urban public spaces - the Lublin case study]. *Studia Regionalne i Lokalne*, *2*(76), 82–109. https://doi.org/10.7366/1509499527605

Milosavljević, M., Spasenić, Ž., & Krivokapić, J. (2023). Bibliometric Review of Participatory Budgeting: Current Status and Future Research Agenda. *International Journal of Financial Studies*, *11*(3), 14. https://doi.org/10.3390/ijfs11030104

Narodowy Bank Polski (b.d.). Historic average exchange rates – table A. https://nbp.pl/statystyka-i-sprawozdawczosc/kursy/archiwum-tabela-a-csv-xls/

Peláez, J., Casado, A.M., Freddy, D., & Alcibar, Y. (2016). A Fuzzy Decision Model To E-Goverment with Mobile Technologies and Linguistic Information, *3rd International Conference on eDemocracy and eGovernment*, ICEDEG 2016, 145–150. https://doi.org/10.1109/ICEDEG.2016.7461712

Šabović, M.T., Milosavljević, M., & Benkovic, S. (2021). Participation of Citizens in Public Financial Decision-Making in Serbia. *Slovak Journal of Political Sciences, 21*(2), 209–229. https://doi.org/10.34135/sjps.210205

Segura-Tinoco, A., Holgado-Sánchez, A., Cantador, I., Cortés-Cediel, M.E., & Rodríguez Bolívar M.P. (2022). A Conversational Agent for Argument-driven E-participation. In L. Hagen, M. Solvak, S. Hwang (Eds.), *Proceedings of the 23rd Annual International Conference on Digital Government Research (dg.o '22). Association for Computing Machinery, New York*, 191–205. https://doi.org/10.1145/3543434.3543447

Soboń, A., (2017). Budżet obywatelski narzędzie rozwoju lokalnego. [Participatory budgeting as a tool of local development]. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, *316*, 172–182.

Sousa, Santos de B. (1998). Participatory budgeting in Porto Allegre: Toward a redistributive democracy. *Politics and Society*, *26*(4), 461–510. https://doi.org/10.1177/0032329298026004003

Svidroňová, M.M., Baláž, M.B., Klimovský, D., & Kaščáková, A. (2024). Determinants of sustainability of participatory budgeting: Slovak perspective. *Journal of Public Budgeting, Accounting & Financial Management*, *36*(1), 60-80. https://doi.org/10.1108/JPBAFM-04-2022-0082

Tatarowska, E., & Furmankiewicz, M. (2018). „Wrocławski Budżet Obywatelski 2014" jako przestrzenna forma identyfikacji potrzeb inwestycyjnych w mieście. ["Wroclaw Civil Budget 2014" as a spatial form of in- vestment needs identification in a city]. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, *504*, 120–134. https://doi.org/10.15611/pn.2018.504.12