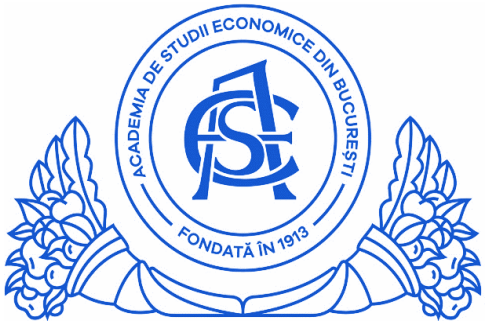
THE BUCHAREST UNIVERSITY OF ECONOMIC STUDIES

The Council for Doctoral Studies

Doctoral School of Economic Informatics



DOCTORAL THESIS

SUMMARY

PhD thesis title:

COMPUTER-BASED SOLUTIONS FOR DEVELOPING RESILIENT SMART CITIES

Author:

Denis-Cătălin G. ARGHIR

PhD supervisor:
Prof.univ.dr. Adriana-Elena DÂRDALĂ

Bucharest, 2024

Table of contents

Acknowledgments 5
Summary/ Synthesis 6
List of abbreviations 11
Introduction 12
Chapter 1: Research context and importance of the topic 13
1.1. Research context 13
1.2. Motivation for choosing the topic 26
1.3. Research questions, hypotheses, and objectives 29
1.4. Thesis structure 33
Chapter 2: Urban development trends and the influence of emerging technologies 35
2.1. The evolution of urban settlements over time 36
2.2. State of knowledge from theoretical and methodological perspective 40
2.2.1. Bibliometric research on key concepts 40
2.2.2. Identification of disruptive factors that can influence urban resilience 50
2.2.3. Definition, dimensions, and characteristics of a resilient smart city 54
2.3. The dynamics of urban reinvention: international and national perspectives 59
2.3.1. Coordination of urban reinvention in international initiatives 59
2.3.2. National initiatives driving the development of smart cities in Romania 62
2.4. Methods for urban resilience assessement 70
2.4.1. Indicators for assessing urban resilience. Study of representative metric systems 70
2.4.2. Comparative analysis of urban resilience indicator systems 77
2.4.3. Limitations of indicator systems 80
2.5. Innovative smart solutions for the development of resilient smart cities 81
2.5.1. Computer-based solutions used at international level for urban resilience 81
2.5.2. Exploring smart solutions used at the national level for urban transformation 89
2.6. Conclusions regarding chapter 2 95
Chapter 3: Paradigms and IT tools used in the smart cities context 97
3.1. Data-driven cities 98
3.2. Urban data sources 100
3.2.1. Official data from digital platforms and government databases 102
3.2.2. Derived data external to the traditional city flow 103
3.2.3. Data collection techniques 107
3.2. Management of big data volumes generated by cities 109
3.2.1. Structured, semi-structured, and unstructured data in the big data context 111
3.3.2. Preprocessing numerical and text data through specific techniques 111

3.2.3. Data streams ... 112
3.3. Data processing and analysis using artificial intelligence ... 113
3.3.1. Machine learning techniques ... 114
3.3.2. Supervised learning models ... 115
3.3.3. Unsupervised learning models ... 122
3.4. Conclusions regarding chapter 3 ... 124
Chapter 4: Practical research for the development of resilient smart cities ... 125
4.1. Development of a microservices-based architecture in the smart cities context ... 126
4.2. Development of data processing modules from official sources ... 130
4.2.1. Identification of primary data sources for the indicators set creation ... 132
4.2.2. Development of a software module for data acquisition from multiple primary sources ... 134
4.2.3. Data integration and construction of indicators for Romanian cities resilience assessement ... 140
4.2.4. Use of principal component analysis for composite indicators construction ... 158
4.2.4. Ranking of cities based on scores obtained through indicators ... 166
4.3. Computer-based components for data acquisition, processing, and analysis from alternative sources in the context of resilient smart cities ... 168
4.3.1. Data collection and integration module from social media APIs ... 169
4.3.2. Advanced preprocessing of social media data ... 174
4.3.3. Data clustering and sentiment analysis from social media real-time flows ... 176
4.3.4. Development of a predictive model architecture for classification of social media data ... 179
4.3.5. Data acquisition and integration module from sensor networks and mobility data sources regarding citizen mobility ... 183
4.3.6. Development of regression models architecture for sensor network and mobility data ... 188
4.4. Integration of computer-based solutions in applied case studies ... 190
4.4.1. Exploration of social media data regarding the immediate impact of the pandemic on consumer behavior ... 190
4.4.2. Exploration of pre-labeled messages regarding the impact of shocks and stressors on urban dynamics ... 198
4.3.3. Exploration of the impact of mobility and meteorological data on air quality ... 208
4.6. Conclusions regarding chapter 4 ... 218
Chapter 5: Conclusions, contributions and research limitations ... 220
Results dissemination ... 224
Bibliographic references ... 228
Annexes and code scripts ... 242
List of tables ... 245
List of figures ... 245
List of annexes ... 246

Summary of the PhD Thesis

Purpose: The doctoral research took place in a sequence of three distinct stages, focusing on the investigation and development of new perspectives related to the main subjects of interest, which represents topics of relevance to both academic discourse and the public agenda at the global, European, and national levels, more precisely, the *study of smart and resilient cities topics*. By adopting a rigorous and systematic approach based on grounded theorizations, empirical analyses, and experimental studies, the foundation of an in-depth understanding of the examined issues was created. This academic investigation not only brings up previously unexplored aspects but also contributes to knowledge advancement in the field.

Statement of the problem: In recent years, there has been a significant increase in public concern about mass mobility, rural exodus, social inequality, public safety, climate change, environmental pollution, access to public services, and so on. The integration of digitization, technological innovation, artificial intelligence (AI), e-governance, renewable energy promotion, sustainable urban development, high-speed internet connectivity, the automation of daily tasks, and data sharing in government clouds have emerged as elements that are reshaping the modern life landscape, making the transition from traditional to smart cities. As cities evolve into dynamic hubs of connectivity and innovation, the integration of sensors, wireless networks, and automated data processing is becoming a transformative force that can improve various aspects of citizens' lives. In the context of recent global events, the most significant being the COVID-19 health crisis, the critical role of modern technologies in strengthening cities' resilience has also been emphasized. The adoption of smart technologies has become vital for re-evaluating and optimizing urban management strategies. The resilience and responsiveness of cities during crises are directly influenced by their ability to capitalize on emerging technologies to quickly respond to various challenges—natural or anthropogenic shocks and stresses.

Methodological approach: The study of intelligent cities, in the research, began by performing an exploratory analysis of the evolution of cities, from the most rudimentary urban forms to contemporary developments, which led to the need for smart and resilient city concepts. Next, a rigorous evaluation of the specialized literature from scientific databases was performed to substantiate the research, including bibliometric analysis to evaluate the authors and the studies that have adopted the concept in the urban context to surprise niches and directions that are not sufficiently explored, as well as the study of frameworks, strategies, and

projects underlying urban development. The next step was to identify and track some metrics systems for measuring and assessing cities' smartness and resilience, useful for measuring the progress and efficiency of interventions in the direction of transforming cities into entities more adapted to contemporary challenges. This approach falls into the current concerns of creating mirror copies of urban systems known as "digital twins". These digital replicas allow real-time monitoring and analysis of urban infrastructure, facilitating better management of resources and services. They also offer the opportunity to simulate scenarios to anticipate and prevent problems, thus optimizing infrastructure operations and interdependence. Thus, urban planning can be improved by integrating precise and updated data, thus contributing to cities' increasing resilience. In parallel, certain relevant computer-based solutions were investigated from various studies or projects conducted in different countries, because digital technologies play a central role in facilitating this process of urban transformation. Following the analysis of these systems, the need to develop a specific set of indicators of evaluation in Romania was concluded. This empirical and personalized approach followed a multidimensional methodology, applying principal component analysis (PCA) to determine an objective manner of assigning weights in the construction of aggregate indicators and using dashboard representations for interactive comparison of urban areas.

The identification of viable and openly exposed data sources was provocative because no centralized platform openly exposes this type of data at the national-level. To facilitate the automation of the data gathering process, a web-based framework was developed. The extracted data were stored in Mongo DB and then subjected to various analyses.

Results: By searching on the basis of relevant keywords, multiple scientific papers were identified and refined to highlight the most relevant ones. These papers provided a clearer understanding of the subject and deepened the current stage of knowledge. Historical open data sources were investigated for the recent time period (which include the period before the outbreak of the global pandemic, the decrease of the pandemic peak and until the time of declaring the medical emergency end), available in various national institutions, to develop an indicator system that matched the purpose of the research. 4 principal dimensions were formed to cover the Social, economic, and well-being of citizens; Infrastructure, public, and technological utilities; Industry and safety governance; and Climate, environment, and green area. In turn, these dimensions include sub-indicators that reveal multiple urban aspects. In addition, computerized modules have been designed for automatic data collection and extraction, including sources that do not expose data in an accessible format, for which mini-software-robots and scraping were used.

Because the indicators offer an overview of the situation from a certain past period, most often with monthly, semi-annual, or annual frequencies, the research considered finding alternative solutions for real-time understanding of the city-level situation by devising proxy measurements and obtaining data from unconventional sources recognized by the academic world. Through computerized procedures, data were analysed in detail and used for their predictive and prescriptive nature, providing better information over time horizons, where the presence of the effects of shocks or stressors has affected various aspects of cities. Data analysis models were built by applying supervised and unsupervised machine learning algorithms to historical sources and real-time streaming data and were continuously optimized to improve prediction performance and accuracy. The most remarkable results obtained from the application to urban areas in Romania (325 small, medium, and large cities) were highlighted following experimental exercises carried out using data from social networks, sensor networks that expose data related to air quality and meteorological indicators, and citizen mobility-related data, obtaining a snapshot that captured the effects of various rapid-impact trigger events on cities.

The validation of the elements identified in the research was achieved by publishing scientific articles and disseminating the findings at internal and international conferences, as well as in peer-reviewed articles indexed in Web of Science, Scopus and other international databases.

Keywords: smart city resilience, evaluation metrics, digital twin, artificial intelligence, machine learning, natural language processing.