

An overview of big data analysis

Fabio Arena, Giovanni Pau

Faculty of Engineering and Architecture, Kore University of Enna, Italy

Article Info

Article history:

Received Feb 4, 2020

Revised Mar 23, 2020

Accepted Apr 15, 2020

Keywords:

Big data

Big data analysis

Internet of Things

ABSTRACT

Big data represents one of the most profound and most pervasive evolutions in the digital world. Examples of big data come from Internet of Things (IoT) devices, as well as smart cars, but also the use of social networks, industries, and so on. The sources of data are numerous and continuously increasing, and, therefore, what characterizes big data is not only the volume but also the complexity due to the heterogeneity of information that can be obtained. The fastest growth in spending on big data technologies is happening within banking, healthcare, insurance, securities and investment services, and telecommunications. Remarkably, three of those industries lie within the financial sector, which has many particularly serviceable use cases for big data analytics, such as fraud detection, risk management, and customer service optimization. In fact, the definition of big data analysis refers to the process that encompasses the gathering and analysis of big data to obtain useful information for the business. This paper focuses on delivering a short review concerning the current technologies, future perspectives, and the evaluation of some use cases associated with the analysis of big data.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Giovanni Pau,
Faculty of Engineering and Architecture,
Kore University of Enna,
Cittadella Universitaria-94100 Enna, Italy.
Email: giovanni.pau@unikore.it

1. INTRODUCTION

Over the past 20 years, new digital technologies have developed to change our way of life radically [1-3]. Ever thinner smartphones, regularly smaller and more powerful processors, appliances that communicate with each other, social media networks that know everything about us and on which we depend, have been developed. The industry has made significant leaps to keep pace in an increasingly demanding market [4-6]. At the same time, companies frequently have to deal with limited resources and constraints imposed by various national and international institutions. For instance, in Europe, there is the purpose of reducing European industry consumption by 20% by 2020 [7]. As an outcome, the need to optimize production processes through investments in the technological development of the factories is becoming increasingly pressing [8]. The Internet of Things (IoT) is a network of physical objects that can be reached through the Internet [9, 10]. These objects are daily things that contain integrated technology able to interact with the external environment. IoT necessitates the development of many technologies currently designed by the leading companies and leads to a list of ever-growing advantages [11, 12]. IoT applications cover many features, from health [13-15] to home automation [16-19], from transport [20-24] to security [25-29]. For instance, by spreading a set of sensors and processors, it is possible to control the windows, the temperature of the house, the lights, and other home appliances wirelessly [30, 31]. Another application case is represented by the cities, which can be developed more intelligent and efficient thanks to the interconnection of smart objects. For instance, thanks to the IoT, the traffic lights can be connected to a camera circuit, disseminated

everywhere in the city, to identify the traffic level and mass movements, then avoiding possible inconveniences preventively [32-34]. As it is possible to perceive, the IoT opens up a world of limitless possibilities, more extensive than those offered a few years ago from the digital age. The creation of the IoT would have been much more complicated if a big data structure had not been followed since the latter allows analyzing vast amounts of data [35-37]. The multifaceted nature of big data is mainly due to the unstructured design of most of the data produced by modern technologies, such as web registers, radio frequency identification (RFID), sensors embedded in devices, machinery, vehicles, Internet searches, social networks, laptops, smartphones, GPS devices, and call center records [38-40].

Several studies have revealed that the use of big data represents one of the qualifying marks of the metamorphosis that is taking place in production processes [41-44]. Among these studies, the most notable are those carried by McKinsey & Company [45], by Boston Consulting, and the Milan Polytechnic Observatory [46]. For instance, from their investigations, it can be perceived that the fourth industrial revolution concentrates on the enactment of fascinating essential technologies labeled as enablers. Furthermore, big data analysis represents those techniques for the management of quite large measures of data, accomplished by open methods that yield forecasts or prognostications. The current competitive business environment is forcing companies to process high-speed data and integrate valuable information into production processes. As stated by Forrester Research [47], through interviews with experts in the field, the use cases of big data analysis mainly fall into three groups:

- Efficiency and management of operational risks; the study of big data practiced to risk minimization in financial processes, asset management, human resources, and the supply chain
- Security and application performance through predictive examination and big data analysis
- Employed in information technologies monitoring; they serve to prevent problems in affording services and managing events in real-time. The analysis models use the datalog produced by servers and network devices to evaluate performance levels, find bottlenecks, and other features
- Knowledge and services to customers: this category includes all the solutions and applications for big data analysis used for marketing and sales projects, for product development, but also for optimizing the digital experience

This paper focuses on delivering a short review concerning the current technologies, future perspectives, and the evaluation of some use cases associated with the analysis of Big Data. Section 2 presents a summary of several case studies that have been published in this research field, while section 3 concludes the paper.

2. CASE STUDIES

The first study taken into account treats the use of big data analysis aimed at improving customer loyalty programs [48]. In the era of cloud computing, cyber-physical systems, the Internet of Things and big data analysis, every user click on search engines, social networks, and e-commerce portals produce data, which represent for the companies an added value [49]. For this reason, the adoption of big data analysis in retention strategies is worthwhile not only for the attainment of new customers but also to enhance the profits generated by pre-existing clients, as it allows the industries to adapt their productive strategies directly to the needs of consumers, thus seeking to maximize revenues. The approach introduced in [48] is based on the use of software as a service (SaaS) to gather data from an e-commerce platform and analyze them to provide the users of the portal a list of recommended products. This solution, shown in Figure 1, is made up of an e-commerce platform, a dashboard accessible to the company's marketing operators, an Event Bus, a NoSQL database, a Streaming Manager, a Machine Learning Cluster, and an Email Marketing Microservice.

The case study is composed of a multinational corporation. Buyers perform purchases through the company's e-commerce platform, producing data on their inclinations. Subsequently, a marketing character delivers a pre-analysis of these data to create a pre-selection of candidates for the loyalty program based on different criteria, which can be the geographical location, the number of orders placed or canceled, and the product reviews made. Subsequently, at each occurrence of a portal event, such as a purchase or review, the corresponding data will be directed into the Event Bus. The latter has the dual task of transferring unstructured information (i.e., user profiles, geolocation data, assessments, and so on) to both a NoSQL database and the Streaming manager. They, in turn, sends them to the Machine Learning Cluster, on which stays the "Recommendation" algorithm, which has the assignment of associating each candidate client with a list of suggested products and is executed for each new entry. These two last components work as batch processes. Furthermore, the interconnection and data transfer between the various parts of the SaaS is carried out through a scalable asynchronous bus.

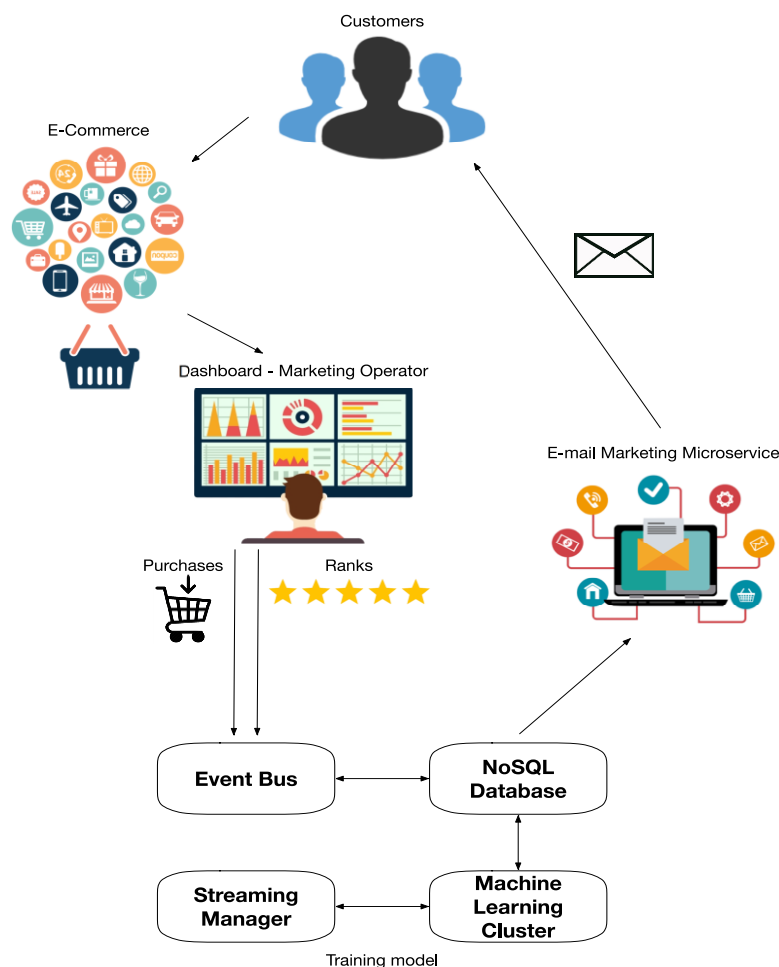


Figure 1. System architecture

The second study under consideration is part of the smart energy sector and concerns the use of data analysis to optimize energy consumption in IoT contexts. The case study proposed by the authors is based on the analysis of the additive manufactory (AM) production process, which uses the selective laser sintering (SLS) technique [50] and which consists of high energy consumption with a high production yield [51]. The energy consumption of the entire production process can be represented by the sum of energy consumption of different machines. For each complete production process, the total energy consumption varies due to the different production conditions and the different work environments. Furthermore, the power of a distinct subsystem, which is different for each subsystem, is difficult to determine with precision in real-time conditions.

In the era of Industry 4.0, among the various key factors that drive industrial development, the attention is focused on the environmental impact that can generate particular production processes. Currently, it is estimated that industrial production activities use around 35% of the entire global electricity supply and that they produce approximately 20% of total carbon emissions. Besides, over the past 20 years, the top five manufacturing countries have experienced a 50% increase in their greenhouse gas emissions. For this reason, in the industrial field, there are strong interests in trying to develop new approaches that allow realizing sustainable production cycles but minimizing energy consumption. The institutions are moving in this direction, such as Europe. The new European energy legislation requires industries to reduce both their gas emissions and their energy use by 20% by 2020 [52]. It is recognized that the energy efficiency of production processes usually is less than 30% and that, for some specific methods, energy losses are dramatically high [53]. Consequently, the use of highly efficient energy through data analysis, in addition to reducing production costs and expanding profit margins, can also have a positive impact on the associated environmental and social problems. This condition would allow industries to fall within the constraints imposed by international regulations.

Another case study introduces an approach whose purpose is to develop an embryonic model for predicting energy consumption in the manufacturing industry. In particular, the problem involves the prediction of the energy consumption of an "Additive Manufacturing" system, because, due to the different correlations between various properties associated to production, the reduction of the energy consumption of this process turns out to be one of the most critical points of the research in this area [54, 55]. Before the coming of Industry 4.0, considering the limited sensors or their apparent unrelatedness, several intangible factors, such as the health conditions of the equipment, the competence of the personnel, the environmental data, the temperature of the workspace and the noise of cut, were not considered for the analysis of production processes.

Nowadays, however, in an industrial environment with large sensors capable of generating massive amounts of data, it is possible to identify the correlation and the causal relationship of multi-source information to reveal important hidden factors in man-machine interactions. For instance, in the presence of the right lighting conditions in the production environment, it is possible to identify some critical latent factors in temperature, environmental humidity, cutting noise and sound field distributions, as these can influence the staff working efficiency, production quality, as well as energy consumption and performance degradation of industrial machines. Recognizing this reasonable goal, an analysis based on large data sets, collected from monitoring several factors, can help to improve the performance degradation prediction model, to identify energy consumption trends and to provide additional knowledge through feedback to achieve a better decision-making system concerning preventive maintenance and optimization of energy consumption, as shown in Figure 2.

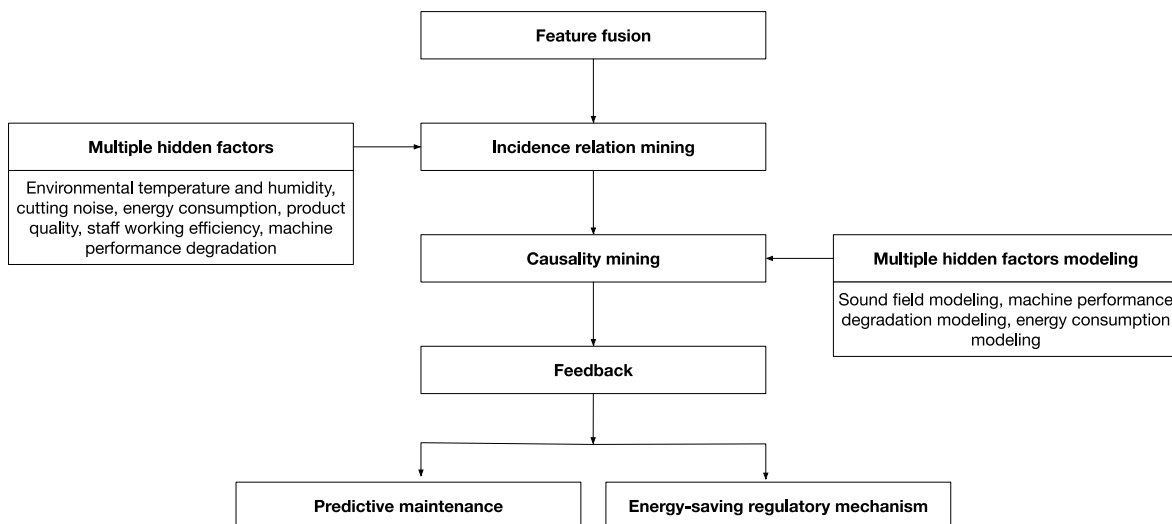


Figure 2. Decision-making scheme relating to preventive maintenance, aimed at saving energy consumption

A final case study concerns the so-called large-scale Data Ingestion of data from various devices in the context of Industry 4.0 [56]. With the fourth industrial revolution, the IoT has taken on a significant role in the development process, but this development still finds the import of heterogeneous and multi-source devices on a large scale as a challenge. In the industrial sector, more and more sensors are being incorporated into intelligent products, production equipment, and production monitoring. Smart production, which has become a vital component of production in the Industry 4.0 era, is facing several challenges mainly related to the following features:

- Heterogeneous data: There are different types of smart devices in a company. Each device has single or multiple parameter sensors, with the consequent creation of heterogeneous data. The first great challenge is to allow interaction among different data;
- Multi-source data: In addition to streaming data in real-time, there are also legacy applications that manage existing devices in companies and that continue to use legacy software despite being dated or no longer supported [57]. Streaming data can be analyzed in real-time. As shown in Figure 3, after the real-time analysis and the original application, the device data can be divided into data streams, file data, or data in relational databases. In this regard, the ingestion of all these heterogeneous data is the second major challenge to be faced;

- Large amounts of data: Recently, there has been an explosive growth in the category, speed, and volume of data [58]. With the development of intelligent manufacturing technology, it is possible to predict that the IoT will increase its data flows to unprecedented levels [59]. As a consequence, there will be the need to develop new approaches that are increasingly effective both for the resolution of the storage and for the querying of records in big data contexts.

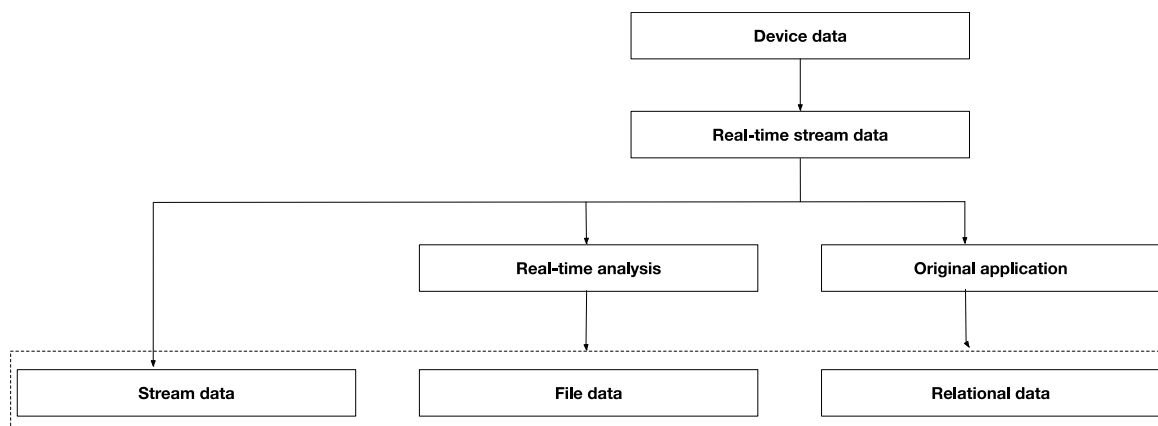


Figure 3. Analysis of heterogeneous data

3. CONCLUSION

This paper aimed to present the advantages introduced by the extensive use of big data through the examination of different approaches in the literature. To this end, some case studies have been explained in which it has been revealed how the employment of data analysis brought countless benefits in investigated scenarios; from energy saving to preventive maintenance, up to the timely adaptation of production through data analysis carried out in the marketing departments of the companies. Initially, a general study focused on Industry 4.0, and the analysis of big data has been evaluated. Afterward, the intentions to be accomplished in the considered scenarios have been reported, as well as the advantages brought by the analysis of big data in some case studies treated in industrial applications. In this sense, some methods introduced in the literature have been acquainted and investigated, also presenting the results to highlight their strengths and weaknesses. It has been found that in the various solutions examined, the analytical approach of big data yielded meaningful profits. Moreover, the adoption of these solutions embodies an added value for all corporations that aspire to prevail competitive on the market.

It is well known that investments in this field are crucial to pursuing this goal. Besides, from the analysis carried out in this work, it is reasonable to emphasize some comprehensive evaluations: concerning big data analysis in marketing strategies, considering that, for instance, every click of a user browsing the web corresponds to a generated data stream, potentially containing his/her preference, the adoption of data analysis could enable companies to identify trends in consumer preferences. As a result, it will be possible to conduct specific marketing and customer loyalty campaigns based on this added value and to adapt the production lines according to the compelling needs promptly; regarding big data analysis in the detection of production anomalies, this procedure could yield the early detection of severe trends and deviations in the products, which could lead to critical drawbacks in terms of economic losses and safety in the use of consumer products; concerning big data analysis in the prediction of the energy consumption of a production plant or a technological plant in green smart cities, it is crucial to emphasize how this component turns out to be intelligent from the environmental point of view and of the economic revenues as it can allow the plant operators to predict the energy consumption of the equipment and, potentially, find various factors that negatively affect the energy expenditure of the industrial machinery, such as failures or errors; in the use of big data analysis in preventive maintenance, the main advantages are measured in terms of economic revenues, since its application in this area may allow for taking into account various factors (environmental and non-environmental) that negatively influence the operation of machinery, through the identification of intangible factors that more often than not escape the qualified technical personnel, transforming themselves, consequently, into more significant revenues and even lower energy consumption; regarding big data analysis in risk management, this method is fundamental in decision-making and in predicting potentially dangerous situations if the study of these data is carried out, for instance, by artificial

intelligence devices which, unlike of human personnel, would allow a safe prediction based on real-time data. Moreover, in the automotive industry, there are more and more discussions about autonomous or driverless vehicles. The possibility of being able to count on automated devices that process enormous amounts of data in real-time can determine immense benefits to reduce any road accidents; in the case of the management of data coming from heterogeneous devices, considering that within a company switched to Industry 4.0 data flows of different types can be generated, their transformation into a unified format can lead to more efficient data management; regarding big data analysis combined with the use of IoT in the monitoring of industrial plants, the advantages consist in the opportunity of being able to implement, even for legacy industrial machinery, a remote monitoring system in real-time, to observe the performance of the production plant at any time and in any part of the world and even receive alarms regarding the occurrence of certain events.

Anyhow, the use of data analysis in an industrial environment can also involve disadvantages. Indeed, the digitization of production also implies potential risks and threats in terms of cyber-attacks and intrusions. Furthermore, the use of dedicated hardware supports, such as distributed computing systems located inside the production plants, affects both economic factors and environmental impact as it will require double energy expenditure due to both power supplies than cooling, weighing on energy costs. In conclusion, it is conceivable to assert that the advantages brought by the use of the ITC in industrial contexts far outweigh the disadvantages and that the companies that have not yet welcomed Industry 4.0, in absolute terms, appear to be little competitive on the market compared to the competition. Hence, there should be more considerable future involvement of industrialists as the potential benefits, as repeatedly stated in this work, are various. Besides, thanks to the increasingly rapid technological development, in the future, the negative factors could be concretely reduced to a minimum, first of all, the environmental impact deriving from industrial production.

REFERENCES

- [1] E. H. D. R. da Silva, J. Angelis, and E. P. de Lima, "In pursuit of digital manufacturing," *Procedia Manufacturing*, vol. 28, pp. 63-69, 2019.
- [2] F. Song, Y-T. Zhou, L. Chang, and H-K. Zhang, "Modeling space-terrestrial integrated networks with smart collaborative theory," *IEEE Network*, vol. 33, no. 1, pp. 51-57, 2019.
- [3] D. N. E. Phon, A. F. Z. Abidin, Mohd F. Ab Razak, S. Kasim, A. H. Basori, and T. Sutikno, "Augmented reality: Effect on conceptual change of scientific," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 4, pp. 1537-1544, 2019.
- [4] O. S. Itani, F. Jaramillo, and B. Paesbrughe, "Between a rock and a hard place: Seizing the opportunity of demanding customers by means of frontline service behaviors," *Journal of Retailing and Consumer Services*, vol. 53, pp. 1-12, 2020.
- [5] Z. Ai, Y. Liu, F. Song, and H. Zhang, "A smart collaborative charging algorithm for mobile power distribution in 5G networks," *IEEE Access*, vol. 6, pp. 28668-28679, 2018.
- [6] Z. Ngadiron, N. H. Radzi, and M. Y. Hassan, "The generation revenue and demand payment assessment for pool based market model in Malaysia electricity supply industry," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 4, pp. 1451-1460, 2019.
- [7] "Energy union and climate action," 2017, [Online] Available at: <http://publications.europa.eu/webpub/com/factsheets/energy/en/>.
- [8] P. Saderholm, H. Hellsmark, J. Frishammar, J. Hansson, J. Mossberg, and A. Sandstrom, "Technological development for sustainability: The role of network management in the innovation policy mix," *Technological Forecasting and Social Change*, vol. 138, pp. 309-323, 2019.
- [9] J. H. Nord, A. Koohang, and J. Paliszkievicz, "The Internet of Things: Review and theoretical framework," *Expert Systems with Applications*, vol. 133, pp. 99-108, 2019.
- [10] F. Song, M. Zhu, Y. Zhou, I. You, and H. Zhang, "Smart collaborative tracking for ubiquitous power IoT in edge-cloud interplay domain," *IEEE Internet of Things Journal to Appear*, pp. 1-9, 2019.
- [11] M. E. Ketzenberg and R. D. Metters, "Adapting operations to new information technology: A failed Internet of Things application," *Omega*, vol. 92, 2020.
- [12] Z-Y. Ai, Y-T. Zhou, and F. Song, "A smart collaborative routing protocol for reliable data diffusion in IoT scenarios," *Sensors*, vol. 18, no. 6, pp. 1-21, 2018.
- [13] M. A. G. Santos, R. Munoz, R. Olivares, P. P. R. Filho, J. Del Ser, and V. H. C. de Albuquerque, "Online heart monitoring systems on the Internet of Health Things environments: A survey, a reference model and an outlook," *Information Fusion*, vol. 53, pp. 222-239, 2020.
- [14] C. A. da Costa, C. F. Pasluosta, B. Eskofier, D. B. da Silva, and R. da R. Righi, "Internet of Health Things: Toward intelligent vital signs monitoring in hospital wards," *Artificial Intelligence in Medicine*, vol. 89, pp. 61-69, 2019.
- [15] Y-W. Chow, W. Susilo, J. G. Phillips, J. Baek, and E. Vlahu-Gjorgievska, "Video games and virtual reality as persuasive technologies for health care: An overview," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 8, pp. 18-35, 2017.

- [16] M. Rahman, A. Basu, T. Nakamura, H. Takasaki, and S. Kiyomoto, "PPM: Privacy policy manager for home energy management system," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 9, no. 2, pp. 42-56, 2018.
- [17] F. Song, Z. Ai, Y. Zhou, I. You, K.-K. R. Choo, and H. Zhang, "Smart collaborative automation for receive buffer control in multipath industrial networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1385-1394, 2019.
- [18] M. Alaa, A. A. Zaidan, B. B. Zaidan, M. Talal, and M. L. M. Kiah, "A review of smart home applications based on Internet of Things," *Journal of Network and Computer Applications*, vol. 97, pp. 48-65, 2017.
- [19] P. S. de Boer, A. J. A. M. van Deursen, and T. J. L. van Rompay, "Accepting the Internet-of-Things in our homes: The role of user skills," *Telematics and Informatics*, vol. 36, pp. 147-156, 2019.
- [20] F. Arena and G. Pau, "An overview of vehicular communications," *Future Internet*, vol. 11, no. 2, pp. 1-12, 2019.
- [21] G. Tesoriere, T. Campisi, A. Canale, Al Severino, and F. Arena, "Modelling and simulation of passenger flow distribution at terminal of Catania airport," *AIP Conference Proceedings*, vol. 2040, no. 1, 2018.
- [22] D. Ticali, M. Denaro, A. Barracco, and M. Guerrieri, "Piezoelectric energy harvesting from raised crosswalk devices," *AIP Conference Proceedings*, vol. 1648, no. 1, pp. 7800061-7800064, 2015.
- [23] F. Arena, G. Pau, and A. Severino, "V2X communications applied to safety of pedestrians and vehicles," *Journal of Sensor and Actuator Networks*, vol. 9, no. 1, pp. 3-13, 2019.
- [24] M. Giliberto, F. Arena, and G. Pau, "A fuzzy-based solution for optimized management of energy consumption in e-bikes," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 10, no. 3, pp. 45-64, 2019.
- [25] P. Thorncharoen Sri, W. Susilo, and J. Baek, "Efficient controlled signature for a large network with multi security-level setting," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 10, no. 3, pp. 1-20, 2019.
- [26] L. Amodu, O. Odiboh, S. Usaini, D. Yartey, and T. Ekanem, "Data on security implications of the adoption of Internet of Things by public relations professionals," *Data in Brief*, vol. 27, pp. 1-4, 2019.
- [27] H. Wang, L. Wang, Z. Zhou, X. Tao, G. Pau, and F. Arena, "Blockchain-based resource allocation model in fog computing," *Applied Sciences*, vol. 9, no. 24, pp. 1-18, 2019.
- [28] F. Song, Y.-T. Zhou, Y. Wang, T.-M. Zhao, I. You, and H.-K. Zhang, "Smart collaborative distribution for privacy enhancement in moving target defense," *Information Sciences*, vol. 479, pp. 593-606, 2019.
- [29] H. Haddadpajouh, A. Dehghantanha, R. M. Parizi, M. Aledhari, and H. Karimipour, "A survey on internet of things security: Requirements, challenges, and solutions," *Internet of Things, article in press*, pp. 1-20, 2019.
- [30] J. Iqbal, M. Khan, M. Talha, H. Farman, B. Jan, A. Muhammad, and H. A. Khattak, "A generic Internet of Things architecture for controlling electrical energy consumption in smart homes," *Sustainable Cities and Society*, vol. 43, pp. 443-450, 2018.
- [31] M. Collotta and G. Pau, "A novel energy management approach for smart homes using Bluetooth low energy," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 12, pp. 2988-2996, 2015.
- [32] G. Pau, T. Campisi, A. Canale, A. Saverino, M. Collotta, and G. Tesoriere, "Smart pedestrian crossing management at traffic light junctions through a fuzzy-based approach," *Future Internet*, vol. 10, no. 2, pp. 1-19, 2018.
- [33] G. Tesoriere, T. Campisi, A. Canale, and A. Severino, "The effects of urban traffic noise on children at kindergarten and primary school: A case study in Enna," *AIP Conference Proceedings*, vol. 2040, 2018.
- [34] M. Collotta, L. L. Bello, and G. Pau, "A novel approach for dynamic traffic lights management based on wireless sensor networks and multiple fuzzy logic controllers," *Expert Systems with Applications*, vol. 42, no. 12, pp. 5403-5415, 2015.
- [35] L. Husamaldin and Saeed, N., "Big Data analytics correlation taxonomy," *Information*, vol. 11, no. 1, 17-29, 2019.
- [36] Y. Cui, S. Kara, and K. C. Chan, "Manufacturing big data ecosystem: A systematic literature review," *Robotics and Computer-Integrated Manufacturing*, vol. 62, 101861, 2020.
- [37] N. Kefalakis, A. Roukounaki, and J. Soldatos, "Configurable distributed data management for the Internet of the Things," *Information*, vol. 10, no. 12, pp. 360-382, 2019.
- [38] P. J. Wu and K. C. Lin, "Unstructured big data analytics for retrieving e-commerce logistics knowledge," *Telematics and Informatics*, vol. 35, no. 1, pp. 237-244, 2018.
- [39] Z. Ai, Y. Liu, L. Chang, F. Lin, and F. Song, "A smart collaborative authentication framework for multi-dimensional fine-grained control," *IEEE Access*, vol. 8, pp. 8101-8113, 2020.
- [40] S. Wang, J. Yuan, X. Li, Z. Qian, F. Arena, and I. You, "Active data replica recovery for quality-assurance Big Data analysis in IC-IoT," *IEEE Access*, vol. 7, pp. 106997-107005, 2019.
- [41] L. Huang, C. Wu, and B. Wang, "Challenges, opportunities and paradigm of applying Big Data to production safety management: From a theoretical perspective," *Journal of Cleaner Production*, vol. 231, pp. 592-599, 2019.
- [42] R. Sahal, J. G. Breslin, and M. I. Ali, "Big Data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case," *Journal of Manufacturing Systems*, vol. 54, pp. 138-151, 2020.
- [43] Y. Zhang, R. Zhang, Y. Wang, H. Guo, R. Y. Zhong, T. Qu, and Z. Li, "Big Data driven decision-making for batch-based production systems," *Procedia CIRP*, vol. 83, pp. 814-818, 2019.
- [44] A. Al-Hamami and A. Flayyih, "Enhancing Big Data analysis by using map-reduce technique," *Bulletin of Electrical Engineering and Informatics*, vol. 7, no. 1, pp. 113-116, 2018.
- [45] "Mc Kinsey & Company," 2020. [Online]. Available at: <https://www.mckinsey.com>.
- [46] Milan Polytechnic Observatory, "Digital innovation research," 2020, [Online], Available at: https://www.osservatori.net/ww_en/.

- [47] "Forrester," 2020, [Online], Available at: <https://www.forrester.com/big-data>.
- [48] Y. Zhang, S. ren, Y. Liu, and S. Si, "A Big Data analytics architecture for cleaner manufacturing and maintenance processes of complex products," *Journal of Cleaner Production*, vol. 142, part 2, pp. 626-641, 2017.
- [49] W. Fan and A. Bifet, "Mining Big Data: Current status, and forecast to the future," *ACM SIGKDD Explorations Newsletter*, vol. 14, no. 2, pp. 1-5, 2014.
- [50] R. Paul and S. Anand, "Process energy analysis and optimization in selective laser sintering," *Journal of Manufacturing Systems*, vol. 31, no. 4, pp. 429-437, 2012.
- [51] M. Baumers, C. Tuck, D. L. Bourell, R. Sreenivasan, and R. Hague, "Sustainability of additive manufacturing: measuring the energy consumption of the laser sintering process," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 225, no. 12, pp. 2228-2239, 2011.
- [52] O. Poiana, "An overview of the European energy policy evolution: From the European Energy Community to the European Energy Union," *On-line Journal Modelling the New Europe*, no. 22, pp. 175-189, 2017.
- [53] A. Dietmair and A. Verl, "A generic energy consumption model for decision making and energy efficiency optimisation in manufacturing," *International Journal of Sustainable Engineering*, vol. 2, no. 2, pp. 123-133, 2009.
- [54] K. Yang and C. Shahabi, "On the stationarity of multivariate time series for correlation-based data analysis," *Fifth IEEE International Conference on Data Mining (ICDM'05)*, p. 4, 2005.
- [55] C. Ji, S. Liu, C. Yang, L. Wu, and L. Pan, "IBDP: An industrial Big Data ingestion and analysis platform and case studies," *2015 International Conference on Identification, Information, and Knowledge in the Internet of Things (IIKI)*, pp. 223-228, 2015.
- [56] A. Galletta et al., "A cloud-based system for improving retention marketing loyalty programs in Industry 4.0: A study on Big Data storage implications," *IEEE Access*, vol. 6, pp. 5485-5492, 2018.
- [57] "Legacy system," 2011, [Online], Available at: https://en.wikipedia.org/wiki/Legacy_system.
- [58] R. Ranjan, "Streaming Big Data processing in datacenter clouds," *IEEE Cloud Comp.*, vol. 1, no. 1, pp. 78-83, 2014.
- [59] J. Qin, Y. Liu, and R. Grosvenor, "Data analytics for energy consumption of digital manufacturing systems using Internet of Things method," *2017 13th IEEE Conference on Automation Science and Engineering (CASE)*, pp. 482-487, 2017.

BIOGRAPHIES OF AUTHORS



Fabio Arena received the Bachelor Degree in Telecommunication Engineer from University of Catania in 2006. He also received the Master Degree in Telecommunication Engineering from University of Catania in 2010. Currently, he is a Ph.D. student at Kore University of Enna. His current research interests include intelligent transportation systems, driverless vehicles and network architecture, big data analysis.



Giovanni Pau is an Associate Professor at Faculty of Engineering and Architecture, Kore University of Enna, Italy. Prof. Pau received his Bachelor degree in Telematic Engineering from University of Catania, Italy; and his Masters degree (cum Laude) in Telematic Engineering and PhD from Kore University of Enna, Italy. Prof. Pau has published more than 65 papers in journals and conferences and authored 1 book chapter. He serves as Associate Editor of several journals. Moreover, he serves/served as a leading Guest Editor in several special issues. He collaborates/collaborated with the organizing and technical program committees of several conferences to prepare conference activities and is serving as a reviewer of several international journals and conferences. His research interests include wireless sensor networks, soft computing techniques, internet of things, and network security.