IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Quality Risk Analysis For Sustainable Smart Water Supply Using Data Perception

G. Snitha *1, I. Sravan*2, D. Ajay*3, Dr. K. Kranthi Kumar*4,
Dr. K.Vijayalakshmi*5

*1 Student, Sreenidhi Institute of Science and Technology (SNIST), Affiliated to JNTUH, ITDepartment, Ghatkesar, Telangana, India

*2 Student, Sreenidhi Institute of Science and Technology (SNIST), Affiliated to JNTUH, ITDepartment, Ghatkesar, Telangana, India

*3 Student, Sreenidhi Institute of Science and Technology (SNIST), Affiliated to JNTUH, ITDepartment, Ghatkesar, Telangana, India

*4 Associate Professor, Sreenidhi Institute of Science and Technology (SNIST),Affiliated to JNTUH,

IT Department, Ghatkesar, Telangana, India.

*5 Professor, Associate Dean, Sreenidhi Institute of Science and Technology (SNIST), Affiliated to JNTUH, IT Department, Ghatkesar, Telangana, India.

ABSTRACT

Water is one of the most essential key aspects of our lives because it is the most necessary thing for all the living things on our planet. Because there is no life without water, we use it for various purposes like cleaning, drinking, and many more. There is no life without water, and in that sense, drinking water is the most essential thing in our lives. Salts like magnesium, sodium, calcium, and potassium are the important salts that we need in our body to be healthy. If any one of those salts is not present in the water, it can cause a lot of health problems like typhoid, polio, cholera, etc. There are so many places where people drink this impure water and are suffering from this type of disease. They are drinking this water because they don’t have enough knowledge about the water quality, so here we will focus on the water quality prediction. This paper suggests Ann, Random Forest, Propose an Adaptive Frequency analysis-based method for earlier prediction of water before drinking it. We predict whether the water is safe to drink or not so that people will be aware of it. This prediction process includes a data set that contains water quality parameters such as D.O(mg/l), PH, conductivity, Nitratio, Fecal Coliform, and Total coliform. And here, first, the data is imported, creating the main window, Defining global variables, Creating GUI elements, Functions for data processing, Configuring the main window, and running the application .

Keywords: water quality, Ann, Random Forest, Propose Adaptive Frequency analysis, creating a main window, Defining global variables, Creating GUI elements, Functions for data processing, Configuring the main window.

INTRODUCTION

Water treatment is a crucial step in maintaining water quality, but surface water and groundwater sources are vulnerable to contamination. The time it takes to identify water quality issues after treatment often hinders timely preventative action. In response, Norway is implementing new national water quality standards. Biological indicators, such as coliform bacteria and Escherichia coli, play a significant role in assessing water quality and are used as the basis for national standards. However, relying solely on coliform bacteria may overlook other harmful organisms. Therefore, there is a need for improved water quality testing methods to ensure a comprehensive assessment and proactive threat identification in modern water distribution networks.

Recent research has explored data-driven approaches to water quality regulation. For example, studies have analyzed water quality indicators and developed generic bacterial sensor cells to monitor changes in water quality. Artificial Neural Networks (ANN) and other algorithms have been used to predict water quality parameters in different regions. These advancements demonstrate the potential of data-driven techniques in predicting and managing water quality.

This project focuses on utilizing data perception techniques to enhance water quality analysis in water supply networks. The study examines data gathered from water supply systems in four Norwegian municipalities: Oslo, Bergen, Strommen, and Aalesund. The objectives include an in-depth examination of water quality issues, identification of research topics and obstacles, and the development of a predictive risk identification strategy based on adaptive frequency analysis. The proposed methodology is compared with traditional approaches like Artificial Neural Networks (ANN) and Random Forests in terms of prediction accuracy and processing time. The project aims to improve risk identification, prediction, and evaluation of water quality indicators by leveraging advanced data analytics and exploring the relationships among various water quality parameters.

The outcomes of this project hold significant implications for the future of Sustainable Smart Cities, as they contribute to improving water quality monitoring through advanced data analysis. By bridging the gap between physical and chemical markers and biological indicators, the study aims to provide a comprehensive understanding of water quality concerns. By utilizing data directly from the manufacturing process, the project addresses challenges related to the reliability and applicability of laboratory data in real-world scenarios. The findings will aid in assessing contamination from commercial and domestic activities in water source locations, facilitating informed decision-making for urban water supply systems. Overall, this project aims to revolutionize water quality monitoring and pave the way for more effective and sustainable management of water resources.

LITERATURE REVIEW

The study conducted by S. Franco, V. Gaetano, and T. Gianni focused on evaluating the combined influence of climate change and urbanization on water quality in an urbanized catchment area in northern Italy. Through a two-year field campaign and integrated modeling analysis, they found that impermeable urban surfaces and rainfall intensity were strong predictors of combined sewer overflows (CSOs) and the subsequent degradation of receiving water bodies. Future scenarios indicated that the phosphorus load from CSOs is expected to double by 2100, significantly impacting river water quality. The findings highlight the importance of reducing impervious surfaces and preventing the creation of new ones to mitigate the adverse effects of climate change and urbanization on water quality.

T. Hak, S. Janoušková, and B. Moldan explored the development of indicators for Sustainable Development Goals (SDGs). They emphasized the need to assess the relevance and quality of indicators to ensure they accurately represent the phenomena under investigation. While a set of global SDGs and indicators has been established, the authors stressed the importance of operationalizing these goals by selecting appropriate indicators and considering their relevance to decision-makers and the general public. They proposed the establishment of a conceptual framework for choosing suitable indicators from existing sets or developing new ones, highlighting the importance of clear and unambiguous communication through indicators.

D. Wu, H. Mohammed, H. Wang, and R. Seidu focused on data analysis using artificial intelligence to track water quality in catchment areas. They highlighted the importance of monitoring and treating water quality, especially in the face of increasing industrial, agricultural, and societal contamination. The researchers developed an advanced data analysis approach that considered all conventional indicators of water quality in a holistic setting. Their study, based on a case study from Oslo's water system, demonstrated the reliability of their models in predicting the evolution process of biological water quality indicators and provided decision support for future control in water supply systems.

J. Liu, A. Shahroudy, D. Xu, and G. Wang explored the application of spatiotemporal LSTM (Long Short-Term Memory) with trust gates for three-dimensional human action recognition. They aimed to analyze human activities based on 3D skeleton data and developed a novel approach that integrated the temporal and spatial dimensions. Their method employed an improved tree-structure-based traversal approach and introduced a gating mechanism within LSTM to mitigate the effects of noise and occlusion in 3D skeleton data. The technique achieved state-of-the-art performance in 3D human action analysis on various benchmark datasets.

METHODOLOGY

The projected water quality uses the water parameters as an input and is a real-time forecast. A block schematic of the water prediction approach is shown in the figure below. The procedure begins with the data being uploaded into the project, after which it is processed and categorized using a variety of algorithms, including ANN, Random Forest, and Adaptive frequency algorithm, from which we may then forecast the data.

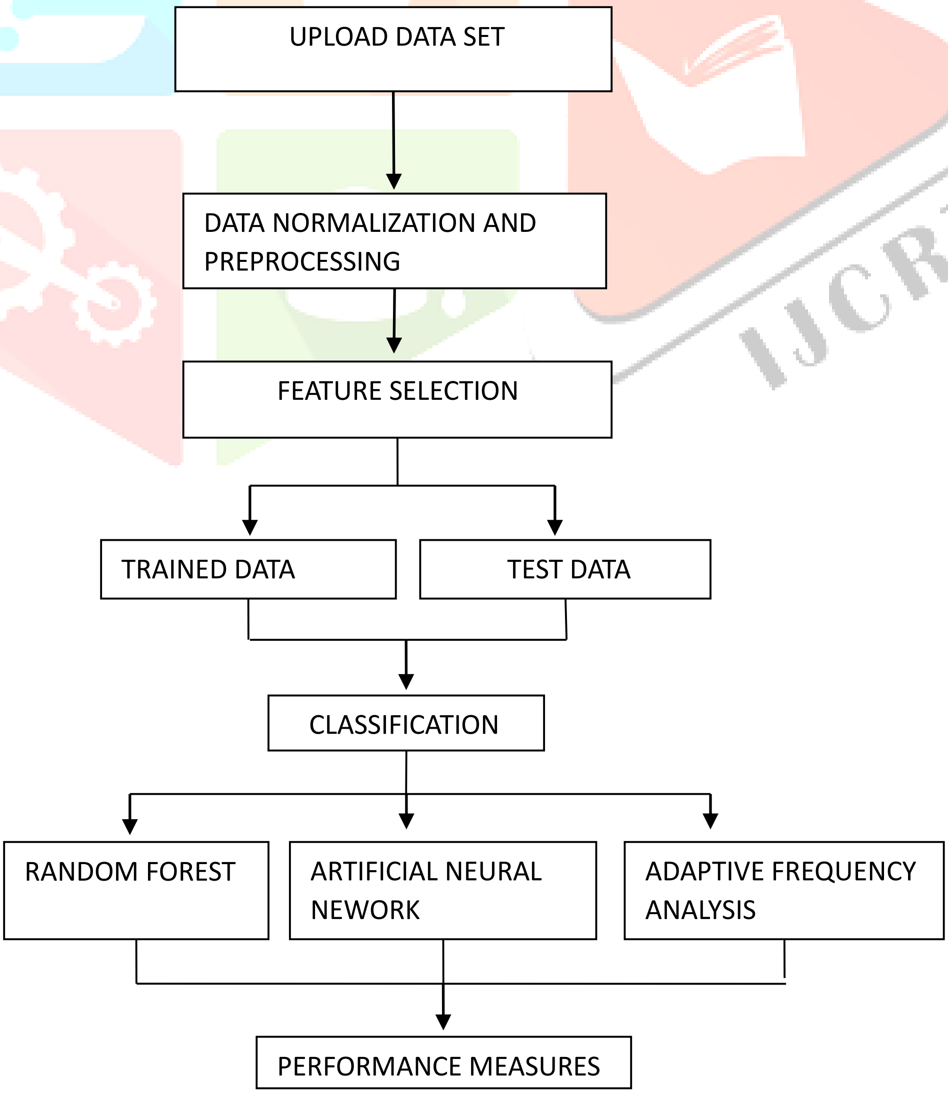

Fig 1

The following factors are included in the water data set: TEMPERATURE, D.O(mg/l), PH, CONDUCTIVITY, B.O.D(mg/l), NITRATED, FAECAL COLIFORM, and TOTAL COLIFORM. These parameters determine whether water is drinkable or not.

Preprocessing & normalized dataset: With these parameters, we can determine if the water has any problems or not. Here, the data is preprocessed and normalized; in general, preprocessing includes managing missing values, eliminating out layers, and encoding categorical variables. Data preprocessing is the act of preparing raw data for analysis or modeling. While normalizing is the process of rescaling a dataset's numeric characteristics to a standardized range, in our case we utilized it to remove unused columns and change data types. whereas normalizing is to rescale the numeric features of a dataset to a standardized range, in this we use Min-Max scaling and z-score scaling.

ANN classification: To train the data and to test the data, we need a classification approach. There are several classification strategies to test and train the model to forecast the issues with the water so that the necessary measures may be taken and the water is clean and drinkable. One of the algorithms we utilized in this model was the ANN (Artificial Neural Network), which we used to train our model. We used this to train our data set, and after that, we made it to predict the test data. The Artificial Neural Network is currently utilized to demonstrate the model's accuracy for validation and training once it has been trained.

Random Forest classification: this is a different classification strategy to train and evaluate the data. With this approach, we can also forecast issues and take preventative measures. Using the same classification approach, Random Forest is used to training the model, and after it has finished training, it is ready to forecast the data and demonstrate the model's accuracy for both validation and training.

Adaptive Frequency Analysis: This is the final classification method we used in our model it was trained using the training dataset and this algorithm. The model then predicts problems and is prepared to show model accuracy for validation and training.

EXISTING SYSTEM

Better ubiquitous sensing technology has an influence on contemporary research, business, and daily life [16]. More environmental indications may be detected, sent, and measured with their help. A sustainable smart water distribution system uses a number of sensors to check water quality and manage resources effectively. Data becomes a potent tool for enhancing our comprehension of modern systems as a result. We can identify changes in our water distribution system by analyzing our data with the right technology. A wide variety of sensors were utilized to detect pH, temperature, conductivity, and other parameters in the water supply locations. New data processing technologies and the enormous amounts of data these technologies generate have significantly increased our ability to regulate water quality.

Drawbacks:

Data scarcity and incomplete indicator collections

Difficulty in data synchronization

Need for specialized modeling

PROPOSED SYSTEM

This study examines the root mean square error (RMSE) performance of many techniques, including Random Forest, Artificial Neural Networks (ANN), and Adaptive Frequency in order to further past studies. None of the methods that are now available will repeatedly refilter the dataset to eliminate extraneous data, improve prediction accuracy, and reduce error rates. After all unnecessary features have been eliminated using DROPOUT procedures, the dataset will be filtered using a function called DENSE, which filtered the dataset by utilizing a predefined number of neurons. The two most popular deep learning algorithms, convolutional neural networks (CNNs) and long short-term memory (LSTMs), both filter knowledge several times. before developing a prediction model, the dataset should be mined for key characteristics. Increasing numbers of companies in the data Because of their better performance and universal acceptance, CNN and LSTM are being used increasingly by the processing sector as their principal classification and prediction models. Each layer of a CNN or LSTM may be designed to receive a distinct set of data filters, and the number of input and output layers can be adjusted. The code below uses CNN, which is implemented in red-colored comments.

RESULTS

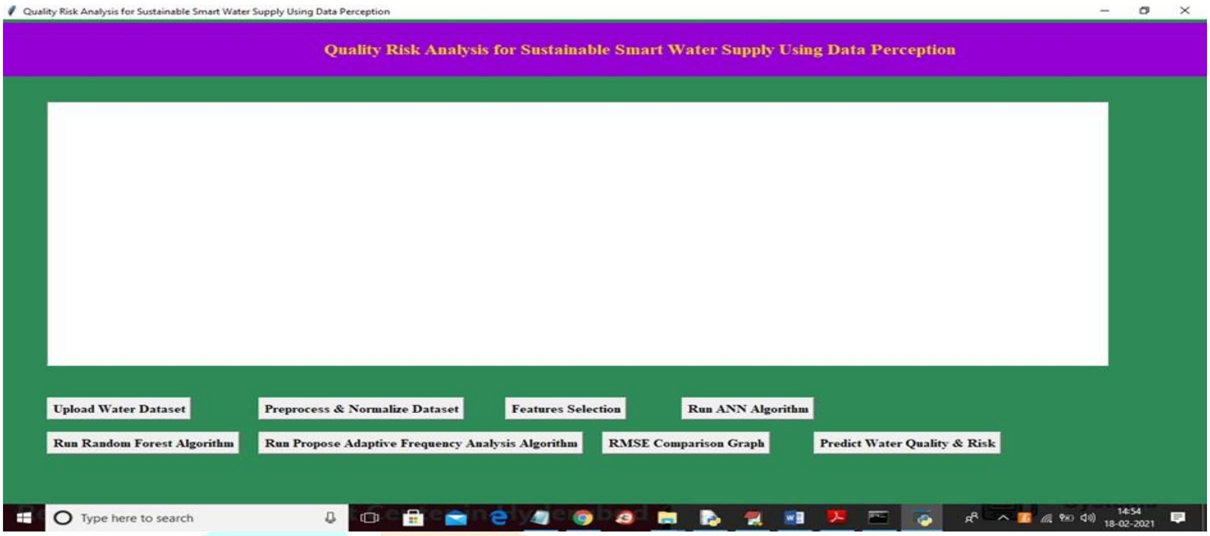

Fig:2

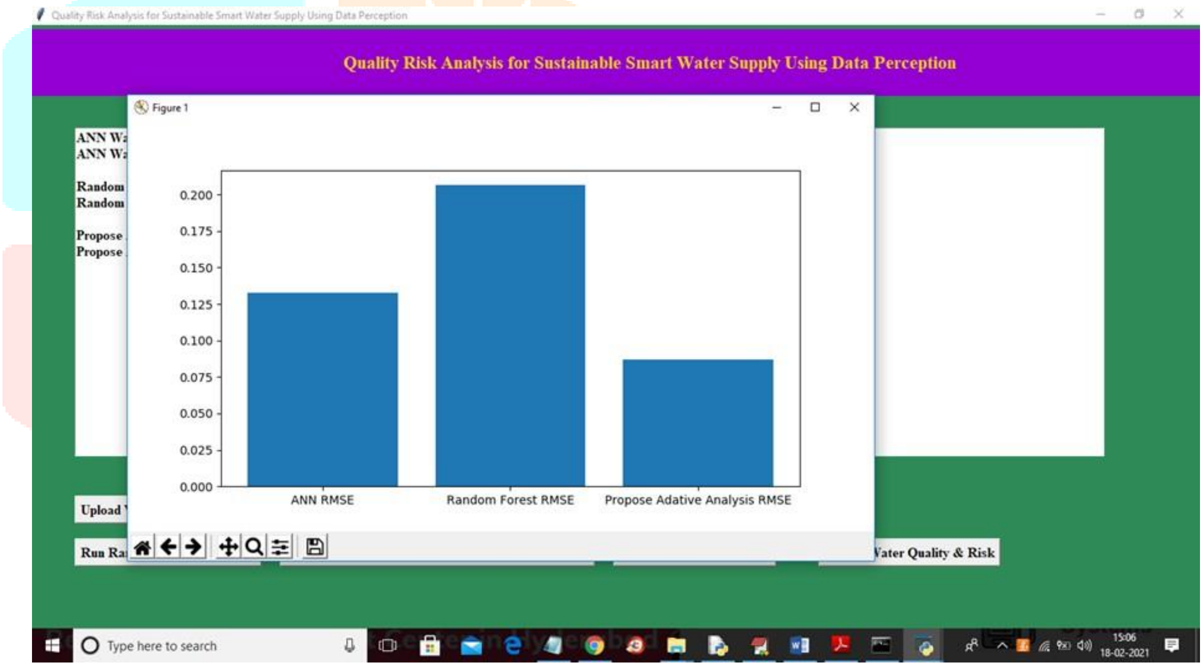

Fig:3

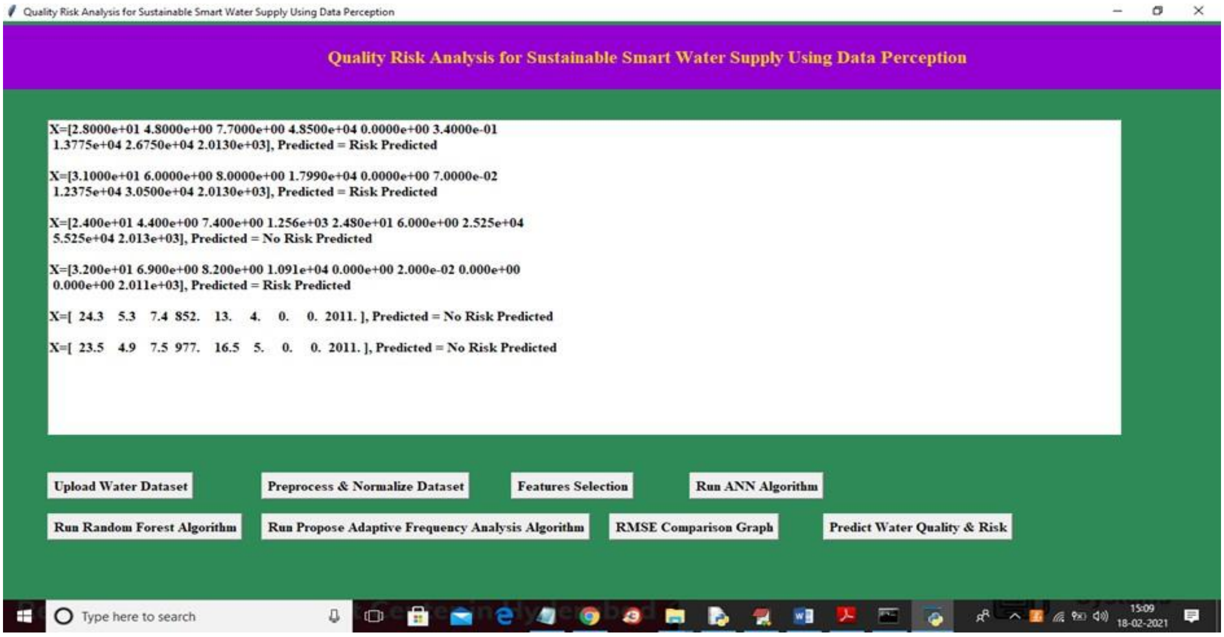

Fig: 4

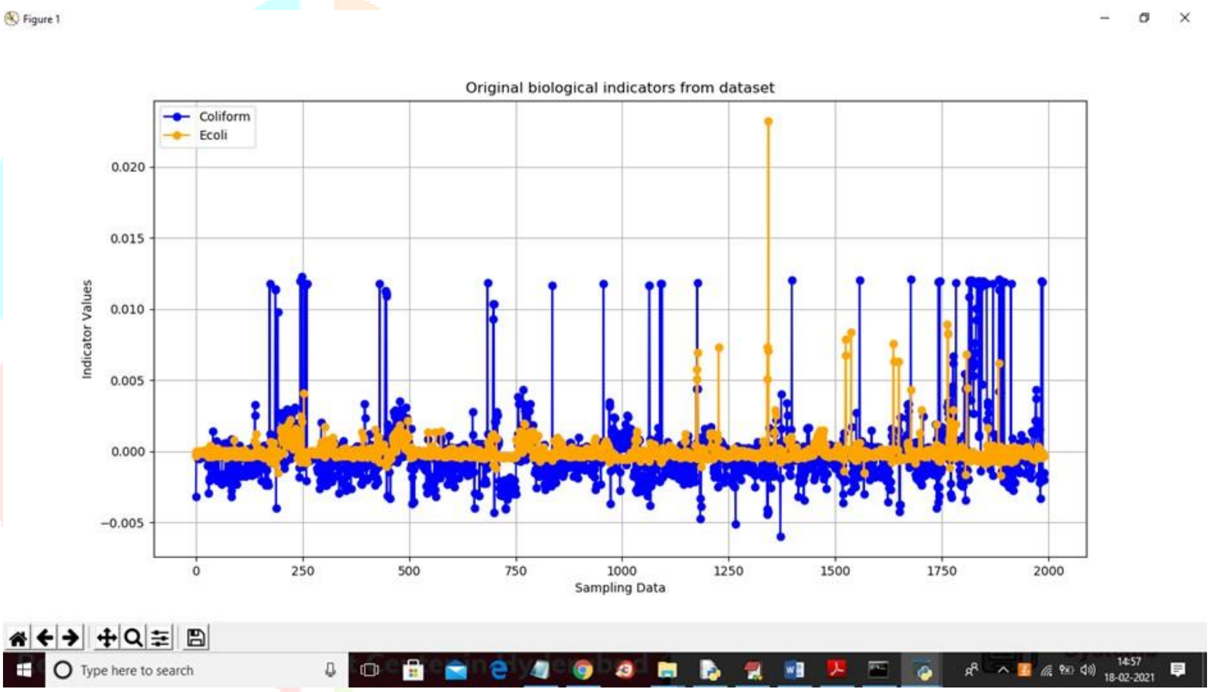

Fig:5

CONCLUSION

The study suggests a reasonably priced early warning system for threats to water quality in metropolitan areas. It offers prompt identification and aid with decisions, enabling proactive measures and better choice-making in the final phases of water delivery. The technique offers a fresh approach to frequency domain analysis by combining numerous indicators, locations, and time spans. In order to promote scalability across all three domains, it evaluates the connection between indicators and their predictive powers. The research is applied to genuine industrial water delivery systems in four Norwegian cities, proving the feasibility, accuracy, and efficacy of the suggested strategy.

REFERENCES

[1] "Hyperconnected Network: A Decentralised Trusted Computing and Networking Paradigm," by H. Yin, D. Guo, K. Wang, Z. Jiang, Y. Lyu, and J. Xing appeared in IEEE Network, vol. 32, no. 1, in January/February 2018.

[2] "LightweightRFIDprotocolformedicalprivacyprotection in IoT," IEEE Trans. Ind. Inform., vol. 14, no. 4, april 2018, pp. 1656–1665.

[3] "Amber: Decoupling user data from Web applications," in Proc. 15th Workshop Hot Topics Operational Systems (HotOS XV), Warth-Weiningen, Switzerland, 2015, pp. 1-6. T. Chajed, J. Gjengset, J. Van Den Hooff, M. F. Kaashoek, J. Mickens, R. Morris, and N. Zeldovich.

[4] "Enhancing selectivity in big data," IEEE Security Privacy, vol. 16, no. 1, pp. 34-42, Jan./Feb. 2018. M. Lecuyer, R. Spahn, R. Geambasu, T.-K. Huang, and S. Sen.

[5] "open PDS: Protecting the privacy of metadata with SafeAnswers," by Y.-A. de Montjoye, E. Shmueli, S. S. Wang, and A. S. Pentland. 2014, Vol. 9, No. 7, Art. No. e98790, PLoS ONE.

[6] C. Perera, R.Ranjan, and L.Wang,"End-to-endprivacyforopenbigdatamarkets," IEEECloud Computing, April 2015, vol. 2, no. 4, pp. 44–53.

[7] "Data linkage in smart Internet of Things systems: A consideration from a privacy perspective," by X. Zheng, Z. Cai, and Y. Li. IEEE Communications Magazine, Sep. 2018, pp. 55–61, vol. 56, no. 9.

[8] "Adaptable blockchain-based systems: A case study for product traceability," Q. Lu and X. Xu. IEEE Softw., Nov./Dec. 2017, vol. 34, no. 6, pp. 21–27.

[9] "Deep learning based inference of private information using embedded sensors in smart devices," by Y. Liang, Z. Cai, J. Yu, Q. Han, and Y. LiIEEENetw.Mag.,vol.32,no.4,pp.8–14,Jul./Aug.2018.

[10] MeDShare: Trust-less medical data sharing among cloud service providers through J. Gao, X. Du, and M. Guizani's paper