# A Management Model of Real-time Integrated Semantic Annotations to the Sensor Stream Data for the IoT

Besmir Sejdiu[1] [a], Florije Ismaili[1] [b] and Lule Ahmedi[2] [c]

[1] *Contemporary Sciences and Technologies, South East European University, Tetovo, Macedonia*

[2] *Faculty of Electrical and Computer Engineering, University of Prishtina, Prishtinë, Kosovo*

Keywords: Internet of Things, Wireless Sensor Networks, Stream Data Management, Semantic Annotations, Water Quality Monitoring.

Abstract: Wireless Sensor Networks (WSNs) are one of the most important components of the Internet of Things (IoT). They produce continuous stream of data and transmit these data to a centralized server. Due to the dramatic increase of streaming data, their management and exploitation has become increasingly important. Furthermore, by adding semantic annotations into sensor stream data, better understanding and more meaningful descriptions is provided, which enables application areas of IoT to become much more intelligent. In this paper, a data stream management model of WSNs for IoT real-time monitoring systems, that supports real-time integration of data from heterogeneous sensors with semantic annotations is presented. To validate the proposed model, an IoT system for real-time water quality monitoring is built, which enables real-time integration of semantic annotations to the sensor stream data in the format of Sensor Observation Service (SOS).

## 1 INTRODUCTION

Nowadays, smart water networks, smart homes, smart cities, smart health, smart grid, intelligent transportation, are infrastructure systems that connect our world more than we ever thought possible. The common vision of such systems is usually associated with one single concept, the *Internet of Things* (IoT) (Yinbiao, 2014). The idea of IoT was developed in parallel to Wireless Sensor Networks (WSNs). Therefore, the main components that enable IoT are WSNs (Lazarescu, 217).

A *Wireless Sensor Networks* is a wireless network consisting of spatially distributed autonomous devices using sensors to monitor physical or environmental conditions. WSNs may be homogeneous or heterogeneous. *Homogeneous sensors* send only one type of information (e.g. the water temperature), while *heterogeneous sensors* send more than one type of information (e.g. temperature and dissolved oxygen). All these sensors send observational data referred to as sensor stream data to a remote server. Therefore, sensory data comes from multiple sensors of different modalities in distributed locations.

Furthermore, sensor stream data is enabled to the web through the *Sensor Web* (SW). SW by incorporating technologies of the Semantic Web creates the *Semantic Sensor Web* (SSW) (Wang, 2020). Therefore, by adding semantic annotations to sensor stream data with concept definitions from domain knowledge (e.g. ontologies), the interpretation and understanding of sensor data and metadata is enabled. The real-time integration of sensor data as dynamic data with semantics is defined as *real-time semantic annotation*, while sensor data that is stored in repository (data store) as static data, and then integrated with semantics is defined as *non-real-time semantic annotation* (Sejdiu, 2020). Organizations like Open Geospatial Consortium (OGC) and World Wide Web Consortium (W3C) have proposed industry standards such as Sensor Web Enablement (SWE), which are aimed at providing unified standards (Shi, 2018).

To manage the real-time integration of semantic annotations into heterogeneous sensor stream data

[a] https://orcid.org/0000-0002-2786-5384
[b] https://orcid.org/0000-0002-3627-0147
[c] https://orcid.org/0000-0003-0384-6952



Sejdiu, B., Ismaili, F. and Ahmedi, L.
A Management Model of Real-time Integrated Semantic Annotations to the Sensor Stream Data for the IoT.
DOI: 10.5220/0010111500590066
In *Proceedings of the 16th International Conference on Web Information Systems and Technologies (WEBIST 2020)*, pages 59-66
ISBN: 978-989-758-478-7

with context in the IoT and by using OGC SOS standards, we introduce a management model in this paper.

The paper is organized as follows: Section 2 provides a discussion on related work for semantic annotations to the sensor stream data. Section 3 is an overview of the data stream models, annotated sensor data streams, and used technologies. Section 4 represents a proposed management model of integrated semantic annotations to the sensor stream data for the IoT. Prototype implementation of the proposed model is presented in Section 5, and system outputs is presented in Section 6. Finally, Section 7 concludes the paper and relevels some of the future perspectives of the proposed model.

# 2 RELATED WORK

Recently, some researchers have already shown up with several investigations related to semantic enrichment of sensor stream data. Lin et al. (Lin, 2019) proposed a semantic annotation method to annotate IoT sensor data through semantics. They used K-means clustering method for knowledge discovery. The proposed mechanism is semiautomatic, because after clustering the data, there are just clusters without knowledge. The knowledge of the clusters is defined by people. Xiaomin el al. (Xiaomin, 2016) proposed an approach for ontology modelling which is used in evaluating the river water quality and its relevant processing knowledge. The presented model consist of the data acquisition layer, diagnosis layer and decision support layer.

Furthermore, in this study, different solutions are identified, (Vera, 2014), (Pradilla, 2016), (Bytyçi, 2019), to semantically annotate sensor stream data. Those proposed solutions used *non-real-time semantic annotation* because the sensor stream data is stored in data store or ontology as static data and then integrated with semantics. However, it is still an open issue on how techniques and models for integration and interpretation of the semantic annotations in real-time in the sensor stream data should be advanced.

The main contributions of our approach are presented as follows:

- A data stream management model of WSNs for real-time IoT monitoring systems has been developed, which support real-time integration of data from heterogeneous sensor with semantic annotations.
- To validate the proposed model, an IoT system for real-time water quality monitoring is built, which enables real-time integration of semantic annotations to the observational data on water quality coming from wireless sensors.
- By incorporating OGC SOS standards, the heterogeneity of data sensor is hidden from arbitrary sources, and a real-time service interface for published enriched sensor stream data is provided with the semantic annotations to display in IoT real-time monitoring systems.

# 3 BACKGROUND

## 3.1 The Data Stream Models

Depending on characteristics, source of data transmission and saving of stream data, those can be modelled in various ways: *Real-time data stream* (data stream in real time is a sequence of data which arrives in order and/or in pre-processed ways), *Stream items* (since data is received in streams, those can be modelled as a sequence in a list of elements), and *Window models* (only a fragment of the streaming data is of interest at any time, and can be classified according to the three criteria: *Fixed sliding window*, *Landmark window*, and *Adaptive window*).

## 3.2 Annotated Sensor Data Streams

WSNs consist of small-scale devices that enable observe various physical phenomenon, which provide sensor data in raw format. Typically, classical IoT applications cannot interpret the sensor data and understand its context. This makes it nearly impossible to get the high-level information of the events and infer additional knowledge to gain situational awareness (Khan, 2015).

To provide meaning (semantics) of raw data, annotated sensor data stream is required. The annotated sensor data becomes more meaningful and understandable, enabling end-users to get high-level details about the real-world situations instead of raw sensor data. This is known as Semantic Sensor Web (SSW). These annotations provide more meaningful descriptions and enhanced access to sensor data than SWE alone, and they act as a linking mechanism to bridge the gap between the primarily syntactic XML-based metadata standards of the Sensor Web Enablement (SWE) and the RDF/OWL-based metadata standards of the Semantic Web.

To encode semantic annotations and data gathered by sensors, SWE is used in this paper, respectively version 2.0 of the Sensor Observation Service (SOS) standard relies on the OGC Observation & Measurement (O&M).

## 3.3 Technologies

The proposed management model of real-time integrated semantic annotations to the sensor stream data for the IoT utilizes:

*Spark Streaming*[4] - is an extension of the Apache Spark which enables to build scalable fault-tolerant IoT applications. Data can be ingested from many sources like Apache Kafka, TCP sockets, Flume, Twitter, etc., and processed data using complex algorithms expressed with high-level functions like map, reduce, join and windows. Finally, processed data can be pushed out to file systems, databases, and live dashboards.

*Apache Kafka*[5] - is a distributed streaming platform which has capabilities to publish and subscribe to streams of records, similar to a message queue or enterprise messaging system.

*Apache Cassandra database*[6] - is a distributed store for structured data that scale-out on cheap. It is designed to handle large amounts of data.

# 4 PROPOSED MODEL

In Figure 1, an overview of the model architecture for processing and managing sensor stream data for IoT real-time monitoring systems, such as: water quality monitoring, air quality monitoring, etc. is presented.

The WSNs are deployed in different locations. They produce a continuous stream of data, and transmit to Apache Kafka in various formats (e.g. binary, JSON, XML, etc). Kafka is utilized to transform them in a specific format that will be processed by Spark Streaming in real-time and parallel. The Spark Streaming enables a real-time integration of semantics into sensor stream data by using association rules, mining data streams, WSNs metadata, and archival data streams, with concept definitions from ontologies or other semantic sources, which provides the understanding and more meaningful descriptions to enable application areas of IoT to become much more intelligent.

The enriched sensor stream data with semantic annotations results are stored in the Cassandra database, and will be displayed in IoT real-time monitoring systems in format of SOS O&M standard by using stakes, such as XLink (without including XPath). A fragment of an output example is presented in Figure 2.

As shown in Figure 1, the proposed data stream management model support real-time data integration of heterogeneous sensors with semantic annotations, continuous queries on streaming data, outlier validation of streaming data, ad-hoc queries, and archive stream data with their semantic annotations for applications that need to answer queries from the archival store (persistent data stored). The proposed model consists of the main components: A) *Input Data Stream*, B) *Stream Processor*, C) *Data Modelling*, D) *OGC standards*, and E) *Ontology*.

Each component of the model is described in details as following:

A. *Input Data Stream* – is implemented in Apache Kafka and accepts in real-time input streaming data sent by the WSNs. Unordered streams come without any kind of pre-processing – as *unordered cash register*. Each stream can provide elements at its own schedule, they do not need to have the same data rates or data types, and the time between elements of one stream need not be uniform.

B. *Stream Processor* is developed in Spark Streaming and contains *Outlier Stream Validator & Classificator*, *Query Process*, *Ad-hoc Queries*, and *Semantic Annotations Stream Process*:

- *Outlier Stream Validator & Classificator* – is part of the stream processor in charge of real-time validating sensor streaming data, which marks stream data with one of the following status, '*valid*' or '*outlier*'. Validated data continues to processing forward, while invalid data are stored in *Invalid Data Streams (IDS)*. A data stream object is considered an outlier if it does not conform to expected behaviour, which corresponds to either noise or anomaly. For example the observed value of the pH sensor is '-2' or '*NULL*', in this case this sensor data will be classified as outlier because the range value of pH phenomena is 0 to 14. Outliers can arise due to different reasons such as mechanical faults, other changes in the system, fraudulent behaviour, instrument error, human error or natural deviation (Yu, 2020). Therefore, the *Outlier Stream Validator & Classificator* provides data quality for IoT real-time monitoring systems.
- *Query Processor* – supports continuous queries for streaming data, and are continuously executed as data streams continue to arrive. The answer of the continuous query is produced over time, always reflecting the stream data seen so far. The answers of our *Query Processor* can include semantic annotated data in the result.
- *Ad-hoc Queries* – queries executed ad-hoc from users; a question asked once about the current state of a stream or streams. Users can also specify ad-hoc queries that integrate streaming data and persistent data stored on *Working Data Streams*,

[4] http://spark.apache.org

[5] https://kafka.apache.org

[6] http://cassandra.apache.org

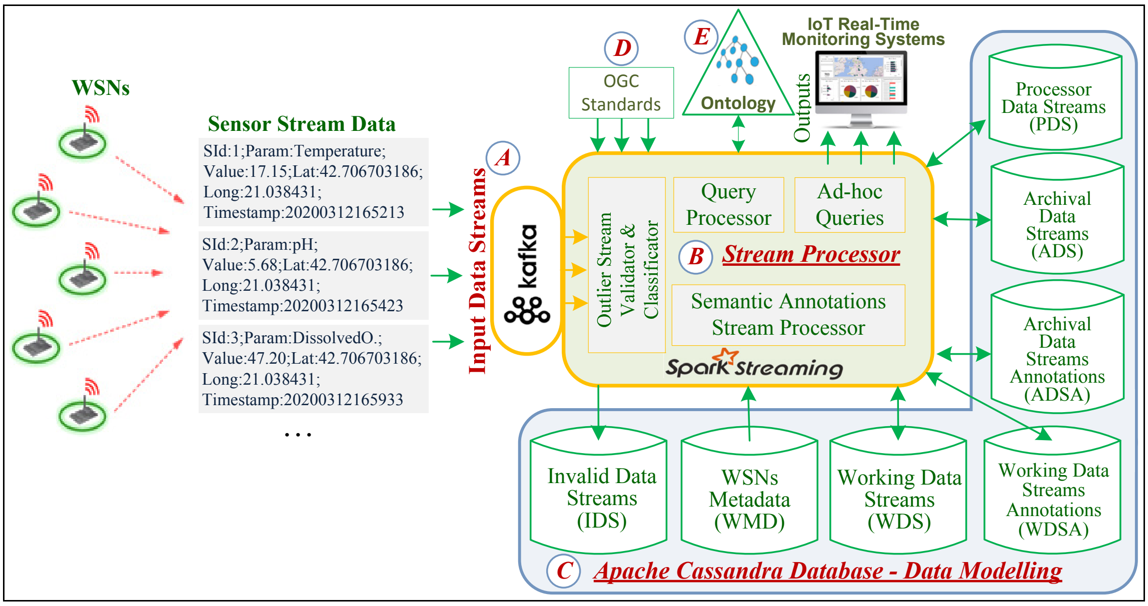


Figure 1: An overview of the model architecture.

*WSNs Metadata* or *Archival Data Streams*. In addition, the result of *Ad-hoc Queries* can include semantic annotated data in the result.

```
<sos:Observation ...="">
  <observationData>
    <om:OM_Observation gml:id="o45125">
      <om:type xlink:href="http://www.opengis.net/def/
observationType/OGC-OM/2.0/OM_Measurement"/>
      <om:phenomenonTime>
        <gml:TimeInstant gml:id="phenomenonTime_1">
          <gml:timePosition> 2020-04-03T14:32:00+09:00
</gml:timePosition>
        </gml:TimeInstant>
      </om:phenomenonTime>
      <om:resultTime xlink:href="#phenomenonTime_1"/>
      <om:procedure xlink:href="http://myserver
/ontologies/ont-core.owl#Sensor3"/>
      <om:observedProperty xlink:href="http://
myserver/ontologies/ont-core.owl#DissolvedOxygen"/>
      <om:featureOfInterest xlink:href="http://
myserver/ontologies/ont-core.owl#Plemetin"/>
     <om:result xsi:type="gml:SemMeasureType" uom="%">
        <value>47.20</value>
        <sem-annotations>
          <annotation embedded:WaterStatus ="High"/>
          <annotation xlink:href="http://myserver/
ontologies/ont-core.owl#WaterStatus_ClassIV"/>
        </sem-annotations>
      </om:result>
    </om:OM_Observation>
  </observationData>
  . . .
</sos:Observation>
```

• *Semantic Annotation Stream Processor* – is part of the Spark Streaming processor which enable a real-time integration of semantic annotations into heterogeneous sensor stream data with context in the Internet of Things. This component can use sensor metadata, archival data streams, mining data streams, association rules (with concept from ontologies), or other sematic sources for adding semantic annotations to the sensor stream data. The semantic annotated of sensor stream data are stored in *Working Data Stream Annotations*.

C. *Data Modelling* is developed in Apache Cassandra database and contains: *Processor Data Streams, Working Data Streams*, *Working Data Stream Annotations*, *Archival Data Streams*, *Archival Data Stream Annotations*, *Invalid Data Streams*, and *WSNs Metadata:*

• *Processor Data Streams (PDS)* – contains a summary of streaming data for Stream Processor, which can be used for answering queries. For each deployed sensor, only a row is saved which includes: *Sensor Id-* an identifier that uniquely identifies a sensor. *Sensor Parameter* - name of the parameter or phenomena that sensor measure (e.g. temperature, humidity, pH, etc). *Sensor Current Value-* current measured value from the sensor. *Sensor Total Rows* - the total number of measurements by the sensor since its deployment. *Sensor Max Value* - the maximum value measured by the sensor since its deployment. *Sensor Min Value* - the minimum value measured by the sensor since its deployment. *Sensor Sum Value* - the sum of values measured by the sensor since its deployment. *Sensor Avg Value* - the average value measured by the sensor since its deployment, which is derived by finding quotient between sum of values measured by the sensor and the total number of measured by sensor (Sensor Sum Value/ Sensor Total Rows). *Sensor Window Max* - the maximum value of sliding window which contains last *n* values, where *n* is a configurable number (e.g. 15 last measured values) sent by the sensor. *Sensor Window Min* - the minimum value of sliding window. *Sensor Window Avg* - the average value of sliding window. *Sensor Current Timestamp* - current measured timestamp sent by the sensor. *Sensor Current Latitude & Longitude* - current latitude and longitude (current geographical position from where sensor has sent data).

• *Working Data Streams (WDS)* – contains streaming data for operation of Stream Processor, which are configurable according the quantity and can be used for answering queries. So, it's a *Fixed Sliding Window* that contains the last sensor streaming data (e.g. 15 last measured values - its configurable number). For each measured values information is stored as described: *Observation Id* - an identifier that uniquely identifies the observation in the *WDS*. *Sensor Id* - an identifier that uniquely identifies a sensor. *Sensor Parameter* - name of the parameter or phenomena (e.g. temperature) measured by the sensor. *Sensed Value* - measured value that is sent by the WSN. *Timestamp* - time when the sensed value has been generated by WSN. *Latitude & Longitude* - geo location, geographical position from where the sensor has sent data. It is especially useful when a sensor is attached to a moving object such as a car, airplane, etc. or in case of mobile sensor to perform monitoring in different ad-hoc selected locations of interest, while in case of static sensor that perform monitoring in the region of interest, the geo location can be NULL because the location of these senor type can be saved as metadata in WSNs Metadata. *Observation Id* - a code which identifies a single sensor node measurement, example in cases where a sensor node measures three parameters, e.g. temperature, turbidity and conductivity at the same time (as a single measurement), then all of three measurements will get the same transaction code, otherwise the transaction code will be NULL. *Entry Timestamp* - date and time when streaming data has arrived to Stream Processor.

• *Working Data Stream Annotations (WDSA)* – stores semantic annotations of sensor streaming data. The *Semantic Annotation Stream Processor* component, is used to tag in real-time integration the sensor data stream with semantic annotations. One measurement that is stored in *Working Data Streams* can obtain several sematic annotations, which includes information: *Annotation Id* - an identifier that uniquely identifies a sematic annotation. *Observation Id* - references to Working Data Streams observation Id. *Annotated Date* - date and time when sensor streaming data has been tagged with sematic annotations. *Annotated Type* - represents the type of annotation, 'Embedded' (only a single value-scalar of semantic annotation) or 'External' (an external resource linked by 'XLink' that point to our ontology '*ont-core.owl*'). *Annotated Value* - stores the semantic annotated value. Example value of 'Embedded' annotation type for water status can be 'Good', 'Moderate', 'Poor', 'Bad' or 'High':
<annotation mbedded:WaterStatus="Bad"/>
Example value of 'External' annotation type can be (for more details see Figure 2): <annotation xlink:href="http://myserver/ontologies/ont-core.owl#WaterStatus_ClassV"/>.

• *Archival Data Streams (ADS)* – archives data streams for generating reports and different statistics. The structure of data modeling of *ADS* is the same as *WDS*.

• *Archival Data Stream Annotations (ADSA)* – archives semantic annotations of sensor stream data for generating reports and different statistics. The structure of data modeling of *ADSA* is same as *WDSA*.

• *WSNs Metadata (WMD)* – the data describing wireless sensor networks itself, its devices and the corresponding site allocation data. This data is named as static data that describes the wireless sensor networks in the field, its configuration which might involve node types, like sensing nodes, gateway nodes, central monitoring node, and description of sensors as devices (sensor name, serial number, manufacturer, and type), as well as data about the deployment sites, like sensors' location, example for water system monitoring, the river basins, municipalities the rivers belong to, etc.

• *Invalid Data Streams (IDS)* – archives invalid sensor stream data that is classified as outlier by *Outlier Stream Validator* & *Classificator*. The data stored in IDS is optional and it's depends on the system requirements.

D. *OGC Standards* - as mentioned above, the enriched sensor stream data with the semantic annotations results will be published to IoT real-time applications in format of OGC standards, respectively version 2.0 of the SOS O&M standard.

E. *Ontology* - an ontology named '*ont-core.owl*' is created, see Figure 3. Here are developed semantic annotations for international regulatory of water quality, such as:

- United Nations Economic Commission for Europe (UNECE) with semantic annotations of water status: Class I, Class II, Class III, Class IV, and Class V.
- Water Framework Directive (WFD) with semantic annotations of water status: Good, Moderate, Poor, Bad, and High.

Details about the working cycle of this model are as follows: streaming data is sent by the wireless sensors networks in *Input Data Streams* (Apache

Kafka). Sensor stream data is an array of different types containing sensor id (sid), name of the parameter, measured value of the sensor, geographical position (latitude and longitude) and timestamp, as seen below:

```
'SId: 1; Parameter: Temperature;
Value: 17.15; Lat: 42.706703186; Long:
21.038431; Timestamp: 20200312165213'
```

Then, the validation of the stream data elements occurs through *Outlier Stream Validator*, in which every sensor stream data takes the status of validity (*true* – data is valid or *false* – data is outlier). When data takes the validation status '*true*', it will be transmitted for further processing in *Semantic Annotated Stream Processor* which makes real-time integration of semantic annotations into those stream data. Then the enriched sensor stream data with the semantic annotations results will be stored in *WDS* and *WDSA* and it will be transformed in SOS O&M format to displayed in IoT real-time monitoring systems.

It is worth mentioning that when a new value is measured by the sensor, it arrives in *WDS* (and their semantic annotations are stored in *WDSA*) then the oldest value is removed from there and goes to *ADS* (respectively *ADSA*) for archiving. So, in *ADS* and *ADSA* are archived data which serve for generating reports and statistics for longer time frames.

## 5 PROTOTYPE IMPLEMENTATION OF PROPOSED MODEL

To validate the proposed model in this study, we implemented a prototype illustrating a water quality monitoring use case, named as Water Quality Monitoring System (WQMS). This system enables the measurement of water quality in real time by applying the latest technology trends, such as wireless sensor networks, which provide continuous monitoring, and consist of nodes referred to as motes that are sensitive to the environment where they are deployed.

WQMS enables water parameters monitoring such as: temperature, potential of hydrogen (pH), and dissolved oxygen (DO). The type, rank, and unit of these parameters are shown in Table 1. The data produced by the sensor is a real number, e.g. 16.5°C for temperature measurement. The sensors are configured in such a way that each node will send data every 10 minutes. So the rate of the data stream is not high, therefore the storing of these data by applying the proposed model, nowadays, will not be a problem for the new technologies.

Table 1: Specification of water parameters.

| Parameter name | Parameter Type | Range value | Unit |
|---|---|---|---|
| Temperature | Thermal conditions | -1 to +50 | °C |
| pH | Acidification | 0 to 14 | pH |
| Dissolved Oxygen | Oxygenation conditions | 0 to 300 | % |

The WQMS system architecture is presented in Figure 4. It mainly consists of the static wireless sensing nodes, mobile wireless sensing nodes, gateway node and the monitoring node. Static wireless sensing nodes are located in a given position and through gateway node, continuously transmit sensed data to the central monitoring node, while mobile wireless sensing nodes can move from one position to the other to measure water parameters.

The sensor data in the central monitoring node are transmitted via 3G/GPRS using web services. The WQMS software is installed in the central monitoring node, in which is implemented the proposed management model.

## 6 SYSTEM OUTPUTS

To display the enriched sensor stream data with semantic annotations, a web based IoT real-time application is developed. The interface of the software is shown in Figure 5, which enables real time water quality monitoring through static and mobile sensors. This software includes modules for system administration (to manage users, user groups, rights of users, and change password), enables definition of the continuous queries, executes of the ad-hoc queries, and configuration of the WSNs metadata.

WQMS software executes continuous queries of the proposed model to display information. The displayed information in the textboxes for each parameter is obtained from *Processor Data Streams* through continuous queries. The information displayed in the charts is obtained from *Working Data Streams*, while the semantic annotations data that indicates the water status is obtained from *Working Data Stream Annotations*. As mentioned above, *Working Data Streams* represent a fixed sliding window which is configured in the WQMS by a certain pre-configured size, say 15. This means that the charts show the last 15 measurements for each sensor. As soon as the data reaches the system from the WSNs, the trigger for execution of queries continuously is activated.

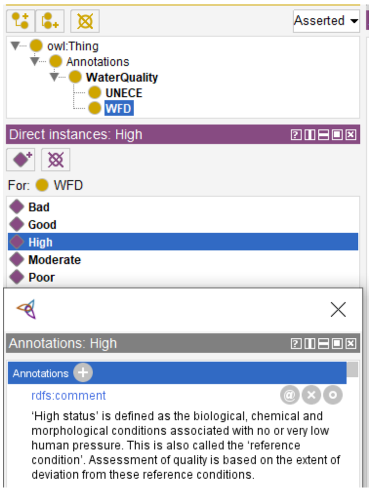


Figure 2: 'ont-core.owl' management model in WQMS.

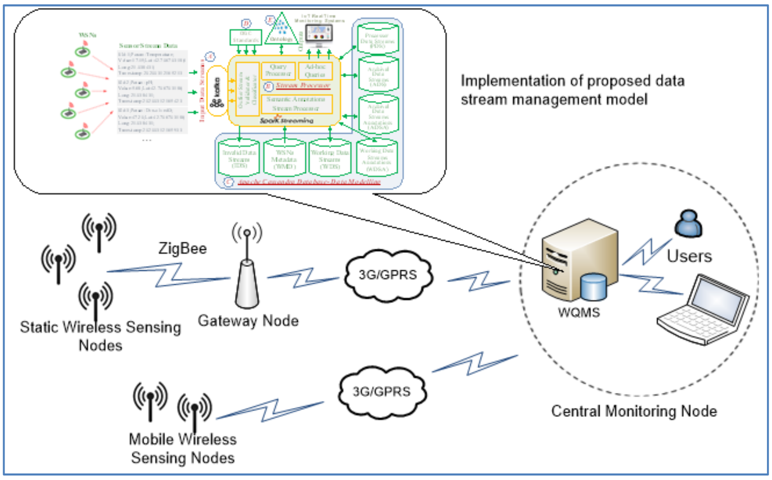


Figure 3: System Architecture: implementation of proposed data stream.

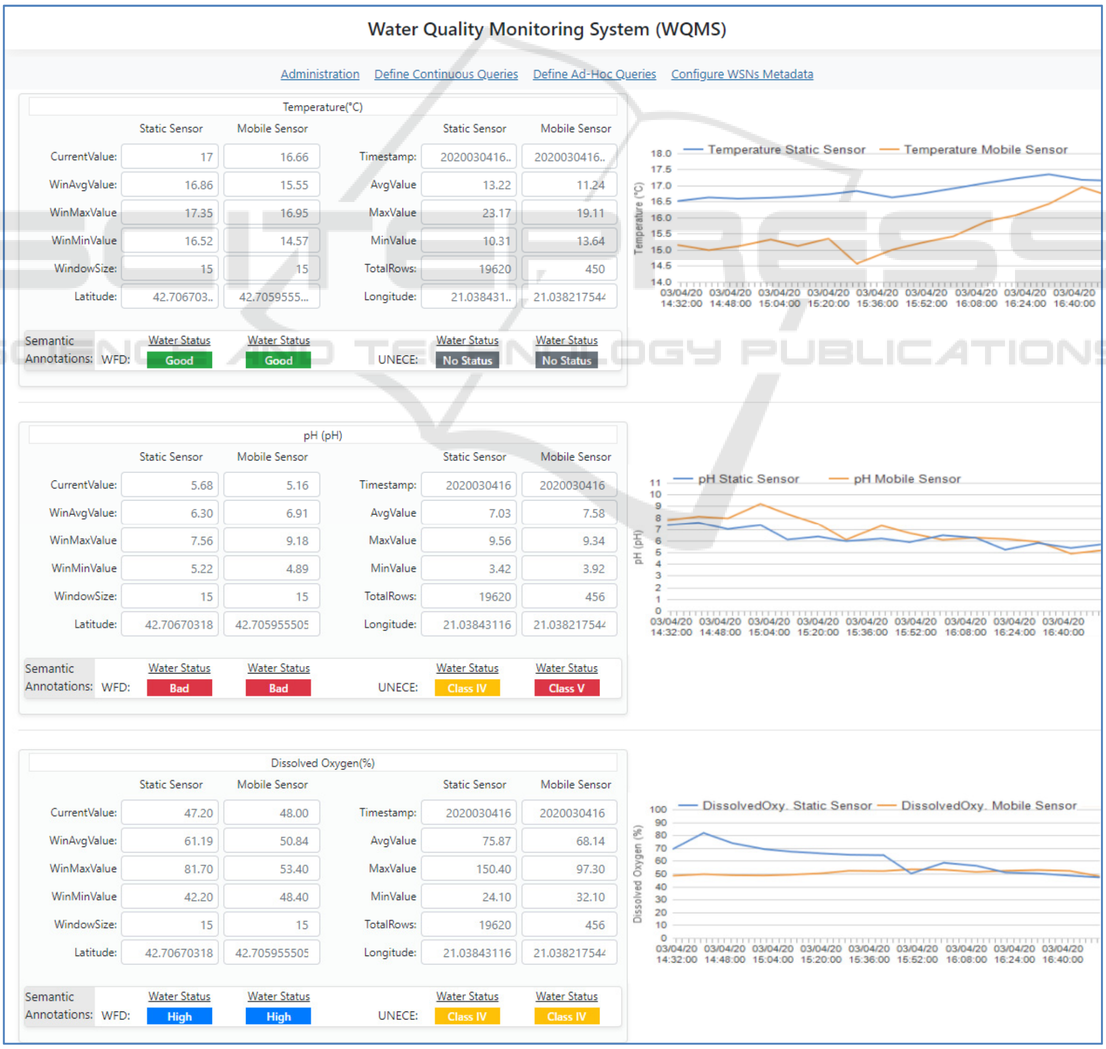


Figure 4: Outputs of proposed model in WQMS.

# 7 CONCLUSION AND FUTURE WORK

In this paper, the proposed model of integrated semantic annotations into the sensor stream data for the Internet of Things is described.

The model supports managing stream data of homogeneous sensors, real-time integration of semantic annotations to the sensor stream data, continuous queries on streaming data, ad-hoc queries, outlier validation of streaming data, archive stream data with semantic annotations for applications that need to answer queries form archival store (persistent data stored).

The model supports the following standards in order to encode semantic annotations and data observed by sensors: Sensor Web Enablement (SWE), respectively version 2.0 of the Sensor Observations Service (SOS) standard that relies on the Open Geospatial Consortium (OGC), Observation & Measurement (O&M).

To validate the proposed conceptual model, we have developed a prototype for water quality monitoring, named Water Quality Monitoring System (WQMS). Applying advanced technologies of the Internet of Things such as WSNs, the WQMS enables water quality monitoring in real time.

Several extensions of the proposed model that can be considered for the future are:

1. To advance annotation techniques, such as XPath, for integration and interpretation of the semantic annotations in real-time into heterogeneous sensor observation data and metadata with context in the Internet of Things.
2. To advance the components *Outlier Stream Validator & Classificator* of the proposed model by implementing some advanced outlier detection algorithms for real time unsupervised anomaly detection.
3. To evaluate the system performance and to compare the proposed model with other existing similar management schemes.

# REFERENCES

Bytyçi, E., Sejdiu, B., Avdiu, A., & Ahmedi, L. (2019). A Semantic Sensor Web Architecture in the Internet of Things. *Semantic Web Science and Real-World Applications* (pp. 75-97). IGI Global.

Elnahrawy E. (2003). Research Directions in Sensor Data Streams: Solutions and Challenges, Rutgers University, Tech. Rep. DCIS-TR-527.

Lazarescu, M. T.. (2017). Wireless Sensor Networks for the Internet of Things: barriers and synergies. *Components and Services for IoT Platforms*. Springer.

Lin, S.Y., Li, J. B., and Yu, Ch. T. (2019). Dynamic Data Driven-based Automatic Clustering. *Sensors and Materials*, Vol. 31, No. 6 (2019) 1789–1801.

Rajaraman, A., Leskovec, J., and Ullman, J. D.. (2014) Mining of Massive Datasets. Cambridge University Press.

Sejdiu, B., Ismaili F., and Ahmedi L., (2020). Integration of semantics into sensor data for the IoT - A Systematic Literature Review. *International Journal on Semantic Web and Information Systems (IJSWIS)*. Volume 16, Issue 4, Article 1.

Sejdiu, B., Ismaili F., and Ahmedi L., (2020). A real-time integration of semantics into heterogeneous sensor stream data with context in the Internet of Things. The 15th International Conference on Software Technologies (ICSOFT 2020). July 07 - 09, 2020, Lieusaint - Paris, France.

Shi, F., Li, Q., Zhu, T., Ning, H. (2018). A Survey of Data Semantization in Internet of Things. *Sensors* 18(1).

Khan I., Jafrin R., Errounda F., Glitho R. (2015). A data annotation architecture for semantic applications in virtualized wireless sensor networks. In *Integrated Network Management*, 2015 IFIP/IEEE International Symposium.

Vera, D., Izquierdo, Á., Vercher, & J., Gómez, L., (2014). A Ubiquitous Sensor Network Platform for Integrating Smart Devices into the Semantic Sensor Web. Sensors 2014, 14, 10725-10752.

Pradilla, J., Palau C., & Esteve, M. (2016). SOSLITE: Lightweight Sensor Observation Service (SOS) for the Internet of Things (IOT). *ITU Kaleidoscope: Trust in the Information Society*, Barcelona.

Wang, X.; Wei, H.; Chen, N.; He, X.; Tian, Z. (2020) An Observational Process Ontology-Based Modeling Approach for Water Quality Monitoring. *Water*, 12, 715.

Xiaomin, Zh., Jianjun, Y., Xiaoci, H., Shaoli, Ch. (2016). An Ontology-based Knowledge Modelling Approach for River Water Quality Monitoring and Assessment. *Procedia Computer Science*, Vol. 96, Pages 335-344.

Yinbiao, S. and Lee, K.. (2014). Internet of Things: Wireless Sensor Networks. *International Electrotechnical Commission (IEC)*, Switzerland.

Yu, K, Shi, W., & Santoro, N. (2020). Designing a Streaming Algorithm for Outlier Detection in Data Mining – An Incremental Approach. *Sensors*.