Machine Learning Techniques and Models for Situation Awareness of IoT based Complex Systems

Domenico Santaniello

UNIVERSITY OF SALERNO

DEPARTMENT OF INDUSTRIAL ENGINEERING

Ph.D. Course in Industrial Engineering
Curriculum in Electronic Engineering - XXXIV Cycle

Machine Learning Techniques and Models for Situation Awareness of IoT based Complex Systems

Supervisor
Prof. Francesco Colace

Ph.D. Student
Domenico Santaniello

Scientific Referees
Prof. Allan Tucker
Prof. Brij B. Gupta
Ing. Alfredo Troiano

Ph.D. Course Coordinator
Prof. Francesco Donsì

Publications List

Journal Articles

Lombardi, M., Pascale, F., Santaniello, D. "Internet of things: A general overview between architectures, protocols and applications" *(2021) Information (Switzerland), 12 (2), art. no. 87, pp. 1-21.* DOI: 10.3390/info12020087

Casillo, M., Colace, F., Conte, D., Lombardi, M., Santaniello, D., Valentino, C. "Context-aware recommender systems and cultural heritage: a survey" *(2021) Journal of Ambient Intelligence and Humanized Computing*. DOI: 10.1007/s12652-021-03438-9

Castiglione, A., Palmieri, F., Colace, F., Lombardi, M., Santaniello, D., D'Aniello, G. "Securing the internet of vehicles through lightweight block ciphers" *(2020) Pattern Recognition Letters, 135, pp. 264-270.* DOI: 10.1016/j.patrec.2020.04.038

Colace, F., De Santo, M., Lombardi, M., Mosca, R., Santaniello, D. "A multilayer approach for recommending contextual learning paths" *(2020) Journal of Internet Services and Information Security, 10 (2), pp. 91-102.* DOI: 10.22667/JISIS.2020.05.31.091

Clarizia, F., Colace, F., De Santo, M., Lombardi, M., Pascale, F., Santaniello, D., Tuker, A. "A multilevel graph approach for rainfall forecasting: A preliminary study case on London area" *(2020) Concurrency and Computation: Practice and Experience, 32 (8), art. no. e5289.* DOI: 10.1002/cpe.5289

Casillo, M., Clarizia, F., D'Aniello, G., De Santo, M., Lombardi, M., Santaniello, D. "CHAT-Bot: A cultural heritage aware teller-bot for supporting touristic experiences" *(2020) Pattern Recognition Letters, 131, pp. 234-243.* DOI: 10.1016/j.patrec.2020.01.003

Casillo, M., Clarizia, F., Colace, F., Lombardi, M., Pascale, F., Santaniello, D. "An approach for recommending contextualized services in e-tourism" *(2019) Information (Switzerland), 10 (5), art. no. 180.* DOI: 10.3390/INFO10050180

Conference Proceedings

Colace, F., Elia, C., Landolfi, E., Lombardi, M., Santaniello, D., Troiano, A. "An IoT-Based Framework for Smart Agriculture Managing and Product Enhancing" *(2022) Lecture Notes in Networks and Systems, 217, pp. 111-119.* DOI: 10.1007/978-981-16-2102-4_10

Casillo, M., Conte, D., Lombardi, M., Santaniello, D., Troiano, A., Valentino, C. "A Content-Based Recommender System for Hidden Cultural Heritage Sites Enhancing" *(2022) Lecture Notes in Networks and Systems, 217, pp. 97-109.* DOI: 10.1007/978-981-16-2102-4_9

Clarizia, F., Colace, F., De Santo, M., Khan, M., Lombardi, M., Mosca, R., Santaniello, D. "An Approach Based on Context and Situation Awareness to Improve Functional Safety in Complex Scenarios" *(2022) Lecture Notes in Networks and Systems, 217, pp. 121-129.* DOI: 10.1007/978-981-16-2102-4_11

Casillo, M., Santo, M.D., Lombardi, M., Mosca, R., Santaniello, D., Valentino, C. "Recommender Systems and Digital Storytelling to Enhance Tourism Experience in Cultural Heritage Sites" *(2021) Proceedings - 2021 IEEE International Conference on Smart Computing, SMARTCOMP 2021, pp. 323-328.* DOI: 10.1109/SMARTCOMP52413.2021.00067

Colace, F., Elia, C., Guida, C.G., Lorusso, A., Marongiu, F., Santaniello, D. "An IoT-based Framework to Protect Cultural Heritage Buildings" *(2021) Proceedings - 2021 IEEE International Conference on Smart Computing, SMARTCOMP 2021, pp. 377-382.*
DOI: 10.1109/SMARTCOMP52413.2021.00076

Casillo, M., Colace, F., Gupta, B.B., Marongiu, F., Santaniello, D. "Decentralized approach for data security of Medical IoT Devices" *(2021) CEUR Workshop Proceedings, 3080.*

Guida, C.G., Gupta, B.B., Lorusso, A., Marongiu, F., Santaniello, D., Troiano, A. "An Integrated BIM-IoT approach to support energy monitoring" *(2021) CEUR Workshop Proceedings, 3080.*

Casillo, M., Colace, F., De Santo, M., Lorusso, A., Mosca, R., Santaniello, D. "VIOT_Lab: A Virtual Remote Laboratory for Internet of Things Based on ThingsBoard Platform" *(2021) Proceedings - Frontiers in Education Conference, FIE, 2021-October.* DOI: 10.1109/FIE49875.2021.9637317

Carbone, M., Colace, F., Lombardi, M., Marongiu, F., Santaniello, D., Valentino, C. "An Adaptive Learning Path Builder based on a Context Aware Recommender System" *(2021) Proceedings - Frontiers in Education Conference, FIE, 2021-October.*
DOI: 10.1109/FIE49875.2021.9637465

Casillo, M., Colace, F., Gupta, B.B., Santaniello, D., Valentino, C. "Fake News Detection Using LDA Topic Modelling and K-Nearest Neighbor Classifier" *(2021) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 13116 LNCS, pp. 330-339.*
DOI: 10.1007/978-3-030-91434-9_29

Casillo, M., Castiglione, A., Colace, F., de Santo, M., Marongiu, F., Santaniello, D. "COVID-19 data sharing and organization through blockchain and decentralized models" *(2021) CEUR Workshop Proceedings, 2991, pp. 128-140.*

Di Filippo, A., Lombardi, M., Marongiu, F., Lorusso, A., Santaniello, D. "Generative design for project optimization" *(2021) Proceedings - DMSVIVA 2021: 27th International DMS Conference on Visualization and Visual Languages, pp. 110-115.*

Casillo, M., Conte, D., Lombardi, M., Santaniello, D., Valentino, C. "Recommender System for Digital Storytelling: A Novel Approach to Enhance Cultural Heritage" *(2021) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12667 LNCS, pp. 304-317.* DOI: 10.1007/978-3-030-68787-8_22

Clarizia, F., De Santo, M., Lombardi, M., Mosca, R., Santaniello, D. "A Contextual Approach for Coastal Tourism and Cultural Heritage Enhancing" *(2021) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12667 LNCS, pp. 318-325.* DOI: 10.1007/978-3-030-68787-8_23

Casillo, M., Colace, F., De Santo, M., Lombardi, M., Santaniello, D. "A Chatbot for Training Employees in Industry 4.0" *(2021) Springer Proceedings in Complexity, pp. 397-409.* DOI: 10.1007/978-3-030-62066-0_30

Clarizia, F., Colace, F., De Santo, M., Lombardi, M., Santaniello, D. "Salerno Telling: An Approach for Cultural Heritage Enhancing" *(2021) Springer Proceedings in Complexity, pp. 103-112.* DOI: 10.1007/978-3-030-62066-0_9

Casillo, M., Colace, F., De Santo, M., Lombardi, M., Mosca, R., Santaniello, D. "A Recommender System for Enhancing Coastal Tourism" *(2021) Springer Proceedings in Complexity, pp. 113-122.* DOI: 10.1007/978-3-030-62066-0_10

Clarizia, F., De Santo, M., Lombardi, M., Santaniello, D. "E-learning and industry 4.0: A chatbot for training employees" (2021) *Advances in Intelligent Systems and Computing, 1184, pp. 445-453.* DOI: 10.1007/978-981-15-5859-7_44

Colace, F., Khan, M., Lombardi, M., Santaniello, D. "A multigraph approach for supporting computer network monitoring systems" *(2021) Advances in Intelligent Systems and Computing, 1184, pp. 470-477.* DOI: 10.1007/978-981-15-5859-7_46

Santo, M.D., Fabbri, L., Mosca, R., Lombardi, M., Romano, A., Santaniello, D. "A Multilevel approach to recommend working paths in industry 4.0" *(2020) Proceedings of 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering, TALE 2020, art. no. 9368327, pp. 754-757.* DOI: 10.1109/TALE48869.2020.9368327

Casillo, M., Colace, F., Fabbri, L., Lombardi, M., Romano, A., Santaniello, D. "Chatbot in industry 4.0: An approach for training new employees" *(2020) Proceedings of 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering, TALE 2020, art. no. 9368339, pp. 371-376.* DOI: 10.1109/TALE48869.2020.9368339

Lombardi, M., Pascale, F., Santaniello, D. "EIDS: Embedded intrusion detection system using machine learning to detect attack over the can-bus" *(2020) Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference, pp. 2028-2035.* DOI: 10.3850/978-981-14-8593-0_5090-cd

Casillo, M., Colace, F., Conte, D., De Santo, M., Lombardi, M., Mottola, S., Santaniello, D. "A Multi-feature Bayesian Approach for Fake News Detection" *(2020) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12575 LNCS, pp. 333-344.* DOI: 10.1007/978-3-030-66046-8_27

Castiglione, A., Palmieri, F., Colace, F., Lombardi, M., Santaniello, D. "On the impact of lightweight ciphers in automotive networks" *(2020) DMSVIVA 2020 - Proceedings of the 26th International DMS Conference on Visualization and Visual Languages, pp. 87-92.* DOI: 10.18293/DMSVIVA20-010

Clarizia, F., Colace, F., Lombardi, M., Santaniello, D. "A chatbot for supporting users in cultural heritage contexts" *(2020) DMSVIVA 2020 - Proceedings of the 26th International DMS Conference on Visualization and Visual Languages, pp. 22-27.* DOI: 10.18293/DMSVIVA20-009

Clarizia, F., Colace, F., Pascale, F., Lombardi, M., Santaniello, D.
"Sentiment analysis in social networks: A methodology based on the latent dirichlet allocation approach" *(2020) Proceedings of the 11th Conference of the European Society for Fuzzy Logic and Technology, EUSFLAT 2019, pp. 241-248.* DOI: 10.2991/eusflat-19.2019.36

Colace, F., De Santo, M., Lombardi, M., Pascale, F., Santaniello, D., Tucker, A. "A Multilevel Graph Approach for Predicting Bicycle Usage in London Area" *(2020) Advances in Intelligent Systems and Computing, 1027, pp. 353-362.* DOI: 10.1007/978-981-32-9343-4_28

Amato, F., Casillo, M., Colace, F., Santo, M.D., Lombardi, M., Santaniello, D. "CHAT: A Cultural Heritage Adaptive Tutor" *(2019) TALE 2019 - 2019 IEEE International Conference on Engineering, Technology and Education, art. no. 9225962.* DOI: 10.1109/TALE48000.2019.9225962

Colace, F., de Santo, M., Lombardi, M., Santaniello, D. "Chars: A cultural heritage adaptive recommender system" *(2019) TESCA 2019 - Proceedings of the 2019 1st ACM International Workshop on Technology Enablers and Innovative Applications for Smart Cities and Communities, co-located with the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, ACM BuildSys 2019, pp. 58-61.* DOI: 10.1145/3364544.3364830

Clarizia, F., Colace, F., De Santo, M., Lombardi, M., Pascale, F., Santaniello, D. "A context-aware chatbot for tourist destinations" *(2019) Proceedings*

- 15th International Conference on Signal Image Technology and Internet Based Systems, SISITS 2019, art. no. 9067869, pp. 348-354. DOI: 10.1109/SITIS.2019.00063

Castiglione, A., Palmieri, F., Colace, F., Lombardi, M., Santaniello, D. "Lightweight Ciphers in Automotive Networks: A Preliminary Approach" *(2019) 2019 4th International Conference on System Reliability and Safety, ICSRS 2019, art. no. 8987693, pp. 142-147.* DOI: 10.1109/ICSRS48664.2019.8987693

Lombardi, M., Pascale, F., Santaniello, D. "An application for Cultural Heritage using a Chatbot" *(2019) 2nd International Conference on Computer Applications and Information Security, ICCAIS 2019, art. no. 8769525.* DOI: 10.1109/CAIS.2019.8769525

Colace, F., Lombardi, M., Pascale, F., Santaniello, D. "A Multilevel Graph Representation for Big Data Interpretation in Real Scenarios" *(2019) Proceedings - 2018 3rd International Conference on System Reliability and Safety, ICSRS 2018, art. no. 8688834, pp. 40-47.* DOI: 10.1109/ICSRS.2018.8688834

Colace, F., Lombardi, M., Pascale, F., Santaniello, D., Tucker, A., Villani, P. "MuG: A Multilevel Graph Representation for Big Data Interpretation" *(2019) Proceedings - 20th International Conference on High Performance Computing and Communications, 16th International Conference on Smart City and 4th International Conference on Data Science and Systems, HPCC/SmartCity/DSS 2018, art. no. 8622971, pp. 1408-1413.* DOI: 10.1109/HPCC/SmartCity/DSS.2018.00233

Lombardi, M., Pascale, F., Santaniello, D.
"A Double-layer Approach for Historical Documents Archiving" *(2018) 2018 IEEE International Conference on Metrology for Archaeology and Cultural Heritage, MetroArchaeo 2018 - Proceedings, art. no. 9089768, pp. 137-140.* DOI: 10.1109/MetroArchaeo43810.2018.13607

Colace, F., Lombardi, M., Pascale, F., Santaniello, D. "A multi-level approach for forecasting critical events in smart cities" *(2018) Proceedings - DMSVIVA 2018: 24th International DMS Conference on Visualization and Visual Languages, pp. 31-35.* DOI: 10.18293/DMSVIVA2018-002

Clarizia, F., Colace, F., Lombardi, M., Pascale, F., Santaniello, D. "A multilevel graph approach for road accidents data interpretation" *(2018) Lecture Notes in Computer Science (including subseries Lecture Notes in*

Artificial Intelligence and Lecture Notes in Bioinformatics), 11161 LNCS, pp. 303-316. DOI: 10.1007/978-3-030-01689-0_24

Clarizia, F., Colace, F., Lombardi, M., Pascale, F., Santaniello, D. "Chatbot: An education support system for student" *(2018) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11161 LNCS, pp. 291-302.* DOI: 10.1007/978-3-030-01689-0_23

Colace, F., Santaniello, D., Casillo, M., Clarizia, F. "BeCAMS: A behaviour context aware monitoring system" *(2017) 2017 IEEE International Workshop on Measurement and Networking, M and N 2017 - Proceedings, art. no. 8078374.* DOI: 10.1109/IWMN.2017.8078374

Summary

Figures Index

Tables Index

Abstract

The current reality is characterized by a solid technological and pervasive component. These elements are expressed through smart devices, which make the environments we live in pervasive and able to exchange information. An example is represented by Smart Cities, complex environments able to leverage large amounts of data from sensors based on the Internet of Things (IoT) paradigm. One of the current challenges is using this information to transform scenarios from complex to helpful for increasing human well-being. This objective can be achieved by acquiring Context-Awareness, analyzing information, and managing the environment through the Situation-Awareness paradigm.

This Thesis aims to introduce a methodology with predictive capabilities and context adaptability for managing complex scenarios. The added value of the proposed approach is the introduction of the semantic value acquired from the Context and Situation Awareness through graph approaches, which, unlike many strategies used, leads to better integration of knowledge, obtaining higher system performance. In particular, a methodology for merging Ontologies, Context Dimension Trees, and probabilistic approaches based on Bayesian Networks will be presented to help experts and end-users handle events and provide suggestions for improving the liveability of smart complex scenarios. The proposed methodology has been validated and applied to several complex scenarios based on the IoT paradigm obtaining promising results.

Introduction

Modern reality is characterized by a vital technological component oriented to provide intelligent and pervasive services. Through modern smartphones, people are always connected to the Internet, accessing and exchanging considerable information. Even the environments and the services are managed, sometimes autonomously, through technological and pervasive systems. This concept has led to modern complex environments, one of which is represented by Smart Cities. Smart Cities arise from the dream of optimizing resources in terms of economic development, environmental efficiency, and stability through integrated technologies. Such modern environments are able to use data, i.e., traffic congestion, power consumption statistics, and public safety events, in order to upgrade the city services. A fundamental paradigm contributing to the generation of these complex scenarios is the Internet of Things (IoT). IoT refers to the concept that objects or "devices" are connected and able to exchange information among them and with humans. One of the goals of the IoT is to provide a digital copy of the real environment. However, this paradigm, oriented to exchanging information, allows information generation according to the Big Data phenomena. In fact, ever-increasing digitization and the attempt of IoT to represent those complex scenarios involve producing specific data that conventional methods cannot process.

This scenario leads through several questions about exploiting this data and the ability to process it to improve the livability of environments. Would it be possible to use such information to turn scenarios from complex to useful for humans?

Properly filtered data can be used to achieve the Context-Awareness. A deep understanding of context is critical to choosing or designing a suitable model, which allows extracting relevant information to process data, model reality, and provide answers. However, describing the context is not enough because it is necessary to perform actions able to improve the environment's livability. For this reason, it is possible to exploit the Situation Awareness paradigm. This concept refers to systems able to extract and understand environmental information and predict the occurrence of certain events. This paradigm can be improved by artificial intelligence techniques oriented to machine learning models, leading systems to perform autonomous decisions. Many approaches in the scientific literature address the management of complex scenarios in which several machine learning and deep learning techniques are employed. However, many of these approaches are complex, and in many cases, even if the algorithms provide understandable results, it is difficult for users to visualize and understand the actual processes that lead to the resolution of the problem. In this scenario, it would be interesting to introduce a methodology that integrates contextual data and manages analysis and prediction processes available to users.

This Thesis aims to develop a methodology with predictive capabilities and context adaptability for the management of complex scenarios. The proposed approach added value would be introducing the semantic value provided by graph approaches, which, unlike many used approaches, could lead to better integration of knowledge, achieving improved system performance.

In addition, these approaches allow the understanding of what is happening in the System at a given time, allowing the manipulation and integration of semantic information. The graph approaches chosen for this purpose are Ontologies, Context Dimension Trees, and Bayesian Networks. Ontologies allow providing a formal representation of a specific domain, which allows humans and machines cooperation. A Context Dimension Tree is a specific model for context representation, representing all possible contexts. Bayesian networks are a probabilistic graph model able to predict specific events' occurrence. In summary, the proposed approach, defined as the Multilevel Graphical approach or MuG approach, exploits three graphical approaches: Ontology, Context Dimension Tree, and Bayesian Networks. The System's goal is to provide answers and applications that can impact the real environment, helping the end-user or expert user manage complex scenarios. The proposed methodology has been validated and applied to several complex scenarios based on the IoT paradigm obtaining promising results.

The document is organized as follows. The first three chapters introduce the fundamental paradigms on which the proposed methodology is based. In particular, the first chapter discusses the paradigm of the Internet of Things, the second and the third Context Awareness and Situation Awareness, respectively. In the fourth chapter, the proposed approach is presented. In particular, the interaction between graph structures is described, and a system's general architecture is presented. The fifth chapter reports case studies. In particular, in the first section, a methodology validation case study related to London Smart City is presented. Subsequently, several case studies in real scenarios of the system interaction with humans are described. Future developments and conclusions are discussed at the end of this work.

Chapter 1
Internet of Things

The Internet of Things (IoT) Paradigm refers to a system of devices, interconnected with each other, with computational capability, identifiable and enabled to transfer data over a network, without necessary human interaction (Aazam, Zeadally and Harras, 2018). The concept behind this paradigm is the pervasive presence of intelligent devices that cooperate with each other and interact with humans to achieve common goals. (Atzori, Iera and Morabito, 2010).

Although this technology has been widely used in recent years, it was already present many years ago, even with theoretical hints. For example, in 1991, Mark Weiser, in an article on Ubiquitous computing, introduced a model of human-computer interaction in which information processing is integrated with everyday objects rather than individual personal computers. (Weiser, 1991). One of the first real applications of a system described above can be found in the industrial sector, where machines are made capable of exchanging information about their state autonomously. These systems were called Machine to Machine (M2M). In this paradigm, the machines can establish a closed system in which the main purpose of information exchange is to monitor and manage machines more efficiently and less expensive. Compared to the current meaning of IoT, there is a lack of awareness that data could provide when reused in a broader context, for example, when aggregated with other systems connected through the Internet (Anand, 2016).

The term "Internet of Things" was first used in 1999 by Kevin Ashton during a Procter & Gamble presentation. (Ashton, 2009). During this presentation, Ashton explained the possible benefits of RFID technology in merchandise management. By equipping goods with specific devices, they could "communicate" information of interest (status, tracking, etc.). In this way, "things" and people could efficiently provide information about their status and the world around them. The actual birth of the IoT dates back, according to Cisco estimates, between 2008-2009, when for the first time, the number of connected objects exceeded the world's population. By 2010, the number of such objects had almost doubled from that period, reaching about

12.5 billion. Since those years, IoT has become more and more widespread in everyday life thanks to continuous technological developments and significant investments by companies. According to IoT-analytics estimates, there are currently about 20 billion connected objects globally, and the IoT industry generates a market of about \$150 billion. By 2024, connected objects will exceed 30 billion, and the market value will be about 1 billion. As with any new technology trend, there are three possible categories of challenges for IoT to overcome: business, society, and technology (Shukla and Munir, 2017; Meneghello *et al.*, 2019).

One of the main challenges in technology to promote IoT systems implementation is defining a reference architecture that supports current functionalities and future extensions. Such architecture must include some key features (Abdmeziem, Tandjaoui and Romdhani, 2016). In particular, it must be scalable, in order to handle the growing number of devices and services without reducing their performance; interoperable, the devices from different vendors can cooperate to achieve common goals; distributive, to allow the creation of a distributed environment in which, after being collected from different sources, data are processed by different entities in a distributed way; able to operate with few resources, since objects generally have little computing power; secure so as not to allow unauthorized access. Currently, there is no single reference architecture, and creating it could be very complicated despite many standardization efforts. The main problem lies in the natural fragmentation of possible applications, each of which depends on many variables and design specifications that are very often different. This problem must be added to the tendency of each vendor to propose its platform for similar applications. (Al-Fuqaha *et al.*, 2015; Čolaković and Hadžialić, 2018). Figure 1.1 shows the most commonly used architectures.

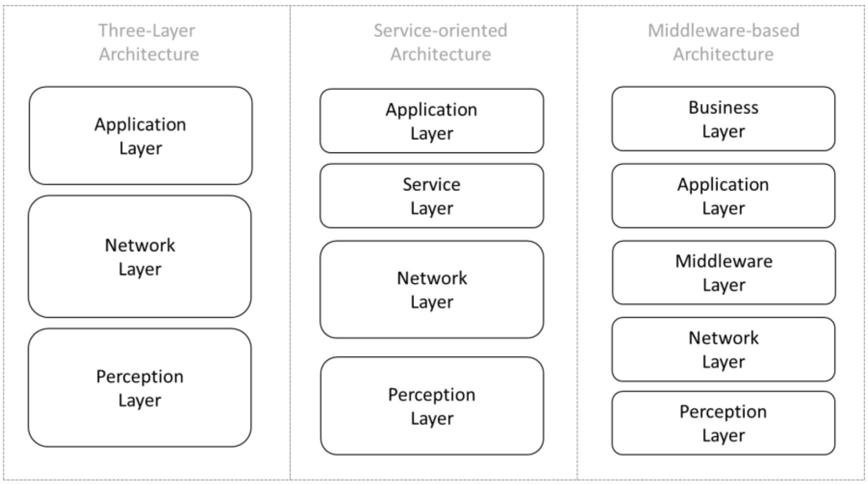


Figure 1.1 *Most common IoT architectures*

The most commonly used architecture is considered the most generic and high-level architecture consisting of three layers: Perception, Network, and Application. (Al-Fuqaha *et al.*, 2015).

The Perception layer represents the physical level of the objects and interacts with the surrounding environment collecting and processing information. This layer includes objects that, being able to interact with the external world and being equipped with computational ability, become somewhat "smart" (or "intelligent"). (Al-Fuqaha *et al.*, 2015; Abdmeziem, Tandjaoui and Romdhani, 2016). These smart objects, which represent the fundamental building blocks on which the IoT is based, can be ordinary objects (a refrigerator, a TV, a car, etc.) or simple devices equipped with sensors and computational capabilities. In general, smart objects are provided with some essential properties such as communication, i.e., the ability to connect to each other and autonomously access resources over the Internet to use data and services, update their status and cooperate to achieve common goals. Identification, i.e., the characteristic of being uniquely identified (Mattern and Floerkemeier, 2010; Taivalsaari and Mikkonen, 2018). Depending on the specific application, it is also possible for smart devices to have one or more properties such as Addressability, Sensing and Actuation, Embedded Information Processing, and User Interface. Addressability refers to the ability of objects to be directly reachable, with the ability to be remotely queried or managed. Sensing and Actuation refer to the ability of objects to collect information about the surrounding environment and manipulate it through the use of sensors and actuators. Embedded Information Processing is represented by the computational capacity to process the results of the sensors and operate the actuators. Localization represents the ability of objects to be aware of their physical location or the possibility of being located. User Interface refers to the ability of the devices, in addition to cooperating, to appropriately communicate with users through displays or other human-friendly interfaces.

The Network layer has the task of transporting the data provided by the perception layer to the application layer. It includes all the technologies and various protocols that make this possible. Figure 1.2 shows some of the most commonly used protocols, grouped according to the TCP/IP model. Each protocol has pros and cons, and its use must be evaluated based on the application. One of the most widely used protocols for data transmission in the application layer is MQTT (Message Queue Telemetry Transport). MTTQ is a lightweight, publish-subscribe messaging protocol designed for situations where low power consumption is required, and the available bandwidth is limited (Thangavel *et al.*, 2014).

In the Network layer, wireless protocols play a crucial role. Wireless sensors can be installed in hard-to-reach environments and require less material and human resources for installation than those requiring cables. Moreover, in a wireless sensor network, the various nodes can be added or removed easily, and their locations can be changed without reconsidering the structure of the entire network. The protocol depends on the size of the network, each node's power consumption, and the throughput required in a given application. However, building a wired network in several applications may be necessary. The latter has more excellent reliability and higher transmission speeds. (Sharma and Gondhi, 2018).

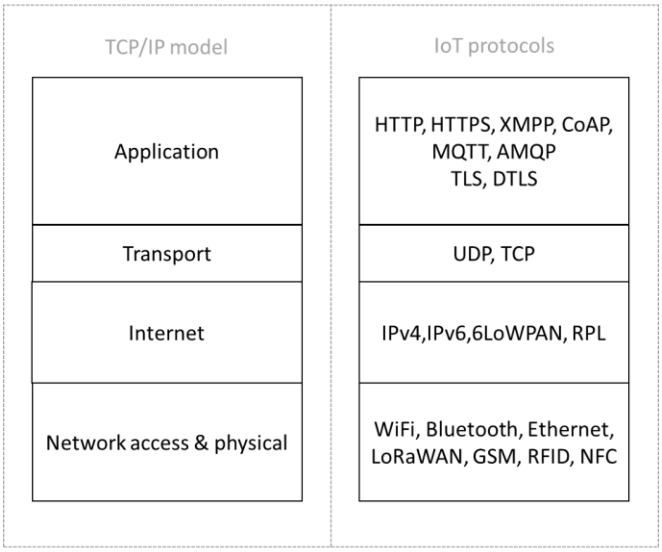


Figure 1.2 *Main protocols used in the IoT environment*

The Application layer includes all the software needed to provide specific services. In this layer, data from previous layers are stored, aggregated, filtered, and processed using databases, analytics software, etc. As a result of this processing phase, the data is made available to different IoT applications (i.e., Smart Cities, Smart Cars, Smart Home, and Smart Agriculture). This phase is often performed using software-defined middleware, which is responsible for hiding the heterogeneity of the underlying layers. Some software technologies currently widely used to manage the large amount of data provided by devices are represented by Cloud Computing and Edge Computing. In Cloud Computing, services such as data storage or processing are provided by a set of pre-existing, configurable, and remotely available resources in the form of a distributed architecture. In Edge computing, data processing is partially distributed on secondary network nodes to increase the performance of IoT systems.

Another type of architecture that is very commonly used is Service-oriented Architecture (SoA). This model is component-based; it can connect different functional applications units through interfaces and protocols (Ilie-Zudor *et al.*, 2011; Xu, 2011). SoA is designed to coordinate services and reuse software and hardware components. SoA can be easily integrated into the IoT architecture by extending the three-layer architecture and adding a new layer between the network, the application layers called the Service layer, which provides services to support the application layer. This layer represents the four-layer SoA-Based IoT architecture, in which there is the perception layer, the network layer, the service layer, and finally, the application layer. The service layer consists of service discovery, composition, management, and service interfaces (Ilie-Zudor *et al.*, 2011; Han *et al.*, 2013). Service discovery is used to discover service requests. Service composition interacts with connected objects and integrates services to obtain requests efficiently. Service management is used to manage and determine trust mechanisms to understand service requests, and service interfaces are used to support interactions between all services provided.

Another essential and prevalent architecture in IoT is the middleware-based IoT architecture or five-layer architecture (Ngu *et al.*, 2017). In recent years, the proposed IoT architecture needs to address many factors such as scalability, interoperability, reliability, QoS, etc. In this regard, middleware based IoT architectures help create applications more efficiently; this layer acts as a link between applications, data, and users. In fact, the development of IoT depends on technological progress and the design of various new applications and business models (Gan, Lu and Jiang, 2011). A five-layer architecture has been proposed to enable these features and many others, composed of five layers: Perception layer, Network layer, Middleware layer, Application layer, and Business layer. In particular, the Middleware layer has some critical functionalities, such as aggregating and filtering data received from hardware devices, performing information discovery, and providing device access control for applications. In general, middleware is a software or programming service that can provide an interposed abstraction between IoT technologies and applications. In middleware, the details of different technologies are hidden, and standard interfaces are provided to allow developers to focus on application development without considering compatibility between applications and infrastructure. Middleware has been gaining more and more importance in recent years because of its central role in simplifying the development of new services and the integration of legacy technologies into new ones.

The Internet of Things (IoT) paradigm is strongly oriented to exchanging data between devices and humans through the network. This permanent exchange of information generates another significant phenomenon in complex scenarios management: Big Data. In fact, an ever-increasing digitalization and the attempt of the IoT to represent all the reality in a digital way (Gubbi *et al.*, 2013) involves the production of a large and valuable amount of data that conventional methods cannot process. (Sun *et al.*, 2016). Developing solutions and algorithms able to interpret and interact with this vast amount of information is a crucial challenge that Big Data poses in today's world. (Sivarajah *et al.*, 2017). Data management has grown along three dimensions: volume, velocity, and variety. The so-called "3Vs" represent key elements regarding the characteristics of Big Data systems (Tole, 2013). Volume refers to the amount of structured or unstructured data generated, which is manipulated and analyzed in order to obtain the desired results. Commonly, this data is generated from heterogeneous sources such as traditional databases, social media, sensors, events, etc. Velocity refers to the speed of data generation and the need for this information to be processed in real-time. Variety deals with the different generated, collected, and used data types. This data, which belongs to the most different sources, suggests using different storage and retrieval approaches. During the last years, other dimensions were introduced, going from 3V to 5V (Demchenko *et al.*, 2013) up to 7V (Sivarajah et al., 2017), adding Variability, Visualisation, Value, and Veracity. Variability is different from variety. Variability refers to data's ability to change constantly, significantly impacting data homogenization. Visualization represents a crucial aspect. Having the ability to correctly represents significant amounts of complex data could be more effective. Value is a key aspect of data, defined by the added value that the collected data can bring to the expected process, activity, or analysis/predictive hypothesis; obviously, this aspect is related to transforming data into Knowledge Database. The Veracity dimension of Big Data includes the consistency and reliability of the data, concerning several factors, including statistical reliability, data origin, processing methods, etc. It is crucial to ensure data reliability, considering that results can be generated on which significant decisions are made. Assigning a veracity index to the data on which the analyses are based is essential to measure the System's overall reliability. The above issues are crucial when dealing with data that can help handle complex scenarios.

The main strength of the combination of IoT and Big Data paradigms represents the high impact on different aspects of the daily life and behavior of potential users (Atzori, Iera and Morabito, 2010). This aspect has led to the idea of managing and interpreting such data to achieve critical goals for humans, such as increased quality of life, economic development, and increased environmental efficiency and stability. The increase of these factors could be achieved by designing urban areas that take advantage of integrated technologies and optimization of resources in order to improve some key objectives such as mobility, communication, economy, work, environment, administration. In 2008 IBM, during the global financial crisis, suggested an intelligent approach to address the problems plaguing economic growth by launching the concept of a smarter planet introducing Smart Cities. Smart Cities are able to use data such as traffic congestion, energy consumption statistics, and public safety events in order to update city services through three basic concepts: instrumented, interconnected, and intelligent supplies. (Harrison *et al.*, 2010). Instrumented refers to data sources from physical or virtual sensors; Interconnected refers to integrating and managing such data in an enterprise computing platform and their communication; Intelligent refers to the ability of complex analysis, modeling, optimization, and visualization to make better operational decisions. Many applications on the smart city concept have been proposed in the literature. Many projects have been presented to make cities smart, such as Padova Smart City (Zanella *et al.*, 2014). This project presents technical solutions, guidelines, and best practices concerning services made possible by integrating the IoT paradigm, such as waste management, noise monitoring, traffic congestion management, energy consumption, etc. There are many other applications in this field, such as Smart Parking, a service that citizens could find helpful and could reduce pollution. (Al-Turjman and Malekloo, 2019). To manage energy consumption in the city, particular attention has been paid to public lighting management through Smart Light services. (Lau *et al.*, 2015; de Paz *et al.*, 2016) but also the management of a Smart Traffic Light system. (Kanungo, Sharma and Singla, 2014). In fact, road traffic congestion management represents a significant issue within the use of the IoT paradigm in Smart Cities (Misbahuddin *et al.*, 2016; Javaid *et al.*, 2018). Many applications have been proposed to monitor and manage road traffic also in an automatic way, exploiting purpose-built devices (Chong and Ng, 2017) smart cameras (Frank, Khamis Al Aamri and Zayegh, 2019) and machine learning algorithms for event prediction. (Tang *et al.*, 2018). Connected to this issue has arisen the need to manage urban public transport through integrated systems that exploit networks of IoT sensors. (Patel, Narmawala and Thakkar, 2019). The diffusion of such systems has generally influenced also the automotive industry. In fact, the last few years have been characterized by a rapid increase of technologies onboard cars both for control and assistance systems and for monitoring and diagnostic systems. In this scenario, IoT offers essential support for the automotive industry. It is possible to say that the Automotive Internet of Things (IoT) is an emerging field of research that applies IoT to intelligent transportation systems (Weyer *et al.*, 2016). With the introduction of smartphones, cloud, edge computing, and mobile Internet, the automotive ecosystem is moving towards the Internet of Vehicles (IoV) (Ji *et al.*, 2020). These technologies have led the car to become a more intelligent vehicle; thus, nowadays, cars are considered an integrated and complicated ecosystem of objects that cooperate autonomously.

In addition to vehicles, which are considered an integral part of the Smart City ecosystem, buildings, and homes, have also been considered. A Smart Home is defined as a modern home, which includes appliances, lighting systems, and electronic devices that can be monitored and managed remotely through mobile applications (Yassine *et al.*, 2019). In addition, fundamental applications have considered the IoT at the service of Smart Grid systems concerning efficiency and energy saving. (Saleem *et al.*, 2019). The Smart Grid aims to use electricity safely and appropriately in which the power supply system can distribute electricity avoiding waste. For these reasons, nowadays, Smart Grid and Smart Micro-Grid are very popular (Reka and Dragicevic, 2018). An example of these applications concerns accurately systems able to manage energy in real-time through the use of different networks of Smart Meters. (Minoli, Sohraby and Occhiogrosso, 2017).

Finally, another critical and fundamental field of applications in Smart City is the production field. Indeed, the advent of the fourth industrial revolution could not but include a stronger synergy between humans and machines (Lu, 2017). In the modern 4.0 industries, all production processes are reconsidered thanks to the use of IoT. In particular, sensors allow monitoring and managing processes in real-time, inferring helpful information for supply, maintenance, and improvement of production (Khan *et al.*, 2020),(Yin, Stecke and Li, 2018). Machines will be increasingly dominant and present in production and cooperation with humans, and systems will be able to make autonomous decisions and provide operators with innovative problem-solving strategies. (Oztemel and Gursev, 2020). In this scenario, also agricultural production has had its development through the advent of IoT. In fact, Smart Agriculture evolved from the Precision Agriculture concept, defining an integrated system of methodologies and technologies designed to increase agricultural production, quality, and productivity in the fields and farms. (Elijah *et al.*, 2018). There are many works in the literature concerning intelligent systems and devices to support agriculture that can limit areas of intervention and suggest the right action at the right time concerning the real needs of crops. Solutions and applications are presented to monitor, manage, and optimize the various processes related to agriculture (Kamilaris and Prenafeta-Boldú, 2018; Muangprathub *et al.*, 2019). Different approaches and devices have been proposed for agricultural fields remote control through Unmanned Aerial Vehicles (UAVs) or other devices oriented to controlling and protecting crops. (Chouhan, Singh and Jain, 2020; Radoglou-Grammatikis *et al.*, 2020). Other studies are focused on the ability of systems to monitor and predict, in cooperation with humans, the progress of crops and the decisions to be taken. (Goap *et al.*, 2018; Zhai *et al.*, 2020).

As can be seen, one of the added values of IoT in all these application areas also lies in the interaction with human beings. Particular attention is paid to those systems that interact with people by increasing the human ability to control, manage, and solve problems in complex environments. (Majid Butt *et al.*, 2020). The importance of the IoT has been established over time, especially for those applications that are natively interconnected, such as mobile applications devoted to the monitoring, management, and processing of pervasive context-based scenarios. (Ponce and Abdulrazak, 2021). In fact, there is often confusion between the concept of IoT and Context-Aware Computing which are closely related. Although these two concepts are often used simultaneously, they remain two distinct concepts. IoT is one of the enabling technologies of context-aware computing, and, nowadays, many pervasive IoT-based systems exploit Context-Awareness as a central feature (Ud Din *et al.*, 2019). The Context-Awareness paradigm will be discussed in the next chapter, representing a fundamental topic for describing complex scenarios, such as those based on the Internet of Things (Sezer, Dogdu and Ozbayoglu, 2018).

Chapter 2
Context Awareness

Technology growth and the advent of the Internet of Things generated complex, data-driven environments. These environments, characterized by the continuous exchange of information, produce Big Data, data that cannot be processed conventionally. The possibility of acquiring a more significant amount of information than in the past has increased the necessity of solving problems such as the interpretation, integration, and processing of Big Data (Chen, Mao and Liu, 2014; Philip Chen and Zhang, 2014; Gandomi and Haider, 2015) through more performant methodologies. In such scenarios, it is necessary to avoid the possible confusion of vast amounts of data coming from different sources and adequately assembled to provide services. Context analysis is instrumental in this field, as it can support analysis by enhancing the filtering of data and services to improve the quality of support provided by applications. Moreover, the ability to model possible usage scenarios represents an added value in Big Data and new information systems such as, for example, mobile systems able to interact and support users (Torralba *et al.*, 2003; Medjahed and Atif, 2007). The use of contextual techniques in user support management scenarios is possible due to integrating smart devices belonging to the Internet of Things paradigm (Atzori, Iera and Morabito, 2010; Gubbi *et al.*, 2013). These devices are able to acquire information and communicate with each other, creating a pervasive network consisting of active elements for the acquisition of helpful information to support users (Zanella *et al.*, 2014). In this way, the user can deal with the so-called "information overload", which consists of the inability to manage the cognitive overload produced by the number of resources available. In this direction, the synergy between contextual techniques and the Internet of Things has allowed a significant step forward in several areas (Xu, He and Li, 2014; Alexopoulos *et al.*, 2016; Peng and Jinqi, 2017; Aceto, Persico and Pescapé, 2020).

Schilit and Theimer first defined context-aware systems as part of their research on distributed mobile computing (Schilit, Adams and Want, 1995). In particular, they defined a context-aware system as a system that "adapts according to the location of use, the collection of nearby people, hosts, and accessible devices, as well as to changes to such things over time". This definition shows the connection of context-aware systems with ubiquitous computing, which refers to human-machine interaction via smart tools (Hightower and Borriello, 2001) and IoT. Starting from Schilit and Theimer's definition, there have been several attempts to define context-aware computing. However, most definitions seem too specific and challenging to use in practice. In this regard, a more generic definition aimed at practical use is provided (Dey, 2001):

"A system is defined as context-aware if it uses the context to provide relevant information and/or services to the user, where the relevance depends on the user's preferences and the tasks the user needs to perform."

Context-aware computing describes the development of technologies and applications capable of detecting data from the surrounding context and reacting accordingly with specific actions, reducing and simplifying the human-computer interaction process. Therefore, context awareness must be identified as technical features capable of adding value to data and services in different application segments. Moreover, functional to this definition, the definition of context is essential, which is not provided precisely and must be adapted to the different fields in which context-aware systems are applied.

Initially, Pascoe et al. (Pascoe, Ryan and Morse, 1999) limited the definition of context to environmental characteristics related, for example, to meteorological factors. Their definition does not consider other factors related to devices capable of acquiring information through users and their activities pointed out by Schmidt et al. (Schmidt *et al.*, 1999). A further step to delineate the concept of context was made by Chen and Kotz (Chen and Kotz, 2000), who emphasized the importance of temporal evaluation such as weekday, year season, or day time. According to Abowd et al., context is any information that can characterize an entity's situation. An entity could be a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves (Abowd *et al.*, 1999a). This last definition is similar to the field of pervasive computing (Satyanarayanan, 2001), closely related to the Internet of Things and Ubiquitous computing. This concept consists of acquiring knowledge through context and providing services to the user based on contextual information.

Moreover, the generality of the definition of Abowd et al. allows specifying new contextual domains, unlike Schilit and Theimer's definition (Perera *et al.*, 2014). However, it needs a classification or a model for concrete development. For this purpose, Abowd and Mynatt (Abowd and Mynatt, 2000) created the method based on 5 W (Who, What, Where, When, Why) to which the 1H (How) component has been added (Pantic *et al.*, 2007; Rudovic, Pavlovic and Pantic, 2015). The objective of the latter model is to determine:

- Who: identity of the user who represents one of the main categories of the context. This concept can be exploited both from an individual point of view and by other users interested in the user situation;
- What: related to the activity in which the user is involved and represents a fundamental category for the different context models identification;
- Where: indicates the location of the device or user. It constitutes the most common and analyzed form of context;
- When: refers to the actual moment of action or duration in terms of the time interval. Temporal context is used to identify patterns in user behaviors and is used to identify changes in the categories that characterize the context.
- Why: the motivations associated with a particular action are explored. For example, it can be related to the reason that led to a particular action or a particular position. It is the most complex context to define and analyze and requires the evaluation of the particular meaning of an action, intention, or sentiment.
- How: manages modes linked to a specific contextual context. It can be linked to the necessary steps to take or to the objectives of a user.

A deep understanding of context is crucial to choosing or designing a suitable model. In fact, the lack of a uniform approach for modeling the information associated with context makes it difficult to understand the requirements to be considered when proposing the adoption of a context model based on its central points. The subdivision of context can be done based on described features (Villegas *et al.*, 2018):

- The physical context: concerns information related to the physical environment, such as, i.e., geographical location (latitude, longitude, and altitude), temperature, humidity, noise, and light;
- The temporal context: concerns information of a temporal nature that may affect a system;
- The social context: concerns the direct or indirect interaction of an entity with people or objects present in the physical or virtual environment;
- The computational context: concerns the resources available to the system, such as computing resources, communication bandwidth, storage resources, etc.;
- The historical context: concerns the historical data that may condition the interpretation of information or the operation of a system;
- The profile: concerns the preferences of an entity for different contextual dimensions.

The use of context within systems allows the development of context-aware systems. The Active Badge Location System provides a first example of the context-aware system proposed by Want et al. (Want *et al.*, 1992), where users' location is exploited to optimize communication strategies in an office environment. There are different ways to classify context-aware systems. A first subdivision is based on the context usage model (Hu, Indulska and Robinson, 2008; Perera *et al.*, 2014):

- No application-level context model: the application is intended to handle all the various steps, including context acquisition and preprocessing necessary to perform specific actions;
- Implicit context model: the context is tied to the specific application and is defined through external resources such as, for example, APIs. The application manages the preprocessing and action management phase, built using standard paradigms.
- Explicit context model: the context and its processing are independently performed outside the system. Context and system are separate entities, and the management of the context for an application is also usable for different applications that need to process contextual information.

The introduced patterns are related to how context is acquired (Baldauf, Dustdar and Rosenberg, 2007). These factors are critical to designing a context-aware system based on the goals to be achieved or the means available to build context-aware systems. The different aspects that need to be considered to define the system tasks are diverse, and Perera et al. list some of them (Perera *et al.*, 2014):

- The ways in which the application is designed, considering the layers that compose it and define the degrees of autonomy of each action performed by the system. In addition, the hardware components that are exploited must be considered. In this phase, it will be fundamental to precisely define the model of introducing the context among those presented previously.
- Hardware and software upgrades planning. It will be necessary to consider the future functionalities of the system based on the number of data to process in order to cope with the scalability issue.
- Define the API (Application Programming Interface) to be used and, therefore, the information resources that the system will use. This phase is crucial for the definition of data preprocessing, where a representative standard for data can be defined.
- Define tools for the recognition of malfunctions within the system.
- Decide how to update contextual scopes within the system. It will have to be predisposed, therefore, if to consider a new context inside

the system based on data. For example, suppose in a system that uses the location context and the time context, it is necessary, based on data or information from experts in the field, the need to introduce a context related to the weather. In that case, the system must be prepared for this upgrade.
- Provide for the support of tools related to the Internet of Things paradigm.

The effectiveness of Context-Aware systems lies in the efficient management of the context, which is a difficult task due to the abstract concept's nature. The context-aware systems present in the literature propose different models that allow the collection and quantitative description of information related to the context in order to be ready for the processing phase. These models, although designed in different domains, have some standard features, such as:
- The subdivision of the context into dimensions or attributes that identify the relevant operating elements of the system such as, for example, time, location, temperature, etc.;
- The possibility to measure several dimensions defining parameters, unit of measure, and range of admissible values.

One of the first basic models for context management is based on an agents model (Batet *et al.*, 2012), where each contextual scope is processed independently and allows for the processing of contextual information appropriately for the specific system in which it is to be applied. This model is not specific to context management; however, it can be used to ensure the independence of different actions of the system through task division. Another feasible model is the "Context Space" which associates to each relevant characteristic of the context ("context attribute") a dimension, forming a multidimensional space that represents all possible scenarios in which the system can be found (Dominici, Pietropaoli and Weis, 2012). The current context, defined by the values assumed by the attributes, can be represented as a geometric point in this space. Besides being extremely intuitive, this model allows the use of geometric tools such as routes and distances to perform reasoning and operations on the system. One of the significant limitations of this approach, which affects its expressive capacity, is the impossibility of ordering the dimensions hierarchically or with a topology. In this way, the model is in stark contrast to ontological models and, in general, to models of knowledge organization. In fact, Ontologies represent a valuable tool for contextual representation (Papagiannakopoulou *et al.*, 2013). Ontology is a formal representation model of reality and knowledge. It represents a data structure that allows describing entities and their relationships in a given domain. An ontology is the explicit formal description of the elements and concepts of a domain; in this sense, it is a model that allows representing entities in the context where they are located (Rhayem, Mhiri and Gargouri, 2020). Furthermore, ontology-based contextual models effectively support context modeling and reasoning in pervasive computing environments (Wang *et al.*, 2004). A further model of effective context representation and management in the graph approach field is represented by the Context Dimension Tree (CDT). In particular, the CDT is a tree composed of a triplet $< r;\ N;\ A >$ where, with r is indicated the root, with N is represented the set of nodes of which it is composed, and with A the set of arcs that join these nodes (Rauseo, Martinenghi and Tanca, 2013). The CDT is used to be able to represent, in a graphical way, all the possible contexts within an application domain. The current context, defined by the values assumed by the various dimensions, can be represented as a subgraph in which each dimension node is matched at most one value or parametric node: it is defined as an AND between different "context elements". Therefore, adopting a hierarchical structure, the CDT allows orthogonally separating the various context dimensions and using different levels of abstraction to specify and represent all possible and admissible contexts in a given application domain (Casillo *et al.*, 2017). An additional context representation approach, exploited especially in particular recommender systems, is tensors. For example, Context-Aware Recommender Systems (Adomavicius *et al.*, 2005, 2011; Verbert *et al.*, 2012) exploit tensors (Kolda and Bader, 2009) in order to support users in making choices in the context of information overload generated by Big Data (Chen, Mao and Liu, 2014). Moreover, these systems are able to personalize the suggestions more than other types of Recommender Systems that do not exploit the context. Other ways to represent context include data structures based on the Key-Value model, Markup Scheme, object-oriented models, and logic models (Baldauf, Dustdar and Rosenberg, 2007).

In recent years, the concept of context has been extended to all aspects that characterize the user concerning the capabilities that an information system can deliver. For this reason, complex and general context models have been proposed to support context-aware applications, which leverage them, for example, to: adapt interfaces, tailor a set of application-relevant data (Orsi, Tanca and Zimeo, 2011), increase the accuracy of information retrieval, discover services and compose services (Furno and Zimeo, 2014), make user interaction implicit, and create intelligent environments (Casillo, Clarizia, D'Aniello, de Santo, *et al.*, 2020).

Consider the example of automated support for museum visitors, provided with a mobile device that reacts to a change in context. Such a system could adapt the user interface according to the different abilities of the visitor; provide different information content based on different visitor profiles and interests (students, journalists, archaeologists, etc.) and location; learn, based on the previous choices made by the visitor, what information the user will be interested in later; provide appropriate services, e.g., to buy a ticket for a temporary exhibition or to reserve a seat for the next exhibition on the life of the favorite author; provide active functions within the various areas of the museum, which indicate to visitors a series of hints and stimuli about what is happening in each particular environment (Colace *et al.*, 2017).

Several other examples of context-aware systems leverage different technologies for context management. In particular, Sarker et al. (Sarker, Kayes and Watters, 2019) examine the context in mobile services by proposing a comparison of supervised machine learning with the goal of classification with graphical techniques for context representation. This work emphasizes the importance of context-aware systems to provide personalized services to users exploiting mobile devices capable of managing contextual domains through acquired information. Ranganathan and Campbell (Ranganathan and Campbell, 2003) propose an infrastructure that acquires contextual information through sensors. The system processes the information through an ontology, which allows the management of the different information acquired. The proposed model allows, therefore, not only the acquisition of contextual information but also the processing of the context automatically. Gu et al. (Gu, Pung and Zhang, 2005) propose context management using ontologies. In particular, the goal is to provide personalized services to users based on processed contextual information. In this way, the SOCAM (Service-Oriented Context-Aware Middleware) architecture is designed that, through an additional property introduced in the ontology, allows the use of contextual information, distinguishing them in: sensed, defined, and deduced. Sensed refers to the context acquired through physical smart devices such as sensors. Defined refers to tailoring the specific user and designing to meet the user's particular needs. Finally, the deduced context is obtained by processing the direct context.

On the other hand, Yoon et al. (Yoon *et al.*, 2015) propose a Bayesian context-based approach that aims to handle multi-object tracking. Stenneth et al. (Stenneth *et al.*, 2011) also exploit a Bayesian approach in the context of transportation mode detection. The location context is exploited to achieve better accuracy for detecting external services, e.g., public transportation. Kooij et al. (Kooij *et al.*, 2014) exploit context-awareness computing in the context of modeling a Dynamic Bayesian Network that aims to predict path per pedestrian. Finally, Al-Sultan et al. (Al-Sultan, Al-Bayatti and Zedan, 2013) propose a Bayesian approach to predict driver behavior in the context of intelligent transportation systems. Specifically, the system captures and then processes contextual information through sensors. This process allows the context-aware system to apply the appropriate solutions based on the specific contextual condition. In particular, the processing is performed by several Bayesian Networks, which, besides analyzing the information, allow the system to manage the evolution of the information over time.

The examples provided by the literature allow us to state that the application of context-aware computing covers many domains (Almusaylim and Zaman, 2019); however, often, such systems are assisted by tools capable of processing the context to perform actions. In fact, simply describing the context is not enough; one must also know the situation and perform actions by modifying the reference environment. For this reason, the next chapter will discuss the Situation Awareness paradigm.

Chapter 3
Situation Awareness

Situation Awareness (SA) or Situational Awareness has been known in the literature, becoming a critical topic, since the 1980' (Endsley, 2000). Ideally, this paradigm is related to the management of complex situations, even of a military nature, as reported by Oswald Boelcke: "the importance of gaining an awareness of the enemy before the enemy gained a similar awareness, and devised methods for accomplishing this." (Gilson, 1995).

Intuitively, SA can be interpreted as an awareness of what is happening in a domain around people (Endsley, 2000). There is no univocal definition of SA. As reported in the scientific literature, several definitions focus on different aspects encompassing many distinct aspects. Table 3.1 provides some of the main accepted definitions of SA (Stanton, Chambers and Piggott, 2001).

The diversity of definitions presented in Table 3.1 is due to different focuses on the concept of Situation Awareness. The first definition focuses on the perception and understanding of the environment; the second focuses on the interaction between the environment and humans; the third focuses on mental models (Stanton, Chambers and Piggott, 2001). In particular, the situation is understandable in reference to the main interpreters: human beings and the external environment. The former can be identified through Projection, Knowledge, Mental Models, Perception, and Reflection. It is evident how the presented definitions focus on different elements of identifying a user (human being) interacting with the external environment.

Although the topic is the same, the way in which Situation Awareness is defined leads to different approaches and theories.

Endsley focuses his theory on a three-level approach (Endsley, 1988, 2000; Endsley and Garland, 2000):

- Level 1 - Perception: the first level is based on the acquisition of information by the user. This level is based on the perception of the surrounding environment by the individual;

- Level 2 - Comprehension: the level in which the individual acquires the information acquired through perception to be aware of the surrounding environment. Moreover, the degree of comprehension depends not only on the information but also on the individual interacting;
- Level 3 - Prediction: the highest level in which the information acquired (level 1) and comprehended (level 2) is processed to predict future actions. Predicting is directly linked to the quality with which the previous levels are performed.

Table 3.1 *Definitions of Situation Awareness.*

N	Definition	Reference
1	"Situational Awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and a projection of their status in the near future."	(Endsley, 1988)
2	"Situational Awareness is the conscious dynamic reflection on the situation by an individual. It provides dynamic orientation to the situation, the opportunity to reflect not only the past, present and future, but the potential features of the situation. The dynamic reflection contains logical-conceptual, imaginative, conscious and unconscious components which enables individuals to develop mental models of external events."	(Bedny and Meister, 1999)
3	"Situational Awareness is the invariant in the agent-environment System that generates the momentary knowledge and behaviour required to attain the goals specified by an arbiter of performance in the environment."	(Smith and Hancock, 1995)

The degree of Awareness increases along with levels until it reaches its maximum in the ability to predict. In practical terms, this indicates the ability to prevent problems. This SA model has been applied more in aviation, and the three levels are related to the pilot's ability to acquire, understand, and process the data acquired during the flight. The model also explains why the understanding degree is related to the user's ability to take the data acquired during the first level. In particular, the Endsley situational awareness model also illustrates several variables that can influence the development and maintenance of situational Awareness, including user, activity, and environmental factors. For example, users vary in their ability to acquire situational Awareness; therefore, the provision of the same System and training does not ensure similar situational Awareness among different individuals. The Endsley model shows how situational Awareness constitutes the primary basis for subsequent decision-making and performance in the functioning of complex and dynamic systems. As shown in Figure 3.1, Situation Awareness is determined by environmental monitoring followed by decision and action phases that influence the surrounding environment.

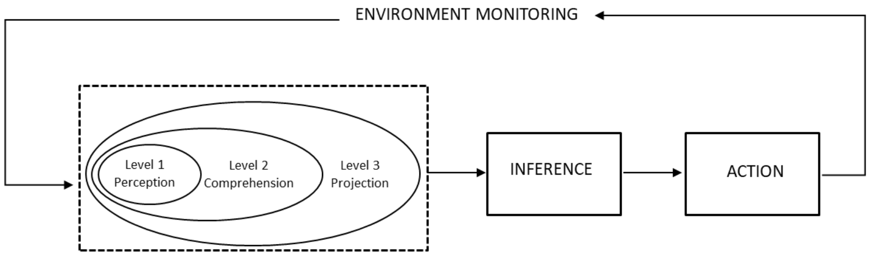


Figure 3.1 *Endsley's model of Situation Awareness*

A second approach to Situation Awareness is proposed by Bedny and Meister (Bedny and Meister, 1999) and is based on eight interacting sub-systems. The sub-systems are functional blocks that aim to situational Awareness through their interaction. The modes of interaction between the functional blocks are described in Figure 3.2 and Table 3.2. Fundamental to the structure are functional blocks 8 (Model), 2 (Image), and 3 (Conditions). The first two blocks are fixed, while the last can be manipulated and adapted to different situations.

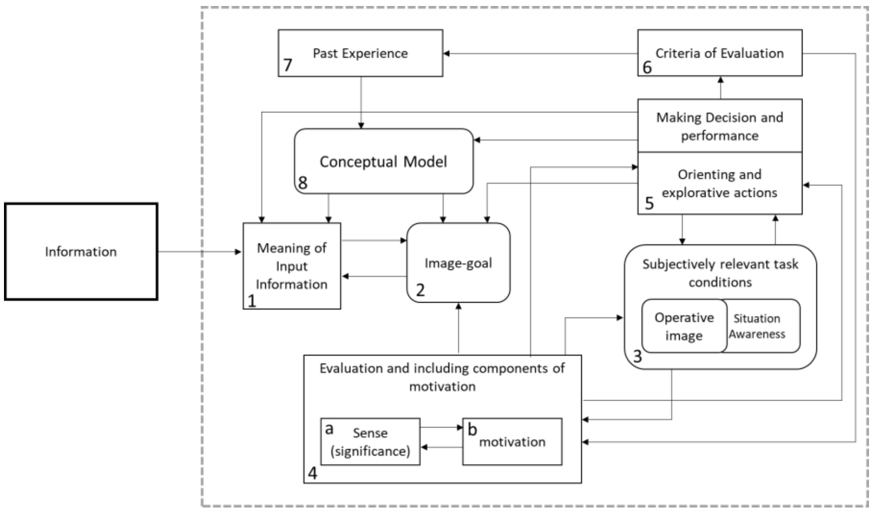


Figure 3.2 *Interactive Sub-Systems approach for Situation Awareness proposed by Bedny and Meister*

Another theory, called the perceptual cycle, was proposed by Smith and Hancock (Smith and Hancock, 1995). On the theory basis, there are perception and acquisition activities of information by the person that allows the state of Awareness to be reached through how the subject processes information. This approach is based neither on the individual nor the environment but on the interaction between the two entities. In particular, in this approach, it is assumed that the possibility of acquiring and anticipating events allows the knowledge of the external environment through the continuous human-environment interaction in a continuous cycle that will lead to the constant new information acquiring.

Table 3.2 *Summary of functional blocks in the Situation Awareness approach proposed by Bedny and Meister*

Block	Function	Input Block	Role
1	Meaning	0-2-5-7	Interpretation of information from world
2	Image	1-4-5-8	Conceptual 'image' of information-task-goal
3	Conditions	4-5	Dynamic reflection of situation and task
4	Evaluation	3-6	Comparing motivation and performance
5	Performance	3-4	Interacting with the world
6	Criteria	4-5	Determining relevant criteria for evaluation
7	Experience	6	Modify experience to interpret new information
8	Model	7	Modify world model to interpret new information

Referring to Endsley's definition (Endsley, 1988, 2000), the Situation Awareness model is based on three levels: Perception, Comprehension, and Prediction. In other words, the SA can be interpreted through three different levels, which start from the information perception to arrive at the comprehension in order to be able to make predictions on what will happen. In this last concept, the introduction of time in Situation Awareness analysis is intuitive. It can be assumed to be a valuable time interval (Best *et al.*, 2010; Lu, Coster and de Winter, 2017) or as system management in the evolution (Jones and Endsley, 2004). Therefore, the temporal aspect mainly concerns the levels relative to comprehension and forecasting. (Endsley, 2000).

Situation Awareness covers many domains; for example, the predictive phase is closely linked to the problem of decision-making (Endsley, 1995; Nibbelink and Brewer, 2018). In particular, the predictions provided through Situation Awareness can be used for prevention. This assumption distinguishes the two concepts: Situation Awareness allows decision-making;

on the other hand, the latter is not included in the awareness process. In fact, it is impossible to make choices without a good knowledge of what is happening in the considered domain, but at the same time, wrong decisions can be performed despite a high level of Situation Awareness.

Other features central to the achievement of Situation Awareness are Attention and Working Memory. Attention is related to appropriately acquiring information during the perception phase (Adams, Tenney and Pew, 1995; Endsley, 2000; Underwood, 2007). Examples include drivers while driving or airplane pilots during a flight. Working memory is a limited human cognitive capacity divided between temporarily holding information and strategically manipulating that information for future action (Baddeley, 1992; Wimisberg, 2007). In fact, Working Memory (Gutzwiller and Clegg, 2013; Leu, Tang and Abbass, 2014) sometimes represents a bottleneck for the development of SA (Endsley and Robertson, 2000; Wimisberg, 2007). There are methods to decrease the amount of memory associated with Situation Awareness, such as chunking, justification of information, or restructuring the environment. (Endsley, 2000).

The themes described allow the definition and improvement of SA, but they are not the only factors that influence the achievement of good Awareness. First, it is necessary to set the objectives to be reached; this phase is fundamental for developing Situation Awareness. (Endsley, 2000). In this way, it will be possible to acquire, during the Perception phase, the appropriate information to be processed in the second level and, therefore, to develop suitable forecasts according to the target.

A further important aspect is an expectation, which, based on contextual information (Abowd *et al.*, 1999b), allows to focus the perception phases anticipating the environmental context belonging to the individual, also facilitating the prediction phase. However, a perception of information anomaly, such as a not-expected event, can seriously undermine the effectiveness of Situation Awareness. Therefore, it is necessary to balance the expected information with a valuable capacity to react to non-predictable phenomena and adjust the prediction capacity associated with Level 3 accordingly.

Automaticity degree can also influence the SA (Endsley, 2000), allowing more accessible model building with a lower level of attention. Nevertheless, at the same time, this lower level of attention can affect the Situation Awareness concept since excessive automaticity can lead to a slow response to external information.

From what has just been described, it is clear that obtaining Situation Awareness is a complex process that requires a considerable amount of theoretical knowledge and incorporates different concepts related to it. Endsley provides a summary of how to develop, and thus improve, situational Awareness through a list of steps (Endsley, 1995; Stanton, Chambers and Piggott, 2001):

1. Decrease the computational cost to users;
2. Data structuring for optimizing the levels associated with understanding and prediction;
3. Select information by integrating user's goals;
4. Leverage system state descriptors for SA improvement;
5. Provide critical signals during significant events;
6. Provide support to the SA through reporting on the operator's objectives;
7. Support through predictions about future events or states;
8. Leverage parallel processing to provide data from multiple sources.

Those steps summarise the complexity associated with achieving Situation Awareness and, at the same time, allow us to understand the potential of a properly executed SA path.

The concept of Situation Awareness covers many application areas and is exploited in diverse scenarios. The use cases are usually complex; one of the most popular is represented by aviation. In particular, in the aviation field, SA assumes a relevant role in the training of pilots (Nguyen *et al.*, 2019). They must acquire information from the surrounding environment adequately, process this information, and predict what will happen in the short and long term. The most widely used formal models developed in SA were conceived in aviation. Gluck et al. propose a *concurrent verbal protocol* supported by eye movement data to determine the degree of identified Situation Awareness (Gluck, Ball and Krusmark, 2012). This approach has also been exploited in other areas, such as road vehicle driving. (Thomas *et al.*, 2015). Hooney et al. (Hooey *et al.*, 2011) construct a protocol for assessing SA in aviation by measuring different aspects of Situation Awareness. In particular, it aims to quantify the following parameters:

- The measure of current SA at the time t_i referred to the task i through the weights of the Situation Elements w_r $r = 1, \dots, m$ and w_d $d = 1, \dots, n$ perception levels p_r $r = 1 \dots, m$ and p_d $d = 1, \dots, n$.

$$SA_{actual}(t_i) = \sum_{r=1}^{m} w_r p_r + \sum_{d=1}^{n} w_d p_d \qquad (3.1)$$

- Measurement of optimal SA

$$SA_{optimal}(t_i) = \sum_{r=1}^{m} w_r + \sum_{d=1}^{n} w_d \qquad (3.2)$$

It is optimal because the levels of perception $p_r\ r = 1 \dots, m$ and $p_d\ d = 1, \dots, n$ belong to the range [0,1].

- A measure of the portion of the Situation Elements of which an individual is aware.

$$SA_{ratio}(t_i) = \frac{SA_a ctual(t_i)}{SA_{optimal}(t_i)} \in [0,1] \qquad (3.3)$$

The developed model is validated through a simulation in which two pilots interact while an airport landing is performed (Hooey *et al.*, 2011).

A model for quantitative analysis of Situation Awareness through probability theory, is defined as *attention allocation model*, which focuses on conditional probabilities of the Bayesian Theorem (Shuang, Xiaoru and Damin, 2014). The method involves defining the following parameters:

- β_i the frequency of occurrence of the Situation Element i −element;
- V_i information priority indicator;
- Sa_i indicates the salient element;
- E_i amount of effort required to acquire the information;

The attention resource A_i associated with the Situation Element SE_i is defined as in (3.4):

$$A_i = \frac{\beta_i V_i Sa_i}{E_i} \qquad (3.4)$$

It is possible to obtain the portion of attention allocation f_i normalizing the element A_i

$$f_i = \frac{A_i}{\sum_{j=1}^{n} A_j} \in [0,1] \qquad (3.5)$$

The value f_i obtained is seen as the probability of occurrence of the event a_i at the time t:

$$p(a_i) = f_i \qquad (3.6)$$

By linking by dependency relations the highest levels of Awareness with the lowest levels, it is possible to define:

- $k_i = p(b_i|a_i)$ the conditional probability on the occurrence of the event a_i that the event b_i occur, where the event b_i indicates that the i-th Situation Element SE_i is not understood.

- $p(c_i|a_i) = 1 - k_i$ the conditional probability on the occurrence of the event a_i, of the occurrence of the event $c_i$or that the Situation Element i-th SE_i is understood.

Thus, the level of knowledge of the i-th Situation Element is obtained from the formula (3.7).

$$\bar{p}_i = \frac{1}{2}p(a_i b_i) + p(a_i c_i) \tag{3.7}$$

Assuming that the coefficient e_i associated with sensitivity, that is the level of understanding of SE_i on Situation Awareness, is equal to the coefficient u_i relative to the importance of each SE_i, it is possible to obtain that at the time t_j the SA is expressed through the coefficient of attention allocation:

$$SA(t_j) = \sum_{i=1}^{n} u_i \bar{p}_i = \frac{\sum_{i=1}^{n}\left(1 - \frac{1}{2}k_i\right)u_i A_i}{\sum_{h=1}^{n} A_h} \tag{3.8}$$

This model was subsequently improved through cognitive process analysis (Liu, Wanyan and Zhuang, 2014), leading to empirical results that determined that the evolution of the SA model can be functional for the design of visualization interfaces that minimize pilot error (Wu *et al.*, 2016; Nguyen *et al.*, 2019). Several models related to the aviation field are based on a qualitative analysis of SA. These models cover flight phases and air traffic control from the ground. In this case, a good level of Situation Awareness can bring benefits by providing additional information about weather conditions or events that pilots cannot access at certain times. Vu et al. propose models in which pilots must avoid areas with adverse weather conditions through instructions provided by air-traffic controllers (Vu *et al.*, 2010). However, in this way, air-traffic controllers are exposed to higher peak workloads, in some cases, than pilots. To cope with this issue, the concept of Situated SA aimed at obtaining and maintaining Situation Awareness through automated tools is introduced in (Chiappe, Vu and Strybel, 2012). These tools, avoiding memory overload, do not have to process complex information, which is functional to maintain SA. Blasch (Blasch, 2013), in this sense, proposes the use of automatic tools able to provide more support to the users through visualization systems of the collected information able to increase the SA.

Like aviation, another widely used field of application of Situation Awareness is the identification of Maritime Anomaly (Riveiro, Pallotta and Vespe, 2018; Tu *et al.*, 2018). The application in maritime safety has been made possible on the one hand by the amount of data collected and on the other by the techniques development data clustering and analysis. (Tu *et al.*, 2018). Among the types of data that can be found and are helpful for Maritime Anomaly Detection, and therefore of Situation Awareness, can be found contextual data related to location, weather conditions, or information related to human activities. Several solutions have been proposed to address the heterogeneity and a large amount of data available. In particular, Rivero et al. (Riveiro, Falkman and Ziemke, 2008) propose a combined methodology of data visualization, interaction, and extraction techniques that filter out anomalies caused by the abundance of information with which operators interact. This methodology is based on two layers: the first one is aimed at the correct data visualization, while the second one allows the interaction with the data through data mining techniques.

In an attempt to collect and organize knowledge efficiently, Brüggemann et al. (Brüggemann *et al.*, 2016) propose an Ontology-based approach for understanding and processing real-time data in order to achieve Maritime Situational Awareness. The Ontology related to the Real-Time Maritime Situation Awareness System processes various data types, such as data stream, static data, and open data. Roy et al. (Roy and Davenport, 2010) present an Ontology approach that leverages expert knowledge and aims to support maritime personnel in anomaly detection. The proposed method has the ultimate goal of having automated reasoning capabilities through maritime Ontology (Roy and Davenport, 2010). The purpose of such studies is to fuse uncertain knowledge with known data in order to achieve Situation Awareness. In this direction, Fischer et al. (Fischer and Bauer, 2010) propose an approach that exploits data from objects to reach the third level of Situation Awareness, i.e., predicting security-threatening events. The System proposed by Fischer et al. is based on acquiring heterogeneous information from sensors. After the data acquisition phase, the representation and prediction model exploits Bayesian techniques, as also in the paper of Fooladvandi et al. (Fooladvandi *et al.*, 2009). Van Den Broek et al. exploit both the information acquired from sensors and the context to identify possible threats and improve Situation Awareness (van den Broek *et al.*, 2011).

In this direction, the advent of the Internet of Things (IoT) (W. E. Zhang *et al.*, 2020) contributes to environmental monitoring, fundamental to the determination of SA, more efficiently. Situation Awareness finds a significant advantage in working with the Internet of Things paradigm. In fact, the increasing use of smart devices and the development of wireless technology can represent an essential advantage for achieving the first two levels of SA naturally and efficiently. Many approaches in the literature leverage the potential of IoT in SA; for example, (Mozzaquatro, Jardim-Goncalves and Agostinho, 2018) and (Kolbe *et al.*, 2017) proposed systems that can combine graph approaches for context representation, such as Ontologies, for knowledge construction. Other approaches can exploit Ontologies for user-defined role construction to increase SA in this domain. In detail, after acquiring heterogeneous information provided by sensors, the data is processed and integrated into the ontology, making the System able to provide personalized alerts based on user-defined rules (Xu *et al.*, 2017). In an attempt to support users in risk scenarios, Krytska et al. (Krytska, Skarga-Bandurova and Velykzhanin, 2017) propose an IoT-based system that can handle dangerous scenarios by suggesting safe sites to people during an emergency. The System is based on IoT technologies that actively participate in the SA process. In this way, Glowacka et al. (Glowacka, Krygier and Amanowicz, 2015) propose a model in which IoT-based Smart Devices are able to acquire Situation Awareness and ready to react to possible threats. This model consists of several modules: Information acquisition and analysis module; inference module that evaluates objects based on observations and recommendations; classification module aimed at the final evaluation; Reaction module that exploits previous modules to enable possible reaction threats.

In this sense, due to the diffusion of the Internet of Things devices in urban areas, other complex scenarios, such as Smart Cities, have been investigated. In particular, Alamgir Hossain et al. propose a framework based on edge computing that is able to provide services of interest to residents and perform an appropriate decision-making phase through situation detection (Alamgir Hossain, Anisur Rahman and Hossain, 2018). This approach is based on multiple consecutive layers. In the primary layers, raw data are collected from smart objects, then the data representation is unified, and the data are aggregated to determine the situation. The last step deals with the representation of the situation. In (Li, Wu and Liu, 2013), an Ontology-based model is proposed to monitor traffic in Smart Cities; Lee et al. (Lee *et al.*, 2020) describe a service, functional to Situation Awareness, based on data from IoT sensors integrated with public applications proper to the Smart Cities paradigm.

In an attempt to integrate the various systems that aim to achieve Context-Awareness and Situation Awareness in IoT-based systems, Alvarez et al. (Alvarez, Morales and Kraak, 2019) propose an approach for information management and processing that has the following characteristics:

- Collaboration between different sources for data acquisition;
- Geographic event detection capability, where the term event is defined as “a formal representation of a relevant geographic event describing a set of attributive, spatial, and temporal characteristics” (Morales and Garcia, 2015);
- Real-time processing capabilities to optimize the quality of service provided by reducing transfer and processing delays;
- Scalability, in order to adequately deal with the amount of computation and processing power required.

Situation Awareness, in the context of Smart Cities, can be leveraged for city planning and decision making. Again, one significant difficulty is working with heterogeneous data, which requires unification, filtering, and pre-processing steps to obtain SA (Eräranta and Staffans, 2015).

As can be seen from the different studies discussed in the previous chapters, the Internet of Things paradigm can be adequately exploited in order to achieve Context and Situation Awareness. The following section will discuss the proposed approach that aims at effective integration of the three paradigms presented in order to manage complex scenarios.

Chapter 4
Situation Awareness in IoT based Complex Systems

This research work aims to investigate new methodologies that can support users in managing complex scenarios based on IoT. This issue could be solved by developing a suitable methodology adaptable to the context and situation to perform predictions. This section will show in detail the proposed approach.

Many of the approaches in the scientific literature rely on Machine Learning techniques to manage IoT-based complex scenarios. Among the most used methodologies include Support Vector Machine, K-Nearest Neighbors, Artificial Neural Network, and all the types related to Deep Learning (Akhter and Sofi, 2021; Ghazal *et al.*, 2021; Kousis and Tjortjis, 2021).

However, these approaches often do not adequately exploit contextual information from the environment and use methodologies that can often be represented as black boxes where the user cannot intervene in decisions and choices (McGovern *et al.*, 2019; Rudin, 2019; Janiesch, Zschech and Heinrich, 2021). In this field, the paradigms presented in the previous chapters related to the Internet of Things, Context Awareness, and Situation Awareness are crucial. In fact, through smart devices belonging to the Internet of Things paradigm, it is possible to acquire a wide range of information related to a user, domain, and environment. The Context Awareness could exploit the specific conditions acquired through the 5W+1H model to personalize further needs and actions related to the user. Finally, the models related to the Internet of Things and Context Awareness are integrated into perception and understanding phases typical of Situation Awareness. The latter provides the prediction assumptions to the System.

In summary, the proposed methodology aims to exploit the IoT paradigm to monitor the reference environment and approaches based on Context-Awareness and Situation Awareness, which are able to formally map the reference domain and provide answers in terms of predictions to support users in managing complex scenarios.

Information management environments, or more generally pervasive data contexts, can be supported by context representation approaches and enhanced by adopting probabilistic approaches such as Bayesian Network (BN) (Zhong *et al.*, 2014; Colace *et al.*, 2015; Goel, Chaudhury and Ghosh, 2017). BNs can offer an analysis framework that can adequately support users through their ability to model data. Some of the advantages of probabilistic approaches are the ability to model complex systems, make predictions and diagnoses, calculate the probability of an event, update probabilities based on evidence, represent multi-modal variables, and offer a user-friendly graphical and compact approach (Weber *et al.*, 2012). As mentioned earlier, a further benefit could be provided by introducing methodologies capable of representing context, such as the Context Dimension Tree (CDT). The CDT represents a valuable tool used for applications that foresee the choice of places of interest. (Colace *et al.*, 2014). Furthermore, the CDT, or more generally context-aware approaches, leads to the rationalization of the data provided to users and the personalized distribution of information (Panigati *et al.*, 2012). A widely used method to represent reality is Ontologies. An Ontology can adequately support pervasive Context-Aware systems (Chen, Finin and Joshi, 2003a). Moreover, there is a strong connection between Ontology and Bayesian networks. (Trifonova *et al.*, 2017). In particular, it is possible to build BNs through Ontologies (Helsper and L. C. Gaag, 2002), and vice versa, the construction of Ontologies is possible automatically through the use of BNs. (Colace and De Santo, 2010).

Addressing the above context representation methodologies and the capability of BNs, which from experimental evidence and through probabilistic approaches are able to identify probable events, it is necessary to introduce techniques and methodologies that can manage the context in real-time, to improve the management of complex scenarios. This work aims to introduce and explain a Multilevel Graph (MuG) Approach, a methodology of fusion between Ontologies, Context Dimension Trees, and Bayesian Networks to help expert and ordinary users manage their needs and provide suggestions to improve people's livability in complex people scenarios.

In the following subsection, a general architecture for the inclusion of the presented System will be presented. Then, the graph structures used, Ontologies, Context Dimension Tree, and the Bayesian Networks will be introduced in detail. In the end, a formalization of the proposed MuG Approach will be reported.

4.1 The Proposed Architecture

Developing an architecture based on IoT systems able to integrate Context and Situation Awareness approaches for the management of complex scenarios is challenging. Based on IoT systems, the architecture must be supported by typical IoT paradigm multilayer architectures. In this case, the proposed architecture will be based on the most generic and high-level architecture consisting of three layers: Perception, Network, and Application (Al-Fuqaha *et al.*, 2015). The proposed architecture must be scalable, interoperable, distributive, and able to operate with low resources devices (Abdmeziem, Tandjaoui and Romdhani, 2016). In addition, it must be able to manage Big Data coming from different heterogeneous sources. Moreover, it must extract and manipulate semantic information using context and situation management tools. To this end, it is possible to refer to Endsley's model shown in figure 3.1 (*Situation Awareness Analysis and Measurement*, 2000), which represents a fundamental starting point concerning the characteristics of the presented System. The purpose of the System, in fact, is to acquire data from the surrounding environment, store, process, and use data by making it available to different categories of users. In particular, users can obtain applications from the System for data visualization and interaction with the System and the external environment. Enclosing the features of the proposed System in a general architecture, this can be represented in Figure 4.1.

The proposed architecture provides several blocks with different functionalities. The top layer of the System is the Environmental Data Acquisition Layer. This layer collects methodologies and techniques for retrieving data from various sources. The purpose, at this stage, is to store and filter as much data as possible, homogenizing them and making them available to the System and users. Data acquisition sources are formed by IoT sensors and actuators and API or Open services. IoT sensor nodes are capable of acquiring information from the reference domain; these data are acquired by specific sensors that can report information about general environmental conditions or the asset under consideration, depending on the application cases. In the IoT sensor network, there are also actuators able to interact with the environment, i.e., interacting with environmental conditioning systems, triggering alarms, activating security actions remotely. The System is able to interact with API and Open data services, integrating data acquired from IoT sensors. These data enrich the System concerning different aspects, i.e., services that can provide information on events, weather, booking ticketing services, and descriptive insights into the assets under study. All data stored in the top layer are acquired, pre-processed, and transferred to the next layer: the Knowledge Database. In order to define the proper use of the acquired data, the information is structured in the Knowledge Database, making it valuable and ready for processing.

The knowledge database represents the information core of the proposed architecture. In fact, the three phases of Situation Awareness proposed by Endsley (*Situation Awareness Analysis and Measurement*, 2000) are performed inside this layer. The KDB collaborates with the Inference Engine to construct and extract knowledge from the graph structures included in the database. In particular, the perception phase is developed through the mentioned graph structures, representing the structured knowledge base. In addition, the comprehension phase is performed by building and exploiting graph structures able to interpret and manage the context, such as Ontologies and CDT. Ontologies allow the description of a domain of interest through a formal, shared, and explicit representation.

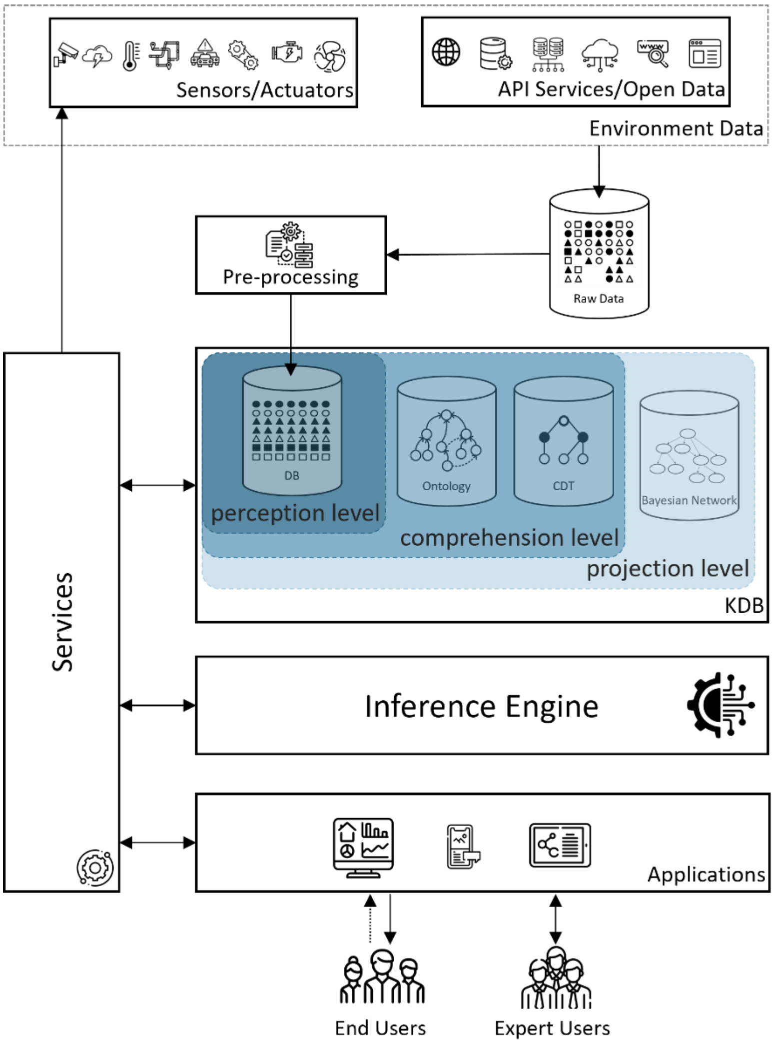


Figure 4.1 *General System Architecture*

More in detail, it is an axiomatic theory expressible in descriptive logic. The ontologies, especially in their graph form, can dialogue with tools such as the CDT and the Bayesian Networks (Pan *et al.*, 2005). The CDT allows managing all the possible contexts through a representation tree. In particular, the nodes present within CDT are divided into two categories, namely dimensional nodes, and conceptual nodes. Dimension nodes describe the possible dimensions of the application domain, and concept nodes collect all the possible values that a dimension can assume. This model is able to interface and query a database efficiently in order to select the right services based on the possible selected context. The third phase of SA, the projection phase, is performed through probabilistic graph structures: Bayesian Networks. This structure gives the System an additional fundamental capability: the predictive ability. The Bayesian Networks are graphical-probabilistic models that represent a set of stochastic variables with their conditional dependencies through a direct acyclic graph (DAG). Such models exploit Bayes' Theorem and are able to predict the probability of occurrence of a given event. Moreover, thanks to their structure, the Bayesian networks are able to interface appropriately with the CDT and Ontologies (Helsper and L. C. van der Gaag, 2002).

To understand in detail the interaction between the graph structures for the MuG Approach operation, it is possible to refer to Figure 4.2, where the workflow of the proposed methodology is shown. The workflow starts from the data from the domain of interest provided by the sensors spread in the environment. From the information collected, and through the contribution and comparison with experts, a Task Ontology can be developed (Ikeda *et al.*, 1998; Abrahao and Hirakawa, 2017). Subsequently, it is necessary to design a CDT that describes all the possible contexts within the reference domain. At this point, the Ontology and the CDT are combined to obtain semantic relations, which constitute semantic constraints valuable to build the Bayesian Network. The semantic constraints in combination with structural learning algorithms (Scutari, Graafland and Gutiérrez, 2019) and the available data develop the Bayesian Network to predict events. To summarize, in this way, the need to perform a decision in a given context can be satisfied by using the correct information provided by the architecture. This information is characterized by innovative elements based on knowledge management and organization, formal context representation, and inferential approaches.

After the knowledge construction and organization phases, in which the KDB and the Inference Engine are involved, the acquired information is exploited by the Application block. The application layer contains all the possible applications that users can use. Different applications suitable for different user categories can coexist in this layer. In detail, expert users can benefit from technical and specific information provided and are able to interact with the System. In contrast, standard users can only access selected contents and cannot interact with the System to perform actions; however, on particular occasions, in order to enrich the KDB, the System is able to ask for feedback from standard users.

All operations between the KDB, Inference Engine, and Applications layer are performed through the Services Module. This module collects all the services able to communicate through the different blocks allowing the various operations of the System, such as processing data between the different layers within the KDB, providing predictions, representing information through applications etc. In addition, this module allows user interaction through actuators.

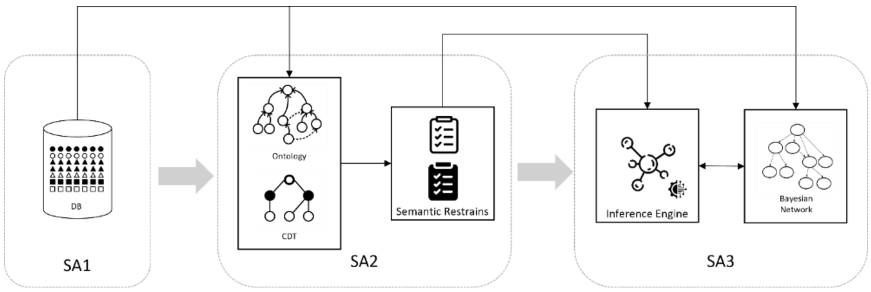


Figure 4.2 *Multilevel Graph Approach Workflow*

The following section will detail the graph structures and algorithms to achieve Context and Situation Awareness, which allow the System to perform predictions.

4.2 Graph structures and Algorithm

In this section, the graph structures used by the System will be presented. In particular, this section introduces the tools used to formalize the Multilevel Graph Approach. Therefore, the graph approaches used to build semantic relations will be introduced, such as Ontologies, Context Dimension Tree (CDT), and Bayesian Networks (BN). These approaches will be presented through their formalisms, present in scientific literature, thanks to which it will be possible to proceed to the formal implementation of the proposed methodology. Moreover, this section will introduce the structural learning algorithm, K2 (Cooper and Herskovits, 1992), which represents the central methodology on which will be based the formalization of the proposed approach for the construction of Bayesian Networks.

4.2.1 Ontologies

Ontology was introduced by Parmenides (505-504 BC) and later defined by Aristotle as the concepts such as existence, being, becoming, and reality. Nowadays, Ontology is a valuable tool to represent information to be interpreted, shared, and reused. In fact, Ontology represents a valuable tool able to limit conceptual or terminological confusion, generating, in other words, a "*shared vocabulary*", although more powerful. Let us consider ontology as a "*shared vocabulary*". It will not only be composed of a list of terms, but it will also contain an unambiguous definition of the word that contains the meaning accessible, understandable, and shareable by anyone.

Furthermore, unlike a "*shared vocabulary*", the Ontology contains the semantic relationships between the entities described. In other words, Ontology represents a classification able to define the relationships between domain elements (Guarino, Oberle and Staab, 2009). Ontological formalism represents the key to formalizing the environment surrounding us. This formalization could lead to communication between human beings and machines (Arp, Smith and Spear, 2016; Guzman, 2020). In this common and shared approach between humans and machines, the reality surrounding us must be formalized univocal and unambiguous. Due to this formalism, it is possible to imagine that each piece of information has a precise meaning related to the context, very similar to the association mechanism of the human mind. In this way, information can be shared among software systems, and, as it happens with human beings, it can be processed.

Providing an elementary and unambiguous explanation of the concept of Ontology is still part of the scientific discussion. Some definitions that are widely accepted are as follows:

- According to Neches et al. (Neches *et al.*, 1991), an ontology identifies the basic terms and relationships of a given domain, defining the vocabulary and the rules in order to combine terms and relationships, going beyond the vocabulary itself.

- An ontology is a set of words capable of describing a domain that can be used as a basis for knowledge (Swartout *et al.*, 1996).

- An ontology is a means of explicitly describing the conceptualization present behind the knowledge, which is represented in a knowledge base (Bernaras, Laresgoiti and Corera, 1996)

The definition best suited to our case and complete is formalized by Gruber, where ontology is defined as a formal and explicit specification of a shared conceptualization (Gruber, 1995) by analyzing this definition, it is possible to infer that:

- **Conceptualization** refers to obtaining an abstract model of a specific real-world phenomenon by identifying the relevant concepts that characterize it.

- **Explicit** means that the concepts, properties, and constraints that characterize the phenomenon are explicitly defined clearly and univocal.

- **Shared** is necessary because the ontology's knowledge is based not on an individual result but shared by a group.

- **Formal** indicates that an ontology is machine-processable.

The so-described Ontology consists of concepts and attributes. The concepts represent the set of elements of reality to be described, and the attributes are related to these concepts, making them unique and unambiguous. Ontology also contains a set of hierarchical and semantic relationships between the various concepts. The formal concept can be summarized through a tuple, which represents a set of attributes X as a function that associates a value of the attribute domain with each attribute belonging to X.

$$O = C, A, H, R_T, A_x \tag{4.1}$$

The formula (4.1) formally describes the Ontology O, where:

- C is the set of concepts. $c \varepsilon C$ expresses a concept, and in every ontology, there is always the root marked as "Thing". Each $c \varepsilon C$ can have a set of descendant concepts (C_{DN}) and a set of ancestor concepts (C_{AN});

- A is the set of attributes. For $c \varepsilon C$ the set of attributes is expressed as $A_C = \{a1, \ldots, an\}$;

- H expresses the hierarchy of concepts. This set contains the "*is-a*" relationships;

- R_T is the set of types of semantic relations. $R_T = R_{TD} \cup R_{TU}$. R_{TD} is the set of predefined relationships (*same_as*, *disjoint_with*, *equivalent*), while R_{TU} is the set of user-defined relationships types.

- A_x is the set of axioms, i.e., the primary relationships that do not consider semantic relations.

Ontologies can be further classified in (Colace, Chang and De Santo, 2010):

- **Lightweight**: it represents taxonomies of concepts, in which there are simple relations between them (such as the specialization "*is-a*") and the properties that describe them;

- **Heavyweight**: adds axioms and limitations to lightweight ontologies.

In the case of lightweight ontologies, there are no R_T and R sets, therefore the Ontology is denoted by O_L and is represented by formula (4.2).

$$O_L = C, A, H \tag{4.2}$$

In those heavyweights, the semantic expressiveness is increased with the addition of A_x axioms, then the Ontology, denotated by O_H, can be described as in formula (4.3).

$$O_H = C, A, H, R_T, R, A_x \quad (4.3)$$

These formalization characteristics and potentials allow us to say that it is possible and convenient to use ontologies whenever necessary to share information. These are useful regardless of the technologies used, the information architecture, and the application domain. Barry Smith and Christopher Welty (Smith and Welty, 2001) identify three significant areas of application of ontologies in computer systems: knowledge engineering, which deals with the design of knowledge bases and systems based on them; conceptual modeling, which covers the initial phases of database design; and software engineering, concerning so-called object-oriented languages. Nowadays, it is possible to say that these applications are used in different fields such as semantic web, shared knowledge organization, collaborative engineering design, development of shared semantic applications, management of complex information environments, etc. (Hachem, Teixeira and Issarny, 2011; Munir and Sheraz Anjum, 2018). Those applications are possible through the versatility and the degree of generality of the Ontology.

4.2.2 Context Dimension Tree

In this section, the Context Dimension Tree (CDT) is presented in detail, as it is the representation model of the context that has been selected for the proposed approach. The CDT can be represented according to formula (4.4).

$$T = \langle N;\ E_T;\ r \rangle \quad (4.4)$$

This graph needs to satisfy the following properties (Bolchini *et al.*, 2006):

- $N = N_D \cup N_C \cup N_A$ such that N_D, N_C, and N_A are pairwise disjoint.

N_D represents the set of dimensions, N_C the set of concepts, and N_A the attributes.

- $r \in N_C$, i.e., the root of the tree is a concept node.
- $E_T = E_R \cup E_A$ such that $E_R \cap E_A = \emptyset$. E_R represents the sub-element of relationships, whereas E_A represents the attribute of relationships.
- $\forall e = \langle n, m \rangle \in E_R$ either $n \in N_D \wedge m \in N_C$ or $n \in N_C \wedge m \in N_D$; i.e., a dimension node has as children concept nodes and a concept node has as children sub-dimension nodes.
- $\forall n \in N_A, \forall e = \langle n, m \rangle \in E_A, m \in N_D \cup N_C$; i.e., an attribute node has a dimension node or a concept node as a father.
- $\forall n \in N_A, \neg\exists e = \langle n, m \rangle \in E_T$; i.e., attribute nodes are leaves.

- $\forall n \in N_A, \forall e = \langle m, n \rangle \in E_A$ if $\exists e_1 = \langle m, n_1 \rangle \in E_T$ then $n = n_1$; i.e., an attribute node is an only child.

- $\forall n \in N_D$ such that $\neg \exists e = \langle n, m \rangle \in E_R$ then $\exists e_1 = \langle n, m_1 \rangle \in E_A$; i.e., dimension nodes without concept children must have an attribute child.

The CDT comprises a root r, a set of nodes N, and edges E. The nodes are divided into the subsets of the nodes dimension N_D, colored black, and into the concept node N_C, colored white, representing the possible values that the dimensions can assume. The set of links between the nodes represents the edge set E. New edges on the Context Dimension Tree represent the set of constraints, which allows us to represent the possible contexts.

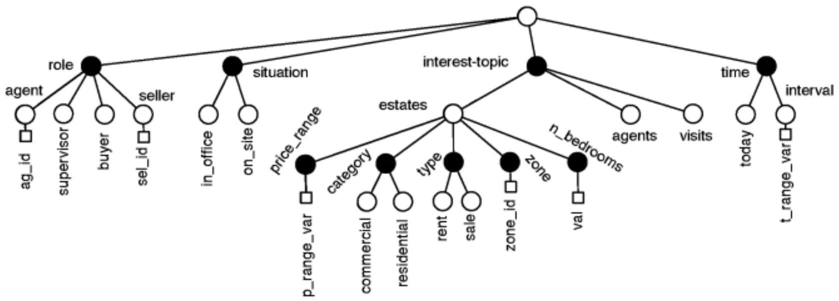


Figure 4.3 *Example of Context Dimension Tree*

The root r is a concept node representing the most general possible context, corresponding to the entire dataset. Direct descendant dimensions of the root are called principal dimensions because they define the user's different characteristics and contexts in which they act. In the example in Figure 4.3, relating to a real estate agency, the main dimensions are the role of the user, the interest topic, the situation, the time, and the place in which it is located. Moreover, each value can be further specialized through sub-dimensions, forming a subtree. (Parent *et al.*, 2007). For example, the "*estates*" interest topic can be analyzed based on price, category, property type, and the number of rooms. Each node of the CDT is characterized by its typology, which can be dimension or concept, and its label; it can be uniquely identified by the only path connecting it to the root. The hypothesis is adopted that each label is unique in a tree; therefore, its label can identify each node. Links between nodes are not labeled. Alternating between dimension nodes and concept nodes allows for the creation of generations, each of which will be composed of nodes of the same color, and each color will be alternated as one proceeds down the tree. Therefore, each dimension node can have concept nodes as its only ones and vice versa.

Furthermore, it is possible to associate one or more parameters to tree leaves' concept nodes and dimension nodes. Each parameter allows to refine the data selection further and select a subset of the particular dataset. For example, the "*val*" parameter associated with the *n_bedrooms* dimension filters properties according to the number of rooms. For each dimension node, it is possible to select only one concept node among its members or, if it does not have any child node, only one parameter must be selected. The use of parameters increases the expressive power of the model, as it makes it easier for the designer to use. The introduction of the parameters is necessary because not all the concepts expressed by a dimension can be enumerated.

4.2.3 Bayesian Network

Bayesian Networks (BNs) belong to the probabilistic graphical models approach. According to Jordan, graphical models are a marriage between probability theory and graph theory (Jordan, 1998). They represent a natural tool for dealing with applied mathematics and engineering problems. In particular, they play an increasingly important role in designing and analyzing machine learning algorithms. In general, a complex system can be built by combining simpler parts. For this purpose is fundamental the notion of modularity. Probability theory provides the glue whereby the parts are combined, ensuring that the system is consistent, and providing ways to interface models to data. A considerable advantage is offered by the graph-theoretic side of graphical models, which provides both an intuitively appealing interface by which humans can model highly interacting sets of variables and a data structure that lends itself naturally to the design of efficient general-purpose algorithms.

Probabilistic graphical models are graphs in which nodes represent random variables, and the arcs represent conditional independence assumptions. Hence, they provide a compact representation of joint probability distributions. Bayesian Networks, also defined as Belief Networks, are direct acyclic graphs. (Friedman, Geiger and Goldszmidt, 1997).

A graph can intuitively be considered as a set of linked points without any metrical description of distance among points (Karin R. SAOUB, 2021). The number of fields in which it is exploited proves the usefulness of graph theory. Indeed, there are applications in recommender systems (Valdiviezo-Diaz *et al.*, 2019), cryptography, and web documents clustering (Majeed and Rauf, 2020).

A graph is defined as a sets couple as in (4.5), where N is the set of graph nodes and E is the set of graph edges. Figure 4.4 provides an example of a generic graph where the set of nodes is $N = \{ 1,2,3,4,5 \}$ and the set of edges is $E = \{a, b, c, d, e, f, g \}$.

$$G = (N, E) \tag{4.5}$$

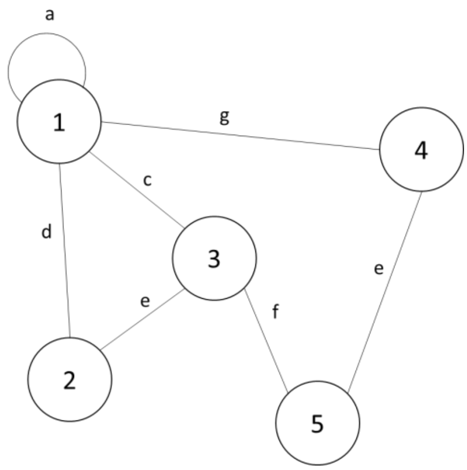


Figure 4.4 *Example of graph*

The following definitions are necessary to clarify the meaning of a direct acyclic graph.

Definition. A *walk* of a graph $G = (N, E)$ is a finite edge sequence

$$e_1 e_2 \cdots e_t \qquad e_i \in E \quad \forall i = 1, \dots, t \tag{4.6}$$

The edge $e_i = (m_i m_{i+1})$ can be identified through the nodes $m_i, m_{i+1} \in N$ that are linked by the edge $e_i \in E$. The *walk* $e_1 e_2 \cdots e_t$ can be denoted with the following notation

$$m_1 \to m_2 \to \cdots \to m_{t+1} \qquad m_j \in N \quad j = 1, \dots, t+1 \tag{4.7}$$

Definition. A *trail* of the graph $G = (N, E)$ is a *walk* in which all edges are different

$$e_i \neq e_j \qquad i \neq j \tag{4.8}$$

Definition. A *path* of the graph $G = (N, E)$ is a route in which all nodes are crossed only once, except at most the first and last node. Indeed, in the path $m_1 \to m_2 \to \cdots \to m_{t+1}$, at most the equal nodes are m_1 and m_{t+1}.

Definition. A *loop* of the graph $G = (N, E)$ is an edge that links the same node

$$(m, m) \in E \quad m \in N \tag{4.9}$$

An example of a *loop* is provided by the edge a in Figure 4.5.

Definition. A *cycle* of the graph $G = (N, E)$ is a *path* $m_1 \to m_2 \to \cdots \to m_{t+1}$ where $m_1 = m_{t+1}$.

Definition. The graph $G = (N, E)$ is an *oriented* (*directed*) graph if the edges (m_1, m_2) and (m_2, m_1) are different elements $\forall m_1, m_2 \in N \quad m_1 \neq m_2$.

Definition. The *oriented* graph $G = (N, E)$ is *acyclic* if there are no cycles.

These provided definitions allow conceptualizing the graph of a Bayesian Network. An example of an oriented acyclic graph is provided by Figure 4.5.

Bayesian Networks, also called Belief Networks, are a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graph (DAG), which take into account the directionality of the arcs.

According to Kevin Murphy, graphical directed models are more prevalent in artificial intelligence and statistics due to several advantages (Murphy, 2002). Most importantly, an arc from A to C indicates that A causes C. This information can be used to construct the graph structure. In addition, directed models can encode deterministic relationships and are easier to learn. In addition to the graph structure, it is necessary to specify the model's parameters. A directed model must specify the Conditional Probability Distribution (CPD) at each node. If the variables are discrete, this can be represented as a Conditional Probability Table (CPT), which lists the probability that the child node takes on each of its different values for each combination of parents' values.

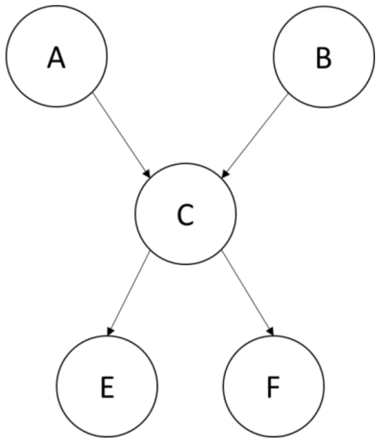


Figure 4.5 *Example of an oriented acyclic graph*

Despite the name, Bayesian networks do not necessarily imply a commitment to Bayesian statistics. Indeed, it is common to use frequentist methods to estimate the parameters of the CPDs. However, they are so-called because they use Bayes' rule for probabilistic inference. Nevertheless, BNs are useful for hierarchical Bayesian models, which form the foundation of applied Bayesian statistics (Lunn *et al.*, 2009). In such a model, the parameters are considered as any other random variable and become nodes in the graph. In order to clarify the theory that allows the construction of Bayesian Networks, some theoretical aspects of probability theory (Kass and Raftery, 1995; Durrett, 2019) are discussed below.

Let be the sample space Ω and the σ-algebra F defined on the sample space; the probability is defined as a function P that has as domain the σ-algebra and as codomain the set of real numbers (4.10).

$$P: A \in F \mapsto P(A) \in \mathcal{R} \tag{4.10}$$

The properties that characterize the probability are:

1. $P(A) \geq 0 \ \ \forall A \in F$;
2. $P(\Omega) = 1$;
3. $P(\cup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} P(A_n)$ for all incompatible events sequence $\{A_n \in F: n \in \mathcal{N}, A_i \cap A_j = \emptyset \ \forall i \neq j\}$.

The tuple (Ω, F, P) is defined probability space. A probability space can be defined as conditional probability as follows.

Definition. In the probability space (Ω, F, P), the probability of the event A conditioned by the event B that has nonzero probability $(P(B) \neq 0)$ is defined as the ratio between the probability of the event $A \cap B$ and the event B.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{4.11}$$

The conditional probability satisfies the properties 1, 2, and 3 related to the definition of probability. This definition is fundamental in order to introduce the Bayes' Theorem that is discussed below

Bayes' Theorem. Let the set of events $\{B_n \in F \ : n = 1, \dots, k\}$ that satisfy the properties:

1. $P(B_n) > 0 \quad n = 1, \dots, k$;
2. $B_i \cap B_j = \emptyset \qquad \forall i, j = 1, \dots, k \quad i \neq j$.

Let the event $A \ \in F$ that satisfies the properties:

1. $P(A) > 0$;
2. $A \subseteq \cup_{n=1}^{k} B_n$;

Then, it is possible to obtain the formula (4.12).

$$P(B_n|A) = \frac{P(B_n)P(A|B_n)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)} \quad \forall n = 1, \dots, k \tag{4.12}$$

The Bayes' Theorem allows taking into account cause-effect connection. In fact, the events $\{B_n \in F \ : n = 1, \dots, k\}$ can be considered the causes of the event $A \ \in F$. The probability $P(B_n)$ that the event B_n happens independently of the occurrence of the event A, is defined '*a priori*' probability and the probability $P(B_n|A)$ that the occurrence of the event B_n conditioned to the occurrence of the event A is defined '*a posteriori*' probability. Then the Bayes'

Theorem links the '*a priori*' probability with the '*a posteriori*' one and allows to obtain the relevant probabilities through a dataset of data (Tucker *et al.*, 2005).

Then, BNs can be seen both as probabilistic and graphical approaches (Das and Ghosh, 2020). In the graph related to a Bayesian Network, the nodes represent the random variables, and the edges represent the probability dependencies. Then, to build the graph related to a Bayesian Network, knowledge of the probability dependencies is needed. The process that allows understanding how to obtain the edges set is defined as structural learning and is composed of two main strategies (Scutari, 2010):

- Constraint-based algorithms. The graph structure is learned by the analysis of the probabilistic relations related to the assumption that each node $m_i \in N$ depend only on its parents in the graph (Markov property) that are the nodes that are linked to m_i. These algorithms are based on Inductive Causation (Inoue, 2020; Okamura *et al.*, 2020), aiming to create a graph skeleton and then correct the skeleton by eliminating cycles.

- Score-based algorithms. These algorithms assign a score to each possible Belief Network built on the nodes considered. The most common algorithm exploited in this field is heuristic (k2, hill-climbing, tabu search) (Cooper and Herskovits, 1992; Scutari, 2010).

The provided dataset allows obtaining the joint probability distribution through the theorem of joint probability. This theorem states that the events $A_1, \ldots, A_n \in F$ such that $P(A_1 \cap \ldots \cap A_n) > 0$, the probability of the intersection events can be obtained through formula (4.13).

$$P(A_1 \cap \ldots \cap A_n) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cap \ldots \cap A_{n-1}) \quad (4.13)$$

The probability $P(A_1 \cap \ldots \cap A_n)$ can be denoted with the notation $P(A_1, \ldots, A_n)$. Applying the Markov property, the formula (4.14) is valid where π_i denotes the parents of the node A_i, that in the Bayesian Network graph are indicated through the edge directed to the node A_i (Pearl, 1988; Kabir and Papadopoulos, 2019).

$$P(A_1, \ldots, A_n) = \prod_{i=1}^{n} P\,(A_i|\pi_i) \quad (4.14)$$

Considering the example proposed in Figure 4.5, which shows a general Bayesian Network containing five variables, the Bayesian Network defines the distribution over those five random variables, the most common task to solve using Bayesian Networks is probabilistic inference. Instead of enumerating all possibilities of combinations of these five random variables, the BN is defined by probability distributions that are inherent to each node, as detailed below:

- The node A and B, noticed which have no incoming arcs, are independent variables, which have the probabilities:

$$P(A) \tag{4.15}$$
$$P(B) \tag{4.16}$$

- The node C is defined by conditional distribution probability, conditioned on A and B:

$$P(C|A,B) \tag{4.17}$$

- The node D and E probabilities are conditioned on C:

$$P(D|C) \tag{4.18}$$
$$P(E|C) \tag{4.19}$$

The joint probability represented by a Bayesian Network is the product of various possibilities defined over individual nodes where each node's probability is only conditioned on the incoming arcs.

$$P(A,B,C,D,E) = P(A) \cdot P(B) \cdot P(C|A,B) \cdot P(D|C) \cdot P(E|C) \tag{4.20}$$

The definition of this joint distribution by using this methodology has a significant advantage. Considering that for simplification purposes, the five variables as Boolean variables, the complete joint distribution of probabilities are represented by 31 probably values.

$$2^5 - 1 = 31 \tag{4.21}$$

According to the probabilities (4.22), the defined Bayesian Network requires only ten probably values, as represented in Table 4.1.

$$P(A) \quad P(B) \quad P(C|A,B) \quad P(D|C) \quad P(E|C) \tag{4.22}$$

Table 4.1 Probability values representation

Probability	$\boldsymbol{P(A)}$	$\boldsymbol{P(B)}$	$\boldsymbol{P(C\|A,B)}$	$\boldsymbol{P(D\|C)}$	$\boldsymbol{P(E\|C)}$	$\boldsymbol{P(A,B,C,D,E)}$
Probability values	1	1	4	2	2	**10**

The compactness of Bayesian Networks leads to significantly better representation at scale, especially for large networks. This feature is a crucial benefit of Bayesian Networks and is the reason they are used for all types of problems so extensively. Recalling the Bayes' Theorem concerning two random variables for which the conditional probability is defined in formula (4.23).

$$P(X|Y) = \frac{P(X,Y)}{P(Y)} = \frac{P(Y|X) \cdot P(X)}{P(Y)} \tag{4.23}$$

Considering a Bayesian Network, in which all Conditional Probability Distributions (CPDs) are defined at each node, it is possible to compute the '*a posteriori*' probability through the Bayes' Theorem. For example, for node E:

$$P(E|C) = \frac{P(C|E) \cdot P(E)}{P(C)} \tag{4.23}$$

4.2.4 Structural Learning Algorithm

The current section presents the structural learning approach developed by Cooper et al. (Cooper and Herskovits, 1992), whose purpose is to determine the most probable belief-network structure based on a known database. Note:

- D the database of m cases;
- Z the set of variables represented by D;
- B_s, B_{s_i}, B_{s_j} several belief networks containing all the Z variables.

The comparison between the various belief networks can be limited to the study of the joint probabilities $P(B_s, D)$ based on formula (4.24).

$$\frac{P(B_{s_i}|D)}{P(B_{s_j}|D)} = \frac{\frac{P(B_{s_i}, D)}{P(D)}}{\frac{P(B_{s_j}, D)}{P(D)}} = \frac{P(B_{s_i}, D)}{P(B_{s_j}, D)} \tag{4.24}$$

In order to calculate the joint probability $P(B_s, D)$ Cooper et al. suppose the following assumptions are valid:

- HYPOTHESIS 1: The variables in database Z are discrete;
- HYPOTHESIS 2: Given a belief network model, the cases are independent;
- HYPOTHESIS 3: There are no cases where the variables have unknown values;
- HYPOTHESIS 4: The density function $f(B_P, B_s)$, which links the vector of conditional probabilities B_P to the belief network structure B_s is uniform.

From hypothesis 1, it is possible to obtain the formula (4.25), with $P(B_s)$ preference bias (Mitchell, 1980; Buntine, 1990), $f(B_P, B_s)$ density function described in assumption 3, $P(D|B_s, B_P)$ probability mass function that can be used instead of the density function $f(D|B_s, B_P)$ since the study is limited to discrete variables. Finally, the integral is a multiple integral since B_P is a vector.

$$\mathrm{P(B_s, D)} = \int_{\mathrm{B_P}} \mathrm{P\,(D|B_s, B_P) f(B_P|B_s) P(B_s) dB_P} \tag{4.25}$$

From hypothesis 2, which assumes case independence among the dataset, the (4.25) becomes formula (4.26), with C_hh-th database case D, $h = 1, \dots, m$.

$$\mathrm{P(B_s, D)} = \int_{\mathrm{B_P}} \left[\prod_{\mathrm{h=1}}^{\mathrm{m}} \mathrm{P\,(C_h|B_s, B_P)} \right] \mathrm{f\,(B_P|B_s) P(B_s) dB_P} \tag{4.26}$$

Exploiting hypothesis 3 and 4, Cooper et al. prove that the probability $P(B_s, D)$ can be calculated through formula (4.27), where:

- n number of discrete variables $X_1, \dots, X_n \in Z$;
- r_i number of possible values assigned to the variable X_i variable among $v_{i1}, \dots, v_{ir_i}$, $i = 1, \dots, n$;
- Defined π_i the list of variables related to X_i e w_{ij} the j-th unique instance of π_i relative to the database D, q_i indicates the number of instances;
- N_{ijk} number of cases in the database D where the variable X_i variable takes value v_{ik} and the list of parents π_i has to instance w_{ij};
- $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$.

$$P(B_s, D) = P(B_s) \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}\,! \tag{4.27}$$

Denoting by Q the set of all belief network structures that contain only the variables of the dataset Z. Returning to the evaluation of the probability of the belief network B_s conditional on the known database D, the formula of such probability will be:

$$P(B_s|D) = \frac{P(B_s, D)}{\sum_{B \in Q} P\,(B, D)} \tag{4.28}$$

4.2.4.1 Exact Method

The number of possible belief network structures grows significantly as the number of variables in the set Z increases. To prevent this increase from leading to high computational cost, Cooper et al. propose the introduction of two additional assumptions:

- HYPOTHESIS 5: It is possible to determine an order for the variables;
- HYPOTHESIS 6: The probability $P(B)$ is equal $\forall B \in Q$.

Denoting with $P(B) = c$ the constant value assumed by the belief network probabilities in Q, it is possible to obtain:

$$\begin{aligned} P(B_s, D) &= P(B_s) \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}\,! \\ &= c \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}\,! \end{aligned} \tag{4.29}$$

Finding the most probable belief network structure involves looking for the element $B_s \in Q$ that maximizes the probability $P(B_s, D)$:

$$\max_{B_s} P\,(B_s, D) = c \prod_{i=1}^{n} \max_{\pi_i} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}\,! \tag{4.30}$$

Finding the most probable network leads back to finding the network for which the instance of the parents is maximized π_i of the variable X_i consistent with the order set on the nodes.

The last presented formula can be generalized by avoiding hypothesis 6 and considering $P(\pi_i^s \to X_i)$ the probability that in the belief-network B_s the variable X_i has as relatives π_i^s. In this way, the preference bias can be calculated as in formula (4.31).

$$P(B_s) = \prod_{i=1}^{n} P\,(\pi_i^s \to X_i) \tag{4.31}$$

Thus the evaluation of the most probable network becomes:

$$\begin{aligned} \max_{B_s} P\,(B_s, D) = \prod_{i=1}^{n} \max_{\pi_i} P\,(\pi_i^s \\ \to X_i) \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}\,! \end{aligned} \tag{4.31}$$

Assuming that the probabilities $P(\pi_i \to X_i)$ e $P(\pi_j \to X_j)$ are marginally independent when $i \neq j$ although this assumption is difficult to justify in practice.

4.2.4.2 Heuristic Method

Cooper et al. also propose a heuristic method for determining the most likely belief network based on the D database. For simplicity of notation, let us denote:

$$g(i, \pi_i) = \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}\,! \qquad (4.32)$$

Let it be, also:

- u the Maximum number of parents that each node can have (fixed '*a priori*');
- $Pred(X_i)$ a function that returns the set of nodes $X_1, \dots, X_{i-1}$ in the given order.

The algorithm on which the heuristic approach is reported below.

K2 Algorithm
Input: database D, set of nodes, order of nodes, u.
Output: a printout of the node's relatives for each node.
for $\boldsymbol{i = 1, \dots, n}$ **do**
$\pi_i = \emptyset$
$P_{old} := g(i, \pi_i)$
$OK - to - Proceed := TRUE$
while $\boldsymbol{(Ok - to - Proceed = TRUE) AND (\lvert\pi_i\rvert \leq u)}$ **do**
Let be $z \in Prec(X_i)$ and be π_i which maximizes $g(i, \pi_i \cup \{z\})$
$P_{new} := g(i, \pi_i \cup \{z\})$
if $\boldsymbol{P_{new} > P_{old}}$ **do**
$P_{old} := P_{new}$
$\pi_i = \pi_i \cup \{z\}$
else
$OK - to - Proceed := FALSE$
write('Node: ', X_i, 'Parents of this node: ', π_i)

4.3 The formal definition of the Multilevel Graph Approach

In this paragraph, the Multilevel Graph Approach will be formalized. In particular, it will be shown how the lists of semantic relations are generated that will correct the algorithm of structural learning exploited to obtain the Bayesian network.

Considering the Ontology general definition in (4.3), we define Ontology O as in formula (4.33), with C set of concepts, A set of attributes, H hierarchical relations, and the sets R^+ and R^- sets denote the relations of axiomatic type (A_x), semantic (R_T) or non-semantic (R) relations, which constitute a dependency or independence link respectively among the concepts of the ontology.

$$O = \{C, A, H, R^+, R^-\} \qquad (4.33)$$

This allows defining the graph G_O associated with the ontology as in formula (4.34), where:

- r is the root of the graph;
- $N_O = C$ set of the nodes of the graph. It coincides with the set of the concepts of Ontology.
- $E_O = \{(n_1, \alpha, n_2) : n_1, n_2 \in N_O, \alpha \in H \cup R^+ \cup R^-\}$ set of arcs. It contains all links between nodes. These links can be hierarchical or specify dependency or independence between concept nodes. From the definition of the set E_O the labels of the arcs are defined through the relations of the given Ontology.

$$G_O =< N_O, E_O, r > \tag{4.34}$$

The Context Dimension Tree (CDT) graphical approach was formally presented in the previous paragraphs. In particular, it was shown that the CDT consists of a graph G_{CDT} defined in formula (4.35).

$$G_{CDT} = < N_{CDT}, E_{CDT}, r > \tag{4.35}$$

In the case of the proposed approach, the Context Dimension Tree provides nodes that are in common with the constructed Ontology. Therefore, both the domain d and the sets of concepts C and attributes A of the Ontology allow to define the sets C_d and A_d as in formulas (4.36) and (4.37).

$$C_d \subseteq C \tag{4.36}$$

$$A_d \subseteq A \tag{4.37}$$

These sets are defined according to properties described in formulas (4.38) and (4.39) with $N_{def} = \{\text{who}, \text{why}, \text{when}, \text{how}, \ldots\}$ set of predefined dimension nodes necessary for the creation of the CDT.

$$N_D \cup N_C = N_{def} \cup C_d \tag{4.38}$$

$$N_A = A_d \tag{4.39}$$

We note, therefore, that the set of nodes of the CDT N_{CDT} is generated through the ontology and the analysis of the domain d:

$$N_{CDT} = N_{def} \cup C_d \cup A_d \tag{4.40}$$

Moreover, the non-default dimension nodes and the concept nodes of the CDT ($N_{CDT} - N_{\text{def}}$) are chosen among the set C elements of the Ontology. In contrast, the parameters are selected among the attributes A of the Ontology.

The set of CDT arcs is partitioned into three sets:

$$E_{CDT} = E_{def} \cup E_C \cup E_A \tag{4.41}$$

where E_{def} is the set of arcs connecting the root to the predefined dimension nodes N_{def} and the arcs connecting the latter to the related concept nodes, E_A is the set of arcs connecting the concept nodes with the relative attributes evaluated in the specific contextual domain, while E_C is the set of the arcs for which there is a connection of hierarchical type between two nodes in the Ontology O.

$$E_C = \{(n_1, \alpha, n_2) \ : n_1, n_2 \in \phi_1^d(C), \alpha \in H\} \tag{4.42}$$

Thus, the Context Dimension Tree defined via the Ontology takes the following form:

$$\begin{aligned} G_{CDT} =& < N_{CDT}, E_{CDT}, r > = \\ & < N_{def} \cup C_d \cup A_d, E_{def} \cup E_C \cup E_A, r > \end{aligned} \tag{4.43}$$

In Ontology (4.33), specific roles have been assigned to concepts, attributes, and hierarchical relations. At this point, instead, a role will be assigned to the relationships of dependence R^+ and independence R^- between ontological concepts. In particular, those relationships will play a central role in designing the lists of semantic relations. To this end, some preliminary definitions are provided, which are helpful in model formalization.

Definition 5.1. Let the generic graph $G \ =< N, E, r >$; it is possible to define *walk* in G as a finite sequence of arcs

$$e_1 e_2 \cdots e_t \qquad e_i \in E \quad \forall i = 1, \dots, t \tag{4.44}$$

since each arc e_i can be denoted by the nodes it connects (m_i, m_{i+1}), a *walk* can be expressed through the notation of the formula (4.45).

$$m_1 \to m_2 \to \cdots \to m_{t+1} \qquad m_j \in N \quad j = 1, \dots, t+1 \tag{4.45}$$

A *trail* is defined as a *walk* in which all arcs are distinct

$$e_i \neq e_j \qquad i \neq j \tag{4.46}$$

A *path* is defined as a *trail* in which there is no repetition of nodes, i.e., all nodes connected by the *path* (except at most the first and last) are distinct. In this way, in the path $m_1 \to m_2 \to \cdots \to m_{t+1}$ at most, it is possible to have $m_1 = m_{t+1}$.

In the case of the graph G_O, the arcs present a label that characterizes the typology of connection among the nodes associated with the Ontology. To such purpose, it is possible to denote a *path* associated with the graph of the Ontology with formula (4.47), in which $m_1, \dots, m_{t+1} \in C$ are graph nodes and $\alpha_1, \dots, \alpha_t \in H \cup R^+ \cup R^-$are arcs labels.

$$m_1 \to_{\alpha_1} m_2 \to_{\alpha_2} \cdots \to_{\alpha_t} m_{t+1} \tag{4.47}$$

Definition 5.2. Let $m_1 \to_{\alpha_1} m_2 \to_{\alpha_2} \cdots \to_{\alpha_t} m_{t+1}$ a *path*. We define in the graph G_O associated to the Ontology a dependency *path*, which is a *path* satisfying the property (4.48).

$$\exists!\, \mathrm{j} \in \{1, \dots, \mathrm{t}\} \ : \alpha_j \in R^+, \quad \alpha_k \in H \ \ \forall k \in \{1, \dots, t\} - \{j\} \tag{4.48}$$

Instead, a *path* of independence is defined as a *path* that satisfies the property (4.49).

$$\exists!\, \mathrm{j} \in \{1, \dots, \mathrm{t}\} : \alpha_j \in R^-, \quad \alpha_k \in H \ \ \forall k \in \{1, \dots, t\} - \{j\} \tag{4.49}$$

Finally, a class *path* is defined as a path that has the property (4.50).

$$\alpha_k \in H \quad \forall k \in \{1, \dots, t\} \tag{4.50}$$

For convenience, it is possible to denote the set of dependency *paths* $P^{(D)}$ and the set of independency *paths* $P^{(I)}$. At this point, it is necessary to define the concepts of dependence and independence between the nodes of the ontology O.

Definition 5.3. Let $m_1, m_2 \in C$ distinct concept nodes of the Ontology, such nodes of define dependent nodes if and only if $\exists\ v\ \in C$ such that m_1 e v are connected by a dependency *path* and m_2 e v are connected through a class *path* or vice versa. Instead, nodes are defined to be independent if and only if $\exists\ v\ \in C$ such that m_1 e v are connected by an independence *path* e m_2 e v are connected by a class *path* or vice versa.

At this point, through the relations just introduced, it is possible to define the set of dependent nodes $R^{(D)}$, as in formula (4.51), and the set of independent nodes $R^{(I)}$, as in formula (4.52). The first one constitutes the list of semantic dependence relations, while the second one constitutes the independent ones.

$$R^{(D)} = \{(m_1, m_2) : m_1, m_2 \text{ dependent nodes}\} \tag{4.51}$$

$$R^{(I)} = \{(m_1, m_2) : m_1, m_2 \text{ independent nodes}\} \tag{4.52}$$

The construction of the Bayesian network is done through the analysis of the ontology O and the Context Dimension Tree. The graph that represents the network is denoted with:

$$G_B =< N_B, E_B > \tag{4.53}$$

where N_B is the set of nodes, E_B is the set of the edges. The nodes of the network are identified through the analysis of the specific context using the CDT, i.e.

$$\exists \psi_d\ N_B = \psi_d(C) \tag{4.54}$$

The generation of the nodes happens through the action of both the Ontology and the Context Dimension Tree. In fact, the latter shares some of the Ontology nodes, as shown above.

The set N_B contains a finite number of nodes, and each can assume a known value within the dataset. This statement is supported by the construction of the set of nodes of the Bayesian network, performed through the synergy between Ontology and Context Dimension Tree.

Considering the hypothesis defined by Cooper et al. (Cooper and Herskovits, 1992), the following assumptions are fulfilled:

- HYPOTHESIS 1: The variables in the database are discrete. The selection of model variables is performed through the data and the cooperation between Ontologies and CDT;
- HYPOTHESIS 2: Given a belief-network model, the cases are independent;
- HYPOTHESIS 3: There are no cases where the variables have unknown values since the nodes are selected through the database and processed by the graph approaches;

An in-depth analysis has to be performed on the possible belief networks through the relations introduced among the nodes. In fact, the goal of the proposed approach is to reduce the number of belief networks to be evaluated and select, among those left available, the most consistent with the dataset exploited through the analysis performed by Ontology and Context Dimension Tree. In particular, the Bayesian network B of which the nodes are already known N_B must be determined through the identification of the set of arcs E_B which will include the arcs obligatorily joining the dependent nodes of the list of dependency relations $R^{(D)}$ and must not present arcs connecting the independent nodes $R^{(I)}$ of the list of independent relations. In particular, by definition, the dependency list $R^{(D)}$ represents the set of oriented arcs joining the dependent nodes, while the independence list $R^{(D)}$ represents the set of arcs connecting the independent nodes.

The target network B will be determined among the belief networks belonging to the set:

$$Q = \{B_i =< N_B, E_i > \ : R^{(D)} \subseteq E_i, R^{(I)} \cap E_i = \emptyset\} \tag{4.55}$$

On this set, the assumptions are assumed to be valid:

- HYPOTHESIS 5: It is possible to determine an order for the variables;
- HYPOTHESIS 6: The probabilities $P(B_i)$ are equal $\forall B_i \in Q$.

The set a priori of conditional probabilities includes the list of semantic dependency relations and excludes the conditional probabilities according to the list of semantic independence relations. The remaining vectors of the possible conditional probabilities B_P respect the semantic relations extracted by the reference domain's Ontology and Context Dimension Tree. The following hypothesis is valid:

- HYPOTHESIS 4: The density function $f(B_P, B_i)$ which links the vector of conditional probabilities B_P to the belief-network structure B_i is uniform.

Denoted with $\overline{\pi_h}$ the instance of the parents of the random variable X_h, this instance will respect both the dependence and independence semantic relations list. In fact, the aleatory variables $X_1, \ldots, X_t$ connected through an arc to the aleatory variable X_h will respect the conditions (4.56) and (4.57).

$$\{(X_1, X_h), \ldots, (X_t, X_h)\} \cap R^{(I)} = \emptyset \tag{4.56}$$

$$\nexists j \notin \{1, \ldots, t\} : \left(X_j, X_h\right) \in R^{(D)} \tag{4.57}$$

In particular, the second relation emphasizes that the relatives $\overline{\pi_\iota}$ of the aleatory variable X_h include all dependency relations set via the list of semantic dependency relations.

Recalling that $P\left(\overline{\pi_h^s} \rightarrow B_i\right)$ represents the probability that in the belief-network B_i variable X_h has as parents $\overline{\pi_\iota^s}$ and that we have previously defined the function

$$g(h, \bar{\pi}_h) = \prod_{j=1}^{q_i} \frac{(r_h - 1)!}{\left(N_{hj} + r_h - 1\right)!} \prod_{k=1}^{r_i} N_{hjk}\,! \tag{5.58}$$

The Bayesian network will be determined as the network that maximizes the belief network probability on data, as shown in the formula (4.59)

$$\max_{B_i \in Q} P\,(B_i, D) = \prod_{h=1}^{n} \max_{\pi_h} P\,(\pi_\iota^s \rightarrow B_i) g(h, \pi_h) \tag{4.59}$$

In the proposed approach, this computation is performed by modifying the heuristic procedure used by Cooper et al. described earlier (Cooper and Herskovits, 1992). In particular, the following pseudocode is executed in which a maximum number of parents u is fixed, and an order on the nodes of the set N_B is selected. The set of possible parents nodes $\overline{\pi_\iota}$ is constructed exploiting the list of dependent semantic relations.

Algorithm
Input: N_B, $R^{(D)}$, $R^{(I)}$, node order, u
Output: a printout of the node's parents for each node.
for $\boldsymbol{i = 1, \ldots, n}$ **do**
$\quad \overline{\pi_\iota} = \{(X, X_i) \in R^{(D)}\} \subseteq R^{(D)}$
$\quad P_{old} := g(h, \bar{\bar{\pi}}_h)$
$\quad OK - to - Proceed := TRUE$
$\quad$ **while** $\boldsymbol{(Ok - to - Proceed = TRUE) AND (\lvert\pi_h\rvert \leq u)}$ **do**
$\qquad$ Let be $z \in Prec(X_h)$ and be $\bar{\bar{\pi}}_h$ which maximizes $g(h, \bar{\bar{\pi}}_h \cup \{z\})$
$\qquad P_{new} := g(h, \pi_h \cup \{z\})$
$\qquad$ **if** $\boldsymbol{P_{new} > P_{old}\ AND\ (z, X_i) \notin R^{(I)}}$ **do**
$\qquad\quad P_{old} := P_{new}$
$\qquad\quad \bar{\bar{\pi}}_h = \bar{\bar{\pi}}_h \cup \{z\}$
$\qquad$ **else**

12	$OK - to - Proceed \coloneqq FALSE$
13	write ('Node: ', X_h, 'Parents of this node: ', $\bar{\pi}_h$)

Once the belief network has been determined based on the lists of dependence and independence semantic relations, it will be possible to perform the training phase of the network to determine its weights. In particular, through the dataset, it is possible to calculate the conditional probabilities associated with the dependence relations that identify the Bayesian network built through the synergy work of the Ontology G_O and Context Dimension Tree G_{CDT}.

Chapter 5
Experimental Results

This section reports the case studies developed using the proposed Multilevel Graph Approach. The first subsection shows a case study used for model validation; the case study is related to the Smart City of London, chosen for reasons of data availability. The second subsection provides other case studies related to the proposed models, which involve different application domains. A case study concerns the application of the model in cybersecurity in the automotive domain. Other case studies are presented, which show the ability of the System to interact and support users. In detail, a case study is presented that involves the System as educational paths recommender in the cultural heritage field, or in supporting expert users and consumers in the Smart Agriculture field, or supporting users in the management of services, and resources, within a Smart Home.

5.1 Model Validation: London Smart City

In this subsection, it is reported the case study that validates the proposed approach. For data availability, this pilot case study is focused on London Smart City, in which, due to open data availability policies, it was possible to obtain a large amount of information. In this scenario, the MuG Approach is exploited to predict different events such as traffic accidents, rainfall, and bike-sharing service availability based on the collected data. To validate the proposed model, several steps were necessary. In particular, this section aims to compare the performance of the proposed approach, which involves the prediction of phenomena through Bayesian Networks built with the support of semantic information derived from the context. The results will be compared with traditional approaches such as the construction of Bayesian Networks with the support of experts in the field or through structural learning algorithms. In detail, the proposed methodology will be validated in an attempt to predict three phenomena in the Smart City scenario: traffic accidents, rainfall, and bike-sharing service availability. For this purpose, the following sections will present the available dataset, the phases of construction of the graphs for the description of the context, the formalization of the experimental phases, and the numerical results.

5.1.1 Dataset

The chosen study area is the city of London; in particular, the complete dataset of information collected refers to the neighborhoods of Islington, Westminster, Hammersmith and Fulham, Lambeth, and Tower Hamlets. Only for these boroughs was it possible to collect enough data to provide an example that allows us to show the ability of the System to predict different events. Depending on the phenomena to be predicted and according to data availability have been considered all areas or only some of them, which correspond to London central areas. The data used come from different sensors, providing information with different formats and time intervals. The data were aggregated at different time intervals for the analysis depending on the event to predict. The full dataset used contains 12 variables and several records aggregated from 2017 to 2019. The one-year observation period exceeds 35000 instances. Details of the data available in the dataset are shown in Table 5.1.

Table 5.1 *London Smart City Case of Study: Dataset details*

Data	Details
DayDate	This data refers to the date of the day with the following format: yyyy-mm-dd, HH:MM:SS
Evapotranspiration	This data refers to the evapotranspiration, expressed in mm.
Pressure	This data refers to the pressure expressed in Bar (hPa).
Radiation	This data refers to the solar radiation expressed in W/m^2.
Rainfall	This data refers to the amount of rain accumulated during a day, expressed in mm.
RainRate	This data refers to the instantaneous measure of precipitation, expressed in mm/h
StormRain	This data refers to the amount of rain accumulated during a storm, expressed in mm.

Temperature	This data refers to the outside temperature, expressed in Celsius degree.
UVIndex	This data refers to the Ultra Violet Index.
WindDirection	This data refers to the direction of the wind, expressed in degrees.
WindSpeed	This data refers to the instantaneous speed of the wind, expressed in m/s.
StartTrips_Islington	This data refers to the number of bicycles rented in the borough.
EndTrips_Islington	This data refers to the number of bicycles returned in the borough.
Accidents.Slight	This data refers to low severity accidents that occurred in the borough.
Accidents.Serius	This data refers to the high severity accidents that occurred in the borough.

5.1.2 Context description

As mentioned above, the Multilevel Graph Approach exploits the semantic relationships that arise from Context-Awareness. Considering the paradigm of Situation Awareness (SA) defined by Endsley, through this phase, the SA 2, namely, the Comprehension phase, is reached through context description, performed by graph structures such as Ontologies and CDT. These graph structures were designed with experts domain supported by scientific literature-defined approaches. In order to pursue the objective of evaluating the proposed approach, it was necessary to deepen the study of the two graph approaches by focusing on scientific literature desk research to design Ontologies and CDTs, which meet specific needs of context representation. This representation will allow the System to integrate semantic relations functional to build constraints between the data. Due to this, it will be possible to design more reliable Bayesian Networks.

The case study for the validation of the model is conducted on data acquired in the city of London. These data come from different sensors placed in the environment and mainly concern weather conditions and transport. Considering the available data, the pilot case study will focus on predicting three phenomena: traffic accidents, rainfall, and availability of bike-sharing services. To this end, to predict phenomena that occur in complex scenarios, such as Smart Cities, involving many aspects and embracing different areas of knowledge, it is reasonable to think of a support service such as Ontologies. A tool such as an ontology, which allows us to represent knowledge in a formal, universally understandable, and reusable way, could be fundamental for several aspects such as assessing, predicting, managing risks and meteorological phenomena. In fact, many scientific studies contribute to the design of ontologies concerning the modeling of environments and services in Smart Cities or concerning the meteorological or hydrological sector in general (e.g. (Chen, Finin and Joshi, 2003b; Scholten *et al.*, 2007; Islam and Piasecki, 2008)).

On modeling Smart Cities and their services, much progress has been made in scientific research of recent years. In particular, in (Troncy *et al.*, 2017), a framework to build a knowledge base includes descriptions of a city's events and activities, places and monuments, transportation facilities, and social activities. About smart transportation, an ontology capable of assisting smart and autonomous vehicles in decision-making is presented in (Juric and Madland, 2020). Moreover, the issue of formal knowledge management concerning IoT and interoperability between different sensors for the development of smart urban mobility services in the Smart Cities context is addressed in several scientific contributions. In particular, in (Fernandez *et al.*, 2016), an ontology-based system to improve safety on roads by providing traffic information is reported, and in (Sotres *et al.*, 2019), a system capable of supporting parking directions and mobility suggestions to users is reported. Furthermore, in (D'Aniello *et al.*, 2020), the Smart City Service System, a knowledge-based system to support decision-making processes in a city, is presented. The considered knowledge base extends Km4City (Bellini *et al.*, 2014) with EventOntology[1] to represent events with more details and technology-oriented ontologies to support the description of services and the representation of sensors and data from them. The integrated information is analyzed to provide situational awareness of the city using inference and classification processes. In (Mohammad *et al.*, 2015), an ontology for road hazards is presented. The presented Ontology provides the representation of concepts and elements such as pedestrians, vehicles, and other elements. This ontology also links weather conditions to hazards in road environments (Mohammad *et al.*, 2015).

Instead, on meteorological phenomena, particularly floods, Raskin & Pan proposed and developed a modular system called SWEET (Semantic Web for Earth and Environmental Terminology), a higher level ontology containing a hierarchy related to weather hazards (Raskin and Pan, 2005; Scholten *et al.*, 2007). Furthermore, according to Scheuler et al., referring to the work mentioned above, it is possible to put this information into an accessible and reusable knowledge base in multi-criteria risk assessment, defining it as crucial for weather hazards management to rely on a correct and comprehensive flood risk assessment (Scheuer, Haase and Meyer, 2013). In the field of weather hazards prediction, according to Agresta et al., the ability to use and share knowledge is the main reason why ontologies are well suited to support the (near) real-time prediction phases of devastating flood events

[1] http://motools.sourceforge.net/event/event.html

and flood warning and emergency management (Agresta *et al.*, 2014). In their work, FloodOntology is developed, an ontology to predict floods related to rainfall events based on continuous measurements of water parameters from sensors and simulation models. In addition, three main domains are addressed: hydrologic, hydraulic, and sensor networks. It can be assumed that weather hazards predictions are centered on simulations of the hydrologic/hydraulic processes involved, based on continuous monitoring from sensor networks. Compton et al. address an interesting work in this field. They propose a Semantic Sensor Network (SSN), which represents an online ontology that can describe sensors and observations in terms of capabilities, measurement processes, observations, and deployments (Compton *et al.*, 2012).

Based on the ontologies, taxonomies, and formal knowledge bases presented in scientific literature, in particular (de Wrachien *et al.*, 2012; Agresta *et al.*, 2014; Mohammad *et al.*, 2015; D'Aniello *et al.*, 2020), it was possible to realize a Task Ontology (Ikeda *et al.*, 1998; Abrahao and Hirakawa, 2017) valid for the presented approach, a small part of which is represented in Figure 5.1.

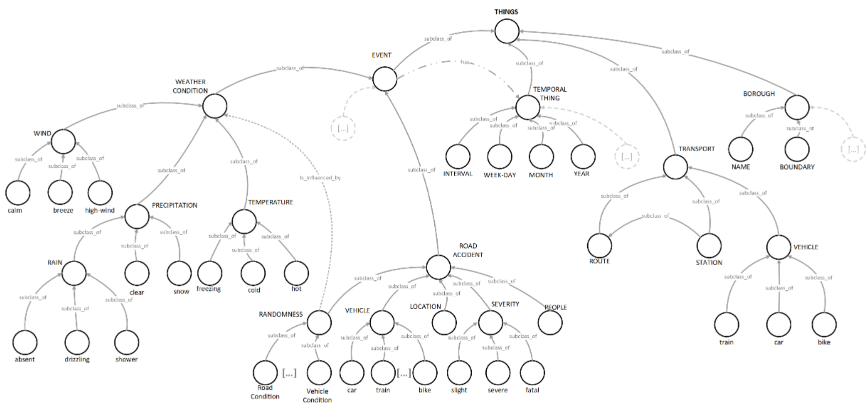


Figure 5.1 *Task Ontology sample on Smart City*

Having designed the task ontology related to the considered scenario, the application of the proposed approach needs the reference context description performed by a Context Dimension Tree. In addition to being useful for our approach, the CDT is a simple, straightforward, and fundamental element for context representation (Annunziata *et al.*, 2016). Various applications of the CDT exist in the literature (Panigati *et al.*, 2012; Colace *et al.*, 2017). However, there does not seem to be a specific version suitable for describing the smart city context in which meteorological or transportation information is present. Therefore, to test the proposed methodology, it was decided to design a specific version, considering that the CDT is a fundamental tool for the proposed methodology applying. According to the proposed methodology, all possible context combinations are researched based on the CDT after the target decision. Based on the ontological correspondences, these context combinations provide the semantic relationships between the data. Those relations are exploited to design a Bayesian prediction network.

Focusing on the importance of this tool, the CDT was designed; a sample is shown in Figure 5.2. As shown in the figure, five upper dimensions (*How, Where, Who, What, Why, When)* help context representation. For example, in the top-dimension *Where*, in addition to the place, representing the geographic location of the event, there is also the type of area that, in this case, may play an essential role in the context. The top-dimension *Who* deals with the factors involved: people or goods. Instead, the top-dimension *Why* describes the events process as something that can be triggered, such as public transportation. The top-dimension *When* instead represents the time interval of the event, this plays a crucial role in predicting the possible damage. Clearly, in the top-dimension *What* are present the targets classified, possibly, according to various danger levels.

The CDT thus designed, even if tailored to the specific proposed methodology, provides a contextual view of the events, representing all possible contexts of the identified Smart City scenario, one of the critical points for applying the proposed method.

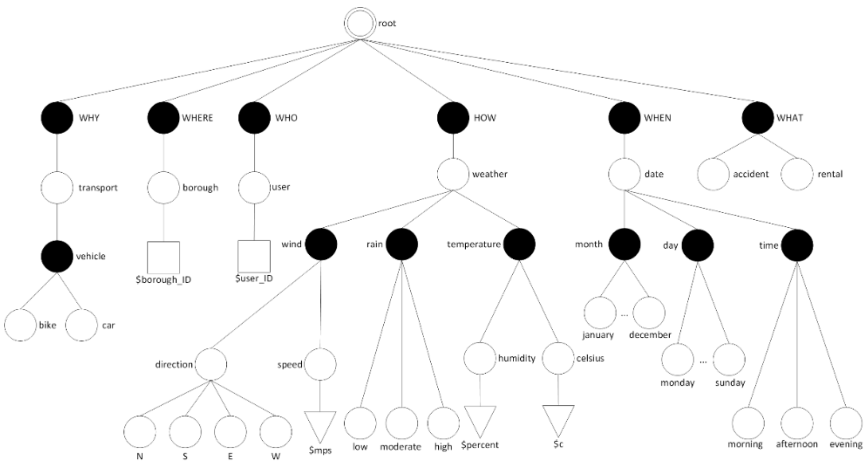


Figure 5.2 *Context Dimension Tree sample on Smart City*

In conclusion, it is possible to argue that it was possible to build the Ontology and CDT view through the support of scientific literature and expert users; those representations are crucial in supporting and applying the proposed innovative approach.

5.1.3 Metrics and Experimental Approach

Appropriate evaluation metrics were selected to validate the proposed approach. In addition, three steps to appropriately compare the proposed methodology were identified.

The metrics chosen are Confusion Matrix, Accuracy, Precision, Recall, and F_1-Score. A Confusion Matrix (Figure 5.3), also known as an error matrix, is a specific table allowing the visualization of the performances of an algorithm indicating how a system is able to classify events.

		Actual Value (as confirmed by experiment)	
		positives	negatives
Predicted Value (predicted by the test)	positives	TP True Positive	FP False Positive
	negatives	FN False Negative	TN True Negative

Figure 5.3 *Example of Confusion Matrix*

Accuracy (A) is the proportion of correct predictions on the total number of cases examined; it represents a measure of the ability of the System to make correct predictions.

$$A = \frac{TP + TN}{TP + FP + FP + FN} \tag{5.1}$$

Precision (P) is the fraction of relevant instances among the retrieved instances, representing the System's ability to predict an event.

$$P = \frac{TP}{TP + FP} \tag{5.2}$$

Recall (R) is the fraction of relevant instances recovered out of the total number of relevant instances, which provides a System's reliability measurement.

$$R = \frac{TP}{TP + FN} \tag{5.3}$$

F_1-Score defines the accuracy of a test, merging Precision and Recall in a single variable, which represents the harmonic mean of Precision and Recall.

$$F_1\text{-Score} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \tag{5.4}$$

Taking these metrics as a reference, three analysis phases were identified to validate the proposed approach appropriately. In order to perform the three analysis phases, it was necessary to divide the available dataset into the Training Dataset (TrD) and the Test Dataset (TeD). Due to this, it was possible to proceed through the three phases identified.

In particular, in the first phase (figure 5.4), the Bayesian network structure is designed through domain experts. The TrD allows learning the Bayesian network weights, which is subsequently validated through the TeD.

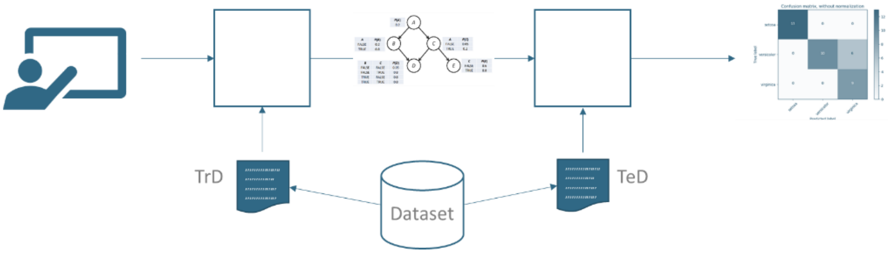


Figure 5.4 *Experimental Approach Step 1*

In the second step (Figure 5.5), the Bayesian network structure is built through a structural learning algorithm. First, the Bayes network is realized through the TrD and the structural learning algorithm. Then the learned Bayesian network is evaluated using the TeD, obtaining the confusion matrix.

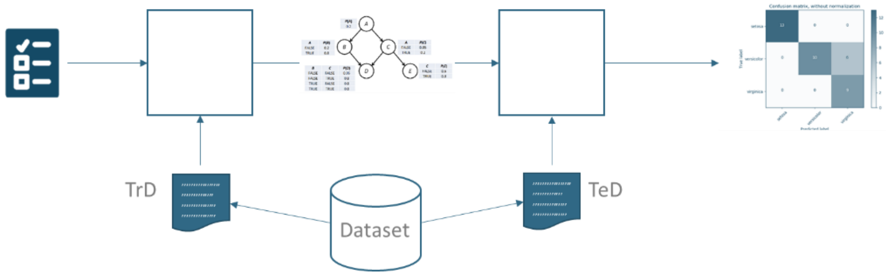


Figure 5.5 *Experimental Approach Step 2*

The proposed Multilevel Graph Approach is applied in the third step (Figure 5.6). First, the Ontological View and the CDT are combined to obtain a list of semantic relations. The list of relations is used with the structural learning algorithm and the training dataset to generate the Bayes network. Finally, the network is tested with the test set obtaining the confusion matrix.

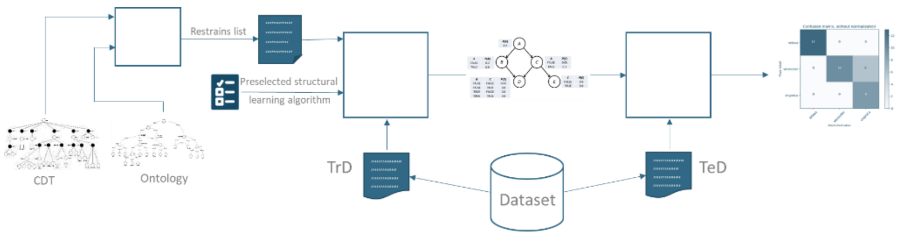


Figure 5.6 *Experimental Approach Step 3.*

5.1.4 Numerical results

The experimental results regarding the first case study were performed to predict three phenomena within the city of London. The phenomena to be predicted were road accidents, rainfall, and the Bicycle Sharing service availability.

In all three cases, the proposed method has been compared using about 75% of the dataset available as training set and about 25% as test set and through the three phases exposed in the paragraph on the experimental approach. In particular, the predictions related to a Bayesian network structure built by a group of experts were evaluated. Using the K2 algorithm (Cooper and Herskovits, 1992), the Bayesian network built by machine learning was evaluated, and then analysis was conducted using the proposed methodology was then analyzed.

However, each application preserved its peculiarities. In the case of traffic accidents, an attempt was made to predict the probability of accident occurrence by classifying it as Low, Medium, and High risk. In this case, the analysis was conducted by observing the data collected for all areas available by collecting aggregate instances every three hours. In particular, among the 7920 events analyzed, 5960 (about 75%) were used for the training set and 1960 (about 25%) for the test set. The results obtained using the test set, in terms of Precision, Recall and F_1-Score are reported in Table 5.2.

In the case of rainfall prediction, the variable under consideration, which measures the accumulation of rain during a rainfall event, defines the event itself. For this reason, the parameter was summarized in a Boolean variable (Yes/No). Moreover, in this case, the prediction was improved using data correlation techniques, which allowed us to find relationships between data with different observation times. This process increased the System's reliability concerning a difficult-to-predict event, such as rainfall forecasting, by exploiting cross-correlated data as precursors or indicators.

Table 5.2 *London Smart City Case of Study: System performance in Accident prediction*

		Low	Medium	High
Step 1	**Precision**	79,02%	87,77%	77,44%
	Recall	91,52%	74,37%	71,43%
	F_1-Score	84,81%	80,52%	74,31%
Step 2	**Precision**	77,25%	54,32%	16,23%
	Recall	61,38%	53,74%	63,64%
	F_1-Score	68,41%	54,03%	25,86%
Step 3	**Precision**	82,78%	89,40%	83,51%
	Recall	93,87%	78,92%	75,94%
	F_1-Score	87,98%	83,83%	79,55%

The results were obtained through the dataset from all the central zones aggregated at hourly intervals. In particular, among the 23760 events analyzed, 17850 (about 75%) were used for the training set and 5910 (about 25%) for the test set. The obtained results, in terms of prediction, recall, and F_1-Score are reported in Table 5.3

Table 5.3 *London Smart City Case of Study: System performance in Rainfall prediction*

		Yes	No
Step 1	**Precision**	59,96%	88,77%
	Recall	73,47%	81,05%
	F_1-Score	66,03%	84,73%
Step 2	**Precision**	45,90%	63,17%
	Recall	42,06%	66,71%
	F_1-Score	43,90%	64,89%
Step 3	**Precision**	85,17%	91,06%
	Recall	83,84%	91,85%
	F_1-Score	84,50%	91,45%

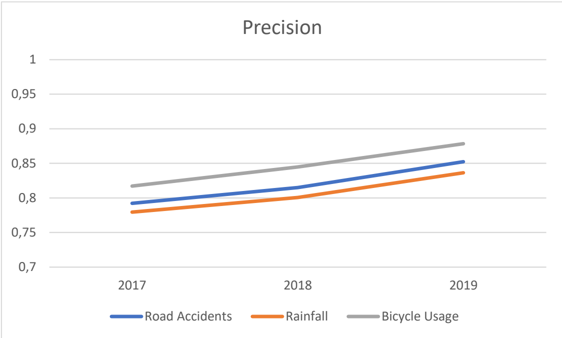
In the third case, the System has been tested on the forecast of the use of Bike Sharing service in the city of London. In this case, two variables representing the beginning (Bicycle Rental) and the end (Bicycle Return) of the service's use were considered to provide users with the probability of availability or unavailability of the service aggregated into three ranges: Low, Medium, and High. In this case, the results were obtained using data from the Westminster area only at 30-minute intervals. In particular, among the 47520 events analyzed, 35747 (about 75%) were used for the training set and 11773 (about 25%) for the test set. In terms of precision, recall, and F_1-Score, the results are reported in table 5.4.

As can be seen from the experimental results shown in Tables 5.2, 5.3, and 5.4, the System was able to predict the variables with varying degrees of accuracy. In particular, the first and the third step bring higher accuracy to the second one, representing the limits of automatic learning algorithms. The step that achieved better results was based on the proposed approach in all three cases. In particular, in all cases, due to graph structures used for context modeling, the proposed methodology was able to correct any associations learned by the automatic structural learning algorithms that work exclusively on data and possibly highlight new hidden semantic relationships.

Table 5.4 *London Smart City Case of Study: System performance in Bicycle usage prediction*

		Bicycle Rental			**Bicycle Return**		
		Low	**Medium**	**High**	**Low**	**Medium**	**High**
Step 1	**Precision**	87,07%	90,08%	83,64%	87,82%	90,52%	80,66%
	Recall	93,84%	85,49%	75,87%	93,41%	86,16%	76,02%
	F_1-Score	90,33%	87,73%	79,57%	90,53%	88,29%	78,27%
Step 2	**Precision**	82,19%	71,43%	52,94%	82,38%	74,05%	53,41%
	Recall	77,64%	72,09%	61,76%	81,55%	69,54%	63,42%
	F_1-Score	79,85%	71,76%	57,01%	81,96%	71,72%	57,99%
Step 3	**Precision**	92,33%	92,78%	86,53%	93,66%	92,70%	86,20%
	Recall	95,69%	91,34%	81,18%	95,58%	92,34%	81,78%
	F_1-Score	93,98%	92,05%	83,77%	94,61%	92,52%	83,93%

From the results obtained, it is possible to observe that, thanks to the innovative approach used, the number of instances classified correctly has increased, and the instances classified incorrectly have decreased. Moreover, it is possible to infer that the application of the MuG Approach increases the reliability of the Bayesian Network. From the tables, it is possible to see that the Precision, Recall, and F_1-Score results related to the proposed approach (step 3) in all cases are better than the others. These results were obtained with a with a few seconds of training time, which is perfectly comparable to the other steps and could represent an advantage in the dynamic training of the system. This encouraging result confirms the MuG approach's benefit in constructing more reliable Bayesian networks. In addition, it can be easily inferred that the System's capacity can increase over time as the available data grows, according to what is shown in Figure 5.7.



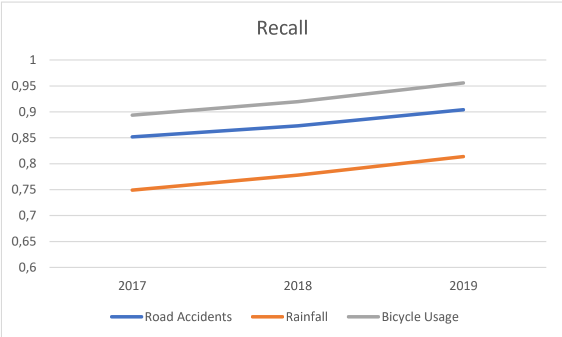


Figure 5.7 *System improvements related to data*

5.2 Further Case Studies

This paragraph reports further case studies related to the Multilevel Graph Approach. In particular, several case studies involving different application domains will be analyzed. A case study will be presented on applying the model in Automotive Cybersecurity, exploiting the System as an Intrusion Detection System to detect malicious attacks. Another case study will show the System interacting with users in Cultural Heritage. In this case, the System is used as a Recommender System to suggest cultural-educational paths within archaeological parks. Other case studies that present the use of the System to support different users concern Smart Agriculture and Smart Home. In particular, in the case of Smart Agriculture, the System can support both expert users, such as farmers, in the management of crops and support simple users such as consumers to enhance agricultural products. In the case study related to the Smart Home environment, the System will be tested to learn and suggest actions for resources management in a proactive and tailored way on the user's behavior.

5.2.1 Cybersecurity in Automotive

This case study focuses on the problem of cybersecurity in IoT (Zarpelão *et al.*, 2017; Colace *et al.*, 2021), with particular attention to Cybersecurity in Automotive (Castiglione *et al.*, 2020), using the Multilevel Graph Approach as an Intrusion Detection System for connected vehicles (Lombardi, Pascale and Santaniello, 2021). Technological development of recent years and the Internet of Things diffusion in our daily lives have allowed the realization of many applications. However, it has also exposed new risks for information security. Information Security refers to protecting assets or resources, ensuring integrity, availability, and confidentiality (Nzabahimana, 2018). Assets can represent logical elements such as web platforms, databases, etc., or physical elements such as computer systems, devices, or people. Threats can originate from cyber attacks, catastrophic natural events, or human errors. Methodologies to protect assets can be implemented through technology, policies, or by increasing the so-called "user awareness" to educate users about information security (Whitman and Mattord, 2011).

In this scenario, connected vehicles are of considerable interest. In fact, in recent years, modern vehicles have been integrated by many smart devices, which communicate with each other through the network and the internet and can offer innovative features and services. Having the possibility to exploit vulnerabilities of devices in cars, it is possible to inflict severe damage to humans. Therefore, it is essential to develop strategies to avoid the described scenario. The main communication channel of vehicles is represented by the CAN-Bus. This system allows communication through different electronic control units (ECUs) (Lin and Sangiovanni-Vincentelli, 2012; Fowler *et al.*, 2017). Due to the diffusion of new technology, modern vehicles are equipped with many functionalities, leading to autonomous or assisted driving. Such functionality often needs to communicate via APIs, WiFi, or cloud systems. However, these communication channels expose the vehicle to several vulnerabilities. Internet access exposes modern vehicles to more significant opportunities for cyberattacks, increasing the vulnerability of the entire System.

The reasons behind a cybersecurity attack could be many and varied. An attacker could, for example, smuggle in personal information, monitor a person's movements, and, in the worst case, take remote control of the vehicle. In detail, the internal network of modern vehicles, called CAN-Bus, consists of about 70 nodes. Each node corresponds to an ECU responsible for controlling a specific vehicle component such as the windows, the ventilation system, or the engine (Koscher *et al.*, 2010; Hoppe, Kiltz and Dittmann, 2011). The ECUs communicate in broadcast mode through an unencrypted communication channel called CAN-bus. If an attacker were able to access it, the security of the entire vehicle would be compromised (Onishi, 2012; Reilly *et al.*, 2015). The CAN-bus security problem has been known for a long time and has already been addressed in various ways, all of them quite effective but not efficient in terms of performance. Many proposals in the literature aim to redesign the CAN standard, making a sort of evolution, both from the hardware and the software side. This kind of solution, especially on the hardware side, does not guarantee the safety of vehicles already on the market. It would be necessary to update all the components involved in communications inside the vehicle, from the ECU to the cables (Li, Liu and Luo, 2008). Other approaches aim to create intrusion detection systems by potential attackers. In this case, however, the computational power required exceeds the microcontroller capacity of today's vehicles. The low computational capacity leads to not even considering MAC-based solutions, depending on the need to keep the information exchanged unchanged (Zalman and Mayer, 2014).

Many research works dealt with using machine learning methodologies as the main topic for detecting cyberattacks on the network. In IoT systems, there has been a focus on identifying what safeguards can be applied. IoT devices are equipped with reduced computational capacity, and it is necessary to consider these limitations finding effective solutions supported by the physical components. In particular, One of the most widely used systems to ensure security over CAN-Bus have been Intrusion Detection Systems (IDS). However, such systems need a high-performance hardware component since using Machine Learning techniques, their efficiency increases as the performance of the hardware components increases.

In many contexts, such as IoT, the devices present low computational capacity, which is crucial for an efficient design (Long, 2018; Xu *et al.*, 2019). As shown in (Buczak and Guven, 2016; Long, 2018), IDS is used for intrusion control in the IoT. Nevertheless, the computational needed are to high compared to the hardware present on ECU microcontrollers. To ensure modern vehicles security many other methodology are used. One approach is performed using randomization tequnique, which cover the CAN-Bus frame identification (Xin *et al.*, 2018). in which randomization of the CAN-Bus frame identification field is generated. In this approach, the set of random identifiers represents a new nodes mapping. However, this random set should be change over time: they can be computed at vehicle startup or modified at well-defined intervals. Other approaches provides security through encryption methodologies. In particular, CAN-Bus security can be increased by using a message authentication code; in this way, masked attacks can be prevented (Azwar *et al.*, 2019). However, this methodology involves finding the tradeoff between bandwidth, real-time usage, and security. As could be intended, one of the main problems related to encryption methodologies and, consequently, security enhancement lies in information transmission delays. This issue can be overcome by assigning different priorities to the encrypted frames, which allow, through an arbitration phase, the sending of priority messages with minor delay (Song, Kim and Kim, 2016).

In this scenario will be interesting to adopt a new methodology to increase the connected vehicle's security. The purpose of this case study is to employ the Multilevel Graph Approach to recognize and report vehicle safety issues. The proposed methodology will be exploited as an advanced Intrusion Detection System capable of contextualizing possible problems and classifying events based on data behavior. In fact, the case study plans to evaluate the proposed methodology, which uses graph approaches to bring semantic value to the System to identify and eventually predict atypical phenomena caused by malicious actors that try to control modern vehicles.

5.2.1.1 Experimental Results

In order to apply the proposed approach, it is necessary to consider that in modern vehicles, all information exchanged is routed over the CAN-Bus. The information concerns both control and diagnostic messages, which are transferred through all the vehicle's electronic control units (ECU). Evidently, all possible attacks, which may come from inside or outside the vehicle, are performed in order to inject potentially dangerous messages through the CAN-Bus.

To apply the proposed methodology, it is necessary to understand the vulnerabilities in the automotive sector, modeling them according to the introduced context-awareness approaches. In particular, the reality was modeled based on several works in the literature describing the automotive industry and attacks through ontologies (Nawir *et al.*, 2017; Klotz, Datta, *et al.*, 2018; Klotz, Troncy, *et al.*, 2018; Syzdykbayev, Hajari and Karimi, 2019). Once the context was identified, it was possible to define a task ontology to identify the System's characteristics and the identification of the parameters to be monitored. Subsequently, a specific Context Dimension Tree was modeled, which analyzes all possible types of context and is able to understand all possible threats and vulnerabilities to which the System is subjected.

Once the models describing the context have been identified, the proposed methodology can be used through the steps described above. First, the possible states of the system were defined. Starting from the EURO-NCAP specification for car safety (van Ratingen, 2017), three possible attacks and one normal status indicated as follows were identified:

- A1) The car walks straight at a constant speed; the malicious user sends a command to turn right or left
- A2) Car drives straight ahead at constant speed; attacker sends a command to brake
- A3) Car is driving straight ahead at constant speed; attacker turns off the cooling System
- R) Condition vehicle regularity.

For simulation purposes, it is assumed that the attacker can have access to the CAN-bus to inject malicious messages, the action can be performed through one of the network access points on modern vehicles (API, Cloud system, Internet services). The attacker, in this way, can force anomalous vehicle behavior. The proposed approach will be tested on the ability to detect intrusions by reading the traffic on the CAN-bus and identifying any irregular behavior.

In order to proceed with the data collection and the use of the proposed methodology, it was necessary to implement a prototype for simulation and data collection. A CARLA-based simulation environment was designed for research purposes (Dosovitskiy *et al.*, 2017). CARLA represents an open-source platform able to perform professional simulation testing for autonomous driving. The platform enables prototype development and vehicle interaction with the surrounding environment and smart objects. The prototype has been designed through a client/server architecture to which different modules and physical components are connected:

- A server that generates the external driving environment;
- A module that converts messages from serial to CAN, to which are connected a steering wheel and pedals for controlling the vehicle;
- A board with an ARM processor, which represents the access point for CAN-bus attacks, simulating a vehicle service with access to the Internet;
- A board with an ARM processor, which represents the intrusion detection system capable of recognizing attacks.

The validation of the system was performed through a simulation dataset. Around 10000 records have been generated were at irregular intervals, the vehicle has been exposed to about 2000 malicious messages. In particular, among the 11800 instances considered, 8890 (about 75%) were used for the training set and 2910 (about 25%) for the test set. The analysis was conducted through the three proposed steps. As can be seen from the experimental results, shown in Table 5.5, the third step, where the proposed methodology is applied, performs better than the others. This case study testifies how the proposed System is able to adapt itself, working as Intrusion Detection System and providing promising results.

This case study aimed to validate the performance of the MuG Approach as an Intrusion Detection System for Cybersecurity in Automotive. The results, even if preliminary, are promising. The System has been able to outperform in comparison to other steps. These results are obtained through approaches that contain the data's semantic value. In fact, the presented model benefits from a probabilistic approach and the relationship between several parameters that allow describing the Context and Situation awareness. The used approach allowed filters able to discard some implausible conditions in advance.

Table 5.5 *Case of Study on Cybersecurity in Automotive: System performance*

		R	A1	A2	A3
Step 1	**Precision**	93,24%	44,63%	51,24%	54,24%
	Recall	93,55%	49,58%	45,42%	53,46%
	F_1-Score	93,39%	46,97%	48,15%	53,85%
Step 2	**Precision**	93,56%	29,72%	37,74%	55,89%
	Recall	91,26%	44,96%	41,15%	39,36%
	F_1-Score	92,40%	35,78%	39,37%	46,19%
Step 3	**Precision**	97,38%	53,51%	87,64%	73,66%
	Recall	97,15%	66,18%	68,66%	74,75%
	F_1-Score	97,26%	59,17%	77,00%	74,20%

5.2.2 E-Learning paths recommender in Archeological Parks

In this case study, the MuG Approach was exploited as a recommender system to suggest learning paths to users visiting archaeological parks (Colace *et al.*, 2020). In the last decades, e-learning has been increasingly enriched with new tools able to improve the educational process. The typical processes of the traditional education world, which are still valid today, are assisted due to the advent of new technologies. Nowadays era is characterized by new intelligent devices capable of exchanging information with each other, contributing to the Internet of Things paradigm. How to take advantage of such technologies to further enhance e-learning? Having the ability to do training activities in a cultural-historical center allows users to connect with historical assets, furthering the training process. However, this is not enough to obtain a good training result. It is necessary to follow a well-structured educational path provided by an expert guide whose objective is not simply to describe the artifacts present as is typically done with tourists but to give a historical-cultural perspective with a formative character. This process is particularly complex. It could be interesting to have a methodology that deals with the automatic design of educational paths in this scenario. This technology could take advantage of smart mobile devices and the large amount of data they produce to build customized learning paths. These training paths could also be developed ad hoc concerning the archaeological sites visited, allowing combining two objectives: promoting learning and enhancing cultural heritage value.

There are many techniques aiming to improve the learning process. As Cicero liked to walk around memorizing by associating physical paths to narration, new systems are increasingly aware of the power of Digital Storytelling in e-learning. Digital Storytelling represents a traditional, modern take on oral Storytelling, combining the ancient tradition of oral Storytelling with today's technological tools. There are many studies in the literature regarding the application of Digital Storytelling techniques in educational pathways (Smeda, Dakich and Sharda, 2013) (Weng, Kuo and Tseng, 2011). In (Smeda, Dakich and Sharda, 2010), guidelines are proposed to develop an advanced framework for e-learning that exploits the Digital Storytelling technique. This is primarily by exploiting the pedagogical-innovative capabilities of this approach that has the potential in engagement that promotes improvement in learning.

Moreover, solutions to offer training models based on Digital Storytelling to different groups of users with different backgrounds (Rossiter and Garcia, 2010) and levels of digital literacy are addressed in the literature to revive such learning models in developing countries (Ungerer, 2019). This technique has been evaluated in different learning domains, including foreign language learning in interdisciplinary projects obtaining exciting results (Yang, Chen and Hung, 2020). Another fascinating field where Storytelling is applied is in e-tourism and museums to enhance Cultural Heritage (Chen, Kao and Kuo, 2014)(Casillo, Clarizia, D'Aniello, De Santo, *et al.*, 2020)(Ioannidis *et al.*, 2013).

Sometimes, Digital Storytelling approaches are complemented with gamification. In (Rossano and Roselli, 2018), a Content Management System for Digital Storytelling to support knowledge acquisition and fruition is proposed. This approach has obtained interesting results in young patients with particular health problems capable of influencing their emotional sphere. Many studies in the literature propose using e-learning systems that are able to use gamification approaches (Amriani *et al.*, 2014; Sanina, Kutergina and Balashov, 2020).

Such approaches aim to consolidate the training path, using the capacity of modern technological systems, which, if well exploited, can adapt to users' needs (Jianu and Vasilateanu, 2017). Based on this, an interesting application is to combine gamification with augmented reality to make the gaming experience more meaningful and enveloping (Bonsignore *et al.*, 2012). This approach has been found in many areas, such as business production (Korn, 2012), social relations (Morschheuser *et al.*, 2017), and e-learning (Saidin, Halim and Yahaya, 2015). In particular, in (Pombo *et al.*, 2020) augmented reality has found excellent feedback in learning paths especially outdoor.

Believing that the use of techniques as such is valid in the e-learning field, another aspect to consider is the use of such methodologies to enhance cultural heritage. These methodologies are exploited to process and interpret personal user information and contextual information. To this end, context can be used to create applications (Dey, 2001; Raento *et al.*, 2005) that can filter relevant data by providing the correct information at the right time and constantly updating (Jin *et al.*, 2014). Modern applications, in addition to personal interests, can adapt to the user's profile (Fink and Kobsa, 2002), distinguishing, for example, between a child and an adult, and can learn from former choices and provide real-time updates concerning the context (Ghiani *et al.*, 2009; Gavalas and Kenteris, 2011). Therefore, the need arises to create a methodology that combines the effectiveness of new technological devices to create training paths in the field. The innovation of the proposed methodology is to exploit the capabilities of the new devices, the amount of data they produce, and the REST services to automatically design context-sensitive training paths valid for different categories of users. These training paths, to be performed in archaeological sites, address education by collecting many innovative techniques in the field of e-learning such as digital Storytelling, augmented reality, and gamification.

This case study aims to use the MuG approach as a recommendation system to suggest educational paths exploiting a high degree of context awareness. This approach is able to combine several methodologies that underlie models working in different domains, such as smart cities (Colace *et al.*, 2018) and cultural heritage enhancement (Casillo *et al.*, 2019). In particular, in concurring with the prefixed objective, that is, to recommend the right educational path to the users according to the context, it is possible to refer to the proposed methodology. This method can bring semantic value to the available data to provide users with illustrated and augmented reality stories in proximity to the visited places and according to different factors able to influence the educational path such as available time, weather conditions, and user's attitudes.

5.2.2.1 Experimental Results

The evaluation process of the proposed methodology, in this case study, was conducted through the development of a prototype based on the proposed architecture. The prototype consists of a server component and a hybrid mobile app. The technologies used were the Ibernate framework, based on Java, to build the Rest API server-side service; the Apache Cordova Framework for developing the mobile app. The prototype is shown in figure 5.8.

The experimental phase was performed in three archaeological parks in southern Italy: The Archaeological Park of Paestum, the Archaeological Park of Herculaneum, and the Archaeological Park of Pompeii. Even if the proposed architecture allows the System to access several open data on the web, the prototype has been modeled with particular attention to the three archaeological sites in this first experimental phase. Several training modules related to the considered archaeological parks have been inserted inside the System, allowing the System to build the related educational paths. The different training modules have been inserted in order to collect the needs of different users (children, adults, and experts) during the training path.

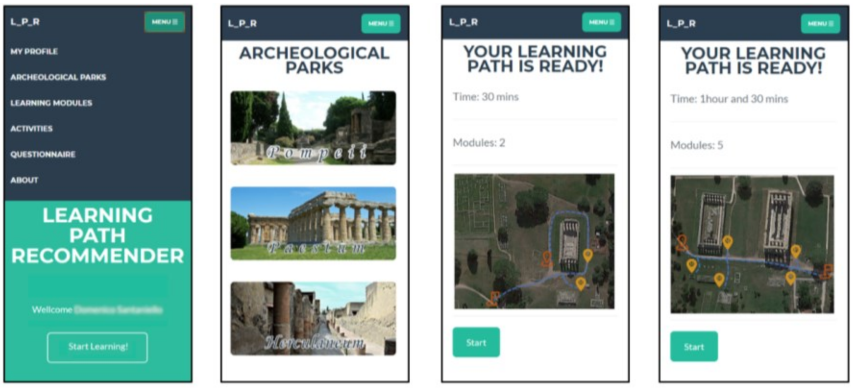


Figure 5.8 *Case of Study on E-Learning paths recommender: Application Prototype*

A total of 230 users were involved, trying to divide them homogeneously by different age groups and characteristics, who were unaware of the purpose of the research. Each participant was equipped with the application prototype, and they enjoyed different training modules within the specific archaeological parks. Users were divided into three different groups composed as follows:

- Group 1: Archaeological Park of Pompeii (95 users)
- Group 2: Paestum Archaeological Park (65 users)
- Group 1: Archaeological Park of Herculaneum (70 users)

Following the experience of using the content, a questionnaire divided into the following sections was proposed to all users:

A. Presentation
  1. The information has been presented appropriately.
  2. The information provided was exhaustive.

B. Reliability
  1. The System provided path suggestions during the entire visit.
  2. The System was able to run adequately during the whole visit.

C. Recommendation
  1. The proposed services and contents have satisfied the needs of the user, based on personal preferences and the current context.
  2. The System has managed to adapt to context changes.

D. Performance
  1. The System was able to show smooth operation and without unexpected jumps.
  2. Response times are adequate.

E. Usability
  1. The system interface is user-friendly.
  2. The System is able to provide suggestions without being unwelcome.

Based on the Likert scale, each section of the questionnaire had two assertions to which five possible responses were associated: totally disagree - TD, disagree - D, Undecided - U, agree - A, totally agree - TA. The responses were collected in Table 5.6.

Table 5.6 *Case of Study on E-Learning paths recommender: Questionnaire answers*

Section	Answer				
	TD	D	U	A	TA
A	18	21	89	187	145
B	20	27	55	230	128
C	18	3	26	239	174
D	11	65	82	203	99
E	37	28	46	205	144

In addition, a smaller number of participants were asked to participate in the experimental phase to evaluate the System's effectiveness in suggesting services. To this end, five training modules were selected for each archaeological park, and users were allowed to indicate whether such a proposed module was relevant according to their needs and context. The participants who took part in the experimental phase for the second time are divided as follows:

- Group 1: Archaeological Park of Pompeii (43 users)
- Group 2: Paestum Archaeological Park (32 users)
- Group 1: Herculaneum Archaeological Park (34 users)

In this experimental phase, the System's knowledge base was augmented by the data that emerged from the experience of previous users. The results, expressing the relevance of the proposed training modules to the context and needs, were collected in the form of a confusion matrix (Figure 5.9, Figure 5.10, Figure 5.11).

Table 5.6 shows the degree of satisfaction of the 230 participants. Users agree that the System is able to provide training modules that are tailored and in line with the context. In Figure 5.12, the results obtained are shown graphically. Users are most satisfied with the ability to recommend the right training path concerning the context. The confusion matrices shown in Figure 5.9, Figure 5.10, and Figure 5.11 show that the System was able to recommend suitable training modules to users based on the profile and time requirements of the users. All the confusion matrices report an overall accuracy greater than 70 %, very encouraging data. Table 5.9 brings back an overall accuracy advanced to 85%; this extraordinary result can be due to two factors. A factor could be the choice of the formative modules that turn out particularly adapted to the selected place. The second factor could be related to the size of the Archaeological Park of Paestum. Unlike other sites, due to its medium-large size and the layout of the archaeological finds, it is better suited to itinerant training and augmented reality. However, all the results obtained are very encouraging and could improve over time by the users' experiences.

Prediction \ Reference	M1	M2	M3	M4	M5
M1	75	5	4	0	2
M2	3	73	5	4	1
M3	1	7	69	6	3
M4	3	0	4	72	7
M5	0	2	5	2	77

Overall Accuracy : 85,12%

Figure 5.9 *Case of Study on E-Learning paths recommender: Confusion Matrix Group 1*

Prediction \ Reference	M1	M2	M3	M4	M5
M1	75	5	4	0	2
M2	3	73	5	4	1
M3	1	7	69	6	3
M4	3	0	4	72	7
M5	0	2	5	2	77

Overall Accuracy : 85,12%

Figure 5.10 *Case of Study on E-Learning paths recommender: Confusion Matrix Group 2*

Prediction \ Reference	M1	M2	M3	M4	M5
M1	75	5	4	0	2
M2	3	73	5	4	1
M3	1	7	69	6	3
M4	3	0	4	72	7
M5	0	2	5	2	77

Overall Accuracy : 85,12%

Figure 5.11 *Case of Study on E-Learning paths recommender: Confusion Matrix Group 3*

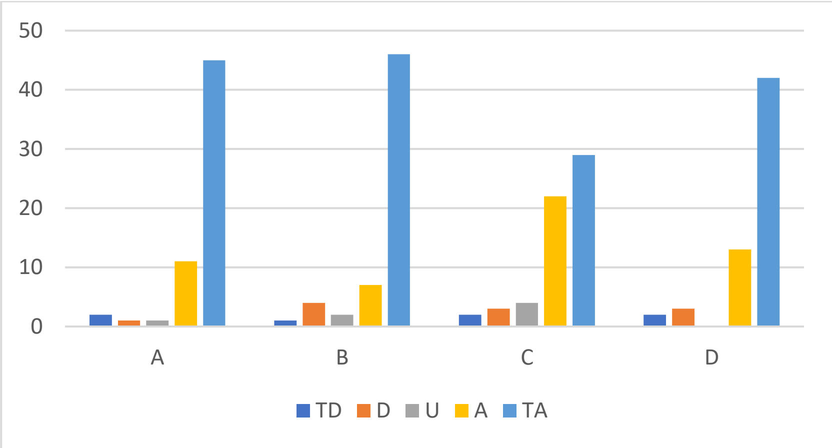


Figure 5.12 *Case of Study on E-Learning paths recommender: Trend of questionnaire answers*

The objective of this case study was to validate the proposed methodology in supporting users to choose training paths within archaeological parks. The aim was to provide tailored training content making the training experience adaptable to the context and the user's needs. This case study confirms that the proposed architecture could be declined in different contexts and mobile applications. The experimental results are promising and encouraging; they show that the System is able to design training paths effectively and that the developed prototype is efficient from several points of view. The degree of reliability the usability of the prototype have been evaluated very positively by the users involved in the experimental campaign. In addition, the recommendation ability of the System reached a high level of accuracy.

5.2.3 Smart Agriculture

This case study concerns applying the MuG Approach in the context of Smart Agriculture (Colace *et al.*, 2022). Nowadays, modern agriculture has become more innovative and smarter (Sishodia, Ray and Singh, 2020), the development of crops occurs with the increasing support of devices able to support our tasks (Bacco *et al.*, 2019). In fact, modern farms are equipped with intelligent devices which can communicate with each other and with farmers contributing to the development of the Internet of Things (IoT) paradigm.

Agriculture has always provided basic needs to humans; however, several problems related to the cultivation phase affect different aspects. For example, specific year periods or sites receive a more significant amount of water than others. Therefore it is necessary to consider different needs according to a selected area. The types of crops are variable. Each of them needs different conditions of temperature, humidity, sun exposure, etc. For this reason, these processes are always managed by experts who ensure the survival of the crops (Gulzar, Abbas and Waqas, 2020). Another problem to be considered is the excessive use of pesticides, insecticides, and fertilizers that create a sort of addiction to the soil (Giri, Dutta and Neogy, 2016). Another problem is related to the product's steps to deliver to the end-user; these steps are expensive and not very productive, lending towards speculation and not aiming to enhance the product generating benefits correctly.

For these reasons, an intelligent agricultural model could cope with the highlighted problems, allowing better management of resources to promote a better profit of crops. In fact, the term Precision Agriculture refers to a business strategy that uses precise and technological information to collect information on spatial and temporal variations within an agricultural field (Calabi-Floody *et al.*, 2019). That information is used to manage agricultural operations to increase production and reduce environmental impact. Indeed, the production strategies employed within this paradigm include variability related to field and site-specific conditions. The primary requirement of this methodology is information, which is considered the heart of Precision Agriculture (Stafford, 2000). Additional requirements are Technology and Management. Electronic and information technologies to serve precision agriculture and agronomic practices (remote sensing, sensors, yield monitoring etc.) can be used individually or in combination. One of the main concepts of Precision Agriculture is to operate only when and where it is needed (according to site-specific logic). Those actions can be performed if a large amount of data is available. In general, the achievement of the paradigm of Precision Agriculture occurs through four phases: data collection (Information), data mapping, decision making, and crop management.

The adoption of Precision Agriculture techniques allows the automation and operational control of activities in the field. The machines assist the operator. According to this model, rainwater harvesting and irrigation can be managed innovatively and remotely. Controlling parameters such as temperature and humidity can be done in real-time. The control of light sources and the detection of pests can be done through special devices and sensors (Elijah *et al.*, 2018). Such information can be acquired and processed to improve the production process more and more. Moreover, the acquired information could be stored and shared for further analysis. Indeed, the life of the specific crops and their biological life can be made available to the customer, which could be used to bring the agricultural product closer to the consumer.

In recent decades, with the advent of the Internet of Things, many efforts have been made in the literature about smart devices aiming to improve the Precision Agriculture paradigm.

Many systems presented in the literature take advantage of Cloud Computing technology (Mekala and Viswanathan, 2017) in order to allow

Precision Agriculture management processes remotely, using lower resources on the fields and more computational resources in the management phase. In particular, in (Prathibha, Hongal and Jyothi, 2017) (Gondchawar and Kawitkar, 2016) (Kodali, Jain and Karagwal, 2016) (Jiang and Moallem, 2018), IoT Based Monitoring Systems are proposed that can measure different parameters through different low-power sensors that can increase the production process. In (Davcev *et al.*, 2018), an IoT agriculture system based on the Low Power Wide Area Network (LoRaWAN) protocol is proposed to achieve significant performance results. In (Muangprathub *et al.*, 2019), an IoT-based hybrid system is presented to leave manual or automatic control for crop management. Such a system is based on a wireless sensor network and is remotely accessible, obtaining promising results. However, although all these systems allow automatic crop management, the System information is only devoted to expert users in all cases.

In this scenario is interesting the application of the MuG Approach that could be able to work on two aspects: management and valorization of agricultural products. The proposed approach, in particular, could help to monitor, manage, and predict what happens concerning crops in terms of data collection that can be analyzed and managed to enhance productivity. Such analyses are helpful to the actors involved in managing the field or system, which we can define as expert users. Moreover, the System could use the available data to allow automatic management or support of the crops. On the other hand, all the data collected, together with Open Data and detailed product information, can be helpful to report general, detailed information and the biological life information about agricultural products that, appropriately shared, could be useful in product enhancement for consumers.

5.2.3.1 Experimental Results

In this case study, an experimental phase was conducted that tested the two fundamental aspects of the proposed methodology. In particular, the ability of the System to support expert users in management through autonomous choices and the ability of the System to present information to consumers was tested. A prototype was developed with a server and client components that can be reached through the web and applications to achieve this goal. Python-based technologies were employed using the Django REST framework for the server part related to the inference engines. The Flutter framework based on the Dart language was used to develop the mobile application.

The dashboard created to support expert users is shown in Figure 5.13 (c), and (d) monitoring a crop includes fundamental steps on which the dashboard is designed:

1. Get an overview of what is happening and identify what needs attention.

2. Focusing on information to understand and compute possible solutions.
3. Trigger the actuators.

The dashboard has been designed to give an overview of what can be done, providing an overview of what is happening and allowing for quick reactions. The environment allows to call attention to valuable items, send notifications when needed, and represent performance measures accurately and directly. The app to support the consumers (Figure 5.13 (a) and (b)) contains all the information that the System is able to supply concerning the products, which could be: curiosity, biological history, nutritional values, and culinary recipes. The app makes it possible to know a product’s etymology, origin, history, varieties, properties, and benefits. There is a section where it is possible to trace the growth of the product, indicating, for each phase, favorable and unfavorable events of the seasons of cultivation. It is also possible to know the nutritional values and a helpful section where possible the possible culinary uses of the product. The app also allows users to interact through comments and feedback.

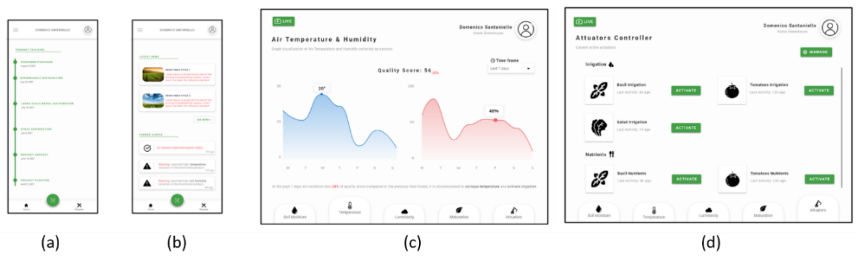


Figure 5.13 *Case of Study on Smart Agriculture: Application Prototype*

The prototype has been connected to a small experimental crop growing environment that includes a greenhouse to cultivate aromatic plants for experimental purposes. The smart growing environment has been realized through a Raspberry Pi 4 board that has been connected to actuators that allow actions such as irrigation, humidification, and ventilation. Various data about temperature, humidity, soil moisture, and images were collected. The monitored parameters come in detail through the sensors shown in Table 5.7. Communications to the central server were made through MQTT and HTTP protocols.

The experimental phase, therefore, was done in two steps. The first step is fundamental to understanding the System's reliability in supporting expert users. To develop this experimentation phase, the data collected by the System concerning about one year of observation has been used. The data collected concern the environmental and meteorological conditions (temperature, humidity, soil moisture, images, etc.) and the actions taken to improve these conditions (ventilation, irrigation, and fertigation). For about one year, from September 2019 to September 2020, collecting more than 17000 instances. In the Irrigation and Ventilation prediction cases, 17280 events were analyzed.

Table 5.7 *Case of Study on Smart Agriculture: Parameters, actions, sensors, and actuators prototype details*

Parameter	Sensor
Soil Moisture	STEMMA Soil Sensor - I2C Capacitive Moisture Sensor
Soil Temperature	STEMMA Soil Sensor - I2C Capacitive Moisture Sensor
Air Temperature	BME280 Sensor Module
Air Humidity	BME280 Sensor Module
Luminosity	SI1145 Digital UV Index / IR / Visible Light Sensor
Growing	Pi Camera Module
Action	**Actuator**
Irrigation	Pump
Fertilization	Pump
Ventilation	Servo Motor

In particular, 12960 (about 75%) instances were used for the training set and 4320 (about 25%) for the test set. In the Fertilization prediction case, the events were aggregated by analyzing 8640 instances. In detail, 6480 (about 75%) instances were used for the training set and 2160 (about 25%) for the test set. Through the training set, the proposed methodology made it possible to learn the Bayesian network structure that was then tested. The test was performed to understand if the learned structure can automatically identify the need for irrigation, fertilization, and ventilation. The results, shown in Figure 5.13 and Table 5.8, show that the System exhibits an Overall Accuracy that consistently exceeds 95% with acceptable Precision, Recall, and F_1-Score.

The second step of the experimental phase concerned the ability of the System to show the information of the product to the consumer. For this purpose, 30 users were involved and shown a prototype app. The app contains all the information the System can provide concerning the product, including its biological history, and allows users to interact with it through comments and feedback.

Irrigation			Fertilization			Ventilation		
Reference			Reference			Reference		
Prediction	Yes	No	Prediction	Yes	No	Prediction	Yes	No
Yes	275	87	Yes	37	12	Yes	287	102
No	105	3659	No	25	2086	No	93	3644
Overall Accuracy:		**95,35%**	Overall Accuracy:		**98,28%**	Overall Accuracy:		**95,27%**

Figure 5.14 *Case of Study on Smart Agriculture: Confusion Matrix*

Table 5.8 *Case of Study on Smart Agriculture: System Performance*

		Yes	**No**
Irrigation	**Precision**	75,97%	97,21%
	Recall	72,37%	97,68%
	F$_1$-Score	74,13%	97,44%
Fertilization	**Precision**	75,51%	98,81%
	Recall	59,68%	99,43%
	F$_1$-Score	66,67%	99,12%
Ventilation	**Precision**	73,78%	97,51%
	Recall	75,53%	97,28%
	F$_1$-Score	74,64%	97,39%

In order to test the System, after the interaction, users were sent the questionnaire shown below. The questionnaire consists of five sections. All participants owned a mobile device and were between 18 and 65 years old. Five possible responses were associated with each statement in a specific section: "I totally disagree" - TD, "I disagree" - D, "Undecided" - U, "I agree" - A, "I totally agree" - TA.

Section A: recommendation

1. The services and content offered have met the consumer's needs.
2. The System was able to effectively show the biological life of the product, providing greater awareness of what was purchased.

Section B: interaction

1. The user is encouraged to use the app, leaving comments and feedback.
2. App notifications are effective and discreet.

Section C: presentation

1. Content and services are readily available and appropriately presented.
2. The information provided is comprehensive.

Section D: usability

1. The system interface is user-friendly.
2. Response times are adequate.

Section E: future developments

1. It would be helpful to introduce storytelling techniques that would bring the product even closer to the consumer.
2. It would be interesting to provide a way for the consumer to be able to show the result of a recipe to all users.

Figure 5.15 shows the results aggregated by section. Looking at all the responses, user satisfaction is high. In addition to the reliability and performance of the System, the ability to present products and usability were tested, which provided us with important feedback on the System's ability to interact.

Although at a preliminary stage, the developed System showed a promising user acceptance index. In particular, significant results have been reached in sections A and E that represent the ability to supply the correct information to the users and the possibility to interact.

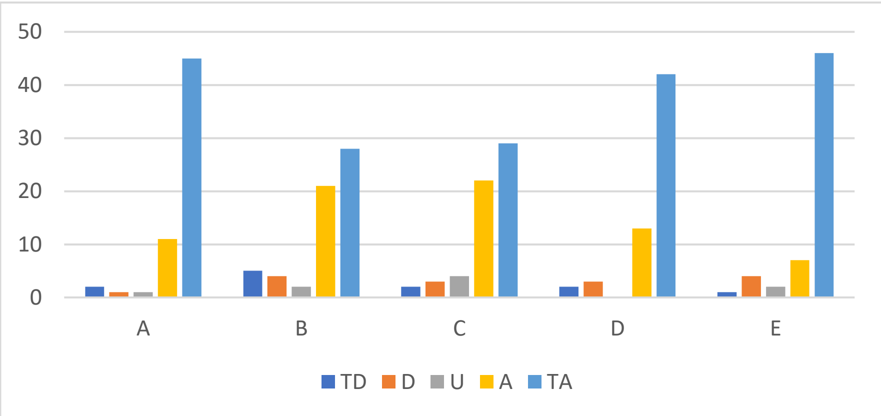


Figure 5.15 *Case of Study on Smart Agriculture: Trend of questionnaire answers*

The objective of this case study was to introduce a framework able to work on two aspects: the management and the valorization of agricultural products. This framework has provided an experimental phase performed through two steps involving the development of a prototype and the ability of the System to learn from the data collected. Although preliminary, the experimental phases have gathered promising results, showing that the System can learn and support expert users and enhance products by bringing them closer to consumer needs.

5.2.4 Smart Home

This case study concerns the management of a Smart Home. In recent years, IoT applications and technologies are becoming increasingly within people's reach, interfacing with their daily activities and occupations even in the home environment (Risteska Stojkoska and Trivodaliev, 2017).

Modern buildings are increasingly becoming an active part of the elements that constitute a Smart City. It is possible to distinguish two approaches, which are often confused in this context. One approach, proposed by the home automation sector (Gill *et al.*, 2009), promotes houses equipped with technologies capable of controlling systems through remote control; this approach is the one with minor complexity. The second approach is based on the Internet of Things, which aims to overcome mere control, introducing the smart component in home management (Alaa *et al.*, 2017). The development of this approach aims to propose systems capable of learning from the actions and controls performed by users to act in a preventive way and optimize the management of the home environment (Al-Ali *et al.*, 2017). In this perspective, the smart home system and the devices work in synergy, exchanging data and information to automate occupants' actions by calibrating them on their preferences.

To be defined as such, smart homes must own specific technological characteristics. The architecture of such systems is based on a communication network between different components and smart devices that allow control, action, and obtaining information for data management. Smart Devices, such as home appliances and different sensors (i.e., temperature and humidity sensor, motion sensor etc.) able to monitor the environment, and actuators able to perform actions are interconnected in a virtual component that generates a Digital Twin of the System (Nazarenko and Camarinha-Matos, 2020). This environment can integrate data processing and management algorithms allowing the various elements to interact with each other, and also allows management and control via interface (Bhat, Bhat and Gokhale, 2007). The smart component can address one or more features of the home. The most common are kitchen equipment, lighting, temperature control, and infotainment systems (Marikyan, Papagiannidis and Alamanos, 2019). Nowadays, the most widely used applications concern those related to video surveillance, access, and presence management (Kodali *et al.*, 2017); in this field, other applications not yet widely used concern surveillance systems for home health and assistance of elderly and sick people (Liu *et al.*, 2016).

Another emerging field of application concerns energy and comfort management (Minoli, Sohraby and Occhiogrosso, 2017); this area is growing due to interaction through voice assistants. Two main functionalities characterize smart home systems. The first one concerns human activity tracking, and the second one is inherent to the acquisition and use of data. These functionalities must be continuous and occur in real-time to respond reactively to the needs of the inhabitants (Mocrii, Chen and Musilek, 2018). The main goal of tracking is to monitor user actions in order to acquire valuable knowledge for possible predictions. The data related to tracking concern the position of the users and the interaction with the interactive devices, such are integrated with other data coming from sensors that can concern energy consumption, environmental conditions such as temperature and humidity etc.

In this scenario, it is interesting to design a system that improves the quality of life by optimizing available resources. In fact, some services allow to schedule tasks and activities and remotely define routine for smart systems, helping to save time in the daily management of the home. Such systems can also help optimize the use of energy resources and help to improve security through more effective control of the environment. The purpose of this case study is to apply the MuG approach in the context of Smart Home. Due to the ability to understand contextual information, the proposed approach could support the sensor network by defining rules and triggering actions that allow the devices to intervene proactively in the occupations and lives of users and propose activities tailored to their preferences.

5.2.4.1 Experimental Results

In this case study, an experimental phase was conducted to test two fundamental aspects of the proposed methodology. In particular, it has been tested the ability of the System to support users in the management of the home environment through suggestions and autonomous choices and the ability of the System, through the application, to present information and provide support to users. To achieve this goal, a prototype has been developed with a server component and client components reachable from the web and through applications. Python-based technologies were employed using the Django REST framework for the server part related to the inference engines. For the development of the mobile application was used the Flutter framework, which is based on the Dart language.

The application was developed in order to support users in managing a smart home. The development of the application aimed to show users an overview of all the key parameters that are monitored in the application and choose System actions. In addition, the application was designed to assist users by providing helpful information about household activities, home care, and food preparation.

To evaluate the proposed approach, and for experimental purposes, the proposed System was connected to two rooms of a corporate building. The smart environment was obtained through Raspberry Pi 4 and Pi 0 boards to which different types of indoor sensors were connected. Using the same type of board (Raspberry Pi 4), the weather parameters were also monitored through a Weather Station as an integration of the collected data. The electrical absorption of a small solar system for experimental purposes was monitored. In particular, the monitored data are coming from the sensors reported in Table 5.9.

The monitored parameters have allowed building a dataset through which the System can acquire indoor-outdoor environmental parameters such as weather conditions, humidity, indoor temperature, air quality, and the presence of people.

Table 5.9 *Case of Study on Smart Home: Parameters, actions, sensors, and actuators prototype details*

Parameter	Sensor
Air indoor/outdoor Temperature	BME680 Sensor Module
Air indoor/outdoor Humidity	BME680Sensor Module
VOC indoor/outdoor	BME680Sensor Module
Luminosity	SI1145 Digital UV Index / IR / Visible Light Sensor
Weather Description	Pi Camera Module
Voltage	Voltage Sensor
Wind Speed	Anemometer
Wind Direction	Wind Vane
Rainfall	Rain Gauge
Action	**Actuator**
Light Control	Dimmer
Ventilation	Fun Motor

The objectives of this experimental campaign were to foresee the need to control the lighting system and to ventilate the building when necessary. This experimental phase was preliminary and aimed at evaluating the System's effectiveness on reduced actions to allow its eventual use in a real environment.

The first step of the experimental phase is fundamental to understanding the System's reliability in supporting users autonomously. To develop this experimentation phase, the data collected by the System for about six months of observation have been used. The data collected concern weather conditions (temperature, humidity, rain, etc.), indoor environmental conditions (temperature, humidity, presence, etc.), conditions of energy production connected to the Solar Panel System, and the actions taken to improve these conditions (lighting, ventilation of environments) for a period of about six months from May to November 2021 collecting about 4000 instances. In particular, among the 3960 instances analyzed, 2970 (about 75%) were used for the training set and 950 (about 25%) for the test set. Through the MuG approach, it was possible to learn the Bayesian network structure that was then tested through the training set. The test evaluated the on and off phase, understanding if the learned structure can identify actions to be performed automatically. In particular, two parameters were identified to be predicted, energy conservation and ventilation. Energy-saving is a reduced lighting condition triggered when there are no people inside the spaces or excessive lighting due to natural brightness is not needed. Ventilation is a practice performed, through exhaust fans, to ventilate and sanitize the environment. Despite the little data available, the experimental results, shown in figure 5.15 and table 5.10, show that the System is able to reach Overall Accuracy measures that consistently exceed 95%, and encouraging results of Precision, Recall, and F_1-Score.

Energy Saving			Ventilation		
	Reference			Reference	
Prediction	Yes	No	Prediction	Yes	No
Yes	136	18	Yes	189	11
No	22	774	No	19	731
Overall Accuracy:		**95,79%**	Overall Accuracy:		**96,84%**

Table 5.10 *Case of Study on Smart Home: Confusion Matrixes*

The second step of the experimental phase concerned the ability of the System to show the information effectively acquired by the System through an app, shown in Figure 5.16. For this purpose, 15 users were involved and were shown a prototype app. The app contains all the information that the System is able to provide concerning the indoor-outdoor environmental conditions and the actions to be taken, such as lighting, heating, and ventilation system control.

In order to test the System, after the interaction, users were sent the questionnaire shown below. The questionnaire consists of five sections. All participants owned a mobile device and were between 25 and 58 years old. Five possible responses were associated with each statement in a specific section: "I totally disagree" - TD, "I disagree" - D, "Undecided" - U, "I agree" - A, "I totally agree" - TA.

Section A: recommendation

1. The proposed services and contents have satisfied the needs of the user, based on personal preferences and the current situation.

2. The System has managed to adapt to context changes.

Section B: interaction

1. Interaction with the app is natural.
2. The app is able to communicate effectively without complications.

Section C: presentation

1. Content and services are presented appropriately.
2. The information provided is comprehensive.

Section D: usability

1. The System interface is user-friendly.
2. Response times are adequate.

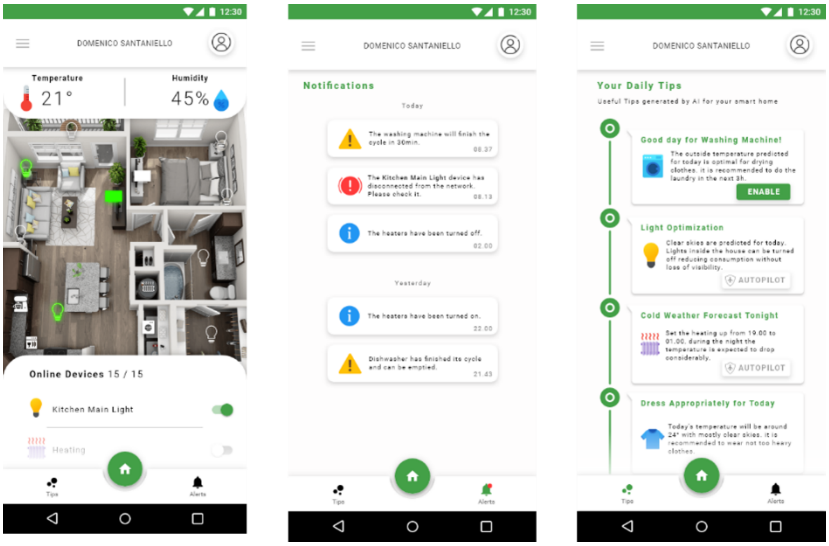


Figure 5. 16 *Case of Study on Smart Home: Application Prototype*

Figure 5.11 shows the results aggregated by section. Looking at all responses, the degree of user satisfaction is high. In addition to the reliability and performance of the System, the ability to present services and the usability of the app were tested, providing us with important feedback on the System's ability to interact.

Although at a preliminary stage, the developed System showed a promising user acceptance index in recommending the right actions to users at the right time. In particular, significant results were achieved in sections A and B, representing the ability to provide the correct information to users and the possibility to interact.

Table 5.11 *Case of Study on Smart Home: System Performance*

		Yes	No
Energy Saving	**Precision**	88,31%	97,24%
	Recall	86,08%	97,73%
	F_1-Score	87,18%	97,48%
Ventilation	**Precision**	94,50%	97,47%
	Recall	90,87%	98,52%
	F_1-Score	92,65%	97,99%

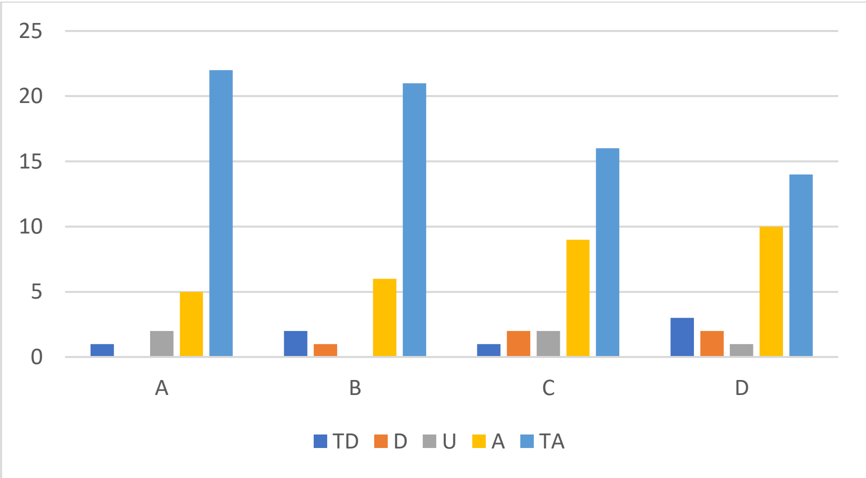


Figure 5.17 *Case of Study on Smart Home: Trend of questionnaire answers*

The objective of this case study was to test the MuG Approach to support users in managing a Smart Home environment. This framework has provided an experimental phase that took place through two steps that involved the development of a prototype and the ability of the System to learn based on the data collected. The experimental phases, although preliminary, have collected promising results. They showed that the System is able to learn and support users. The application has achieved exciting results that allow the System to be used to manage working and cooperative environments. Given the recent cases of the pandemic, the System, with an appropriate expansion of the available sensors and with an appropriate data collection, could be used to ensure the sanitation and safe use of shared work environments.

Chapter 6
Future Trends

This research introduced the Multilevel Graph Approach, an innovative approach to managing complex IoT-based scenarios. The added value of the proposed system is to exploit graph approaches for context modeling, which allow the extraction of useful semantic information. The graph approaches chosen for context modeling are Ontologies and Context Dimension Tree, assisted through a probabilistic graph approach, such as Bayesian Networks, which allows for event prediction. The integration of the three graph approaches leads the system toward better performance, allowing users to understand the relationships between entities.

The validation of the proposed model has been developed through a preliminary case study on data acquired in the city of London, obtaining valid results. In addition, the system has been employed in several real-based complex scenarios obtaining promising results. From the case studies, it can be seen that the system's performance improves with the increase of available data. Moreover, a further peculiarity of the system lies in its ability to communicate with other systems, which use the same formalisms for context modeling, in order to improve its performance further.

However, although the results obtained in all the case studies are encouraging, the system could be improved by investigating the use of other probabilistic graph methods, e.g., Markov chains, and exploiting boosting strategies to improve classification and probabilistic prediction capability.

Markov chains (Rabiner, 1989) are stochastic processes, representable by graphs, that describe the links between all possible configurations of a system. According to stochastic systems, the evolution of a phenomenon is described through random variables that have value on a set of defined spaces of states, dependent on a parameter belonging to an ordered set of times. The peculiarity of Markov chains consists in having a set of times equal to $\mathcal{N}_0$ and space of states S finite or numerable.

$$\{X_n : \Omega \mapsto S \ : \mathrm{n} \in \mathcal{N}_0\} \tag{6.1}$$

In particular, the Markov chains satisfy the condition (6.2) defined Markov property. This relation states that the probability of the random variable associated with instant n depends only on the value assumed by the process at the previous instant.

$$P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_1 = x_1) = P(X_n = x_n | X_{n-1} = x_{n-1}) \quad (6.2)$$

In the Context Awareness field, stochastic processes and, more specifically, Markov chains can allow model enrichment through temporal context, helpful in analyzing phenomena (Boytsov, Zaslavsky and Synnes, 2009). In addition, other approaches in the scientific literature exploit Markov chains in supporting Situation Awareness models for handling complex IoT-based scenarios (Liang *et al.*, 2008; P. Zhang *et al.*, 2020). However, according to the scientific literature, there are still not many approaches that employ Markov chains to support the management of complex scenarios; hence it remains an open question to establish the support capability of such a model in the prediction phase.

A further improvement to the system in classification and probabilistic prediction capability could lie in exploiting boosting strategies. Boosting is a general method that attempts to "boost" the accuracy of any learning algorithm (Schapire, 1999). In the context of probabilistic graph approaches, such as Bayesian Networks, the goal is to make more accurate conditional probabilities, which can be affected by errors related to the dataset's structure. Such techniques are handy for improving the classification of events or combining with structural learning models, allowing the improvement of prediction ability (Viaene, Derrig and Dedene, 2004; Jing, Pavlović and Rehg, 2008). In addition, the study of such techniques in combination with Markov Chains could be investigated further to improve the proposed system (Jing, Pavlović and Rehg, 2008).

In addition, other future developments may include the introduction of new methodologies to increase the capabilities and reliability of the system. In particular, it would be helpful to use methodologies that allow a natural interface with users, such as Chatbots and approaches that can increase data security and reliability through decentralized certification techniques. For this purpose, two modules could be included in the proposed architecture that can provide improvements to the proposed system.

- **Chatbot Module.** A chatbot is a software designed to simulate a conversation with a human being (Shum, He and Li, 2018). The dialogue can be reproduced by text or in a vocal way. The development and accuracy of chatbots have increased over the years and have continued to grow with artificial intelligence, machine learning, and improved algorithms for natural language processing (Pérez-Soler, Guerra and de Lara, 2018). The simplest and most popular types of chatbots can interact automatically, recognizing relevant parts of speech and retrieving answers; other types can take advantage of machine learning and Deep Learning techniques to provide generated responses according to a training process (Ning *et al.*, 2019). These virtual agents are based on the understanding of natural language and are able to learn the user's preferences, adapting searches and responses based on them (Casillo, Clarizia, D'Aniello, de Santo, *et al.*, 2020).

A chatbot represents an innovative user interaction model. In such a scenario, a chatbot module aims to manage interactions with users naturally, making the system capable of maintaining a conversation and interacting if the system provides suggestions to users according to the acquired information.

- **Blockchain Module.** Decentralized systems become fundamental due to flexibility and particular characteristics, aiming to guarantee security. A decentralized system is composed of a peer-2-peer (P2P) network in which several nodes constantly communicate with each other to exchange information. Ideally, each node is perfectly synchronized with others in the network sharing the same knowledge. This approach leads to systems where information is distributed over the network through thousands of replicas (Wang *et al.*, 2019). Blockchains are a particular decentralized system that developed around 2009 to provide an alternative to highly centralized financial systems. The knowledge that nodes in the P2P network share is structured in logically connected blocks to form a chain. Over the years, this technology has evolved to support more and more functionality distant to financial purposes (Zheng *et al.*, 2017). Specifically, Blockchains rely on robust hashing techniques and public key cryptography to maintain the security and consistency of information across the hundreds of thousands of nodes scattered across the network (Yakavenka, 2018).

  The idea behind a Blockchain is to distribute knowledge making it immutable. Such a feature can be crucial to act correctly in a decision-oriented system like the proposed one. In this scenario, the purpose of the Blockchain module is to share information with other nodes guaranteeing the originality of the transmitted data (Casillo *et al.*, 2021).

Conclusions

The increase in technology leads to the creation of increasingly complex and sophisticated scenarios that bring advantages to modern society if managed efficiently. This research aimed to deepen the technological aspects, degree and probabilistic approaches to propose a methodology for the management of complex scenarios based on IoT. An innovative model capable of exploiting Machine Learning techniques in order to predict events in complex scenarios has been presented and validated. The added value of the proposed system is to exploit graph approaches able to include and manage semantic information. Such approaches allow the system to perform better while also allowing users to understand the relationships between entities. The performance of the system improves as the available data increases. In addition, a further advantage of the system lies in the ability to interfacing with other systems that use the same formalisms to improve its performance further.

In different application scenarios, experimental results show the system's ability to be effective. In particular, in the model validation related to the case study of London Smart City, the performance of the system proved to be particularly profitable in the application related to the prediction of rainfall phenomena. In this case, in fact, the system was able to provide reliability to the forecasts, significantly improving the performance compared to traditional methodologies. Moreover, in real case studies, the system has always shown remarkable forecast reliability due to semantic relations arising from the analysis of the context and situation. These relationships, which arise from the context modeling graphs, potentially allow the proposed methodology to interact with other systems by expanding its knowledge and exploiting knowledge graphs already known in the literature. However, this strength may be a limitation in attempting to use the proposed approach in a new domain, where the construction of the knowledge graphs, which represents a crucial part of the proposed approach, could be demanding.

Future activities include insights through other graph methods applicable to the system, such as Markov Chains, and leveraging boosting strategies to improve classification and probabilistic prediction capabilities. In addition, other future enhancements include the development of new modules to increase the capabilities and reliabilities of the system. Specifically, it could be interesting to develop a module that integrates the proposed system with a Chatbot, an agent able to provide a natural language interface for users. Another fascinating development lies in including modules to increase the security and reliability of the data managed by the model through certification techniques that exploit the Blockchain paradigm.

References

Aazam, M., Zeadally, S. and Harras, K.A. (2018) "Deploying Fog Computing in Industrial Internet of Things and Industry 4.0," *IEEE Transactions on Industrial Informatics*, 14(10), pp. 4674–4682. doi:10.1109/TII.2018.2855198.

Abdmeziem, M.R., Tandjaoui, D. and Romdhani, I. (2016) "Architecting the internet of things: State of the art," in *Studies in Systems, Decision and Control*. doi:10.1007/978-3-319-22168-7_3.

Abowd, G.D. *et al.* (1999a) "Towards a better understanding of context and context-awareness," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/3-540-48157-5_29.

Abowd, G.D. *et al.* (1999b) "Towards a better understanding of context and context-awareness," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/3-540-48157-5_29.

Abowd, G.D. and Mynatt, E.D. (2000) "Charting Past, Present, and Future Research in Ubiquitous Computing," *ACM Transactions on Computer-Human Interaction*, 7(1). doi:10.1145/344949.344988.

Abrahao, E. and Hirakawa, A.R. (2017) "Task ontology modeling for technical knowledge representation in agriculture field operations domain," in *Proceedings - 2017 2nd International Conference on Information Systems Engineering, ICISE 2017*. doi:10.1109/ICISE.2017.18.

Aceto, G., Persico, V. and Pescapé, A. (2020) "Industry 4.0 and Health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0," *Journal of Industrial Information Integration*. doi:10.1016/j.jii.2020.100129.

Adams, M.J., Tenney, Y.J. and Pew, R.W. (1995) "Situation awareness and the cognitive management of complex systems," *Human Factors*. doi:10.1518/001872095779049462.

Adomavicius, G. *et al.* (2005) "Incorporating contextual information in recommender systems using a multidimensional approach," *ACM*

Transactions on Information Systems, 23(1), pp. 103–145. doi:10.1145/1055709.1055714.
Adomavicius, G. *et al.* (2011) "Context-Aware Recommender Systems," *AI Magazine*, 32(3), p. 67. doi:10.1609/aimag.v32i3.2364.
Agresta, A. *et al.* (2014) "An Ontology Framework for Flooding Forecasting," in *Computational Science and Its Applications –ICCSA 2014*. Springer, Cham, pp. 417–428. doi:10.1007/978-3-319-09147-1_30.
Akhter, R. and Sofi, S.A. (2021) "Precision agriculture using IoT data analytics and machine learning," *Journal of King Saud University - Computer and Information Sciences*. doi:10.1016/j.jksuci.2021.05.013.
Alaa, M. *et al.* (2017) "A review of smart home applications based on Internet of Things," *Journal of Network and Computer Applications*. doi:10.1016/j.jnca.2017.08.017.
Al-Ali, A.R. *et al.* (2017) "A smart home energy management system using IoT and big data analytics approach," *IEEE Transactions on Consumer Electronics*, 63(4). doi:10.1109/TCE.2017.015014.
Alamgir Hossain, S.K., Anisur Rahman, M. and Hossain, M.A. (2018) "Edge computing framework for enabling situation awareness in IoT based smart city," *Journal of Parallel and Distributed Computing*, 122. doi:10.1016/j.jpdc.2018.08.009.
Alexopoulos, K. *et al.* (2016) "A concept for context-aware computing in manufacturing: the white goods case," *International Journal of Computer Integrated Manufacturing*, 29(8). doi:10.1080/0951192X.2015.1130257.
Al-Fuqaha, A. *et al.* (2015) "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," *IEEE Communications Surveys and Tutorials*, 17(4). doi:10.1109/COMST.2015.2444095.
Almusaylim, Z.A. and Zaman, N. (2019) "A review on smart home present state and challenges: linked to context-awareness internet of things (IoT)," *Wireless Networks*, 25(6). doi:10.1007/s11276-018-1712-5.
Al-Sultan, S., Al-Bayatti, A.H. and Zedan, H. (2013) "Context-aware driver behavior detection system in intelligent transportation systems," *IEEE Transactions on Vehicular Technology*, 62(9). doi:10.1109/TVT.2013.2263400.
Al-Turjman, F. and Malekloo, A. (2019) "Smart parking in IoT-enabled cities: A survey," *Sustainable Cities and Society*. doi:10.1016/j.scs.2019.101608.
Alvarez, M.G., Morales, J. and Kraak, M.J. (2019) "Integration and exploitation of sensor data in smart cities through event-driven applications," *Sensors (Switzerland)*, 19(6). doi:10.3390/s19061372.
Amriani, A. *et al.* (2014) "An empirical study of gamification impact on e-Learning environment," in *Proceedings of 2013 3rd International Conference on Computer Science and Network Technology, ICCSNT 2013*. doi:10.1109/ICCSNT.2013.6967110.

Anand, P. (2016) "Towards evolution of M2M into Internet of Things for analytics," in *2015 IEEE Recent Advances in Intelligent Computational Systems, RAICS 2015*. doi:10.1109/RAICS.2015.7488447.

Annunziata, G. *et al.* (2016) "Appoggiomarino: A context Aware app for e-citizenship," *ICEIS 2016 - Proceedings of the 18th International Conference on Enterprise Information Systems*, 2, pp. 273–281.

Arp, R., Smith, B. and Spear, A.D. (2016) "What Is an Ontology?," in *Building Ontologies With Basic Formal Ontology*. doi:10.7551/mitpress/9780262527811.003.0001.

Ashton, K. (2009) "That 'internet of things' thing," *RFID journal*, 22(7), pp. 97–114.

Atzori, L., Iera, A. and Morabito, G. (2010) "The Internet of Things: A survey," *Computer Networks*, 54(15), pp. 2787–2805. doi:10.1016/j.comnet.2010.05.010.

Azwar, H. *et al.* (2019) "Intrusion Detection in secure network for Cybersecurity systems using Machine Learning and Data Mining," in *2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences, ICETAS 2018*. doi:10.1109/ICETAS.2018.8629197.

Bacco, M. *et al.* (2019) "The Digitisation of Agriculture: a Survey of Research Activities on Smart Farming," *Array*, 3–4, p. 100009. doi:10.1016/j.array.2019.100009.

Baddeley, A. (1992) "Working memory: The interface between memory and cognition," *Journal of Cognitive Neuroscience*. doi:10.1162/jocn.1992.4.3.281.

Baldauf, M., Dustdar, S. and Rosenberg, F. (2007) "A survey on context-aware systems," in *International Journal of Ad Hoc and Ubiquitous Computing*. doi:10.1504/IJAHUC.2007.014070.

Batet, M. *et al.* (2012) "Turist@: Agent-based personalised recommendation of tourist activities," *Expert Systems with Applications*, 39(8). doi:10.1016/j.eswa.2012.01.086.

Bedny, G. and Meister, D. (1999) "Theory of Activity and Situation Awareness," *International Journal of Cognitive Ergonomics*, 3(1). doi:10.1207/s15327566ijce0301_5.

Bellini, P. *et al.* (2014) "Km4City ontology building vs data harvesting and cleaning for smart-city services," *Journal of Visual Languages and Computing*, 25(6). doi:10.1016/j.jvlc.2014.10.023.

Bernaras, A., Laresgoiti, I. and Corera, J. (1996) "Building and Reusing Ontologies for Electrical Network Applications," in Wahlster, W. (ed.) *Proceedings of the 12th European Conference on Artificial Intelligence (ECAI'96)*. Chichester, UK: John Wiley and Sons, pp. 298–302.

Best, D.M. *et al.* (2010) "Real-time visualization of network behaviors for situational awareness," in *ACM International Conference Proceeding Series*. doi:10.1145/1850795.1850805.

Bhat, O., Bhat, S. and Gokhale, P. (2007) "Implementation of IoT in Smart Homes," *International Journal of Advanced Research in Computer and Communication Engineering ISO*, 3297(December).

Blasch, E. (2013) "Enhanced air operations using JView for an air-ground fused situation awareness udop," in *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*. doi:10.1109/DASC.2013.6712597.

Bolchini, C. *et al.* (2006) *Context-ADDICT*, *Technical Report 2006.044*. Politecnico di Milano.

Bonsignore, E.M. *et al.* (2012) "Mixed reality games," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*. doi:10.1145/2141512.2141517.

Boytsov, A., Zaslavsky, A. and Synnes, K. (2009) "Extending context spaces theory by predicting run-time context," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/978-3-642-04190-7_2.

van den Broek, A.C. *et al.* (2011) "Improving maritime situational awareness by fusing sensor information and intelligence," in *Fusion 2011 - 14th International Conference on Information Fusion*.

Brüggemann, S. *et al.* (2016) "Ontology-based data access for maritime security," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/978-3-319-34129-3_45.

Buczak, A.L. and Guven, E. (2016) "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," *IEEE Communications Surveys and Tutorials*, 18(2). doi:10.1109/COMST.2015.2494502.

Buntine, W. (1990) "Myths and Legends in Learning Classification Rules," in *Proceedings of the Eighth National Conference on Artificial Intelligence - Volume 2*.

Calabi-Floody, M. *et al.* (2019) "Optimization of wheat straw co-composting for carrier material development," *Waste Management*, 98, pp. 37–49. doi:10.1016/j.wasman.2019.07.041.

Casillo, M. *et al.* (2017) "Context-aware computing for improving the touristic experience: A pervasive app for the Amalfi coast," in *2017 IEEE International Workshop on Measurement and Networking, M and N 2017 - Proceedings*. doi:10.1109/IWMN.2017.8078373.

Casillo, M. *et al.* (2019) "An Approach for Recommending Contextualized Services in e-Tourism," *Information*, 10(5), p. 180. doi:10.3390/info10050180.

Casillo, M., Clarizia, F., D'Aniello, G., de Santo, M., *et al.* (2020) "CHAT-Bot: A cultural heritage aware teller-bot for supporting touristic experiences," *Pattern Recognition Letters*, 131, pp. 234–243. doi:10.1016/j.patrec.2020.01.003.

Casillo, M., Clarizia, F., D'Aniello, G., De Santo, M., *et al.* (2020) "CHAT-Bot: a Cultural Heritage Aware Teller-Bot for supporting touristic experiences," *Pattern Recognition Letters*. doi:10.1016/j.patrec.2020.01.003.

Casillo, M. *et al.* (2021) "COVID-19 data sharing and organization through blockchain and decentralized models," in *CEUR Workshop Proceedings*.

Castiglione, A. *et al.* (2020) "Securing the internet of vehicles through lightweight block ciphers," *Pattern Recognition Letters*, 135, pp. 264–270. doi:10.1016/j.patrec.2020.04.038.

Chen, G. and Kotz, D. (2000) "A survey of context-aware mobile computing research. Technical Report TR2000-381," *Time*.

Chen, H., Finin, T. and Joshi, A. (2003a) "An ontology for context-aware pervasive computing environments," *The knowledge engineering review*, 18(3), pp. 197–207.

Chen, H., Finin, T. and Joshi, A. (2003b) "An ontology for context-aware pervasive computing environments," *Knowledge Engineering Review*, 18(3). doi:10.1017/S0269888904000025.

Chen, H.C., Kao, Y.F. and Kuo, C.L. (2014) "A multimedia storytelling in a rural village: The show Taiwan e-tourism service using tablet technologies," in *Proceedings - 2014 IIAI 3rd International Conference on Advanced Applied Informatics, IIAI-AAI 2014*. doi:10.1109/IIAI-AAI.2014.109.

Chen, M., Mao, S. and Liu, Y. (2014) "Big Data: A Survey," *Mobile Networks and Applications*, 19(2), pp. 171–209. doi:10.1007/s11036-013-0489-0.

Chiappe, D., Vu, K.P.L. and Strybel, T. (2012) "Situation Awareness in the NextGen Air Traffic Management System," *International Journal of Human-Computer Interaction*, 28(2). doi:10.1080/10447318.2012.634768.

Chong, H.F. and Ng, D.W.K. (2017) "Development of IoT device for traffic management system," in *Proceedings - 14th IEEE Student Conference on Research and Development: Advancing Technology for Humanity, SCOReD 2016*. doi:10.1109/SCORED.2016.7810059.

Chouhan, S.S., Singh, U.P. and Jain, S. (2020) "Applications of Computer Vision in Plant Pathology: A Survey," *Archives of Computational Methods in Engineering*, 27(2). doi:10.1007/s11831-019-09324-0.

Colace, F. *et al.* (2014) "A Context-Aware Framework for Cultural Heritage Applications," in *2014 Tenth International Conference on Signal-Image Technology and Internet-Based Systems*. IEEE, pp. 469–476. doi:10.1109/SITIS.2014.14.

Colace, F. *et al.* (eds) (2015) *Data Management in Pervasive Systems*. Cham: Springer International Publishing (Data-Centric Systems and Applications). doi:10.1007/978-3-319-20062-0.

Colace, F. *et al.* (2017) "A Context-Aware Mobile Solution for Assisting Tourists in a Smart Environment," in *Proceedings of the 50th Hawaii*

International Conference on System Sciences. doi:10.24251/HICSS.2017.061.

Colace, F. *et al.* (2018) "MuG: A Multilevel Graph Representation for Big Data Interpretation," *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pp. 1408–1413. doi:10.1109/HPCC/SmartCity/DSS.2018.00233.

Colace, F. *et al.* (2020) "A multilayer approach for recommending contextual learning paths," *Journal of Internet Services and Information Security*, 2(May), pp. 91–102. doi:10.22667/JISIS.2020.05.31.091.

Colace, F. *et al.* (2021) "A Multigraph Approach for Supporting Computer Network Monitoring Systems," in *Advances in Intelligent Systems and Computing*, pp. 470–477. doi:10.1007/978-981-15-5859-7_46.

Colace, F. *et al.* (2022) "An IoT-Based Framework for Smart Agriculture Managing and Product Enhancing," in *Lecture Notes in Networks and Systems*. doi:10.1007/978-981-16-2102-4_10.

Colace, F., Chang, S.-K. and De Santo, M. (2010) "SINMS: A Slow Intelligence Network Manager based on SNMP Protocol.," in *DMS*, pp. 47–52.

Colace, F. and De Santo, M. (2010) "Ontology for E-learning: A Bayesian approach," *IEEE Transactions on Education*, 53(2), pp. 223–233.

Čolaković, A. and Hadžialić, M. (2018) "Internet of Things (IoT): A review of enabling technologies, challenges, and open research issues," *Computer Networks*. doi:10.1016/j.comnet.2018.07.017.

Compton, M. *et al.* (2012) "The SSN ontology of the W3C semantic sensor network incubator group," *Journal of Web Semantics*, 17, pp. 25–32. doi:10.1016/j.websem.2012.05.003.

Cooper, G.F. and Herskovits, E. (1992) "A Bayesian Method for the Induction of Probabilistic Networks from Data," *Machine Learning*, 9(4), pp. 309–347. doi:10.1023/A:1022649401552.

D'Aniello, G. *et al.* (2020) "Knowledge-Based Smart City Service System," *Electronics*, 9(6). doi:10.3390/electronics9060965.

Das, M. and Ghosh, S.K. (2020) "Standard Bayesian network models for spatial time series prediction," in *Studies in Computational Intelligence*. doi:10.1007/978-3-030-27749-9_2.

Davcev, D. *et al.* (2018) "IoT agriculture system based on LoRaWAN," in *2018 14th IEEE International Workshop on Factory Communication Systems (WFCS)*. IEEE, pp. 1–4. doi:10.1109/WFCS.2018.8402368.

Demchenko, Y. *et al.* (2013) "Addressing big data issues in Scientific Data Infrastructure," in *2013 International Conference on Collaboration Technologies and Systems (CTS)*. IEEE, pp. 48–55. doi:10.1109/CTS.2013.6567203.

Dey, A.K. (2001) "Understanding and Using Context," *Personal and Ubiquitous Computing*, 5(1), pp. 4–7. doi:10.1007/s007790170019.

Dominici, M., Pietropaoli, B. and Weis, F. (2012) "Experiences in managing uncertainty and ignorance in a lightly instrumented smart home," *International Journal of Pervasive Computing and Communications*. doi:10.1108/17427371211262635.

Dosovitskiy, A. *et al.* (2017) "CARLA: An Open Urban Driving Simulator," in Levine, S., Vanhoucke, V., and Goldberg, K. (eds) *Proceedings of the 1st Annual Conference on Robot Learning*. PMLR (Proceedings of Machine Learning Research), pp. 1–16. Available at: https://proceedings.mlr.press/v78/dosovitskiy17a.html.

Durrett, R. (2019) *Probability: Theory and Examples*, *Probability: Theory and Examples*. doi:10.1017/9781108591034.

Elijah, O. *et al.* (2018) "An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges," *IEEE Internet of Things Journal*. doi:10.1109/JIOT.2018.2844296.

Endsley, M.R. (1988) "Design and Evaluation for Situation Awareness Enhancement," *Proceedings of the Human Factors Society Annual Meeting*, 32(2). doi:10.1177/154193128803200221.

Endsley, M.R. (1995) "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), pp. 32–64. doi:10.1518/001872095779049543.

Endsley, M.R. (2000) "Theoretical underpinnings of situation awareness: A critical review," in *Situation awareness analysis and measurement*.

Endsley, M.R. and Garland, D.J. (2000) *Situation Awareness Analysis and Measurement*, *Situation Awareness Analysis and Measurement*. doi:10.1201/b12461.

Endsley, M.R. and Robertson, M.M. (2000) "Training for situation awareness in individuals and teams," in *Situation awareness analysis and measurement*.

Eräranta, S. and Staffans, A. (2015) "From situation awareness to smart city planning and decision making," in *CUPUM 2015 - 14th International Conference on Computers in Urban Planning and Urban Management*.

Fernandez, S. *et al.* (2016) "Ontology-based architecture for intelligent transportation systems using a traffic sensor network," *Sensors (Switzerland)*, 16(8). doi:10.3390/s16081287.

Fink, J. and Kobsa, A. (2002) "User modeling for personalized city tours," *Artificial Intelligence Review*. doi:10.1023/A:1016383418977.

Fischer, Y. and Bauer, A. (2010) "Object-oriented sensor data fusion for wide maritime surveillance," in *2010 International Waterside Security Conference, WSS 2010*. doi:10.1109/WSSC.2010.5730244.

Fooladvandi, F. *et al.* (2009) “Signature-based activity detection based on Bayesian networks acquired from expert knowledge,” in *2009 12th International Conference on Information Fusion, FUSION 2009*.
Fowler, D.S. *et al.* (2017) “Towards a Testbed for Automotive Cybersecurity,” in *Proceedings - 10th IEEE International Conference on Software Testing, Verification and Validation, ICST 2017*. doi:10.1109/ICST.2017.62.
Frank, A., Khamis Al Aamri, Y.S. and Zayegh, A. (2019) “IoT based smart traffic density control using image processing,” in *2019 4th MEC International Conference on Big Data and Smart City, ICBDSC 2019*. doi:10.1109/ICBDSC.2019.8645568.
Friedman, N., Geiger, D. and Goldszmidt, M. (1997) “Bayesian Network Classifiers,” *Machine Learning*. doi:10.1002/9780470400531.eorms0099.
Furno, A. and Zimeo, E. (2014) “Context-aware composition of semantic web services,” *Mobile Networks and Applications*, 19(2). doi:10.1007/s11036-014-0494-y.
Gan, G., Lu, Z. and Jiang, J. (2011) “Internet of things security analysis,” in *2011 International Conference on Internet Technology and Applications, iTAP 2011 - Proceedings*. doi:10.1109/ITAP.2011.6006307.
Gandomi, A. and Haider, M. (2015) “Beyond the hype: Big data concepts, methods, and analytics,” *International Journal of Information Management*, 35(2). doi:10.1016/j.ijinfomgt.2014.10.007.
Gavalas, D. and Kenteris, M. (2011) “A web-based pervasive recommendation system for mobile tourist guides,” *Personal and Ubiquitous Computing*, 15(7), pp. 759–770. doi:10.1007/s00779-011-0389-x.
Ghazal, T.M. *et al.* (2021) “IoT for smart cities: Machine learning approaches in smart healthcare—A review,” *Future Internet*. doi:10.3390/fi13080218.
Ghiani, G. *et al.* (2009) “UbiCicero: A location-aware, multi-device museum guide,” *Interacting with Computers*. doi:10.1016/j.intcom.2009.06.001.
Gill, K. *et al.* (2009) “A ZigBee-based home automation system,” *IEEE Transactions on Consumer Electronics*, 55(2). doi:10.1109/TCE.2009.5174403.
Gilson, R.D. (1995) “Situation Awareness - Special Issue Preface,” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1).
Giri, A., Dutta, S. and Neogy, S. (2016) “Enabling agricultural automation to optimize utilization of water, fertilizer and insecticides by implementing Internet of Things (IoT),” in *2016 International Conference on Information Technology (InCITe) - The Next Generation IT Summit on the Theme - Internet of Things: Connect your Worlds*. IEEE, pp. 125–131. doi:10.1109/INCITE.2016.7857603.
Glowacka, J., Krygier, J. and Amanowicz, M. (2015) “A trust-based situation awareness system for military applications of the internet of things,” in

IEEE World Forum on Internet of Things, WF-IoT 2015 - Proceedings. doi:10.1109/WF-IoT.2015.7389103.

Gluck, K.A., Ball, J.T. and Krusmark, M.A. (2012) "Cognitive Control in a Computational Model of the Predator Pilot," in *Integrated Models of Cognitive Systems*. doi:10.1093/acprof:oso/9780195189193.003.0002.

Goap, A. *et al.* (2018) "An IoT based smart irrigation management system using Machine learning and open source technologies," *Computers and Electronics in Agriculture*, 155. doi:10.1016/j.compag.2018.09.040.

Goel, D., Chaudhury, S. and Ghosh, H. (2017) "An IoT approach for context-Aware smart tra€ic management using ontology," in *Proceedings - 2017 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2017*. doi:10.1145/3106426.3106499.

Gondchawar, N. and Kawitkar, R.S. (2016) "IoT based smart agriculture," *International Journal of Advanced Research in Computer and Communication Engineering*.

Gruber, T.R. (1995) "Toward principles for the design of ontologies used for knowledge sharing?," *International Journal of Human-Computer Studies*, 43(5–6), pp. 907–928. doi:10.1006/IJHC.1995.1081.

Gu, T., Pung, H.K. and Zhang, D.Q. (2005) "A service-oriented middleware for building context-aware services," *Journal of Network and Computer Applications*, 28(1). doi:10.1016/j.jnca.2004.06.002.

Guarino, N., Oberle, D. and Staab, S. (2009) "What Is an Ontology?," in *Handbook on Ontologies*. doi:10.1007/978-3-540-92673-3_0.

Gubbi, J. *et al.* (2013) "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, 29(7). doi:10.1016/j.future.2013.01.010.

Gulzar, M., Abbas, G. and Waqas, M. (2020) "Climate Smart Agriculture: A Survey and Taxonomy," in *2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*. IEEE, pp. 1–6. doi:10.1109/ICETST49965.2020.9080695.

Gutzwiller, R.S. and Clegg, B.A. (2013) "The role of working memory in levels of situation awareness," *Journal of Cognitive Engineering and Decision Making*, 7(2). doi:10.1177/1555343412451749.

Guzman, A. (2020) "Ontological Boundaries between Humans and Computers and the Implications for Human-Machine Communication," *Human-Machine Communication*, 1. doi:10.30658/hmc.1.3.

Hachem, S., Teixeira, T. and Issarny, V. (2011) "Ontologies for the Internet of Things," in *Proceedings of the 8th Middleware Doctoral Symposium, MDS'11 of the 12th ACM/IFIP/USENIX International Middleware Conference*. doi:10.1145/2093190.2093193.

Han, C. *et al.* (2013) "A cross-layer communication module for the Internet of Things," *Computer Networks*, 57(3). doi:10.1016/j.comnet.2012.10.003.

Harrison, C. *et al.* (2010) “Foundations for Smarter Cities,” *IBM Journal of Research and Development*, 54(4), pp. 1–16. doi:10.1147/JRD.2010.2048257.

Helsper, E.M. and Gaag, L.C. (2002) “Building Bayesian networks through ontologies,” in *Proceedings of the 15th European Conference on Artificial Intelligence*. IOS Press, pp. 680–684.

Helsper, E.M. and Gaag, L.C. van der (2002) “Building Bayesian networks through ontologies,” *ECAI2002, Proceedings of the 15th European Conference on Artificial Intelligence*, pp. 680–684.

Hightower, J. and Borriello, G. (2001) “Location systems for ubiquitous computing,” *Computer*, 34(8). doi:10.1109/2.940014.

Hooey, B.L. *et al.* (2011) “Modeling Pilot Situation Awareness,” in *Human Modelling in Assisted Transportation*. doi:10.1007/978-88-470-1821-1_22.

Hoppe, T., Kiltz, S. and Dittmann, J. (2011) “Security threats to automotive CAN networksPractical examples and selected short-term countermeasures,” in *Reliability Engineering and System Safety*. doi:10.1016/j.ress.2010.06.026.

Hu, P., Indulska, J. and Robinson, R. (2008) “An autonomic context management system for pervasive computing,” in *6th Annual IEEE International Conference on Pervasive Computing and Communications, PerCom 2008*. doi:10.1109/PERCOM.2008.56.

Ikeda, M. *et al.* (1998) “Task ontology: Ontology for building conceptual problem solving models,” *Proceedings of ECAI98*.

Ilie-Zudor, E. *et al.* (2011) “A survey of applications and requirements of unique identification systems and RFID techniques,” *Computers in Industry*. doi:10.1016/j.compind.2010.10.004.

Inoue, K. (2020) “Application of Bayesian causal inference and structural equation model to animal breeding,” *Animal Science Journal*. doi:10.1111/asj.13359.

Ioannidis, Y. *et al.* (2013) “One object many stories: Introducing ICT in museums and collections through digital storytelling,” in *Proceedings of the DigitalHeritage 2013 - Federating the 19th Int’l VSMM, 10th Eurographics GCH, and 2nd UNESCO Memory of the World Conferences, Plus Special Sessions fromCAA, Arqueologica 2.0 et al.* doi:10.1109/DigitalHeritage.2013.6743772.

Islam, A.S. and Piasecki, M. (2008) “Ontology based web simulation system for hydrodynamic modeling,” *Simulation Modelling Practice and Theory*, 16(7), pp. 754–767. doi:10.1016/j.simpat.2008.04.019.

Janiesch, C., Zschech, P. and Heinrich, K. (2021) “Machine learning and deep learning,” *Electronic Markets*, 31(3), pp. 685–695. doi:10.1007/s12525-021-00475-2.

Javaid, S. *et al.* (2018) "Smart traffic management system using Internet of Things," in *International Conference on Advanced Communication Technology, ICACT*. doi:10.23919/ICACT.2018.8323770.

Ji, B. *et al.* (2020) "Survey on the Internet of Vehicles: Network Architectures and Applications," *IEEE Communications Standards Magazine*, 4(1). doi:10.1109/MCOMSTD.001.1900053.

Jiang, J. and Moallem, M. (2018) "Development of Greenhouse LED System with RedlBlue Mixing Ratio and Daylight Control," in *2018 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, pp. 1197–1202. doi:10.1109/CCTA.2018.8511374.

Jianu, E.M. and Vasilateanu, A. (2017) "Designing of an e-learning system using adaptivity and gamification," in *2017 IEEE International Symposium on Systems Engineering, ISSE 2017 - Proceedings*. doi:10.1109/SysEng.2017.8088270.

Jin, J. *et al.* (2014) "An information framework for creating a smart city through internet of things," *IEEE Internet of Things Journal*. doi:10.1109/JIOT.2013.2296516.

Jing, Y., Pavlović, V. and Rehg, J.M. (2008) "Boosted Bayesian network classifiers," *Machine Learning*, 73(2). doi:10.1007/s10994-008-5065-7.

Jones, D.G. and Endsley, M.R. (2004) "Use of real-time probes for measuring situation awareness," *International Journal of Aviation Psychology*. doi:10.1207/s15327108ijap1404_2.

Jordan, Michael Irwin (1998) *Learning in Graphical Models*. Edited by Michael I. Jordan. Dordrecht: Springer Netherlands. doi:10.1007/978-94-011-5014-9.

Juric, R. and Madland, O. (2020) "Semantic Framework for Creating an Instance of the IoE in Urban Transport: A Study of Traffic Management with Driverless Vehicles," in *Proceedings of the 2020 IEEE International Conference on Human-Machine Systems, ICHMS 2020*. doi:10.1109/ICHMS49158.2020.9209469.

Kabir, S. and Papadopoulos, Y. (2019) "Applications of Bayesian networks and Petri nets in safety, reliability, and risk assessments: A review," *Safety Science*. doi:10.1016/j.ssci.2019.02.009.

Kamilaris, A. and Prenafeta-Boldú, F.X. (2018) "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*. doi:10.1016/j.compag.2018.02.016.

Kanungo, A., Sharma, A. and Singla, C. (2014) "Smart traffic lights switching and traffic density calculation using video processing," in *2014 Recent Advances in Engineering and Computational Sciences, RAECS 2014*. doi:10.1109/RAECS.2014.6799542.

Karin R. SAOUB (2021) *Graph Theory : an introduction to proofs, algorithms, and applications*, *Textbooks in Mathematics*.

Kass, R.E. and Raftery, A.E. (1995) "Bayes factors," *Journal of the American Statistical Association*, 90(430). doi:10.1080/01621459.1995.10476572.

Khan, W.Z. *et al.* (2020) “Industrial internet of things: Recent advances, enabling technologies and open challenges,” *Computers and Electrical Engineering*, 81. doi:10.1016/j.compeleceng.2019.106522.

Klotz, B., Datta, S.K., *et al.* (2018) “A car as a semantic web thing: Motivation and demonstration,” in *2018 Global Internet of Things Summit, GIoTS 2018*. doi:10.1109/GIOTS.2018.8534533.

Klotz, B., Troncy, R., *et al.* (2018) “VSSo: The Vehicle Signal and Attribute Ontology,” in *SSN@ISWC*.

Kodali, R.K. *et al.* (2017) “IoT based smart security and home automation system,” in *Proceeding - IEEE International Conference on Computing, Communication and Automation, ICCCA 2016*. doi:10.1109/CCAA.2016.7813916.

Kodali, R.K., Jain, V. and Karagwal, S. (2016) “IoT based smart greenhouse,” in *2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*. IEEE, pp. 1–6. doi:10.1109/R10-HTC.2016.7906846.

Kolbe, N. *et al.* (2017) “Enriching a situation awareness framework for IoT with knowledge base and reasoning components,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/978-3-319-57837-8_4.

Kolda, T.G. and Bader, B.W. (2009) “Tensor decompositions and applications,” *SIAM Review*. doi:10.1137/07070111X.

Kooij, J.F.P. *et al.* (2014) “Context-based pedestrian path prediction,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/978-3-319-10599-4_40.

Korn, O. (2012) “Industrial playgrounds: How gamification helps to enrich work for elderly or impaired persons in production,” in *EICS’12 - Proceedings of the 2012 ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. doi:10.1145/2305484.2305539.

Koscher, K. *et al.* (2010) “Experimental security analysis of a modern automobile,” in *Proceedings - IEEE Symposium on Security and Privacy*. doi:10.1109/SP.2010.34.

Kousis, A. and Tjortjis, C. (2021) “Data mining algorithms for smart cities: A bibliometric analysis,” *Algorithms*. doi:10.3390/a14080242.

Krytska, Y., Skarga-Bandurova, I. and Velykzhanin, A. (2017) “IoT-based situation awareness support system for real-Time emergency management,” in *Proceedings of the 2017 IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2017*. doi:10.1109/IDAACS.2017.8095228.

Lau, S.P. *et al.* (2015) “A traffic-aware street lighting scheme for Smart Cities using autonomous networked sensors,” *Computers and Electrical Engineering*, 45. doi:10.1016/j.compeleceng.2015.06.011.

Lee, R. *et al.* (2020) “Making IoT data ready for smart city applications,” in *Proceedings - 2020 IEEE International Conference on Big Data and Smart Computing, BigComp 2020*. doi:10.1109/BigComp48618.2020.00020.

Leu, G., Tang, J. and Abbass, H. (2014) “On the role of working memory in trading–off skills and situation awareness in sudoku,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/978-3-319-12643-2_69.

Li, L., Wu, W.G. and Liu, N. (2013) “Ontology model for situation awareness of city tunnel traffic,” in *Applied Mechanics and Materials*. doi:10.4028/www.scientific.net/AMM.347-350.715.

Li, R., Liu, C. and Luo, F. (2008) “A design for automotive CAN bus monitoring system,” in *2008 IEEE Vehicle Power and Propulsion Conference, VPPC 2008*. doi:10.1109/VPPC.2008.4677544.

Liang, Y. *et al.* (2008) “A novel stochastic modeling method for network security situational awareness,” in *2008 3rd IEEE Conference on Industrial Electronics and Applications, ICIEA 2008*. doi:10.1109/ICIEA.2008.4582951.

Lin, C.W. and Sangiovanni-Vincentelli, A. (2012) “Cyber-security for the Controller Area Network (CAN) communication protocol,” in *Proceedings of the 2012 ASE International Conference on Cyber Security, CyberSecurity 2012*. doi:10.1109/CyberSecurity.2012.7.

Liu, L. *et al.* (2016) “Smart homes and home health monitoring technologies for older adults: A systematic review,” *International Journal of Medical Informatics*. doi:10.1016/j.ijmedinf.2016.04.007.

Liu, S., Wanyan, X. and Zhuang, D. (2014) “Modeling the situation awareness by the analysis of cognitive process,” in *Bio-Medical Materials and Engineering*. doi:10.3233/BME-141044.

Lombardi, M., Pascale, F. and Santaniello, D. (2021) “Two-Step Algorithm to Detect Cyber-Attack Over the Can-Bus: A Preliminary Case Study in Connected Vehicles,” *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg*, 8(3). doi:10.1115/1.4052823.

Long, J. (2018) “Automobile Electronic Control Network Design Based on CAN Bus,” in *Proceedings - 3rd International Conference on Intelligent Transportation, Big Data and Smart City, ICITBS 2018*. doi:10.1109/ICITBS.2018.00010.

Lu, Y. (2017) “Industry 4.0: A survey on technologies, applications and open research issues,” *Journal of Industrial Information Integration*, 6, pp. 1–10. doi:10.1016/j.jii.2017.04.005.

Lu, Z., Coster, X. and de Winter, J. (2017) “How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving,” *Applied Ergonomics*, 60. doi:10.1016/j.apergo.2016.12.003.

Lunn, D. *et al.* (2009) “The BUGS project: Evolution, critique and future directions,” *Statistics in Medicine*, 28(25), pp. 3049–3067. doi:10.1002/sim.3680.
Majeed, A. and Rauf, I. (2020) “Graph theory: A comprehensive survey about graph theory applications in computer science and social networks,” *Inventions*. doi:10.3390/inventions5010010.
Majid Butt, M. *et al.* (2020) “Agent-Based Modeling for Distributed Decision Support in an IoT Network,” *IEEE Internet of Things Journal*, 7(8). doi:10.1109/JIOT.2020.2976802.
Marikyan, D., Papagiannidis, S. and Alamanos, E. (2019) “A systematic review of the smart home literature: A user perspective,” *Technological Forecasting and Social Change*, 138. doi:10.1016/j.techfore.2018.08.015.
Mattern, F. and Floerkemeier, C. (2010) “From the internet of computers to the internet of things,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/978-3-642-17226-7_15.
McGovern, A. *et al.* (2019) “Making the black box more transparent: Understanding the physical implications of machine learning,” *Bulletin of the American Meteorological Society*, 100(11). doi:10.1175/BAMS-D-18-0195.1.
Medjahed, B. and Atif, Y. (2007) “Context-based matching for Web service composition,” *Distributed and Parallel Databases*, 21(1). doi:10.1007/s10619-006-7003-7.
Mekala, M.S. and Viswanathan, P. (2017) “A Survey: Smart agriculture IoT with cloud computing,” in *2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS)*. IEEE, pp. 1–7. doi:10.1109/ICMDCS.2017.8211551.
Meneghello, F. *et al.* (2019) “IoT: Internet of Threats? A Survey of Practical Security Vulnerabilities in Real IoT Devices,” *IEEE Internet of Things Journal*, 6(5). doi:10.1109/JIOT.2019.2935189.
Minoli, D., Sohraby, K. and Occhiogrosso, B. (2017) “IoT Considerations, Requirements, and Architectures for Smart Buildings-Energy Optimization and Next-Generation Building Management Systems,” *IEEE Internet of Things Journal*, 4(1). doi:10.1109/JIOT.2017.2647881.
Misbahuddin, S. *et al.* (2016) “IoT based dynamic road traffic management for smart cities,” in *2015 12th International Conference on High-Capacity Optical Networks and Enabling/Emerging Technologies, HONET-ICT 2015*. doi:10.1109/HONET.2015.7395434.
Mitchell, T.M. (1980) “The Need for Biases in Learning Generalizations,” *Readings in Machine Learning*, (CBM-TR-117).
Mocrii, D., Chen, Y. and Musilek, P. (2018) “IoT-based smart homes: A review of system architecture, software, communications, privacy and security,” *Internet of Things (Netherlands)*. doi:10.1016/j.iot.2018.08.009.

Mohammad, M.A. *et al.* (2015) "Ontology-based framework for risk assessment in road scenes using videos," in *Procedia Computer Science*. doi:10.1016/j.procs.2015.08.300.

Morales, J. and Garcia, M. (2015) "GeoSmart cities: Event-driven geoprocessing as enabler of smart cities," in *2015 IEEE 1st International Smart Cities Conference, ISC2 2015*. doi:10.1109/ISC2.2015.7366207.

Morschheuser, B. *et al.* (2017) "How games induce cooperation? A study on the relationship between game features and we-intentions in an augmented reality game," *Computers in Human Behavior*. doi:10.1016/j.chb.2017.08.026.

Mozzaquatro, B.A., Jardim-Goncalves, R. and Agostinho, C. (2018) "Situation awareness in the Internet of Things," in *2017 International Conference on Engineering, Technology and Innovation: Engineering, Technology and Innovation Management Beyond 2020: New Challenges, New Approaches, ICE/ITMC 2017 - Proceedings*. doi:10.1109/ICE.2017.8279988.

Muangprathub, J. *et al.* (2019) "IoT and agriculture data analysis for smart farm," *Computers and Electronics in Agriculture*, 156, pp. 467–474. doi:10.1016/j.compag.2018.12.011.

Munir, K. and Sheraz Anjum, M. (2018) "The use of ontologies for effective knowledge modelling and information retrieval," *Applied Computing and Informatics*. doi:10.1016/j.aci.2017.07.003.

Murphy, K. (2002) *Dynamic Bayesian Networks: Representation, Inference and Learning*, *Chemistry & …*.

Nawir, M. *et al.* (2017) "Internet of Things (IoT): Taxonomy of security attacks," in *2016 3rd International Conference on Electronic Design, ICED 2016*. doi:10.1109/ICED.2016.7804660.

Nazarenko, A.A. and Camarinha-Matos, L.M. (2020) "The Role of Digital Twins in Collaborative Cyber-Physical Systems," in *IFIP Advances in Information and Communication Technology*. doi:10.1007/978-3-030-45124-0_18.

Neches, R. *et al.* (1991) "Enabling Technology for Knowledge Sharing," *AI Mag.*, 12(3), pp. 36–56. doi:10.1609/aimag.v12i3.902.

Ngu, A.H. *et al.* (2017) "IoT Middleware: A Survey on Issues and Enabling Technologies," *IEEE Internet of Things Journal*, 4(1). doi:10.1109/JIOT.2016.2615180.

Nguyen, T. *et al.* (2019) "A Review of Situation Awareness Assessment Approaches in Aviation Environments," *IEEE Systems Journal*, 13(3). doi:10.1109/JSYST.2019.2918283.

Nibbelink, C.W. and Brewer, B.B. (2018) "Decision-making in nursing practice: An integrative literature review," *Journal of Clinical Nursing*. doi:10.1111/jocn.14151.

Ning, Y. *et al.* (2019) "Review of deep learning based speech synthesis," *Applied Sciences (Switzerland)*. doi:10.3390/app9194050.

Nzabahimana, J.P. (2018) “Analysis of security and privacy challenges in Internet of Things,” in *Proceedings of 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies, DESSERT 2018*. doi:10.1109/DESSERT.2018.8409122.

Okamura, T. *et al.* (2020) “Inferring phenotypic causal structure among farrowing and weaning traits in pigs,” *Animal Science Journal*, 91(1). doi:10.1111/asj.13369.

Onishi, H. (2012) “Paradigm change of vehicle cyber security,” in *19th Intelligent Transport Systems World Congress, ITS 2012*.

Orsi, G., Tanca, L. and Zimeo, E. (2011) “Keyword-based, context-aware selection of natural language query patterns,” in *ACM International Conference Proceeding Series*. doi:10.1145/1951365.1951390.

Oztemel, E. and Gursev, S. (2020) “Literature review of Industry 4.0 and related technologies,” *Journal of Intelligent Manufacturing*. doi:10.1007/s10845-018-1433-8.

Pan, R. *et al.* (2005) “A bayesian network approach to ontology mapping,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/11574620_41.

Panigati, E. *et al.* (2012) “Aspects of pervasive information management: An account of the green move system,” in *Proceedings - 15th IEEE International Conference on Computational Science and Engineering, CSE 2012 and 10th IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, EUC 2012*. doi:10.1109/ICCSE.2012.93.

Pantic, M. *et al.* (2007) “Human computing and machine understanding of human behavior: A survey,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/978-3-540-72348-6_3.

Papagiannakopoulou, E.I. *et al.* (2013) “Leveraging Semantic Web Technologies for Access Control,” *Emerging Trends in ICT Security*, pp. 493–506. doi:10.1016/B978-0-12-411474-6.00030-X.

Parent, C. *et al.* (eds) (2007) *Conceptual Modeling - ER 2007*. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science). doi:10.1007/978-3-540-75563-0.

Pascoe, J., Ryan, N. and Morse, D. (1999) “Issues in developing context-aware computing,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/3-540-48157-5_20.

Patel, P., Narmawala, Z. and Thakkar, A. (2019) “A Survey on Intelligent Transportation System Using Internet of Things,” in *Advances in Intelligent Systems and Computing*. doi:10.1007/978-981-13-5953-8_20.

de Paz, J.F. *et al.* (2016) “Intelligent system for lighting control in smart cities,” *Information Sciences*, 372. doi:10.1016/j.ins.2016.08.045.

Pearl, J. (1988) “BOOK - Pearl - Probabilistic reasoning in intelligent systems: : networks of plausible inference (Morgan kaufmann series in representation and reasoning),” *Morgan Kaufmann Publishers, San Mateo, Calif.*

Peng, T. and Jinqi, P. (2017) “A Recommendation System for Collaborative Visualization Platforms,” in *Proceedings - 2017 5th International Conference on Enterprise Systems: Industrial Digitalization by Enterprise Systems, ES 2017*. doi:10.1109/ES.2017.17.

Perera, C. *et al.* (2014) “Context aware computing for the internet of things: A survey,” *IEEE Communications Surveys and Tutorials*, 16(1). doi:10.1109/SURV.2013.042313.00197.

Pérez-Soler, S., Guerra, E. and de Lara, J. (2018) “Collaborative Modeling and Group Decision Making Using Chatbots in Social Networks,” *IEEE Software*. doi:10.1109/MS.2018.290101511.

Philip Chen, C.L. and Zhang, C.Y. (2014) “Data-intensive applications, challenges, techniques and technologies: A survey on Big Data,” *Information Sciences*, 275. doi:10.1016/j.ins.2014.01.015.

Pombo, L. *et al.* (2020) “Evaluation of a mobile augmented reality game application as an outdoor learning tool,” *International Journal of Mobile and Blended Learning*. doi:10.4018/IJMBL.2019100105.

Ponce, V. and Abdulrazak, B. (2021) “Intention as a Context: An Activity Intention Model for Adaptable Development of Applications in the Internet of Things,” *IEEE Access*, 9, pp. 151167–151185. doi:10.1109/ACCESS.2021.3126000.

Prathibha, S.R., Hongal, A. and Jyothi, M.P. (2017) “IOT Based Monitoring System in Smart Agriculture,” in *2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT)*. IEEE, pp. 81–84. doi:10.1109/ICRAECT.2017.52.

Rabiner, L.R. (1989) “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition,” *Proceedings of the IEEE*, 77(2). doi:10.1109/5.18626.

Radoglou-Grammatikis, P. *et al.* (2020) “A compilation of UAV applications for precision agriculture,” *Computer Networks*, 172. doi:10.1016/j.comnet.2020.107148.

Raento, M. *et al.* (2005) “ContextPhone: A prototyping platform for context-aware mobile applications,” *IEEE Pervasive Computing*. doi:10.1109/MPRV.2005.29.

Ranganathan, A. and Campbell, R.H. (2003) “An infrastructure for context-awareness based on first order logic,” *Personal and Ubiquitous Computing*, 7(6). doi:10.1007/s00779-003-0251-x.

Raskin, R.G. and Pan, M.J. (2005) “Knowledge representation in the semantic web for Earth and environmental terminology (SWEET),” *Computers & Geosciences*, 31(9), pp. 1119–1125. doi:10.1016/j.cageo.2004.12.004.

van Ratingen, M.R. (2017) "The Euro NCAP Safety Rating," in. doi:10.1007/978-3-658-18107-9_2.

Rauseo, A., Martinenghi, D. and Tanca, L. (2013) "Contextual Data Tailoring Using ASP," in, pp. 99–117. doi:10.1007/978-3-642-36008-4_5.

Reilly, J. *et al.* (2015) *On cybersecurity of freeway control systems: Analysis of coordinated ramp metering attacks*. Washington DC, United States.

Reka, S.S. and Dragicevic, T. (2018) "Future effectual role of energy delivery: A comprehensive review of Internet of Things and smart grid," *Renewable and Sustainable Energy Reviews*. doi:10.1016/j.rser.2018.03.089.

Rhayem, A., Mhiri, M.B.A. and Gargouri, F. (2020) "Semantic Web Technologies for the Internet of Things: Systematic Literature Review," *Internet of Things (Netherlands)*, 11. doi:10.1016/J.IOT.2020.100206.

Risteska Stojkoska, B.L. and Trivodaliev, K. v. (2017) "A review of Internet of Things for smart home: Challenges and solutions," *Journal of Cleaner Production*. doi:10.1016/j.jclepro.2016.10.006.

Riveiro, M., Falkman, G. and Ziemke, T. (2008) "Improving maritime anomaly detection and situation awareness through interactive visualization," in *Proceedings of the 11th International Conference on Information Fusion, FUSION 2008*.

Riveiro, M., Pallotta, G. and Vespe, M. (2018) "Maritime anomaly detection: A review," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. doi:10.1002/widm.1266.

Rossano, V. and Roselli, T. (2018) "Game-Based Learning as Effective Learning Method: An Application of Digital Storytelling," in *2018 22nd International Conference Information Visualisation (IV)*. IEEE, pp. 542–546. doi:10.1109/iV.2018.00100.

Rossiter, M. and Garcia, P.A. (2010) "Digital storytelling: A new player on the narrative field," *New Directions for Adult and Continuing Education*. doi:10.1002/ace.370.

Roy, J. and Davenport, M. (2010) "Exploitation of maritime domain ontologies for anomaly detection and threat analysis," in *2010 International Waterside Security Conference, WSS 2010*. doi:10.1109/WSSC.2010.5730278.

Rudin, C. (2019) "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence*. doi:10.1038/s42256-019-0048-x.

Rudovic, O., Pavlovic, V. and Pantic, M. (2015) "Context-sensitive dynamic ordinal regression for intensity estimation of facial action units," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(5). doi:10.1109/TPAMI.2014.2356192.

Saidin, N.F., Halim, N.D.A. and Yahaya, N. (2015) "A review of research on augmented reality in education: Advantages and applications," *International Education Studies*. doi:10.5539/ies.v8n13p1.

Saleem, Y. *et al.* (2019) "Internet of Things-Aided Smart Grid: Technologies, Architectures, Applications, Prototypes, and Future Research Directions," *IEEE Access*, 7. doi:10.1109/ACCESS.2019.2913984.

Sanina, A., Kutergina, E. and Balashov, A. (2020) "The Co-Creative approach to digital simulation games in social science education," *Computers and Education*. doi:10.1016/j.compedu.2020.103813.

Sarker, I.H., Kayes, A.S.M. and Watters, P. (2019) "Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage," *Journal of Big Data*, 6(1). doi:10.1186/s40537-019-0219-y.

Satyanarayanan, M. (2001) "Pervasive computing: Vision and challenges," *IEEE Personal Communications*. doi:10.1109/98.943998.

Schapire, R.E. (1999) "A brief introduction to boosting," in *IJCAI International Joint Conference on Artificial Intelligence*.

Scheuer, S., Haase, D. and Meyer, V. (2013) "Towards a flood risk assessment ontology – Knowledge integration into a multi-criteria risk assessment approach," *Computers, Environment and Urban Systems*, 37, pp. 82–94. doi:10.1016/j.compenvurbsys.2012.07.007.

Schilit, B., Adams, N. and Want, R. (1995) "Context-aware computing applications," in *Mobile Computing Systems and Applications - Workshop Proceedings*. doi:10.1109/wmcsa.1994.16.

Schmidt, A. *et al.* (1999) "Advanced Interaction in Context," in, pp. 89–101. doi:10.1007/3-540-48157-5_10.

Scholten, H. *et al.* (2007) "A methodology to support multidisciplinary model-based water management," *Environmental Modelling & Software*, 22(5), pp. 743–759. doi:10.1016/j.envsoft.2005.12.025.

Scutari, M. (2010) "Learning Bayesian networks with the bnlearn R Package," *Journal of Statistical Software*, 35(3). doi:10.18637/jss.v035.i03.

Scutari, M., Graafland, C.E. and Gutiérrez, J.M. (2019) "Who learns better Bayesian network structures: Accuracy and speed of structure learning algorithms," *International Journal of Approximate Reasoning*, 115. doi:10.1016/j.ijar.2019.10.003.

Sezer, O.B., Dogdu, E. and Ozbayoglu, A.M. (2018) "Context-Aware Computing, Learning, and Big Data in Internet of Things: A Survey," *IEEE Internet of Things Journal*. doi:10.1109/JIOT.2017.2773600.

Sharma, C. and Gondhi, N.K. (2018) "Communication Protocol Stack for Constrained IoT Systems," in *Proceedings - 2018 3rd International Conference On Internet of Things: Smart Innovation and Usages, IoT-SIU 2018*. doi:10.1109/IoT-SIU.2018.8519904.

Shuang, L., Xiaoru, W. and Damin, Z. (2014) "A Quantitative Situational Awareness Model of Pilot," *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 3(1). doi:10.1177/2327857914031019.

Shukla, R.M. and Munir, A. (2017) “An efficient computation offloading architecture for the Internet of Things (IoT) devices,” in *2017 14th IEEE Annual Consumer Communications and Networking Conference, CCNC 2017*. doi:10.1109/CCNC.2017.7983224.

Shum, H. yeung, He, X. dong and Li, D. (2018) “From Eliza to XiaoIce: challenges and opportunities with social chatbots,” *Frontiers of Information Technology and Electronic Engineering*. doi:10.1631/FITEE.1700826.

Sishodia, R.P., Ray, R.L. and Singh, S.K. (2020) “Applications of Remote Sensing in Precision Agriculture: A Review,” *Remote Sensing*, 12(19), p. 3136. doi:10.3390/rs12193136.

Situation Awareness Analysis and Measurement (2000) *Situation Awareness Analysis and Measurement*. doi:10.1201/b12461.

Sivarajah, U. *et al.* (2017) “Critical analysis of Big Data challenges and analytical methods,” *Journal of Business Research*, 70. doi:10.1016/j.jbusres.2016.08.001.

Smeda, N., Dakich, E. and Sharda, N. (2010) “Developing a framework for advancing e-learning through digital storytelling,” in *Proceedings of the IADIS International Conference e-Learning 2010, Part of the IADIS Multi Conference on Computer Science and Information Systems 2010, MCCSIS 2010*.

Smeda, N., Dakich, E. and Sharda, N. (2013) “The effectiveness of digital storytelling in the classrooms: A case study,” in *Proceedings - 2013 IEEE 13th International Conference on Advanced Learning Technologies, ICALT 2013*. doi:10.1109/ICALT.2013.157.

Smith, B. and Welty, C. (2001) “Ontology: Towards a New Synthesis,” in *Formal Ontology in Information Systems*. New York: ACM Press.

Smith, K. and Hancock, P.A. (1995) “Situation awareness is adaptive, externally directed consciousness,” *Human Factors*. doi:10.1518/001872095779049444.

Song, H.M., Kim, H.R. and Kim, H.K. (2016) “Intrusion detection system based on the analysis of time intervals of CAN messages for in-vehicle network,” in *International Conference on Information Networking*. doi:10.1109/ICOIN.2016.7427089.

Sotres, P. *et al.* (2019) “Breaking vendors and city locks through a semantic-enabled global interoperable internet-of-things system: A smart parking case,” *Sensors (Switzerland)*, 19(2). doi:10.3390/s19020229.

Stafford, J. V. (2000) “Implementing Precision Agriculture in the 21st Century,” *Journal of Agricultural Engineering Research*, 76(3), pp. 267–275. doi:10.1006/jaer.2000.0577.

Stanton, N.A., Chambers, P.R.G. and Piggott, J. (2001) “Situational awareness and safety,” *Safety Science*, 39(3). doi:10.1016/S0925-7535(01)00010-8.

Stenneth, L. *et al.* (2011) “Transportation mode detection using mobile phones and GIS information,” in *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*. doi:10.1145/2093973.2093982.

Sun, Y. *et al.* (2016) “Internet of Things and Big Data Analytics for Smart and Connected Communities,” *IEEE Access*, 4. doi:10.1109/ACCESS.2016.2529723.

Swartout, B. *et al.* (1996) “Toward Distributed Use of Large-Scale Ontologies,” *Proc. of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems*, pp. 138–148.

Syzdykbayev, M., Hajari, H. and Karimi, H.A. (2019) “An Ontology for Collaborative Navigation among Autonomous Cars, Drivers, and Pedestrians in Smart Cities,” in *2019 4th International Conference on Smart and Sustainable Technologies, SpliTech 2019*. doi:10.23919/SpliTech.2019.8783045.

Taivalsaari, A. and Mikkonen, T. (2018) “A Taxonomy of IoT Client Architectures,” *IEEE Software*, 35(3). doi:10.1109/MS.2018.2141019.

Tang, F. *et al.* (2018) “An Intelligent Traffic Load Prediction-Based Adaptive Channel Assignment Algorithm in SDN-IoT: A Deep Learning Approach,” *IEEE Internet of Things Journal*, 5(6). doi:10.1109/JIOT.2018.2838574.

Thangavel, D. *et al.* (2014) “Performance evaluation of MQTT and CoAP via a common middleware,” in *IEEE ISSNIP 2014 - 2014 IEEE 9th International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Conference Proceedings*. doi:10.1109/ISSNIP.2014.6827678.

Thomas, M. *et al.* (2015) “Can we talk about Speed? The Effect of Verbal Protocols on Driver Speed and Perceived Workload,” *Procedia Manufacturing*, 3. doi:10.1016/j.promfg.2015.07.608.

Tole, A.A. (2013) “Big data challenges,” *Database systems journal*, 4(3), pp. 31–40.

Torralba, A. *et al.* (2003) “Context-based vision system for place and object recognition,” in *Proceedings of the IEEE International Conference on Computer Vision*. doi:10.1109/iccv.2003.1238354.

Trifonova, N. *et al.* (2017) “Predicting ecosystem responses to changes in fisheries catch, temperature, and primary productivity with a dynamic Bayesian network model,” *ICES Journal of Marine Science*. doi:10.1093/icesjms/fsw231.

Troncy, R. *et al.* (2017) “3cixty: Building comprehensive knowledge bases for city exploration,” *Journal of Web Semantics*, 46–47. doi:10.1016/j.websem.2017.07.002.

Tu, E. *et al.* (2018) “Exploiting AIS Data for Intelligent Maritime Navigation: A Comprehensive Survey from Data to Methodology,” *IEEE Transactions* *on Intelligent Transportation Systems*, 19(5). doi:10.1109/TITS.2017.2724551.

Tucker, A. *et al.* (2005) "A spatio-temporal Bayesian network classifier for understanding visual field deterioration," *Artificial Intelligence in Medicine*, 34(2). doi:10.1016/j.artmed.2004.07.004.

Ud Din, I. *et al.* (2019) "The Internet of Things: A Review of Enabled Technologies and Future Challenges," *IEEE Access*. doi:10.1109/ACCESS.2018.2886601.

Underwood, G. (2007) "Visual attention and the transition from novice to advanced driver," *Ergonomics*, 50(8). doi:10.1080/00140130701318707.

Ungerer, L. (2019) "Digital Storytelling: Possible Applications in an Open Distance E-learning Environment," in *Storytelling: Global Reflections on Narrative*. Brill | Rodopi. doi:10.1163/9789004396401_012.

Valdiviezo-Diaz, P. *et al.* (2019) "A Collaborative Filtering Approach Based on Naïve Bayes Classifier," *IEEE Access*, 7. doi:10.1109/ACCESS.2019.2933048.

Verbert, K. *et al.* (2012) "Context-aware recommender systems for learning: A survey and future challenges," *IEEE Transactions on Learning Technologies*. doi:10.1109/TLT.2012.11.

Viaene, S., Derrig, R.A. and Dedene, G. (2004) "A case study of applying boosting naive bayes to claim fraud diagnosis," *IEEE Transactions on Knowledge and Data Engineering*, 16(5). doi:10.1109/TKDE.2004.1277822.

Villegas, N.M. *et al.* (2018) "Characterizing context-aware recommender systems: A systematic literature review," *Knowledge-Based Systems*, 140. doi:10.1016/j.knosys.2017.11.003.

Vu, K.P.L. *et al.* (2010) "Pilot and controller workload and situation awareness with three traffic management concepts," in *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*. doi:10.1109/DASC.2010.5655346.

Wang, W. *et al.* (2019) "A Survey on Consensus Mechanisms and Mining Strategy Management in Blockchain Networks," *IEEE Access*, 7. doi:10.1109/ACCESS.2019.2896108.

Wang, X.H. *et al.* (2004) "Ontology based context modeling and reasoning using OWL," in *Proceedings - Second IEEE Annual Conference on Pervasive Computing and Communications, Workshops, PerCom*. doi:10.1109/percomw.2004.1276898.

Want, R. *et al.* (1992) "The Active Badge Location System," *ACM Transactions on Information Systems (TOIS)*, 10(1). doi:10.1145/128756.128759.

Weber, P. *et al.* (2012) "Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas," *Engineering Applications of Artificial Intelligence*. doi:10.1016/j.engappai.2010.06.002.

Weiser, M. (1991) "The Computer for the 21st Century," *Scientific American*, 265(3). doi:10.1038/scientificamerican0991-94.

Weng, J.F., Kuo, H.L. and Tseng, S.S. (2011) "Interactive storytelling for elementary school Nature Science education," in *Proceedings of the 2011 11th IEEE International Conference on Advanced Learning Technologies, ICALT 2011*. doi:10.1109/ICALT.2011.104.

Weyer, S. *et al.* (2016) "Future Modeling and Simulation of CPS-based Factories: an Example from the Automotive Industry," in *IFAC-PapersOnLine*. doi:10.1016/j.ifacol.2016.12.168.

Whitman, M.E. and Mattord, H.J. (2011) *Principles of Information Security*, *Cengage Learning*.

Wimisberg, J. (2007) "Situation Awareness in Dynamic Decision Making: Effects of Practice and Working Memory," *Journal of Cognitive Engineering and Decision Making*, 1(1). doi:10.1177/155534340700100103.

de Wrachien, D. *et al.* (2012) "Ontology for flood management: A proposal," *WIT Transactions on Ecology and the Environment*, 159, pp. 3–13. doi:10.2495/FRIAR120011.

Wu, X. *et al.* (2016) "Pilot situational awareness modeling for cockpit interface evaluation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/978-3-319-40030-3_46.

Xin, Y. *et al.* (2018) "Machine Learning and Deep Learning Methods for Cybersecurity," *IEEE Access*, 6. doi:10.1109/ACCESS.2018.2836950.

Xu, G. *et al.* (2017) "Network Security Situation Awareness Based on Semantic Ontology and User-Defined Rules for Internet of Things," *IEEE Access*, 5. doi:10.1109/ACCESS.2017.2734681.

Xu, L. da (2011) "Enterprise systems: State-of-the-art and future trends," *IEEE Transactions on Industrial Informatics*, 7(4). doi:10.1109/TII.2011.2167156.

Xu, L. da, He, W. and Li, S. (2014) "Internet of things in industries: A survey," *IEEE Transactions on Industrial Informatics*. doi:10.1109/TII.2014.2300753.

Xu, T. *et al.* (2019) "Voltage Based Authentication for Controller Area Networks with Reinforcement Learning," in *IEEE International Conference on Communications*. doi:10.1109/ICC.2019.8761744.

Yakavenka, V. (2018) "Blockchain impact on food supply chains," *MIBES Transactions*.

Yang, Y.-T.C., Chen, Y.-C. and Hung, H.-T. (2020) "Digital storytelling as an interdisciplinary project to improve students' English speaking and creative thinking," *Computer Assisted Language Learning*, pp. 1–23. doi:10.1080/09588221.2020.1750431.

Yassine, A. *et al.* (2019) "IoT big data analytics for smart homes with fog and cloud computing," *Future Generation Computer Systems*, 91. doi:10.1016/j.future.2018.08.040.
Yin, Y., Stecke, K.E. and Li, D. (2018) "The evolution of production systems from Industry 2.0 through Industry 4.0," *International Journal of Production Research*, 56(1–2). doi:10.1080/00207543.2017.1403664.
Yoon, J.H. *et al.* (2015) "Bayesian multi-object tracking using motion context from multiple objects," in *Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015*. doi:10.1109/WACV.2015.12.
Zalman, R. and Mayer, A. (2014) "A secure but still safe and low cost automotive communication technique," in *Proceedings - Design Automation Conference*. doi:10.1145/2593069.2603850.
Zanella, A. *et al.* (2014) "Internet of things for smart cities," *IEEE Internet of Things Journal*, 1(1), pp. 22–32. doi:10.1109/JIOT.2014.2306328.
Zarpelão, B.B. *et al.* (2017) "A survey of intrusion detection in Internet of Things," *Journal of Network and Computer Applications*. doi:10.1016/j.jnca.2017.02.009.
Zhai, Z. *et al.* (2020) "Decision support systems for agriculture 4.0: Survey and challenges," *Computers and Electronics in Agriculture*. doi:10.1016/j.compag.2020.105256.
Zhang, P. *et al.* (2020) "Situational Awareness Model of IoV Based on Fuzzy Evaluation and Markov Chain," in *Communications in Computer and Information Science*. doi:10.1007/978-3-030-63119-2_44.
Zhang, W.E. *et al.* (2020) "The 10 Research Topics in the Internet of Things," in *Proceedings - 2020 IEEE 6th International Conference on Collaboration and Internet Computing, CIC 2020*. doi:10.1109/CIC50333.2020.00015.
Zheng, Z. *et al.* (2017) "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends," in *Proceedings - 2017 IEEE 6th International Congress on Big Data, BigData Congress 2017*. doi:10.1109/BigDataCongress.2017.85.
Zhong, C. *et al.* (2014) "Inferring building functions from a probabilistic model using public transportation data," *Computers, Environment and Urban Systems*, 48. doi:10.1016/j.compenvurbsys.2014.07.004.