|  |  |
| --- | --- |
| cetlogo ***CHEMICAL ENGINEERING TRANSACTIONS***  ***VOL. , 2025*** | A publication of  aidiclogo_grande |
| The Italian Association  of Chemical Engineering  Online at www.cetjournal.it |
| Guest Editors:  Copyright © 2025, AIDIC Servizi S.r.l. **ISBN** 979-12-81206-XX-X; **ISSN** 2283-9216 | |

Mapping the Convergence of Digital Social Innovation, Air Quality Management, and Air Pollution Management: Insights from a Bibliometric Analysis

Leonardo Trianaa, Helinä Melkasb,\*, Anne Pässiläc, Eugenio Morellod, Mikko Happoe

a, b, c Lappeenranta-Lahti University of Technology LUT, Mukkulankatu 19, Lahti, Finland

d Politecnico di Milano, via Bonardi, 3, Milano, Italy

e Ramboll Finland Oy, Puutarhakatu 9 Kuopio, Finland

leonardo.triana.cuesta@lut.fi

Effective management of air quality and pollution control requires innovative solutions. Digital Social Innovation (DSI) offers valuable insights for improving air quality management by integrating technology and fostering community engagement. This study examines the connections between DSI and Air Quality/Pollution Management, through bibliometric analysis and a text dimensionality reduction technique. The findings reveal a convergence in the literature, which can be categorized into two types: Soft Convergence, emphasizing the role of community involvement in policy development, and Hard Convergence, focusing on technological innovations in modeling and air pollution management. These findings underscore DSI's potential to support both social and technological advancements in environmental management.

* 1. Introduction

Air pollution presents one of the most pressing environmental and public health challenges of the modern era, contributing to a significant burden of disease, environmental degradation, and economic loss worldwide (Bulto et al., 2024). To address these challenges effectively, innovative solutions are required that leverage both technological advances and active community engagement (Mehmood & Imran, 2021). Digital Social Innovation (DSI) and Air Quality/Pollution Management (AQPM) represent two promising and complementary approaches. DSI, through digital platforms and community-centered technologies, enables greater public involvement in environmental initiatives (Millard, 2021), whereas AQPM offers tools and policies aimed at monitoring, analyzing, and controlling pollutant sources (Kaginalkar et al., 2022). However, these fields remain largely unintegrated in current research, which limits our ability to create responsive, effective solutions for pollution control and air quality improvement.

The research question guiding this study is: To what extent and in which direction is the literature on DSI and AQPM converging? To address this question, we pursue two main objectives: (1) a bibliometric analysis to assess how frequently DSI and AQPM are referenced together in academic literature, offering insights into the depth of current integration, and (2) applying a dimensionality reduction technique to identify key research topics and their trends. The findings aim to outline pathways for advancing both policy and practice, potentially fostering a more inclusive, data-driven approach in AQPM.

* 1. Materials and methods

A structured bibliometric approach combined with Uniform Manifold Approximation and Projection (UMAP) was employed to map the research intersections between DSI and AQPM, aiming to capture both thematic interplay and topics within these domains. UMAP is a dimensionality reduction technique that transforms high-dimensional data into lower-dimensional spaces while preserving intrinsic structures, enabling the visualization of patterns, clusters, and relationships (Christensen et al., 2023). UMAP facilitated clustering and classification by preserving text structures and improving feature representations (Armstrong et al., 2021).

Web of Science (WoS) provides a reliable and well-structured document repository, enabling comprehensive and detailed analysis. The search strategy uses a specific [WoS search equation](https://www.webofscience.com/wos/woscc/summary/53e8630a-b77e-4f5a-99c4-1bc14e54b5e3-01191b9649/relevance/1) that includes the terms "Digital Social Innovation," "Air Quality Management," and "Air Pollution Management." To ensure precision and relevance, these terms are applied exclusively to titles. The search is restricted to English-language documents to maintain linguistic consistency and includes various publication types, such as articles, conference proceedings, book chapters, review articles, and books.

The bibliometric analysis was performed using Bibliometrix, an open-source software developed in the R language (Aria & Cuccurullo, 2017). This tool has been widely used in scientific mapping in the literature (Belfiore et al., 2022). The choice of Bibliometrix was based on its compatibility with different databases, the multiple analytical functions it offers, its accessibility, and especially because it facilitates the identification of trends (Aria et al., 2024). In this study, R version 4.3.3 and Python version 3.11.11 were utilized. Python’s pyBibX library supported the dimensionality reduction technique UMAP (Pereira, 2022).

* 1. Results and discussions
     1. Bibliometric results

The dataset analyzed spans from 1975 to 2024 and consists of 506 documents authored by researchers affiliated with 525 institutions across 67 countries. These documents were published in 152 sources and collectively cite 15,327 references. The dataset includes various types of publications, with articles comprising the majority at 285, followed by 168 proceedings papers, 19 reviews, 30 book chapters, and 4 books. Of the total documents, 71 were single-authored, while 435 involved multiple authors, resulting in a collaboration index of 3.66, which underscores the prevalence of collaborative research in the dataset. In terms of citations, the dataset has accumulated 5927 citations. On average, there are 4.09 citations per author, 11.71 citations per document, 37.06 citations per source, and 11.29 citations per institution.

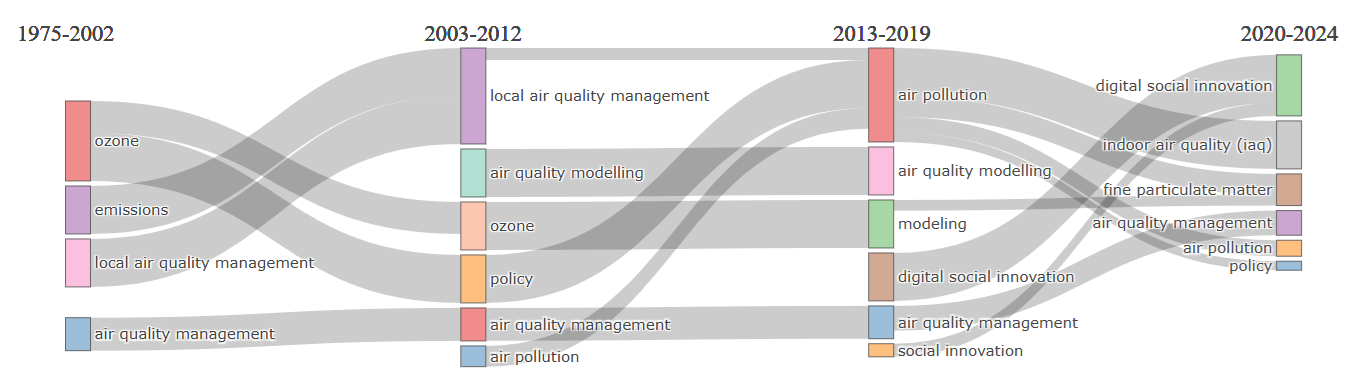
A close-up of a diagram

Description automatically generated

Figure 1: Thematic network map based on author keywords

The thematic network map (Figure 1) highlights distinct areas of convergence within the interdisciplinary domain of air pollution management, air quality management, and their intersection with DSI.

DSI demonstrates strong convergence with Information and Communication Technologies (ICT), big data analytics, smart cities, and social innovation. Smart cities further show convergence with machine learning and air pollution, establishing a clear technological pathway that links air-related topics with DSI. Additionally, there is notable soft convergence between DSI and policy. Air Quality and Air Pollution exhibit significant convergence in their connections to modeling strategies, featuring terms such as air quality modeling, air pollution, particulate matter emissions, dispersion modeling, emission inventory, and decision support systems. Air quality shows strong convergence with uncertainty and management, while air pollution highlights convergence with climate change. Both air quality and air pollution demonstrate strong convergence with policy, emphasizing the central role of governance in addressing environmental challenges.



*Figure 2: Evolution of research topics (1975–2024)*

Figure 2 illustrates the evolution of research topics in DSI and AQPM from 1975 to 2024, highlighting shifts in focus across four distinct periods. Early research (1975–2002) concentrated on emissions and local air quality, while the 2003–2012 period emphasized air quality modeling, policy, and optimization. Between 2013 and 2019, the introduction of digital social innovation and the growing importance of indoor air quality in the air quality management field became evident. From 2020 to 2024, trending topics included fine particulate matter and indoor air quality.

A graph of a graph

Description automatically generated with medium confidence

*Figure 3: Trending author’s keywords*

Figure 3 further details temporal patterns across key terms, with foundational topics like modeling, GIS, and emission peaking between 2002 and 2012. Emerging concepts such as local air quality management gained traction between 2006 and 2016, while terms like uncertainty, ozone, and air quality marked the period from 2010 to 2019, showcasing an increasing focus on nuanced and complex assessments. Post-2015 shows the introduction of socio-technological approaches, as evidenced by the rise of social innovation and digital social innovation. From 2020 onward, concentrated interest in urban air quality management, urban planning, smart cities, and ICT underscores the integration of digital tools in addressing air quality challenges, aligning with broader urban sustainability trends.

* + 1. UMAP results

Uniform Manifold Approximation and Projection (UMAP) helps identify six primary clusters, each representing distinct thematic areas in air quality management and digital social innovation research. These clusters (Figure 4) highlight the evolution of research topics over time (Figure 5).

A diagram of different colored circles

Description automatically generatedFigure 4: Cluster visualization of research themes using UMAP

The identification of the clusters’ main topics was achieved by analyzing the 20 most frequent n-grams in the titles and abstracts of each document, focusing on terms consisting of one to three words. Document frequency, representing the number of documents in which a specific n-gram appears, provides insights into the thematic focus of each cluster. Common terms such as "air," "quality," and "management," which are highly frequent across all clusters (100% except for the DSI cluster), were excluded as primary naming candidates to ensure specificity. Below are the descriptions of the six identified clusters:

The "Urban Air Quality Management" cluster comprises 121 documents and highlights challenges associated with urban environments. The analysis revealed terms such as "urban," appearing in 67 documents, which accounts for 55.4% of the cluster, "pollution," found in 85 documents, representing 70.2%, and "emissions," occurring in 39 documents, or 32.2%.

The "Digital Social Innovation" cluster consists of 42 documents and focuses on the use of digital tools to drive social innovation. Main terms include "digital social innovation" and "social innovation," each present in 40 documents, corresponding to 95.2% of the cluster, alongside "DSI," which appears in 20 documents, or 47.6%.

The "Emissions and Health" cluster is composed of 112 documents and investigates the impacts of emissions on public health. Key terms include "pollution" occurring in 40 documents, representing 35.7%, "emissions" appearing in 35 documents, or 31.3%, and "health" found in 29 documents, accounting for 25.9%.

The "Local Air Quality Management" cluster comprises 91 documents and examines region-specific strategies and governance for air quality management. Frequently used terms such as "local" appearing in 65 documents, representing 71.4% of the cluster, "authorities" found in 40 documents, or 44.0%, and "process," also present in 40 documents, emphasize a focus on localized approaches. Additionally, phrases like "local air quality" and "local authorities" occur in 40 and 36 documents, respectively, further reinforcing the geographic specificity of this cluster.

The "Indoor Air Quality Management" cluster includes 90 documents and focuses on air quality issues within confined spaces. Terms such as "indoor", "concentrations" and "control" each appearing in 25 documents, account for 27.8% of the cluster. These terms highlight the research emphasis on managing air quality in indoor environments, setting this cluster apart from those addressing broader outdoor concerns.

The "Programming Decision Models" cluster is composed of 50 documents and emphasizes computational and methodological approaches to decision-making in air quality management. Frequently occurring terms include "decision" present in 37 documents, representing 74.0%, "model" appearing in 28 documents, or 56.0%, and "programming" found in 19 documents, which accounts for 38.0%.

The "Digital Social Innovation" cluster shares conceptual similarities with AQPM clusters, particularly in its use of "data" and "model" as recurring n-grams. While DSI shows a document frequency of "data" (12%) and "model" (14%), clusters such as Urban Air Quality Management (39% and 27%, respectively) and Indoor Air Quality (39% and 31%, respectively) exhibit higher frequencies. This shared emphasis highlights a reliance on data-driven insights and modeling for analysis and decision-making. Similarly, the "Programming Decision Models" cluster, where "model" dominates with a document frequency of 56%, reflects a strong methodological connection to DSI, as both focus on developing frameworks to optimize complex systems.

The “Digital Social Innovation” cluster demonstrates similarities with other clusters through its focus on "policy" highlighting its connection to governance and societal frameworks. With a document frequency of 17% for "policy" DSI aligns closely with clusters like Emissions and Health (18%) and Urban Air Quality Management (26%), where policy development plays a significant role in addressing challenges. The strongest similarity emerges with Local Air Quality Management, which emphasizes "policy" (33%). This connection suggests that DSI allows policy development (Novak et al., 2018) and policy acts as a critical mechanism for air pollution management (Badach et al., 2020).

Figure 5 illustrates the evolution of each cluster using histograms with 4-year intervals, spanning the period from 1970 to 2024. The vertical axis represents the relative proportion of documents within each cluster, expressed as probabilities, providing a visualization of how these clusters have changed over time.

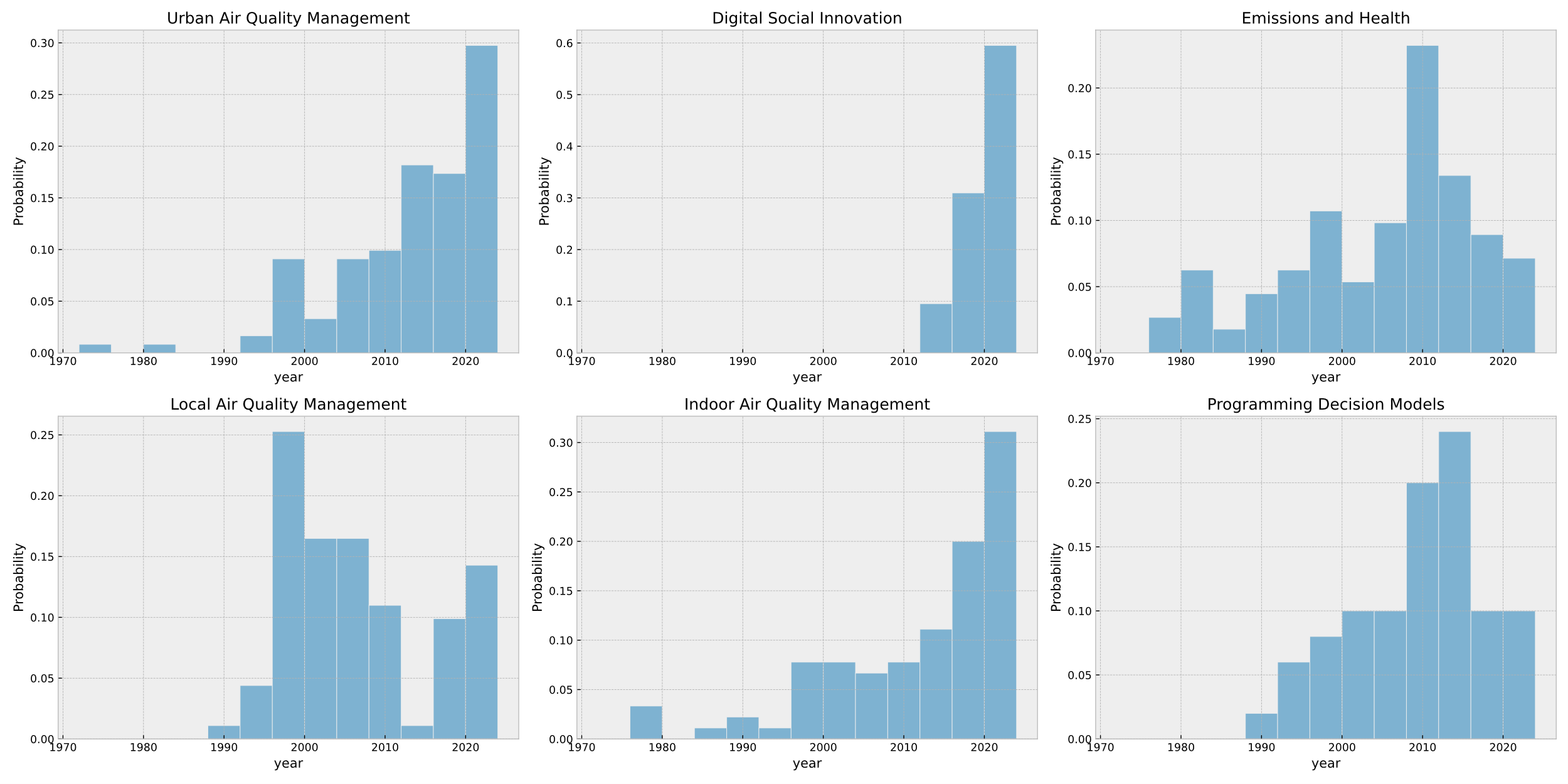


Figure 5: Temporal evolution of research UMAP clusters

The research trends across the six clusters reveal varied trajectories, reflecting the evolution of focus in distinct areas. Urban Air Quality Management and Indoor Air Quality exhibit significant long-term growth, with marked accelerations in recent years, underscoring their growing relevance to environmental and public health concerns, particularly considering challenges such as the COVID-19 pandemic. DSI, as a relatively new field, has demonstrated exponential growth since 2012, highlighting the increasing role of digital technologies in addressing societal challenges. In contrast, Emissions and Health experienced early and sustained growth, peaking in 2008-2012 but subsequently declining, suggesting a shift in focus or maturation of foundational research. Local Air Quality Management shows a pattern of early growth and subsequent fluctuation, with a recent resurgence reflecting renewed interest in localized strategies. Finally, Programming Decision Models displays steady growth until 2016, followed by stabilization, likely indicating the field’s maturation and integration into broader applications.

* 1. Conclusion

Bibliometric and UMAP analyses reveal two types of convergence: "Soft Convergence," emphasizing community-driven policy development, and "Hard Convergence," focusing on technological innovations for air pollution modeling. Key clusters, such as "Digital Social Innovation" and "Urban Air Quality Management," highlight shared themes like policy, data, and models, showcasing an integrated research trajectory. This convergence suggests the need for an integrative approach that combines social innovation and technological development to address air quality challenges. The exponential growth of DSI research since 2012, particularly in areas like ICT, aligns with AQPM's emphasis on data-driven models and emission control. Cluster evolution shows consistent growth in urban and indoor air quality research, reflecting a continued interest in technologically sophisticated strategies for environmental management. Together, DSI and AQPM can provide a complementary framework for addressing complex air quality issues through innovative, community-focused solutions.

Acknowledgments

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them. This study has received funding from the European Union through Marie Skłodowska-Curie actions under the Grant Agreement no 101081466: SEED - Systems and Engineering Science Doctorate number.

References

Allaoui M., Kherfi M.L., Cheriet A., 2020, Considerably improving clustering algorithms using UMAP dimensionality reduction technique: A comparative study, 317–325.

Aria M., Cuccurullo C., 2017, bibliometrix: An R-ability for comprehensive science mapping analysis, Journal of Informetrics, 11(4), 959–975.

Aria M., Cuccurullo C., D’Aniello L., Misuraca M., Spano M., 2024, Comparative science mapping: A novel conceptual structure analysis with metadata, Scientometrics, 1–27.

Armstrong G., Martino C., Rahman G., Gonzalez A., Vázquez-Baeza Y., Mishne G., Knight R., 2021, Uniform manifold approximation and projection (UMAP) reveals composite patterns and resolves visualization artifacts in microbiome data, MSystems, 6(5), 10–1128.

Badach J., Dymnicka M., Baranowski A., 2020, Urban vegetation in air quality management: A review and policy framework, Sustainability, 12(3). MDPI. https://doi.org/10.3390/su12031258

Belfiore A., Cuccurullo C., Aria M., 2022, IoT in healthcare: A scientometric analysis, Technological Forecasting and Social Change, 184, 122001.

Bulto T.W., Chebo A.K., Werku B.C., Debele K.N., Kloos H., 2024, Visualization and analysis of urban air quality management using bibliometric techniques and social network analysis for the period 1975 to 2022: A review, Environmental Health Insights, 18. SAGE Publications Ltd. https://doi.org/10.1177/11786302241252733

Christensen A.J., Sen Gupta A., Kirsteins I., 2023, Analysis of braid manifolds using uniform manifold approximation and projection, The Journal of the Acoustical Society of America, 153(3\_supplement), A178–A178.

Kaginalkar A., Kumar S., Gargava P., Kharkar N., Niyogi D., 2022, SmartAirQ: A big data governance framework for urban air quality management in smart cities, Frontiers in Environmental Science, 10. Frontiers Media SA. https://doi.org/10.3389/fenvs.2022.785129

Mehmood A., Imran M., 2021, Digital social innovation and civic participation: Toward responsible and inclusive transport planning, European Planning Studies, 29(10), 1870–1885. Routledge Journals, Taylor & Francis Ltd. https://doi.org/10.1080/09654313.2021.1882946

Millard J., 2021, Networks, communities and value chains in digital social innovation for social services, In D. Davide, A. Gaggioli, & G. Misuraca (Eds.), Perspectives for Digital Social Innovation to Reshape the European Welfare Systems, 13, 88–114. IOS Press. https://doi.org/10.3233/STPC200006

Novak J., Becker M., Grey F., Mondardini R., 2018, Citizen engagement and collective intelligence for participatory digital social innovation, In S. Hecker, M. Haklay, A. Bowser, Z. Makuch, J. Vogel, & A. Bonn (Eds.), Citizen Science: Innovation in Open Science, Society and Policy, 124–145. Univ Coll London Press - UCL Press.

Pereira V., 2022, Project: pyBibX, File: Pbibx. Py. Retrieved from GitHub repository.