ESTIDAMAA
Vol. (2025), 2025, pp. 21-28
ISSN: 3078-428X

Research Article

Natural Language Processing-Driven Communication for Improved Citizen Alignment in Smart Cities

Weiwei Jiang[1,*], Mohammed Almaayah[2], Rami Shehab[3]

[1] Beijing University of Posts and Telecommunications, Beijing, China

[2] Fellowship Researcher, INTI International University, Nilai 71800, Malaysia

[3] Vice-Presidency for Postgraduate Studies and Scientific Research, King Faisal University, Al-Ahsa 31982, Saudi Arabia

ARTICLE INFO

Article History
Received 2 Nov 2024
Revised: 20 Dec 2024
Accepted 20 Jan 2025
Published 7 Feb 2025

Keywords
Citizen Alignment,
Natural Language,
Smart City,
Internet Search,
Internet Quality of Service.

ABSTRACT

This study aims to explore the integration of smart buildings (SBs) into smart cities (SCs) and assess their potential to improve urban performance. It uses AI tools, such as OpenAI's ChatGPT and Google Bard, to evaluate 26 SB service criteria across five SC domains: energy, mobility, water, and waste management. The study employs an AI-driven methodology to quantify the cost-effectiveness, resilience, and sustainability of SCs. The framework involves five processes: (1) scraping smart city announcements from the web, (2) typescript pre-processing, (3) using a latent Dirichlet allocation model for strategic topics, (4) measuring intertopic similarities with Hellinger distance, and (5) assessing communication trends in real-world urban environments. The study deconstructs top-down discourses of smart cities to uncover political programs aimed at aligning citizens and improving urban outcomes. Insights gained from the analysis reveal the impacts of SB services on SC domains, identifying areas for further integration and enhancement. This study presents a novel AI-driven approach to evaluate the integration of SBs into SCs, providing a comprehensive assessment of their performance. The findings offer valuable insights into the potential benefits and challenges of incorporating smart buildings into urban environments, highlighting opportunities for improved cost-effectiveness, resilience, and sustainability.

1. INTRODUCTION

A smart city is an innovation-based urban system based on ICT. The goal is to use ICT to improve urban services that are being caused by urbanization without properly establishing policies regarding well-being. In a smart city, citizens' quality of life is improved. To successfully manage complex government-to-government cooperation processes, in particular with citizens, the Author [1] suggests that a new model of governance is needed. Developing a new governance model involves government, citizens, and other social actors rethinking their roles, developing new processes, and developing new government structures [2]. Urbanization has become one of the most prominent global developments in recent times, with over half of the world's population now residing in cities [3], [4]. Managing urban environments presents numerous challenges such as traffic congestion, environmental pollution, public safety concerns, and social disparities [5]. Sustainable urban development also introduces critical concerns, including the consumption of natural resources, climate-related risks, and the need for social equity and inclusion. Addressing these complex, cross-disciplinary issues requires innovative approaches in urban policy and governance research [6], [7], [8], [9]. In this context, urban sustainability has emerged as a vital and evolving domain within the broader field of urban governance studies [3], [10].

Several emerging technologies, especially smart and artificial intelligence, offer promising solutions to these challenges [11], [12]. In smart city governance, cutting-edge technologies such as Big Data and AI are used to develop innovative management methods, management models, and management concepts in urban management[13], [14]. It is not without challenges, however, to implement smart city governance [15], [16]. Using AI for urban planning poses many ethical and practical challenges, including ensuring that algorithms are transparent, fair, andaccurate [13]. People may also discover more specific risks associated with smart cities as they explore their boundaries. Due to these deficiencies, we looked for ways to complement and improve this manual, bureaucratic effort that requires a large amount of human intervention and

*Corresponding author email: w.jiang83@gmail.com
DOI: https://doi.org/10.70470/ESTIDAMAA/2025/003

automate parts of the current system using Artificial Intelligence (AI). Specifically, we examine the role opinion mining plays in decision-making by applying Natural Language Processing (NLP). The goal is ultimately to be able to extract residents' current viewpoints from social media and improve a smart city's decision-making process by incorporating them into the tools provided. In addition to Calgary, the system developed in this research can be applied to other cities.

This study introduces a novel method for consistently assigning land use categories to spatial development plans by analyzing their textual descriptions. ML methods are used for classifying and clustering text using natural language processing (NLP) [17]. The currently used methods of harmonizing spatial planning data require manual mapping between land use categories. Mapping occurs when plans are harmonized, such as when they are reduced to a single conceptual framework. Spatial data is harmonized using ETL (Extract, Transform, Load). Land use is considered an attribute of the spatial object. An area's functions are classified according to a standard, e.g. a commune, and mapping is performed accordingly. There are difficulties arising when different land use classifications are applied to each plan. After that, multiple classifications need to be mapped to the source classification. Using NLP methods, we present our approach to integrate spatial development plans using well-known methods. The textual arrangements of many plans with different land use classifications can be automatically classified in this manner, allowing us to integrate them. There has not been a literature study that discusses how artificial intelligence algorithms can integrate planning documents, especially in the area of natural language processing.

2. RELATED WORK

In a smart city, infrastructure, attributes, and themes all come together. As well as attributes smart cities have characteristics as well. As part of the planning for a smart city, themes are often referred to as pillars since they contribute to the city's ongoing progress. An essential component of any smart city is its infrastructure, which facilitates its operation. In this section, we examine the features mentioned earlier from the perspective of a generic smart city deployment. Smart cities are multi-attribute, combining a number of factors. The majority of smart city proposals emphasize sustainability, urbanization, and smartness as the four main attributes [18]. There are a few sub attributes under each attribute. Sub attributes of sustainability include governance, pollution, energy, climate change, and social issues[19]. Efforts to keep ecosystems in balance while providing services and performing city operations are known as sustainability. There has been an improvement in urban citizens' emotional and financial well-being. It examines the transition from a rural environment to an urban environment from a technological, economic, infrastructure, and governing perspective. The desire to improve social, economic, and environmental standards is what defines smartness in a city.

There is abundant literature on monitoring social media data. Researchers have used event detection to detect citizen dissatisfaction with local authorities, identify hate speech, evaluate misinformation, and optimize energy conservation in the past. In other words, machine learning has been used, but its paradigms, particularly for real-time event detection, are not as effective as our combination of NLP techniques. A smart project will be identified within selected cities as the first stage of the research, followed by the identification of SC as the second stage. Smart cities and real estate projects were assessed at both levels of the selection criteria in order to identify areas for improvement and set attainable goals. It was determined that six factors and ten key characteristics would determine a smart city's real estate standards. In spite of this, the methodological framework's wide range of categories and indicators may make its implementation challenging. Research also suggests that the framework may not cover all aspects of intelligence and integration, including those at the urban and project levels.

To improve the human-building interface, a number of barriers need to be removed. Human comfort and preferences can be collected continuously and non-invasively using NLP research in the context of building occupancy. In addition to data feeds, this information contributes to FL control, ML and NN training data sets. This review highlights the connections between AI and smart building topics, as well as the importance of natural language processing for obtaining occupant data. Also, if a greater number of these data are available, it will greatly benefit additional research in this field. In spite of the lack of corresponding research into the field, the authors believe that NLP within Smart Buildings deserves a far greater level of attention since it has the potential to drastically transform how humans interact with buildings, as well as the potential to enable all other areas discussed in this paper.

Multimedia e-learning requires personal computers (desktops, laptops, and handhelds), platform-as-a-services (PaaS), infrastructure-as-a-services (IaaS), and software-as-a-services (SaaS). Several tools are used for distance learning, including e-tutoring, self-assessment, and communication apps, such as Zoom, chat apps, forums, and video call applications. Aspects of mobile e-learning include technological mobility, pervasive learning environments emphasizing social interaction within their activity spheres, and organizations' relationships with individuals. There is a merger between mobile learning (M-learning) and pervasive learning (P-learning) that results in ubiquitous learning (U-learning).

Education technology is characterized by the combination of information systems, computer hardware, communication networks, software, lifeware, and educational theory into electronic learning systems (E-learning). By formulating and implementing technological developments and implementing informative resources, educational technology improves individual intellectual performance. Mobile computing devices, personal computers, smartphones, tablets, and other devices are being used in e-learning programs to allow students to conduct their learning outside of a traditional classroom. Smart multimedia technologies have enabled e-learning education in the 21st century, including online, virtual, distance, and automated libraries. E-library, also known as virtual library, saves and makes multimedia digital files accessible over large computer networks, allowing users to access information from anywhere in the world.

3. PROPOSED METHODOLOGY

The framework illustrated in Figure 1 was designed to evaluate the prominence of strategic themes in top-down smart city governance, aiming to enhance alignment with citizens and analyze their interconnections. This approach helps in understanding how top-down governance influences urban performance and whether such communication strategies are effective and suitable for smart city development. Each case study follows a consistent five-step process: (1) collecting official smart city announcements via web scraping; (2) preprocessing the extracted textual data; (3) applying Latent Dirichlet Allocation (LDA) to identify thematic priorities; (4) assessing topic similarity using the Hellinger distance along with a comparative analysis; and (5) evaluating the effectiveness of top-down communication by comparing it with actual urban performance outcomes.

3.1 Top-down data Extraction

Several text datasets related to urban management were analyzed in this study, including those from Boston. The smart city strategies we chose instead of digital notices from government portals provide top-down messages that inform and align citizens with urban policy. Duplicate notices and notices with fewer than 40 words were eliminated. In addition, all notice tags have been retained, including Analytics Team, Arts and Culture, Budget, Environment, Food Access, and Public Safety. Although some announcements occasionally convey unrelated information, mass web scraping was conducted to create voluminous datasets that, when combined with topic modelling, enabled us to identify the major trends in smart city strategies despite their sporadically unrelated nature. The topic modelling stage could be used to eliminate non-smart city-specific content through unsupervised LDA.

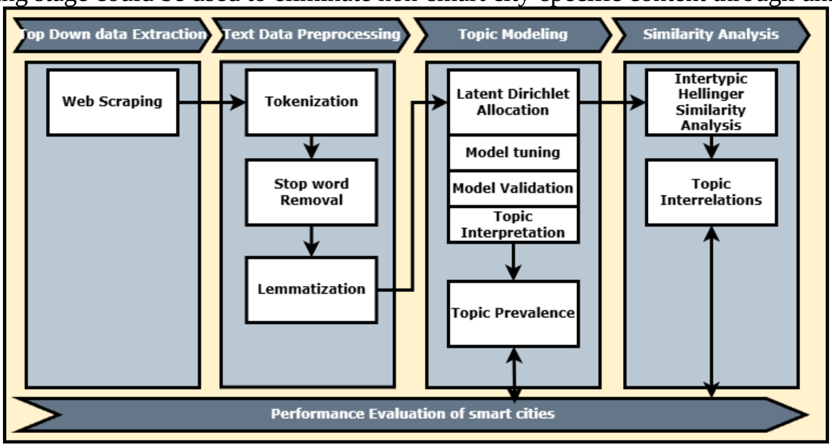


Fig.1. An overview of the research framework.

A smart city announcement and a smart initiative announcement would typically be retained. A smart city is often perceived as focusing on digital content rather than human-centred communication, which is ambiguous and subjective [21],[22]. Using Web scraping, you can collect and convert unstructured Web data, usually in HTML format, into structured data that can be analyzed locally. Common HTML templates are typically used to encode government announcements on websites for a given city. Through analysis of the HTML documents' structure, a Python Beautiful Soup program was developed to navigate and extract textual information from them.

3.2 Text data pre-processing

An efficient text analysis requires appropriate data preparation; it can eliminate unnecessary characters and inaccuracies that may occur during web data extraction. In text data preparation, stopwords are removed, bigrams are inserted, and lemmatization is performed. Text analysis units in NLP are tokens or strings of encoded bytes representing text. Tokenization is a method of separating individual words from punctuation marks in text data. In the end, words were lemmatized to determine their root forms. This pipeline is based on Python and executes these tasks sequentially. Two objects were created for each city using the Gensim Python NLP library: a dictionary for converting distinct words into unique numerical IDs and a list for mapping articles into "word_id" and "word_frequency" lists."

3.3 Topic modelling

Those who produce language produce unstructured data that is intended for other people to understand. There is a linguistic property to text data that makes it easy for others to understand and for computers to process [23]. The statistical topic model provides a powerful framework for extracting and summarizing text corpora contents, which is built using standard Latent Semantic Analysis (LSA) and is further generalized and extended. In contrast to traditional keyword matching methods, topic modelling automates content analysis and reduces text dimensions without affecting topic discovery. Topic modelling using LDA, as presented, is becoming increasingly popular in engineering and project management. From unlabeled documents, this method extrapolates the core topics using an unsupervised machine-learning technique. Documents, d, are treated as probability distributions, θ_d, ended a set of K topics, and topics, $k \in \{1, \dots, K\}$, are treated as probability distributions, φ_k, over keywords in the vocabulary. A matrix with rows representing documents and columns representing topics is shown in Figure 2. $\theta_{d.k}$ represents the probability of document d containing topic k. The structure of φ is similar, with rackets representing themes and pilasters representing arguments. Figure 3 illustrates the LDA process simplistically.

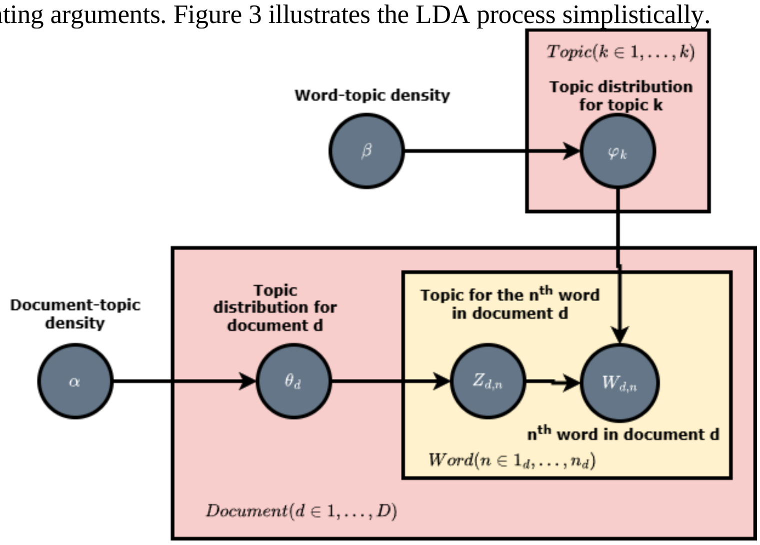


Fig. 2. The latent Dirichlet allocation method is represented graphically.

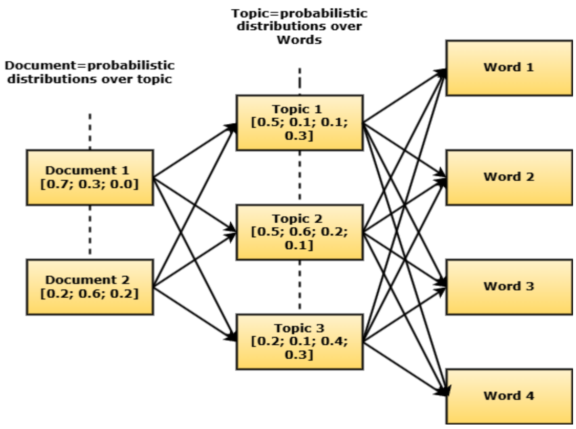


Fig. 3. The latent Dirichlet allocation process is represented in a simplified manner.

Topics, while observable variables in LDA, should be statistically inferred so that dimensions can be reduced. The LDA algorithm allows each document d is connected to each of its N constituent words through a thread, and threads are probabilistically clustered into topics based on the α and β parameters, as well as the Dirichlet Compound Multinomial (DCM) distribution [12]. Thus, $\varphi_{k,w}$ represents the chance of a given word w falling into topic k, while $\varphi_{k,w}$ signifies the chance of a topic k being contained in document d. By leveraging these distributions, the algorithm incorporates a latent layer of topics that links words to documents. As outlined in the next section, topic coherence metrics are employed to evaluate the statistical relevance of the extracted K topics.
The Latent Dirichlet Allocation (LDA) model was applied using the Gensim library in Python on the preprocessed textual data. A core topic extraction algorithm was implemented after implementation. A validation study has been

conducted on metrics used to determine the coherence of words within topics [12]. Therefore, these metrics are effective for assessing how well LDA models align with human judgments [24]. A C_{UMass} and C_V coherence indicator is considered in this study. $t, C_{UMass}(t; W^{(t)})$ In Eq (1), the pairwise score logarithmic function is averaged to measure $t, C_{UMass}(t; W^{(t)})$ internal coherence. A topic coherence score, such as CU, is based on the average of topic coherence scores across the topics in a topic model, as described in Eq (2).

$$C_{UMass}(t; W^{(t)}) = \frac{2}{N(N-1)} \sum_{i<j} log \frac{D(w_i^{(k)} + \epsilon)}{D(w_i^{(k)})} \quad (1)$$

$$C_{UMass} = \frac{1}{K} \sum_{k=1}^{K} C_{UMass}(k; W^{(k)}) \quad (2)$$

A document with a word $w_i; D(w_i; w_j)$ contains the wordsw_i and w_j; while a document with word $W^{(k)} = (w_1^{(k)}, \ldots, w_N^{(k)})$ contains the words most likely to be associated with the topic k [24]. Additionally, $\varepsilon = 1$ is used in Eq. (1) so that no logarithm is required. As a result of its logarithmic function, C_{UMass} resulting in negative values to 0 indicating topics that are more interpretable by humans. C_V was proposed as another method for measuring coherence. Four main components make up the framework: (1) word-pair segmentation of topics, (2) a Boolean sliding window is used to estimate the probability of words or word pairs, (3) NPMI and cosine similarity are then combined to calculate topic confirmation measures, (4) aggregating the final coherence into C_V by arithmetic mean [24], [25]. When coherence is high, topic models are more interpretable, and the coherence indicator returns a value between [0, 1]. While C_V requires a lot of computational effort, it outperforms all existing coherence measures, includingC_{UMass}, which still has the advantage of being faster

The optimal combination of hyperparameters (topic number, k, document-topic density,α, and word-topic density,β) was determined using grid search optimization, Initially, we computed coherence metrics C_V and C_{UMass}, based on models trained at $\alpha = 0.01$ and $\beta = 0.11$ values within [3; 26] for determining the optimal number of topics k. Due to the likelihood of C_V scores increase with more topics, and optimal K values are determined by figuring out at what point the benefit of increasing returns becomes uneconomical. C_{UMass} fluctuations were kept in mind when making these heuristic determinations. For each case study, LDA models were tuned to optimize coherence metrics. The outputs were visualized using a web-based Python package that verified their human interpretability; LDA topics were plotted on a two-dimensional graph using the (t-SNE dimensionality reduction method. According to Eq (3), terms w are relevant to their respective topics if λ is a weight parameter.

$$r(w, k|\lambda) = \lambda log\phi_{kw} + (1-\lambda) log \frac{\phi_{kw}}{p_w} \quad (3)$$

It represents the probability that the word w would occur in the corpus, where ϕ_{kw} is the probability that word w would occur for topic $k \in \{1, \ldots, K\}$. There is a challenge involved when interpreting topics when $\lambda = 1$, since, in this case, the most commonly occurring and insignificant words are ranked first according to their probability. The same holds when $\lambda = 0$ is listed, as these words are essentially topic-specific. It led to the decision to set λ to 0.3 to improve topic interpretation.

3.4 Similarity Analysis

To explore the relationships among the four categorized text corpora, we employed the Hellinger distance as a measure of similarity. Various metrics are available for comparing topic distributions produced by LDA, including the Kullback-Leibler (KL) divergence, the Jensen-Shannon divergence—a symmetric and smoothed adaptation of KL—and the Hellinger distance [26]. Among these, the Hellinger distance was selected due to its consistency when dealing with probability distributions that have different supports. As expressed in Equation (4), this metric quantifies the similarity between two distributions, p and q, making it suitable for analyzing topic interactions in this study.

$$HD(p, q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{T} \left(\sqrt{p_i} - \sqrt{q_i}\right)^2} \quad (4)$$

In Hellinger's distance, lower values indicate greater similarity, while higher values indicate greater differences. Hellinger distance is a useful method for analyzing LDA topics because topics are modelled as probability distributions over words. As a measure of similarity, the Hellinger distance between topics was subtracted from one in this study. For the four smart cities, intertopic Hellinger similarity matrices were computed to identify differences and complementarities. Seaborn Python's diagonal heatmaps were used to display the results.

4. EXPERIMENT AND RESULTS

We used a three-point scale for expert evaluations when artificial intelligence-generated ratings differed from expert evaluations: 0 (no impact), 1 (moderate impact), and 2 (significant impact). Smart city domains were evaluated according to a five-point Likert scale. The framework was evaluated based on its clarity, comprehensiveness, constraints, and possible improvements. An 80% consensus was sought in Round 2, based on the results of Round 1. A personalized feedback session was held with experts who were unable to reach a consensus on scores. The nine remaining variables were all agreed upon by 80% at the end of Round 2, demonstrating the strength of agreement on the evaluation framework.

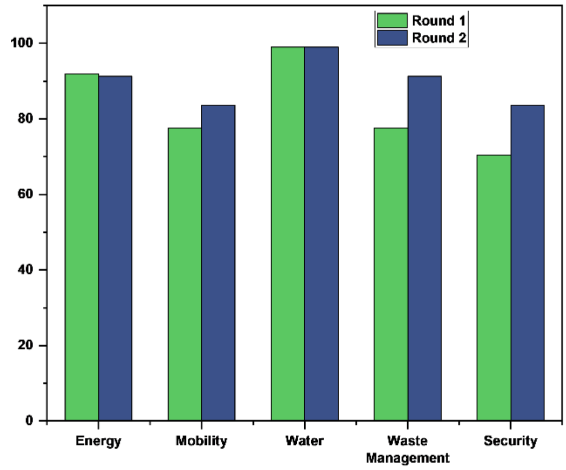


Fig. 4. Amount of agreement on the factors related to smart buildings.

In order for smart cities to reach their full potential, some areas still need to be developed. A 54.32% score indicates moderate progress towards ideal integration, which is aided by energy storage, smart EV charging, and energy usage monitoring. Resilience scores are 54%, which suggests a moderate level of preparedness for emergencies. Off-grid capability, rainwater collection, and smart water meters make the building more environmentally friendly. An overview of all case studies is presented in Figure 5.

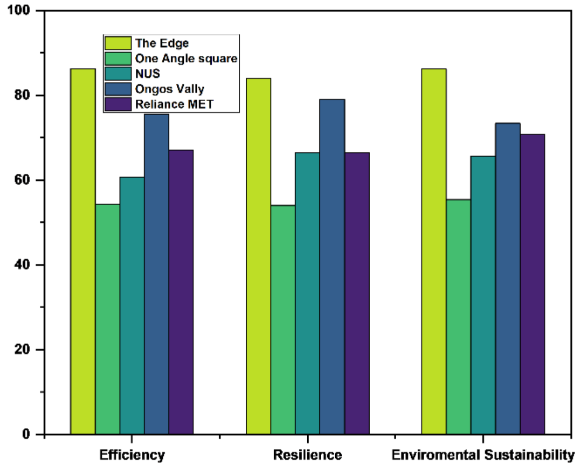
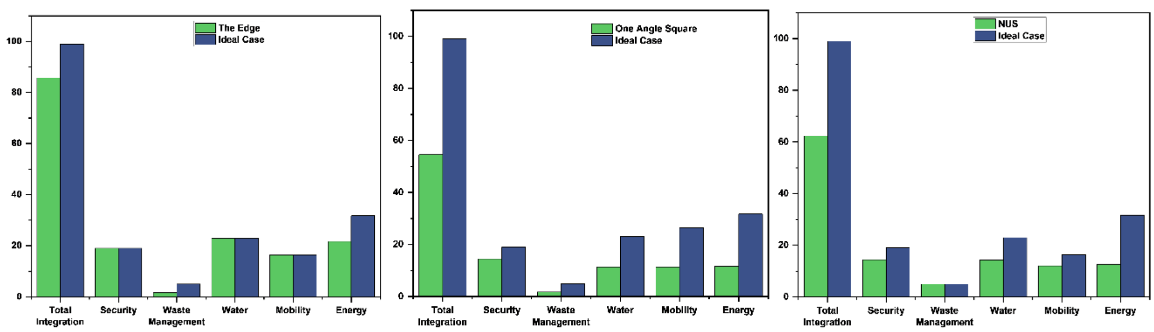


Fig. 5. The integration of smart buildings into the performance of smart cities.

In comparison to the ideal model, Edge smart buildings have high integration scores. Integrating mobility, water, and security to deliver exceptional performance. Although it improves in some areas, it falls short in others, such as energy and waste management. Although this may be the case, The Edge's overall performance is consistent with Amsterdam's smart city goals, underscoring the important role that integration plays in achieving sustainable urban growth. Several smart buildings and cities have been inspired by the success it has had in integrating smart services.



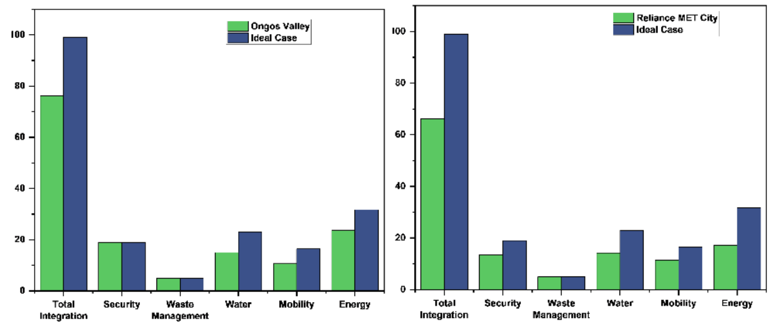


Fig. 6. Scores for the case studies with respect to integration.

5. CONCLUSION

In a smart city, smart buildings (SB) and smart cities (SCs) are integrated to improve performance and development. Through the use of AI-based methodologies, such as OpenAI's ChatGPT and Google Bard, this study demonstrates SB services' importance to efficiency, resilience, and environmental sustainability in key SC infrastructure domains. A proposed framework, which uses web scraping, topic modelling, and comparative analysis, provides valuable insight into how urban performance aligns with strategic communication trends. Smart city integration is a complex process, and further advances in fields like energy and waste management are needed to realize the full benefits of smart cities. In addition to demonstrating SB's ability to support SC infrastructure, the study also suggests pathways to improve integration in order to achieve sustainable urban growth.

Funding:

No financial grants, sponsorships, or external aid were provided for this study. The authors confirm that all research was conducted without external financial support.

Conflicts of Interest:

The authors declare that there are no conflicts of interest regarding this publication.

Acknowledgment:

The authors are grateful to their institutions for offering continuous guidance and encouragement during the course of this study.

References

[1] A. Boeri and C. Testoni, "Smart Governance: urban regeneration and integration policies in Europe. Turin and Malmö case studies," Int. J. Sci. Eng. Res., vol. 6, no. 3, pp. 527–533, 2015.
[2] J. R. Gil-Garcia, "Towards a smart State? Inter-agency collaboration, information integration, and beyond," Inf. Polity, vol. 17, no. 3–4, pp. 269–280, Dec. 2012, doi: 10.3233/IP-2012-000287.
[3] L. Chen et al., "Green construction for low-carbon cities: a review," Environ. Chem. Lett., vol. 21, no. 3, pp. 1627–1657, Jun. 2023, doi: 10.1007/s10311-022-01544-4.
[4] B. Bhola et al., "Quality-enabled decentralized dynamic IoT platform with scalable resources integration," IET Commun., 2022.
[5] F. O. Asogwa, J. I. Amuka, A. A. Igwe, and C. N. Nkalu, "Dynamics of population, urban agglomeration, and economic growths in SUB-SAHARAN Africa: Evidence from panel data," J. Public Aff., Sep. 2020, doi: 10.1002/pa.2447.
[6] J. Gutberlet, "More inclusive and cleaner cities with waste management co-production: Insights from participatory epistemologies and methods," Habitat Int., vol. 46, pp. 234–243, Apr. 2015, doi: 10.1016/j.habitatint.2014.10.004.
[7] F. Alessandria, "Inclusive city, strategies, experiences and guidelines," Procedia-Soc. Behav. Sci., vol. 223, pp. 6–10, 2016.
[8] J. Castor, K. Bacha, and F. Fuso Nerini, "SDGs in action: A novel framework for assessing energy projects against the sustainable development goals," Energy Res. Soc. Sci., vol. 68, p. 101556, Oct. 2020, doi: 10.1016/j.erss.2020.101556.
[9] I. H. Shah, L. Dong, and H.-S. Park, "Tracking urban sustainability transition: An eco-efficiency analysis on eco-industrial development in Ulsan, Korea," J. Clean. Prod., vol. 262, p. 121286, Jul. 2020, doi: 10.1016/j.jclepro.2020.121286.
[10] G. Ansari, P. Rani, and V. Kumar, "A novel technique of mixed gas identification based on the group method of data handling (GMDH) on time-dependent MOX gas sensor data," in Proc. Int. Conf. Recent Trends Comput. (ICRTC 2022), Springer, 2023, pp. 641–654.
[11] S. Modgil, S. Gupta, U. Sivarajah, and B. Bhushan, "Big data-enabled large-scale group decision making for circular economy: An emerging market context," Technol. Forecast. Soc. Change, vol. 166, p. 120607, May 2021, doi: 10.1016/j.techfore.2021.120607.
[12] D. Mimno, H. Wallach, E. Talley, M. Leenders, and A. McCallum, "Optimizing semantic coherence in topic models," in Proc. 2011 Conf. Empir. Methods Nat. Lang. Process., 2011, pp. 262–272. [Online]. Available: https://aclanthology.org/D11-1024.pdf

[13] Z. Allam and Z. A. Dhunny, "On big data, artificial intelligence and smart cities," Cities, vol. 89, pp. 80–91, Jun. 2019, doi: 10.1016/j.cities.2019.01.032.
[14] M. Angelidou, "Smart cities: A conjuncture of four forces," Cities, vol. 47, pp. 95–106, Sep. 2015, doi: 10.1016/j.cities.2015.05.004.
[15] L. Kørnøv, I. Lyhne, and J. G. Davila, "Linking the UN SDGs and environmental assessment: Towards a conceptual framework," Environ. Impact Assess. Rev., vol. 85, p. 106463, Nov. 2020, doi: 10.1016/j.eiar.2020.106463.
[16] P. Rani and R. Sharma, "Intelligent transportation system for internet of vehicles based vehicular networks for smart cities," Comput. Electr. Eng., vol, 105, p. 108543, 2023.
[17] I. Kaczmarek, A. Iwaniak, A. Swietlicka, M. Piwowarczyk, and F. Harvey, "Spatial planning text information processing with use of machine learning methods," ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci., vol. VI-4/W2-2020, pp. 95–102, Sep. 2020, doi: 10.5194/isprs-annals-VI-4-W2-2020-95-2020.
[18] S. P. Mohanty, U. Choppali, and E. Kougianos, "Everything you wanted to know about smart cities: The Internet of things is the backbone," IEEE Consum. Electron. Mag., vol. 5, no. 3, pp. 60–70, 2016.
[19] P. Rani and R. Sharma, "An experimental study of IEEE 802.11 n devices for vehicular networks with various propagation loss models," in Int. Conf. Signal Process. Integr. Netw., Springer, 2022, pp. 125–135.
[20] P. Rani and R. Sharma, "Intelligent transportation system performance analysis of indoor and outdoor internet of vehicle (IoV) applications towards 5G," Tsinghua Sci. Technol., vol. 29, no. 6, pp. 1785–1795, Dec. 2024, doi: 10.26599/TST.2023.9010119.
[21] S. E. Bibri and J. Krogstie, "Smart sustainable cities of the future: An extensive interdisciplinary literature review," Sustain. Cities Soc., vol. 31, pp. 183–212, 2017.
[22] C. Martin, J. Evans, A. Karvonen, K. Paskaleva, D. Yang, and T. Linjordet, "Smart-sustainability: A new urban fix?," Sustain. Cities Soc., vol. 45, pp. 640–648, 2019.
[23] B. Bengfort, R. Bilbro, and T. Ojeda, Applied Text Analysis with Python: Enabling Language-aware Data Products with Machine Learning. O'Reilly Media, Inc., 2018. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=lK1fDwAAQBAJ
[24] M. Röder, A. Both, and A. Hinneburg, "Exploring the space of topic coherence measures," in Proc. 8th ACM Int. Conf. Web Search Data Mining, Shanghai, China: ACM, Feb. 2015, pp. 399–408, doi: 10.1145/2684822.2685324.
[25] S. Syed and M. Spruit, "Full-text or abstract? Examining topic coherence scores using latent Dirichlet allocation," in 2017 IEEE Int. Conf. Data Sci. Adv. Analytics (DSAA), IEEE, 2017, pp. 165–174. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8259775/
[26] V. Rus, N. Niraula, and R. Banjade, "Similarity measures based on latent Dirichlet allocation," in Comput. Linguist. Intell. Text Process., vol. 7816, A. Gelbukh, Ed., Berlin, Heidelberg: Springer, 2013, pp. 459–470, doi: 10.1007/978-3-642-37247-6_37.