DOI: 10.47743/eufire-2024-1-20

SOCIALLY RESPONSIBLE USE OF ARTIFICIAL INTELLIGENCE IN SMART CITIES: AN ETHICAL IMPERATIVE

DANIELA POPESCUL
Alexandru Ioan Cuza University of Iasi
Iasi, Romania
rdaniela@uaic.ro

LAURA-DIANA RADU
Alexandru Ioan Cuza University of Iasi
Iasi, Romania
glaura@uaic.ro

Abstract
Currently, Artificial Intelligence (AI) is promisingly complementing the technological landscape associated with smart cities, consisting of contextual combinations of solutions based on the Internet of Things (IoT), mobile devices, mobile communications via satellites or short-range, data platforms, Virtual and Augmented Reality, etc. The benefits associated with the use of AI relate to improving services and optimizing resource utilization, with the stated aim of ensuring a better life for citizens in urban areas. The top-down approach in the design and implementation of technological solutions associated with smart cities remains dominant, as a consequence of the power and impact of major stakeholders (technology producers and policymakers), and the social sustainability of technological initiatives is not sufficiently explored. Against this background, this paper aims to analyze the ethical challenges associated with the use of AI in smart cities development projects. The "black-box" functioning of AI tools and the related privacy concerns will be discussed, with a focus on identifying sound solutions for an ethical use of AI in smart cities. Our results show that the eXplainable AI (XAI) approach enhances transparency, trust, control, efficiency, fairness, ethical compliance, accountability, bias mitigation, privacy, and security of AI in smart cities, and a responsible AI implementation implies diverse privacy techniques like encryption, differential privacy, anonymization, edge computing, multi-party computation, federated learning, machine learning, and blockchain.
Keywords: *smart cities; AI ethics; social sustainability; XAI; privacy.*
JEL Classification: O33, M15, O35.

1. INTRODUCTION

The growth of smart cities is significantly driven by technological contributions. Traditionally, a top-down approach of urban development has been used, with city administrators selecting technologies based on proposals from influential producers. To create a more balanced development process, a bottom-up perspective has been introduced, emphasizing societal needs, equity, inclusion, and community sustainability (Monfaredzadeh and Krueger, 2015; Trivellato, 2017; Marsal-Llacuna, 2017; Rebernik *et al.,* 2020; Zavratnik *et al.,* 2020). In addition, (Hendawy and Da Silva, 2023) present an in-between perspective, in which the integration of technology is carried out in parallel with urban social needs. In the same spirit, social sustainability advocates for a combined top-down and bottom-up approach in developing contemporary cities. This view fosters four key aspects: social capital, social cohesion, social inclusion, and social equity, while also recognizing and valuing the diverse needs and desires of people in the spaces they inhabit (Hemani and Das, 2016). In a socially sustainable smart city, ICT-based solutions must be developed with and for users, in order to ensure equitable and inclusive access for all citizens to available resources.

The range of technologies currently implemented in urban environments is vast and varied. The *enfant terrible* of our times, Artificial Intelligence (AI), complements it, benefiting from the abundance of extensive data collected through Internet of Things (IoT), the increasingly significant processing capacity of computing systems, and the enhanced transmission capabilities of communication networks. AI-driven cities aim to enhance efficiency, environmental protection, and quality of life. However, they face challenges such as technological limitations, ethical concerns, and regulatory complexities (Javed *et al.,* 2023), generated by the black-box functioning of AI algorithms, the signaled biases in their use, and the unpredictability of their development (Tegmark, 2019; Bellagarda and Abu-Mahfouz, 2022; Jagatheesaperumal *et al.,* 2022). In the context of accelerating AI development, empowering citizens in their interactions with AI is needed for fostering truly sustainable and socially beneficial smart cities (Toribio-Roura, 2020; Kelleher and Kerr, 2020). The role of technology developers becomes essential in ensuring these innovations are implemented responsibly, as the integration of AI, the IoT, and other advanced systems in smart cities necessitates a thorough consideration of ethical responsibilities to safeguard the interests and rights of citizens.

2. LITERATURE REVIEW

Living in a smart city offers benefits such as a better quality of life, access to diverse services and facilities, job opportunities, and higher incomes, yet also presents challenges including busy traffic, pollution, and housing and services affordability. Besides the inhabitants, the city administrators face issues such as costly infrastructure improvements and provocative management due to the increased urban complexity. This duality extends to the technological domain, where advancements can enhance urban service efficiency but introduce issues like imperfect security and protection, data privacy concerns, deficits in transparency of services, ethical considerations (Lytras and Visvizi, 2018), digital divide, and issues with compliance and accountability.

Currently, AI significantly contributes to progress within the realm of smart cities, driving substantial technological innovation and providing valuable data-driven insights (Dos Santos *et al.,* 2024) but, on the other side, profoundly influencing human behavior and lifestyle. Some of the AI mechanisms and technologies used in smart cities are Machine Learning (ML), defined as the development of algorithms and statistical models that enable computers to perform tasks without explicit instructions, Deep Learning (DL), a subset of ML that uses neural networks with many layers (known as deep neural networks) to model complex patterns in data, and Natural Language Processing, in which machines are able to understand, interpret, and generate human language in a meaningful way (Dos Santos *et al.,* 2024). Computer Vision technology, which enables machines to interpret and understand visual information from the world, is also used in smart cities. It has numerous applications, including critical infrastructure protection, sanitation and waste management, transportation, traffic management, pandemic control, security, smart water management, and disaster management (Ramalingam *et al.,* 2023). According to (Morel, 2021), ML and DL are prevalent in smart cities, with a recent trend towards a merging of these two AI trends known as neuro-symbolic AI. The author mentions that these connectionists AI operate without explicit rational reasoning, as demonstrated by their ability to identify objects in both static and dynamic scenes, and considers that such tools are essential in smart cities for the rapid analysis of the extensive images, videos, and other large datasets generated by sensors and various devices.

AI technologies are essential for smart city applications but also bring challenges to cities and citizens, such as the need to address risks associated with wider AI utilization and examine upcoming disruptions of AI in cities and societies (Yigitcanlar *et al.,* 2020). In a comprehensive analysis of the application and barriers of AI across various domains within smart cities (smart mobility, environmental management, governance, economic initiatives, living standards, and community empowerment), Wolniak and Stecuła address specific challenges and solutions associated with AI implementation in each dimension, emphasizing the necessity for customized approaches to ensure the responsible and equitable deployment of AI solutions in these diverse areas (Wolniak and Stecuła, 2024).

Allam and Dhunny show that the rise of smart cities, powered by IoT and massive data generation, necessitates AI-driven big data management to enhance system efficiency, while ensuring technology supports human-centric dimensions of livability, resilience, safety, and sustainability in alignment with the 11th Sustainable Development Goal introduced by United Nations in 2015, which is to “make cities and human settlements inclusive, safe, resilient, and sustainable” (Allam and Dhunny, 2019).

3. RESEARCH METHODOLOGY

Based on the aspects above, we formulated two research questions:

RQ1. What concerns have been raised on the AI use in smart cities development?

RQ2. What solutions have been formulated to address the challenges of AI use in smart cities development?

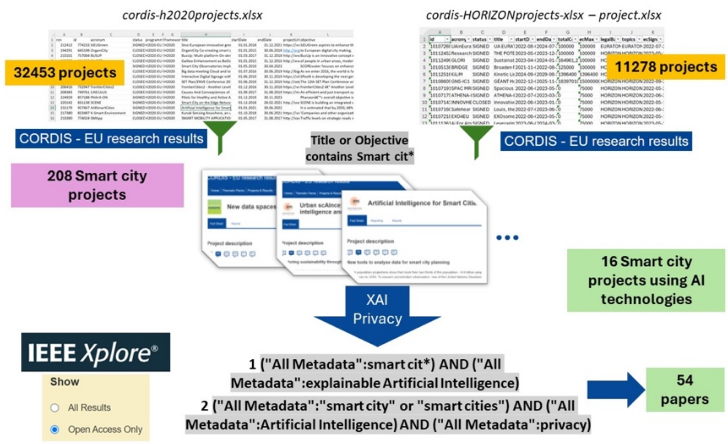

Figure 1. Research methodology

We selected as a basis for our analysis the smart cities projects funded by the European Union under the Horizon 2020, Urban Action Initiative and Horizon Europe programs. The mentioned programs were chosen because of their considerable contribution to the smart, sustainable and inclusive development of the European Union. The highly innovative initiatives funded by the programs are carried out in partnership and support the creation and dissemination of excellent knowledge and top technologies. Using CORDIS, we downloaded the data sets containing information about research projects under Horizon 2020 (2014-2020) (European Union, 2022) and Horizon Europe (2021-2027) (European Union, 2023). CORDIS, the COmmunity Research and Development Information Service, serves as the European Commission’s primary source of research results obtained in the projects funded by the European Union framework programs for research and innovation. We filtered the projects that implemented/implement smart cities initiatives and use AI-based solutions. Firstly, we extracted projects with titles and/or objectives that include the term 'smart cities' or its variations. Subsequently, we analyzed the project fact sheets on the European Commission's website to determine whether the projects have utilized or are utilizing AI, and to identify any concerns raised regarding AI usage.

The fact sheets provide *General information* about projects and their objectives, *Results in brief*, and *Reporting and results* documents (publications obtained as outcomes of these projects). To complement our perspective on the subject, following the identification of key ethical aspects associated with AI use in smart city projects, we searched for articles on these themes in the open-access version of IEEE Xplore. This database provides "full text access to the world's highest quality technical literature in engineering and technology" (IEEE, 2024). This additional bibliographic research was necessary because most concerns regarding the responsible and ethical use of AI in smart city development have been highlighted in projects funded by Horizon Europe, which are still in early stages as the program was launched only in 2021, and few results are currently available.

4. RESULTS

4.1 Smart cities projects using AI technologies in Horizon 2020 and Horizon Europe

From the total number of 32453 projects financed in Horizon 2020, respectively 11278 projects financed by far in Horizon Europe, after filtering by the keywords "smart city" or "smart cities" in the *Title* or *Objective* fields in the above-mentioned CORDIS *xlsx* files, 208 (157+51) projects were identified. By analyzing the fact sheets of these projects on the European Commission's website, we found that 16 of them use various AI technologies in a sufficient extend to be considered in our analysis. Table 1 presents a selection of smart cities projects that use/intent to use AI technology for sustainable development, social assistance, better urban mobility management, improved environmental protection, security and privacy.

Table 1. Smart cities projects using AI technologies in Horizon 2020 and Horizon Europe

Project	Scope	AI Use
Responsible use of AI in smart cities		
Urban scAInce - Why and how cities transform through Artificial Intelligence and their associated technologies **(scAInce, CORDIS, 2024)**	Based on the presumption that more sustainable cities are smarter, the urban scAInce theory investigates how AI and its associated technologies has, can, and will alter urban systems.	Privately-driven AI solutions impact on urban sustainability will be evaluated.
Assistance		
TOTEM SPOON - Interactive Digital Signage with emotional intelligence	A screen-based avatar named SPooNy was designed for deployment in public spaces to	SPooN was designed as an enabler for the AI ecosystem in public spaces, based on an

Project	Scope	AI Use
for smart cities **(TOTEM SPOON, CORDIS, 2019)**	provide welcoming and assistance. Voice-based interaction was implemented to mitigate the digital divide.	*Ethics by Design* approach.
GAVIUS - From reactive to proactive public administrations **(Urban Innovative Actions, 2024)**	The project utilizes AI to enhance social assistance efficacy and modernize public administration. It aims to streamline citizen interactions with municipal services, support data-driven decision-making by governmental teams, and increase efficiency in managing civic procedures, serving as a model for modernization in the public sector.	Using ML methods, the system aims to achieve three main goals: streamline municipal tasks and improve interactions with citizens, empower citizens to easily access public services, and assist municipal leaders in effectively allocating resources for social services planning.
Mobility		
SPECK - Smart Pedestrian Crosswalk for Increased Traffic Safety at Uncontrolled Crossings **(SPECK, CORDIS, 2022)**	The project focuses on enhancing safety for vulnerable road users (pedestrians, cyclists, and motorcyclists) by proposing an intelligent roadside unit designed for installation at uncontrolled crossings.	Narrow AI algorithms coupled with sensor fusion techniques are used to predict and avoid situation-based traffic conflicts.
MobiSpaces - New data spaces for green mobility **(MobiSpaces, CORDIS, 2024)**	The data governance platform provided by MobiSpaces for utilizing mobility data aims to minimize the environmental impact of mobility by establishing an efficient, dependable, equitable, trustworthy, and privacy-preserving infrastructure.	eXplainable AI (XAI) techniques are to be applied at the level of data management and Machine Learning, supporting the creation of comprehensive and interpretable prediction models.
EMERALDS - Extreme-scale Urban Mobility Data Analytics as a Service - **(EMERALDS, 2024) (EMERALDS, CORDIS, 2024)**	The project proposes the creation of an urban data-oriented Mobility Analytics as a Service toolset, in order to benefit from the big data generated by the urban mobility ecosystem.	Extensive and diverse spatio-temporal data will be analyzed using real-time responsive AI/ML algorithms. Privacy preservation techniques, including Active and Federated Learning and XAI, will be implemented across all data modalities and throughout all layers of the Mobility Analytics as a Service architecture.
Environmental protection		
DIME - Distributed Inference for Energy-efficient Monitoring at the Network Edge **(DIME, CORDIS, 2024)**	The project aims to minimize the redundancy of data samples generated by the IoT sensors, redundancy that creates an “excessive and unjustified	The project proposes a general modelling framework, scheduling algorithms and sampling techniques that minimize the Total System

Project	Scope	AI Use
	carbon footprint of these systems".	Energy. Compute-intensive ML models, specifically Deep Neural Networks (DNNs), are used for providing inference.
GreenInCities **(GreenInCities, CORDIS, 2024)**	Inequities in urban areas, characterized by disparities in infrastructure and resources between neglected and affluent neighborhoods, motivate this project to prioritize societal awareness, innovate beyond traditional greening approaches, and integrate advanced technologies.	The project develops tools for collaborative urban planning in deprived areas, focusing on climate action.
Security and privacy		
AIAS - Innovative security platform against adversarial AI attacks **(AIAS, CORDIS, 2024)**	In response to the actions of attackers who target ML and DL systems, AIAS project aims to conduct research on adversarial AI to empower security teams, fortifying AI systems against potential attacks.	The developed innovative security platform for organizations will employ adversarial AI defense methods, deception mechanisms, and XAI solutions.
PrivacyForDataAI - A privacy layer to power all research and AI workflows **(PrivacyForDataAI, CORDIS, 2024)**	A privacy-preserving data solution for all research and AI projects will enable researchers and data scientists to leverage private data assets without seeing them.	The solution acts as protection layer between the data source and the scientist, guaranteeing the privacy of any AI-based computation.

4.2 Concerns on AI-use in smart cities projects

The exploration of AI in smart cities reveals both potential and challenges. Kaseens-Noor *et al.* (2022) explores the concept of scAInce in the context of smart cities, presenting both utopian and dystopian visions of their transition from "smart" to "intelligent". The paper identifies smart city megaprojects and large infrastructures as early adopters and generators of big data collections interpretable by machines. AI-driven megaprojects target significant goals such as sustainability and resilience, offering flexibility in how AI systems achieve these objectives and facilitating sustainable changes through evolving algorithms. For instance, in transportation, an AI algorithm can optimize travel times to reduce emissions, altering mobility patterns while maintaining the city's physical infrastructure. However, the black-box nature of AI poses challenges in understanding and controlling its decision-making processes. Additionally, while AI excels at analyzing historical data, it struggles with future predictions. Implementing urban AI megaprojects necessitates overcoming skepticism towards AI, particularly with autonomous vehicles due to their involvement in accidents. Moreover, technological products tested individually but operating within ecosystems create unpredictability, especially in new and unexpected contexts. Concerns about job losses and increased dependency on technology further exacerbate these challenges. The authors propose an open science approach, where citizens and residents of smart cities actively participate in the planning, design, construction, and management of AI projects to achieve sustainability and resilience.

A review published as a result of the MobiSpaces project highlights how the black-box nature of DL algorithms impacts location privacy concerning individuals' movement data in cities and presents several technical solutions. Synthesizing trajectories through various algorithms effectively ensures the anonymization of spatial-temporal movement data of subjects, allowing for their sharing/publication without compromising privacy (Graser *et al.,* 2024). In the same project, with reference to mobility data, Makridis *et al.* (2024) propose a unified model that harmonizes various essential XAI techniques. In the EMERALDS project, Jalali *et al.* (2023) emphasize the need to ensure transparency and explainability in the field of Geospatial AI, due to the social and environmental implications that these applications can have.

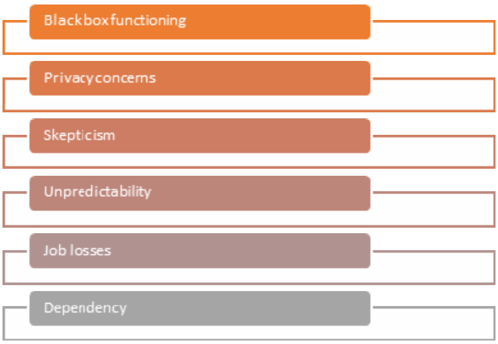

Figure 2. AI-related concerns

As illustrated in Figure 2, the primary concerns mentioned refer to the opaque functioning of AI technologies and the privacy of sensitive data regarding the identity, location, and movement of smart city inhabitants and visitors. The proposed solutions in the analyzed projects include the implementation of eXplainable AI (XAI), Federated Learning, and privacy assurance measures for sensitive data.

4.3 eXplainable AI (XAI)

Our analysis of open access IEEE eXplore papers on smart cities and eXplainable AI (XAI) revealed a diversity of perspectives on this topic, including: Cyber-Physical systems (Hoenig *et al.,* 2024), IoT (García-Magariño *et al.,* 2019; Jagatheesaperumal *et al.,* 2022), Distributed Ledger Technology

(Bellagarda and Abu-Mahfouz, 2022), cybersecurity (Houda *et al.,* 2022), crowd monitoring (Samarajeewa *et al.,* 2024), air pollution (Song *et al.,* 2022), prediction of free parking spots (Bilotta *et al.,* 2023), or early assessment of certification chances in MOOC courses (Kostopoulos *et al.,* 2021).

In a comprehensive review on the role of XAI in Cyber-Physical systems (Hoenig *et al.,* 2024) state that XAI helps in the achievement of the CPS' efficiency goals (such as optimization, security, cyber-resilience and safety) while "maintaining fairness, proper use of AI, and ethical and legal compliance". The authors identify a series of smart city services that require a high degree of transparency and trust, such as CCTV-based monitoring of individuals for safety-assurance purposes. The training of AI algorithms in the functioning of such CPS systems can still be exposed to biases caused by the manual manipulation of input data or other factors, and the classifications and decisions made by the algorithms can have a negative impact on people's reputation and lives. Standardized methods of assessing the understandability and performance of XAI are necessary.

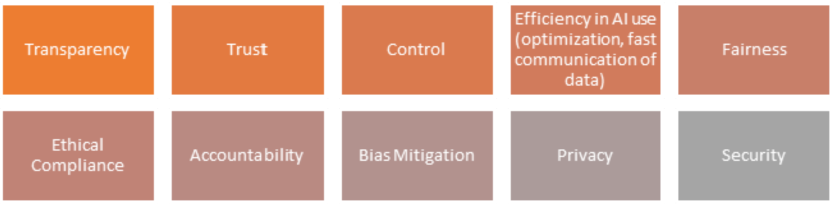

Figure 3. XAI benefits

Jagatheesaperumal *et al.* (2022) examine the role of AI in obtaining big data-based insights through proper analytics and enhancing the capabilities of Internet of City Things, with a focus on smart buildings and smart mobility infrastructure. The authors show that XAI contributes to high-speed communication and trustworthy data transfer, and to energy savings, an important aspect in greening the urban agglomerations. Opposed to the conventional AI-systems, which function as opaque black boxes, interpretable and human-understandable ML algorithms can assure an increased transparency, interpretability, trustworthiness and control on data, processing models and derived classifications and decisions. The authors assert that XAI contributes to enhancing cybersecurity through encryption, but, on the other hand, they point out that the advancement in the use of XAI in smart cities is hindered by the inherent limitations of interconnected urban objects, primarily due to their reduced processing capacity.

In Bellagarda and Abu-Mahfouz (2022), the authors show that the ML algorithms, especially deep neural networks, often produce decisions that are not easily understood. XAI aims to address the black box issue by using tools and frameworks to clarify AI model predictions and provide justifiable reasoning for decisions. Distributed Ledger Technology (DTL) can track AI data, provide an audit trail, and enhance transparency and security of AI systems, by explaining decisions and providing a secure, transparent ledger for AI metadata. Another use of blockchain DTL in combination with ML is presented in Khan *et al.* (2024**)**. The authors propose a model of increasing remote sensing data privacy through encryption, experimenting with data of land surface in smart cities and concluding that the following aims were achieved: "minimal resource consumption costs, high privacy protection during data travel, optimal pathway, and quick data exchange capabilities".

García-Magariño *et al.* (2019) discuss the human-centric AI (HAI), an emerging field supported by the European Union, in IoT systems. In the authors' view, HAI allows users to understand and verify AI decisions, improving supervision and trust, and XAI is closely related to HAI, focusing on generating explanations for AI algorithms. The use of XAI in cybersecurity is discussed in Houda *et al.* (2022), where the authors show how interpreting and understanding predictions made by DL-based Intrusion Detection Systems can help in making security measures more transparent to the experts. The paper discusses the necessity of new security mechanisms for IoT networks, highlighting the effectiveness of Intrusion Detection Systems enhanced by DL for predicting attacks. It introduces a new XAI-based framework integrating DL and explainable AI techniques (such as LIME, SHAP, RuleFit) to improve real-time intrusion detection, transparency, and trust for both model users and cybersecurity experts.

4.4 Privacy-related concerns and solutions

Data used by smart city technologies are often highly sensitive. In Badidi *et al.* (2023), the case of video data is discussed. Video analytics significantly enhances various fields such as traffic management, security, healthcare and retail, by intelligently analyzing data to uncover hidden patterns and correlations, thereby improving decision-making, efficiency, and the ability to predict future events. Edge AI devices, which use sensors connected to a microcontroller unit loaded with ML models, can independently process data and make decisions without needing internet connectivity. The models are trained for typical scenarios, and the device can continuously learn from new situations. The AI's responses can include physical actions nearby or notifications to users or the cloud for further analysis. Smart cities use edge computing and AI for local decision-making, enhancing safety, managing resources, and improving efficiency while reducing latency, network congestion, and data privacy risks. Integrating edge AI and video analytics enhances search and monitoring by precisely identifying license plates, faces, pedestrians, and ensuring workplace safety. Privacy-preserving techniques in edge video analytics include encryption, differential privacy, anonymization, edge computing, and multi-party computation to ensure data security.

Federated Learning (FL) offers a distributed AI approach suitable for modern IoT networks by enabling AI training on distributed devices without data sharing. FL is an AI algorithm that trains models using distributed data on local devices without transferring the data, ensuring privacy and preventing leakage (Sepasgozar and Pierre, 2022). In Nguyen *et al.* (2021), the authors survey the applications of FL in IoT, covering advancements, integration, and potential services, such as data sharing, attack detection, and privacy. Leveraging recent advances in mobile hardware and addressing privacy concerns, FL enables AI training at the network edge on IoT devices, ensuring user data remains local. This approach benefits both network operators and users by saving network resources and enhancing privacy, making FL a strong alternative to traditional centralized AI and facilitating large-scale IoT deployment.

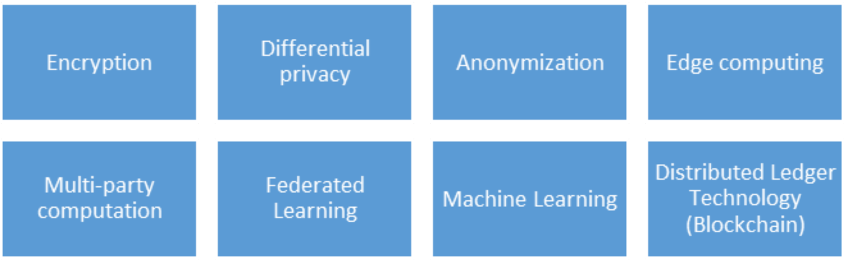

Figure 4. Privacy-preserving solutions

Analyzing the case of using collaborative drones and IoT in smart cities, Alsamhi *et al.* (2019) highlight the sensitivity of data collected in this context and show that ML can be used to keep the network secure from cyber-attacks, detect attacks during training and identify security vulnerabilities using adversarial setups. Another good response of ML algorithms to the privacy restrictions is presented in Santana *et al.* (2020). The authors detail the design, implementation, and validation of a lightweight, privacy-focused AI-based crowd-management system in real-world deployments in Santander.

In Sheraz *et al.* (2024), it is shown that AI in wireless networks faces challenges due to high data needs and limited computing resources, which can lead to inaccurate predictions. Also, AI is essential for data security in wireless networks, but it requires diverse datasets and continuous updates. Digital Twin Networks (DTN) offer a safe testing environment for these updates.

The smartness of a city depends on the richness of data that feeds the urban services. Beyond the concerns related to assuring the privacy for sensitive data, its collection from different IoT sensors generates energy consumption (Alsamhi *et al.,* 2019), and addition of security measures such as encryption or blockchain technology is "computationally expensive" (Bernal Bernabe *et al.,* 2019). In (Azzaoui *et al.,* 2020), the authors propose a framework that combines blockchain and AI for secure, efficient, and energy-saving 5G networks.

4.5 Relevant themes and topics in the analyzed paper collection. The most influential authors

To investigate the relationship between the most visible AI-related topics in the European Union funded projects for the smart city development and the themes or topics in the collection of IEEE eXplore articles we used Biblioshiny() from Bibliometrix 4.0.1, a powerful tool for literature analysis. In our case, the thematic map is based on "keyword plus" that results from literature analysis and utilizes Walktrap algorithm. This algorithm is designed to identify clusters of closely related topics based on the frequency and patterns of keyword co-occurrences. Each cluster represents a group of related topics or keywords. The most often used metrics are centrality and density. Density signifies the strength of internal connections among all keywords used to describe the research theme, while centrality represents the strength of external connections to other themes by leveraging the field of authors' keywords (Xiao *et al.,* 2022). There are six clusters grouped in four categories:

Cluster 1 and 2 include *basic themes* – system, networks and classification, respectively blockchain and IoT. These themes are transversal and general, being essential for a good understanding of the field.

Cluster 3 – *motor themes* include research related to internet, privacy, AI, challenges, security etc. They have a high density and centrality, being the main themes in the field with a strong connection between concepts. These themes are well-developed and important for research.

Cluster 4 – *emerging or declining themes* – includes more generic themes related to models and technologies. They could gain more centrality or more density becoming the new trends or development in the field.

Cluster 5 and 6 – *niche themes* – includes communication and wireless networks, respectively data collection, digital twin and resource allocation. These themes are more specialized and peripheral, but highly developed and have insignificant external ties.

It can be noticed that the topics investigated in the current paper (XAI and privacy) are motor themes, of primary interest among researchers. As stated above, motor themes are characterized by both high centrality and high density (Herrera-Viedma *et al.,* 2020). These themes can drive innovation and future research directions. Researchers focus on these themes to advance knowledge, develop new theories, and create novel applications. Understanding motor themes helps researchers, policymakers, and funding agencies identify critical areas for investment and development. It also helps in setting research agendas and priorities.

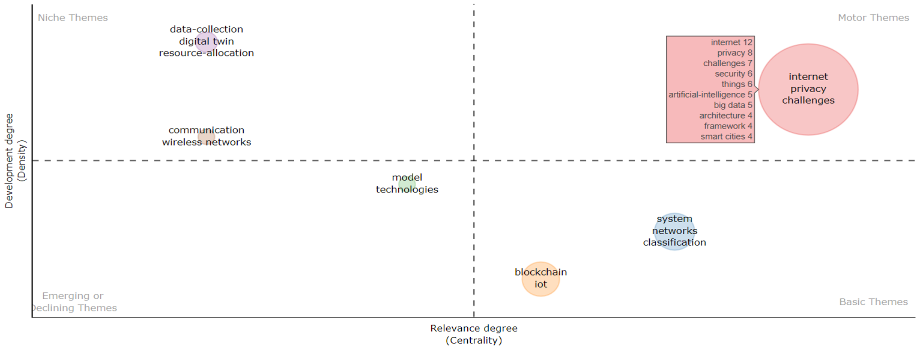

Figure 5. Thematic map

Table 2 presents the most 10% influential authors based on the most cited documents from the data set. All the papers were published in journals from quartile 1 (IEEE Communications Surveys & Tutorials and IEEE Internet of Things Journal) and quartile 2 (IEEE Access) according to Web of Science classification reflecting their popularity and central importance. These journals are also in Zone 1 and Zone 2 in Bradford's distribution (according to classification derived by Bibliometrix 4.0.1). This is based on Bradford's Law, which assumes that a relatively small number of journals publish most significant scientific papers.

Table 2. Most cited documents

Authors	Title	Keywords	Total Citation	Total Citation per Year
(Nguyen *et al.*, 2021)	Federated Learning for Internet of Things: A Comprehensive Survey	Federated Learning, IoT, AI, ML, privacy	347	86.75
(Zhang and Tao, 2020)	Empowering Things with Intelligence: A Survey of the Progress, Challenges, and Opportunities in Artificial Intelligence of Things	IoT, AI, DL, Cloud/Fog/Edge Computing, security, privacy, sensors, biometric recognition, 3D, speech recognition, machine translation, causal reasoning, Human-Machine Interaction, smart city	264	66.00
(Alsamhi *et al.*, 2019)	Survey on Collaborative Smart Drones and Internet of Things for Improving Smartness of Smart Cities	smart city, energy consumption, smart drone, IoT, pollutions, gathering data, Internet of Drones disaster, public safety, security and privacy,	182	30.33

Authors	Title	Keywords	Total Citation	Total Citation per Year
		collaborative drone.		
(Yu *et al.*, 2020)	When Deep Reinforcement Learning Meets Federated Learning: Intelligent Multi-Timescale Resource Management for Multi-access Edge Computing in 5G Ultra Dense Network	Multi-access Edge Computing, Computation Offloading, Service Caching, Ultra Dense Network, Blockchain, Deep Reinforcement Learning, Federated Learning	153	38.25
(Yu *et al.*, 2021)	A Blockchain-Based Shamir's Threshold Cryptography Scheme for Data Protection in Industrial Internet of Things Settings	Industrial IoT, data protection, blockchain, Shamir secret sharing	81	27.00

5. CONCLUSIONS

The paper analyzes a set of research projects in the field of smart cities, funded under European Union programs from 2014 to 2024, with the aim of identifying the challenges and AI-related solutions encountered and proposed within them. After identifying the black box nature of AI technologies and their impact on the privacy of personal data, the location and movement data of residents and visitors in smart cities, the paper focuses on these aspects, investigating them within the IEEE Xplore database of scientific papers. The reason for selecting this database is the technical nature of the published works.

The benefits of using the XAI approach are related to an increase in transparency, trust, control, efficiency, fairness, ethical compliance, accountability, bias mitigation, privacy and security of AI technology in smart cities. All these characteristics contribute to a more responsible, socially oriented implementation of AI. Regarding the privacy, we noticed that a significant variety of techniques are used for its assurance, such as the "classic" encryption, differential privacy, anonymization, supplemented by edge computing, multi-party computation, Federated Learning, Machine Learning and Distributed Ledger Technology (blockchain).

References

1) Allam, Z. and Dhunny, Z.A. (2019). On big data, artificial intelligence and smart cities. *Cities,* 89, pp. 80-91.
2) Alsamhi, S., Ma, O., Ansari, M.S. and Almalki, F.A. (2019). Survey on Collaborative Smart Drones and Internet of Things for Improving Smartness of Smart Cities. *IEEE Access,* 7, pp. 128125-128152.
3) Azzaoui, A.E., Singh, S.K., Pan, Y. and Park, J.H. (2020). Block5GIntell: Blockchain for AI-Enabled 5G Networks. *IEEE Access,* 8, pp. 145918-145935.

4) Badidi, E., Moumane, K. and Ghazi, F.E. (2023). Opportunities, Applications, and Challenges of Edge-AI Enabled Video Analytics in Smart Cities: A Systematic Review. *IEEE Access,* 11, pp. 80543-80572.
5) Bellagarda, J.S. and Abu-Mahfouz, A.M. (2022). An Updated Survey on the Convergence of Distributed Ledger Technology and Artificial Intelligence: Current State, Major Challenges and Future Direction. *IEEE Access*, 10, pp. 50774-50793.
6) Bernal Bernabe, J., Canovas, J.L., Hernandez-Ramos, J.L., Torres Moreno, R. and Skarmeta, A. (2019). Privacy-Preserving Solutions for Blockchain: Review and Challenges. *IEEE Access,* 7, pp. 164908-164940.
7) Bilotta, S., Ipsaro Palesi, L.A. and Nesi, P. (2023). Predicting Free Parking Slots via Deep Learning in Short-Mid Terms Explaining Temporal Impact of Features. *IEEE Access*, 11, pp. 101678-101693.
8) Dos Santos, S.C., Vilela, J.F.F., Carvalho, T.H., Rocha, T.C., Candido, T.B., Bezerra, V.S. and Silva, D.J. (2024). *Artificial Intelligence in Sustainable Smart Cities: A Systematic Study on Applications, Benefits, Challenges, and Solutions*. Angers: SciTePress, pp. 644-655.
9) EMERALDS, CORDIS (2024). *Extreme-scale Urban Mobility Data Analytics as a Service.* [online] Available at: https://cordis.europa.eu/project/id/101093051 [Accessed 14.01.2024].
10) EMERALDS (2024). *Name of the document or webpage* [online] Available at: https://emeralds-horizon.eu/ [Accessed 01.05.2024].
11) European Union (2022). *European Data - CORDIS - EU research projects under Horizon 2020 (2014-2020).* [online] Available at: https://data.europa.eu/data/datasets/cordish2020projects [Accessed 14.01.2024].
12) European Union (2023). *European Data CORDIS - EU research projects under Horizon Europe.* [online] Available at: https://data.europa.eu/data/datasets/cordis-eu-research-projects-under-horizon-europe-2021-2027 [Accessed 14.01.2024].
13) García-Magariño, I., Muttukrishnan, R. and Lloret, J. (2019). Human-Centric AI for Trustworthy IoT Systems With Explainable Multilayer Perceptrons. *IEEE Access*, 7, pp. 125562-125574.
14) Graser, A., Jalali, A., Lampert, J., Weißenfeld, A. and Janowicz, K. (2024). MobilityDL: a review of deep learning from trajectory data. *Geoinformatica*.
15) GreenInCities, CORDIS (2024). *GreenInCities.* [online] Available at: https://cordis.europa.eu/project/id/101139730 [Accessed 12.04.2024].
16) Hemani, S. and Das, A.K. (2016). Humanizing urban development in India: call for a more comprehensive approach to social sustainability in the urban policy and design context. *International Journal of Urban Sustainable Development*, 8(2), pp. 144-173.
17) Hendawy, M. and Da Silva, I. (2023). Hybrid Smartness: Seeking a Balance Between Top-Down and Bottom-Up Smart City Approaches. In: Goodspeed, R., Sengupta, R., Kyttä, M., Pettit, C. (eds). *Intelligence for Future Cities*. Cham:Springer, pp. 9-27.
18) Herrera-Viedma, E., López-Robles, J.-R., Guallar, J. and Cobo, M.-J. (2020). Global trends in coronavirus researchat the time of Covid-19: A general bibliometric approach and contenta nalysis using SciMAT. *El profesional de la información*, 29(3), pp. 290-322.

19) Hoenig, A., Roy, K., Acquaah, Y.T., Yi, S. and Desai, S.S. (2024). Explainable AI for Cyber-Physical Systems: Issues and Challenges. *IEEE Access*, 12, pp. 73113-73131.
20) Houda, Z.A., Brik, B. and Khoukhi, L. (2022). "Why Should I Trust Your IDS?": An Explainable Deep Learning Framework for Intrusion Detection Systems in Internet of Things Networks. *IEEE Open Journal of the Communications Society,* 3, pp. 1164-1176.
21) IEEE (2024). *IEEE Xplore*. [online] Available at: https://ieeexplore.ieee.org/Xplore/home.jsp[Accessed 16.05.2024].
22) Jagatheesaperumal, S.K., Pham, Q.V., Ruby, R., Yang, Z., Xu, C. and Zhang, Z. (2022). Explainable AI Over the Internet of Things (IoT): Overview, State-of-the-Art and Future Directions. *IEEE Open Journal of the Communications Society*, 3, pp. 2106-2136.
23) Jalali, A., Graser, A. and Heistracher, C. (2023). Towards eXplainable AI for Mobility Data Science. *Arxiv - Cornell University*.
24) Javed, A., Ahmed, W., Pandya, S., Maddikunta, P.K.R., Alazab, M. and Gadekallu, T.R. (2023). A Survey of Explainable Artificial Intelligence for Smart Cities. *Electronics,* 12, p. 1020.
25) Kaseens-Noor, E., Darcy, K., Cojocaru, A., Rzepecki, R., Jang, S., Jiang, W., Monzert, T., Cai, M. and Crittenden, M. (2022). Proposing the foundations of scAInce by exploring the future of artificially intelligent, sustainable, and resilient megaprojects. *Journal of Mega Infrastructure & Sustainable Development,* 2(1), pp. 5–20.
26) Kelleher, J. and Kerr, A. (2020). Finding Common Ground for Citizen Empowerment in the Smart City. *Ethics & Politics*, XXII(2), pp. 33-66.
27) Khan, A., Laghari, A.A., Alroobaea, R., Baqasah, A., Alsafyani, M., Baccara, R. and Alsayaydeh, J. (2024). Secure Remote Sensing Data With Blockchain Distributed Ledger Technology: A Solution for Smart Cities. *IEEE Access,* 12, pp. 69383-69396.
28) Kostopoulos, G., Panagiotakopoulos, T., Kotsiantis, S., Pierrakeas, C. and Kameas, A. (2021). Interpretable Models for Early Prediction of Certification in MOOCs: A Case Study on a MOOC for Smart City Professionals. *IEEE Access*, 9, pp. 165881-165891.
29) Lytras, M. and Visvizi, A. (2018). Who uses smart city services and what to make of it: Toward interdisciplinary smart cities research. *Sustainability,* 10(6), 1998.
30) Makridis, G., Koukos, V., Fatouros, G., Separdani, M.M. and Kyriazis, D. (2024). *Enhancing Explainability in Mobility Data Science Through a Combination of Methods*. Cham: Springer.
31) Marsal-Llacuna, M.-L. (2017). Building Universal Socio-cultural Indicators for Standardizing the Safeguarding of Citizens' Rights in Smart Cities. *Social Indicators Research*, 130(2), pp. 563-579.
32) MobiSpaces, CORDIS (2024). *MobiSpaces - New data spaces for green mobility.* [online] Available at: https://cordis.europa.eu/project/id/101070279 [Accessed 14.04.2024].

33) MODELAIR, CORDIS (2024). *Groundbreaking tools and models to reduce air pollution in urban areas*. [online] Available at: https://cordis.europa.eu/project/id/101072559 [Accessed 12.04.2024].
34) Monfaredzadeh, T. and Krueger, R. (2015). Investigating social factors of sustainability in a smart city. *Procedia Engineering*, 11(8), pp. 1112 – 1118.
35) Morel, G. (2021). *Neuro-symbolic A.I. For the smart city*. Lausanne: IOP Publishing Ltd.
36) Nguyen, D.C., Ding, M., Pathirana, P.N., Seneviratne, A., Li, J. and Vincent Poor, H. (2021). *Federated* Learning for Internet of Things: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*, 23(3), pp. 1622-1658.
37) PrivacyForDataAI, CORDIS (2024). *A privacy layer to power all research and AI workflows*. [online] Available at: https://cordis.europa.eu/project/id/101145303 [Accessed 12.04.2024].
38) Ramalingam, M., Chemmalar, S., Nancy, V., Rajeswari, C., Sweta, B., Praveen, K.R.M., Duehee, L., Jalil Neelu, K., Gokul, Y. and Thippa, R.G. (2023). A Comprehensive Analysis of Blockchain Applications for Securing Computer Vision Systems. *IEEE Access,* 11, pp. 107309-107330.
39) Rebernik, N., Szajczyk, M., Bahillo, A. And Golicnik Maruši, B. (2020). Measuring Disability Inclusion Performance in Cities Using Disability Inclusion Evaluation Tool (DIETool). *Sustainability*, 12(4), 1378.
40) Samarajeewa, C., De Silva, D., Manic, M., Mills, N., Rathnayaka, P. and Jennings, A. (2024). Explainable Artificial Intelligence for Crowd Forecasting Using Global Ensemble Echo State Networks. *IEEE Open Journal of the Industrial Electronics Society,* 5, pp. 415-427.
41) Santana, J.R., Sánchez, L., Sotres, P., Lanza, J., Llorente, T. and Muñoz, L. (2020). A Privacy-Aware Crowd Management System for Smart Cities and Smart Buildings. *IEEE Access,* 8, pp. 135394-135405.
42) scAInce, CORDIS (2024). *Urban scAInce: why and how cities transform through artificial intelligence and their associated technologies*. [online] Available at: https://cordis.europa.eu/project/id/101087218 [Accessed 12.04.2024].
43) Sepasgozar, S.S. and Pierre, S. (2022). Fed-NTP: A Federated Learning Algorithm for Network Traffic Prediction in VANET. *IEEE Access,* 10, pp. 119607-119616.
44) Sheraz, M., Chuah, T.C., Lee, Y.L., Alam, M.M., Al-Habashna, A. and Han, Z. (2024). A Comprehensive Survey on Revolutionizing Connectivity Through Artificial Intelligence-Enabled Digital Twin Network in 6G. *IEEE Access*, 12, pp. 49184-49215.
45) Song, S., Bang, S., Cho, S. and Han, H.L.S. (2022). Attentive Multi-Task Prediction of Atmospheric Particulate Matter: Effect of the COVID-19 Pandemic. *IEEE Access,* 10, pp. 10176-10190.
46) SPECK, CORDIS (2022). *Smart Pedestrian Crosswalk for Increased Traffic Safety at Uncontrolled Crossings*. [online] Available at: https://cordis.europa.eu/project/id/887775 [Accessed 12.04.2024].
47) Tegmark, M. (2019). *Life 3.0 (Viața 3.0 - Omul în epoca inteligenței artificiale).* Bucharest: Humanitas.

48) Toribio-Roura, E. (2020). Living with Machines. Ethical Implications and Imaginative Agency as Local Tactic[s] of Dwelling and Resistance in Everyday *Interactions with Artificial Intelligence. Ethics & Politics*, XXII, pp. 63-87.
49) TOTEM SPOON, CORDIS (2019). *Interactive Digital Signage with emotional intelligence for smart cities.* [online] Available at: https://cordis.europa.eu/project/id/867995 [Accessed 12.04.2024].
50) Trivellato, B. (2017). How can 'smart' also be socially sustainable? Insights from the case of Milan. *European Urban and Regional Studies*, 24(4), pp. 337-351.
51) Urban Innovative Actions (2024). *GAVIUS.* [online] Available at: https://www.uia-initiative.eu/en/uia-cities/gava [Accessed 12.04.2024].
52) Wolniak, R. and Stecuła, K. (2024). Artificial Intelligence in Smart Cities—Applications, Barriers, and Future Directions: A Review. *Smart Cities,* 7(3), pp. 1346-1389.
53) Xiao, Z., Qin, Y., Xu, Z., Antucheviciene, J. and Zavadskas, E.K. (2022). The journal Buildings: A bibliometric analysis (2011–2021). *Buildings*, 12(1), pp. 37.
54) Yigitcanlar, T., Desouza, K. C., Butler, L. and Roozkhosh, F. (2020). Contributions and risks of artificial intelligence (AI) in building smarter cities: Insights from a systematic review of the literature. *Energies,* 13(6), pp. 1473.
55) Yu, K., Tan, L., Yang, C., Choo, K.K.R., Bashir, A.K., Rodrigues, J.J. and Sato, T. (2021). A blockchain-based Shamir's threshold cryptography scheme for data protection in industrial Internet of Things settings. *IEEE Internet of Things Journal,* 9(10), pp. 8154-8167.
56) Yu, S., Chen, X., Zhou, Z., Gong, X. and Wu, D. (2020). When Deep Reinforcement Learning Meets Federated Learning: Intelligent Multi-Timescale Resource Management for Multi-access Edge Computing in 5G Ultra Dense Network. *IEEE Internet of Things Journal,* 8(4), pp. 2238-2251.
57) Zavratnik, V., Podjed, D., Trilar, J., Hlebec, N., Kos, A. and Duh, E.S. (2020). Sustainable and Community-Centred Development of Smart Cities and Villages. *Sustainability,* 12(1), 3961.
58) Zhang, J. and Tao, D. (2020). Empowering Things with Intelligence: A Survey of the Progress, Challenges, and Opportunities in Artificial Intelligence of Things. *IEEE Internet of Things Journal,* 8(10), pp. 7789-7817.