ENABLING CITY DIGITAL TWINS THROUGH URBAN LIVING LABS

Petar O. Hristov¹, Desislava Petrova-Antonova¹, Sylvia Ilieva¹, Radoslav Rizov²

¹GATE Institute, Sofia University “St. Kliment Ohridski”, Sofia, Bulgaria – (petar.hristov, desislava.petrova, sylvia)@gate-ai.eu
²Telelink Business Services, Sofia, Bulgaria – radoslav.rizov@tbs.tech

Commission I, WG I/6

KEY WORDS: City digital twin, sensor network, real-time data platform, urban living labs.

ABSTRACT: The population density in urban areas is rapidly rising, leading to a constant need for new infrastructure and services for citizens. To reduce the time to implementation and optimise the monetary cost of various solutions, the plans and policies of local authorities and stakeholders would benefit from undergoing a series of virtual stress tests. To this end, prescriptive and predictive technologies are widely adopted to optimise city planning and to understand the urban processes and environment such as air pollution and transportation. Nevertheless, holistic sandboxes tightly integrated with cities are still largely lacking. The city digital twin is a promising concept that provides a tool for exploration of new solutions in a controlled environment before their deployment. The digital twin is a virtual replica of the real city, which collects data from the infrastructure, processes and services using not only the available systems, but also purposely built connected devices and sensors. In this context, the establishment of urban living labs facilitates the monitoring and understanding of urban processes and enriches the digital twin with highly-relevant data. This paper presents an urban living lab, under deployment in the district of Lozenets in Sofia, Bulgaria. It is part of a larger initiative for developing a city digital twin of Sofia to support the design, exploration, and experimentation of different solutions. The living lab is equipped with sensors for monitoring air quality, atmospheric parameters, noise pollution and pedestrian flows. In addition, a Light Detection and Ranging (LiDAR) system is realised as an edge computing facility at one of the busiest intersections of the district. Along with the equipment, the paper describes the architecture and components of the platform for data collection, storage, processing, and visualization. Finally, high-priority studies are presented, and their demographic and economic impact is discussed.

1. INTRODUCTION

Cities are the largest and most complex human artifacts, setting the physical frames for an increasing majority of the world’s population. They already consume most of the global natural resources and produce a large portion of the planet’s pollution and waste, giving them a key role in climate change. Yet, according to the United Nations, close to another 2.5 billion people are estimated to relocate in cities in the next 30 years (Unite Nations, 2019a; United Nations, 2019b).

1.1 Motivation and Challenges

The development of future cities that are liveable, sustainable, and resilient and yet accommodate such a rapid population rise requires careful strategic planning. The problem is further complicated by the need to ensure the expansion will not jeopardise the health and wellbeing, livelihood, and prospects of the existing population, and will not compromise on the solution of immediate problems. To achieve this, the urban environment, its processes, and characteristics must be analysed and understood. Virtually all areas of science and engineering which require the investigation of complex phenomena, resort to the use of computer models. However, the complexity and interconnectedness of different modules of the urban environments, which may traditionally be described in separate models, have dictated the use of city digital twins (Curetan and Dunn, 2021; Raes et al., 2021; DTCC, n.d.).

Digital twins (Wagg, 2020) consist of a chain of coupled computational and statistical models, enriched by live data, which must present a faithful reflection of a physical asset, in this case the urban environment. The carefully constructed virtual system allows analysts, designers, and engineers to make adequate and credible inferences about the city, without the need for expensive or infeasible physical experimentation. A key aspect of digital twins, enabling the high-quality representation of the physical counterpart, is the stream of information between the physical and virtual domains (Ketzler et al., 2020). Given the complexity of the urban environment, a city digital twin doubles down on the requirement for the availability of relevant and representative data and opens up new challenges. In contrast to traditional modelling, performing synthetic experimentation in a closely controlled setting is unlikely to provide a faithful depiction of urban processes, including the variety and scale of data required to analyse the urban environment. To ensure the necessary veracity of information, an urban living lab is required, where city life itself is the subject of experimentation (Talari et al., 2017).

Such a facility comes with two principal challenges. Firstly, the design and implementation must consider the processes of interest and plan for the provision of equipment to adequately measure relevant phenomena. Measurements must possess sufficient accuracy and reliability, if they are to be used in decision making, which is the primary goal of the city digital twin. The lab must also possess a modular structure and be designed in a way which allows for future integrations to be carried out easily. Secondly, the software platform powering the lab’s network must be capable of process large volumes of variable quality data coming in different formats, including numerical measurements, point clouds and images, and transferring it from individual sensor stations to the final user. Moreover, the data is likely to carry a complex suite of uncertainties, such as range values, approximations, and downright unknowns (Gubbi et al., 2013).

¹ Corresponding author
1.2 Related Work

The importance of understanding the processes that take place in cities and using this knowledge to improve the quality of urban life is evidenced by the number of national and international initiatives aiming at developing city-scale living laboratories (Hancock et al., 2012).

This drive has resulted in the term urban living lab having different meanings (Bergvall-Kåreborn et al., 2009). In the larger sense, the term encompasses facilities and infrastructure open to various collaborators wishing to develop and test their solutions for smarter urban areas. Such living labs often comprise built facilities such as those part of the Internet of Things in Amsterdam², participant cities in The Urban Lab of Europe initiative², or the ISCAPE Living Labs³ with locations in the UK, Ireland, Italy, Germany, Belgium and Finland. Alternatively, programmes such as the Living Lab Bus⁴ deployed in the city of Helsinki, provide a technological testbed on the go. The general idea of these living labs is to encourage sustained growth in real-world conditions (Zalokar, 2020).

The meaning this paper adopts is the strictly technical one of a collection of laboratory devices (sensors) used in the live urban environment (Schumacher, Feurstein, 2007). Such facilities exist on different scales and vary in their purpose and architecture (Tokes, 2017). For example, the DOLL Living Lab (Larsen, Hammer 2020) provides air quality, weather, and noise sensing, as well as innovative lighting solutions. Another example is the UK Collaboratorium for Research in Infrastructure & Cities (UKCRIC)⁵ which consists of networks of weather, air quality, and energy efficiency analysis equipment in five major UK cities. The UrbanSense platform⁶, which comprises a set of air quality and solar irradiation sensors deployed at important locations around the city of Porto, Portugal, and the urban air monitoring stations of the Array of Things project⁷ in the city of Chicago, serve as further examples of the increased interest towards monitoring urban life. In addition to these programmes, there are many other initiatives focusing on citizen science and experimentation, providing low-cost equipment which non-specialists can use to provide data about their local environment. Despite the variety of applications all these labs inherently employ the Internet of Things (IoT) philosophy (Zanella et al., 2014), also fundamental to the work described in this paper.

There are several initiatives in Sofia particularly focused on the quality of the atmospheric air and weather conditions around the city. These include several stations operated by the Executive Environmental Agency, the municipal platform AirThings⁸, the INNOAIR⁹ project, as well as the citizen science tools Luftdaten¹⁰ and AirBG¹¹.

1.3 Outline of Contributions

This paper describes the planning and design phases of the construction of an urban living lab in Sofia, Bulgaria. When constructed, this lab will serve two main purposes. Firstly, it will provide verified and open data about the air quality, noise pollution, weather conditions, people and traffic flows and more to interested stakeholders, enabling the monitoring and evaluation of different processes and the implementation of solutions. Secondly, the lab and more precisely the information it will provide, will form part of the larger initiative for developing a digital twin of the city, to enable its sustainable growth. The paper outlines some of the envisioned uses of the data collected and how this fits in with efforts for improved urban quality of life.

The remainder of this paper is organised as follows. Section 2 describes the hardware of the Lozenets living lab, while Section 3 describes its software platform. Several high-priority use cases for the living lab, in various stages of maturity are outlined in Section 4. The paper concludes in Section 5 with some final remarks and points for future work.

2. SENSOR NETWORK

The urban living lab consists of 12 air quality monitoring stations, 60 noise measurement stations, 50 pedestrian counting radars and a LIDAR system for vehicle counting and traffic analysis. The network is designed in a way which provides good coverage of Lozenets and places sensors in locations that are key to understanding important metrics of its complex environment.

2.1 Description of the Pilot Area

The district of Lozenets in Sofia is selected as a pilot area for development of the digital twin and the Living Lab in particular. It is located on a hill, south of the old town of Sofia and extends to the northern foothills of Vitosha Mountain. It covers an area of 9.24 km², almost 30% of which is covered by forests, with two small rivers crossing the district. As shown in Figure 1, the area contains a blend of old and new architecture, residential, office, healthcare and educational buildings, shopping centres, and the city’s zoo.

![Figure 1. Map of the district of Lozenets with main features.](https://example.com/map.png)

The district is covered by a network of busy, high-speed roads and small neighbourhood scale streets and cul-de-sacs. Various different modes of public transportation, including internal combustion engine, hybrid-electric and electric buses, and trams are also part of the urban landscape. Moreover, despite Lozenets being the greenest part of Sofia, many challenges arise due to the rapidly expanding population of the district.

---

³ https://livinglabs.iscapeproject.eu/
⁴ http://livinglabus.fi/use-cases/
⁵ https://www.ukrcic.com/
⁶ https://futurecities.up.pt/hybrid-sensor-networking-testbed/
⁷ https://arrayofthings.github.io/index.html
⁸ https://platform.airthings-project.com/
¹⁰ https://sensor.community/en/
¹¹ https://www.airbg.info/
These features make Lozenets an ideal candidate host for an urban living lab, focused on gathering and using data in parametric urban planning and the analysis and simulation of air quality, pollution dispersion and wind comfort, to provide decision support to architects and government officials.

2.2 Lab Sensors and Connectivity

The air quality stations comprise an atmospheric composition analysis module (ACAM) and a meteorological module (MetM). The ACAM houses four electrochemical gas concentration (CO, NO2, O3, SO2) sensors (Williams, 2020) and a particulate matter sensor using laser-scattering technology (Bohren, Huffman, 2008), to detect particles from 0.35 μm to 40 μm (main interest is in PM1, PM2.5 and PM10). The detection ranges for each of the sensors were chosen in accordance with the typical levels of pollutant concentration in an urban environment. The ACAM is equipped with a proprietary active climatisation system which maintains optimal conditions for the module to operate in, improving the accuracy of its measurements.

The MetM provides measurements of ambient temperature, relative humidity, pressure, CO2 concentration, levels of precipitation, wind speed and direction, and levels of noise. The entire station, shown in Figure 2 is packaged in a weather resistant (IP65) case and can be powered by either its own solar panel-battery system, or via 240 volts AC from the grid. Each station is supplied with a GPS module which facilitates the synchronisation of separate stations in an integrated network. All measurements are collected autonomously and transmitted via standard wireless communication protocols, together with metadata about the state of health of the station.

Figure 2. Air quality station in "summer" configuration.

The noise measurement stations are equipped with a dual microphone and an electronic compass for determining the strength and direction of noise. Each station is enclosed in an IP54-rated container. It is powered via a dedicated solar panel and has an integrated emergency battery (14-day endurance under normal transmission rates). Each station transmits noise measurements and metadata autonomously via long range protocols.

The pedestrian counting radar modules (IP67 casing), are equipped with a two-channel transceiver which provides counting of pedestrians crossing the sensor field of view (see Figure 3) from left to right and from right to left within a 10-meter range. The modules have a built-in 16-bit memory for storing pedestrian counts for periods from 1 minute to 24 hours before transmission. Data is transmitted through a Long-Range Wide Area Network (LoRaWAN) controller. Each module is powered by a solar panel and contains a rechargeable nickel-metal-hydrate (NiMH) battery, as a back-up.

Figure 3. Operational logic of the pedestrian radar module.

The Light Detection And Ranging (LiDAR) system part of the urban living lab is realised as an edge computing system at one of the busiest intersections in Lozenets. The system infrastructure is composed of four layers, shown in Figure 4, primarily differentiated by the volume fraction of original data present at each layer.

Figure 4. LiDAR system connection and data reduction.

The first layer comprises four high-resolution LiDAR sensors, generating point cloud data. The sensors are positioned at each corner of the intersection and are mounted on traffic light poles to maximise their field of view.

The second layer of the infrastructure is a set of four data consolidation node boxes, one for each sensor. These devices are responsible for the initial data processing and reduce the data to approximately 90% of its original volume. The node boxes also act as a regulated power supply for the respective sensor.

The third layer of the system is a fusion augmented LiDAR box, responsible for collecting and amalgamating the information from the four system branches. The fusion box combines, further process and thus reduces the data to approximately 5% of its original volume. In this state the data is sent to the fourth layer.

The fourth layer of the infrastructure is the host computer, receiving the data for further processing and analyses. The communication between the system and the host computer is implemented through a virtual private network (VPN) to ensure the secure transmission of data.

The entire system is powered by 240 volts AC from the grid and is equipped with an uninterruptable power supply (UPS) for increased operational reliability.

2.3 Complementary Infrastructure and Services

There are some devices already in place in the city of Sofia, primarily focused on air quality monitoring. The Bulgarian...
Executive Environmental Agency (EEA) operates six stations in key locations around the city. These stations provide high-quality measurements, but the data is not freely available for stakeholder interaction – one of the principal aims of the urban living lab. Another air quality monitoring suite is that operated under the AirThings initiative, comprising 22 stations across Sofia. The data can be accessed both as visual summaries and in real time through an API. No coordinated activities for noise measurement, pedestrian counting, or LiDAR measurements exist in Sofia or Bulgaria, to the best of the authors’ knowledge.

3. DATA PLATFORM

This section presents the architecture and components of the data platform developed for data collection, transfer, processing, and visualisation. The data platform is based on micro-services and adopts Docker containers technology. Applications and pipelines can be run in Docker containers in an isolated, self-contained manner. Thus, Docker containers allow the efficient and portable distribution and execution across a wide range of computing platforms (Gerlach et al., 2014).

3.1 Platform Architecture

The architecture of the data platform, shown in Figure 5, consists of distributed micro-services, which is packaged in Docker containers based on Kubernetes. Kubernetes provides core concepts for Azure Kubernetes Service (AKS) and allows a reliable scheduling of fault-tolerant application workloads. Two databases are implemented as follows: (1) Timeseries database, using InfluxDB, and (2) Relational database for data storage, using PostgreSQL.

3.2 Platform Components

The data platform implements 9 micro-services, described in this section. The IoT Message Hub Service provides a message queue, which is used for message routing. It allows a bidirectional communication with the integrated sensors and devices for sending configuration and control commands and for transfer of timeseries data. The timeseries data in the message queue is transferred to the Timeseries database by the Data Transfer Service. This service performs data normalisation and calibration. As a result, three types of data are stored – raw data, normalized data, and calibrated data.

The Event Stream Analytics Service processes the calibrated data, correlating it with previous measurements and specific rules for identifying potential errors. The Device Management Service sends configuration and control commands to the integrated sensors and devices using the IoT Message Hub. It implements a REST interface for abstraction and sending commands. The Telemetry Service provides an abstract layer, which is used for access to the data in the database. It delivers a telemetric data flow for real-time visualisation of data in the user interface. The Maps Service performs a geolocation of data, which is used when the data is shown on a map.

The Persistent Service provides an abstract layer, which implements data access mechanism used to persist and get data from the relational database. Thus, the relational database can be easily changed without the need for a change in the platform’s architecture. The Persistence Service enables the atomicity, consistency, isolation, and durability of read/write operations in the relational database.

The authentication of the platform’s users is performed by an Authentication Service. If the user is successfully authenticated with username and password, the service returns a JWT token. The token is used subsequently to authorize any request made by the user. The Logging Service logs the platform’s behavior by collecting and centralising log files from any origin. It supports the diagnostics and root cause analysis of failures of the platform.

4. DATA UTILISATION

The complex urban morphology of the district of Lozenets presents an abundance of potential applications for the data gathered from the sensor network described in Sections 2 and 3 (Rathore et al., 2016). In this section some high-priority studies are presented, and their demographic and economic impact is discussed. Furthermore, several ways in which data from the urban living lab will support other efforts in building the city digital twin are examined. A key idea behind these use cases and the digital twin in general, is that any relevant uncertainty in measurements and consequent analyses will be considered. More details on this are given in the Section 4.4.

4.1 Spatial-temporal analysis of air pollution

The immediate value of the air quality monitoring stations is that their measurements will be used to observe and analyse air quality and air pollution trends. Work is already ongoing to develop approaches for using point measurements of key gas concentrations and meteorological parameters to create a digital, user-oriented map of the air quality in Lozenets. Statistical and machine learning methods will underpin the discovery of important features in both time and space. In addition to real time measurements, the platform will be able to generate statistical predictions in time to inform and assess the effectiveness of air management policies. It is envisioned that the map operates at the resolution of individual addresses to enhance the understanding of phenomena driving the air quality (Jensen et al., 2016).

4.2 Noise level quantification

The network of noise measurement sensors will be used to explore areas of high intensity noise, that can affect human

Figure 5. Architecture of the data platform.
health and overall wellbeing. This data can be used to assess noise levels for compliance with national and international requirements and suggest actions for improvement. The continual monitoring of noise levels around Lozenets can be used to determine steps to reroute heavy traffic away from residential, educational and healthcare facilities and identify areas for developing recreational zones. The uniform placement of noise sensors around the district can help inform the design of an extended network of bicycle routes and the construction of future buildings in appropriate areas. The collocation of some of the noise sensors with the air quality stations will also allow trends between noise (e.g., generated from traffic or the lack of noise) and air quality to be analysed.

4.3 Traffic and people flow investigations

As outlined in Section 1, the district of Lozenets includes several parks, shopping centres and other amenities, which imply heavy people flows. The network of people counting radars and noise sensors, as well as the LiDAR system will be used to map the movement of vehicles and people and analyse any emerging trends with noise and air pollution levels. Based on this information, recommendations can be made for improving the public transport (such as unserved routes, suboptimal placements of stops and deficient schedules) and privately to business owners interested in maximising the exposition of their services. Advanced transportation analyses informing the allocation of new parking spaces, optimised park-and-ride services, and the use of alternative modes of transport can be performed (Li et al., 2008).

4.4 Simulation validation and calibration

The development of a city digital twin relies heavily on modelling and simulation of various physical and operations processes. In order to be useful, any model must conform well to reality. The data from the urban living lab can be used to assess the accuracy of models for different features of the city. That is, sensor data can be used to validate the computer simulations and to indicate their predictive capability. Beyond fixed-model assessment, the collected and aggregated data can help with understanding different physical phenomena and provide a means to perform model calibration - adjusting the models so that they become a more faithful representation of reality. Both validation and calibration must acknowledge the various types of uncertainty present in models and data if the inferences are to be trustworthy and meaningful. Current initiatives that are taking place, which will directly benefit from observations from the urban living lab include multiscale air pollution dispersion modelling, wind and thermal comfort simulations and computational noise studies.

4.5 Validation of the urban living lab concept

Despite the fact that components for some of the equipment (air quality and noise measuring stations) have already been fielded in other applications, one of the important problems to be investigated is the operability, maintenance, and calibration of devices in the urban living lab as a unified network. This will include the critical assessment of the quality of the acquired data, the sustainability of the concept as a whole and its applicability to larger portions of Sofia and other cities.

5. CONCLUSION

The paper introduces the sensor network and data platform of an urban living lab, which is under development in the district of Lozenets in Sofia, Bulgaria. The living lab is instrumented with sensor stations for air quality and weather conditions monitoring, noise levels measurement and pedestrian counting. One of the busiest crossroads of the district is equipped with advance LiDAR system for vehicle monitoring. The data collected in the living lab will be linked with other data and information to enable development of a city digital twin of Sofia. A variety of potential specialised applications are considered, which allows for increased planning reliability and reduced risk through analysis and simulation of urban interventions.

The synergy between the city digital twins and urban living labs delivers a complete showcase of technologies and services for delivering more liveable and sustainable cities. Enriching the city digital twin with data from the urban living lab gives an opportunity to create a more realistic replica of the city that better supports the decision making of city authorities. The presented work will be constantly extended and refined. It provides a testing ground for start-ups and companies to launch new projects and to create innovative solutions. Results and experience from the work in the district of Lozenets will be used to inform the design of future urban living labs throughout Sofia and other urban regions.

ACKNOWLEDGEMENTS

This research was funded by the Horizon 2020 WIDESPREAD-2018-2020 TEAMING Phase 2 programme under grant agreement no. 857155, by Operational Programme Science and Education for Smart Growth under Grant Agreement No. BG05M2OP001-1.003-0002-C01 and by the National Scientific Program “Petar Beron i NIE” under the AUDiT project, no. KP-06-DB/3.

REFERENCES


Tekes, 2017. Smart City Solutions from Finland. Accessible at: https://www.businessfinland.fi/globalassets/julkaisut/Smart-City-Solutions-from-Finland.pdf


