IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 25, 2019, accepted September 18, 2019, date of publication October 7, 2019, date of current version October 17, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2945124

Extending Data Quality Management for Smart Connected Product Operations

SUNHO KIM[1], RICARDO PÉREZ DEL CASTILLO[2], ISMAEL CABALLERO[2], JIMWOO LEE[3], CHANGSOO LEE[4], DOWNGWOO LEE[5], SANGYUB LEE[5], AND ALEJANDRO MATE[6]

[1]Department of Industrial and Management Engineering, Myongji University, Seoul 449-728, South Korea
[2]Information Technologies and Systems Institute (ITSI), University of Castilla-La Mancha, 13071 Ciudad Real, Spain
[3]2e-Consulting, Seoul 150-010, South Korea
[4]Department of Industrial, Information and Management Engineering, Gangneung-Wonju National University, Gangneung 210-702, South Korea
[5]GTOne, Seoul 07299, South Korea
[6]Lucentia Lab, University of Alicante, 03690 Alicante, Spain

Corresponding authors: Sunho Kim (shk@mju.ac.kr) and Ismael Caballero (Ismael.Caballero@uclm.es)

We would like to acknowledge the financial support provided by the Korea-Spain joint R&D program of MOTIE (Project No. N0002610), CDTI (DQIoT, Project No. INNO-20171086) and EUREKA (Project No. E!11737). Also, we would like to acknowledge the ECLIPSE (RTI2018-094283-B-C31) and GEMA (SBPLY/17/180501/000293) projects.

ABSTRACT Smart connected product (SCP) operation embodies the concept of the internet of things (IoT). To increase the probability of success of SCP operations for customers, the high quality of the IoT data across operations is imperative. IoT data go beyond sensor data, as integrate some other various type of data such as timestamps, device metadata, business data, and external data through SCP operation processes. Therefore, traditional data-centric approaches that analyze sensor data and correct their errors are not enough to preserve, in long-term basis, adequate levels of quality of IoT data. This research provides and alternative framework of data quality management as a process-centric approach to improve the quality of IoT data. The proposed framework extends the process reference model (PRM) for data quality management (DQM) defined in ISO 8000-61, and tailored to fully adapt to the special requirements of the IoT data management. These involve several adaptations: first, the scope of the SCP operations for data quality management is determined, and the processes required for SCP operations are defined following the process description format of ISO 8000-61. Second, the relationship between the processes and the structure of the processes in the technology stack of the SCP operations are described to cover the actual nature of the IoT data flows. Finally, a new IoT DQM-PRM is proposed by integrating the processes for the SCP operations with DQM-PRM. When these processes are executed in the organization, the quality of IoT data composed of data of various types can be continuously improved and the utilization rate of SCP operations is expected to increase.

INDEX TERMS IoT, Internet of Things, SCP, smart connected product, data quality, data quality management, process reference model, ISO 8000-61, DQM, PRM.

I. INTRODUCTION

In recent years, the Internet of Things (IoT) has evolved as one of the trendier technologies to support Industry 4.0, among others. In this sense, by 2020, it is expected that 20 billion IoT products will be used as dedicated-function objects such as aviation engines, weapons systems, automobiles, gas turbines, beverage vending machines, ATM machines, medical equipment, and vacuum cleaners [1]. According to the report by ITU-T Y.2060 [2], the IoT is defined as *a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies*. Smart connected products (SCP) embody the concept of the IoT. A typical SCP consists of 1) physical components that comprise the product's mechanical and electrical parts, 2) smart components that comprise sensors, microprocessors, data storage, controls, software, and, typically, an embedded operating system and enhanced user interface, and 3) connectivity components that comprise ports, antennae, and protocols enabling wired or wireless connections between the product and the product cloud [3]. In a previous study [2], physical components are referred to as physical things, and the integrated parts of smart and connectivity components are referred to as devices.

The associate editor coordinating the review of this manuscript and approving it for publication was Praveen Gunturi.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

The SCP generates, uses and processes a large amount of IoT data during the operational phase. The IoT data are fast, massive, and complex as they consist of various types of data such as sensor data, metadata, external data, and business data. If the IoT data do not have adequate levels of quality, the SCP may malfunction or even stop working. Therefore, the performance of SCP operations highly relies on the quality of the IoT data.

To solve this issue, multiple studies have been conducted to identify characteristics of the IoT data and data quality ('characteristics' are also called 'dimensions' or 'criteria') and improve the data quality in IoT or wireless sensor network (WSN) environments [7]–[11], [22]–[25]. However, most of the studies have focused on improving the quality of sensor data rather than IoT data that include data of various types. It is not desirable to improve the quality of the sensor data only because the sensor data are used in conjunction with metadata, business data, and external data in the SCP operation phase.

In general, data quality improvement can rely on two approaches, namely data-centric and process-centric approaches [5]. The data-centric approach assesses data nonconformities and corrects them, and the data related to them (i.e. data cleansing). This data centric approach can be supported by some international standards like ISO 25012 [20] and ISO 19157 [19] that define data quality characteristics in software quality and geographic information, respectively. This approach has the advantage of rapidly correcting the data nonconformities and quantitatively suggesting changes to attribute values. However, in this data-centric approach, the quality of the data may be degraded over time because the same kind of data nonconformity occurs repeatedly. As a result, this approach works well in short-term basis, while root-causes, i.e., underlying processes that manage data are ignored.

In order to mitigate the drawbacks of the data-centric approach, the process-centric approach that can sustain high data quality is used. This approach maintains data quality management processes that include functions of the data-centric approach, and thereby, continuously improves data quality. This approach avoids the systematic recurrence of the same data nonconformity by eliminating the root causes through data cleansing and business process improvements [6], [26]. Analogously, this approach is also supported by international standards the ISO 8000-60 series specifically for data quality [6], [17], the ISO 9000 series for product quality [18], the ISO 33000 series for process assessment in information technology [21], the CMMI model for software quality [27].

Most of the studies for data quality improvement in the IoT or WSNs have been performed with the data-centric approaches we mentioned. As a common task in data-centric approaches, data cleansing rules or data analytics algorithms are being suggested in various domains [22]–[24]. The QoDID framework [22] proposes a procedure that checks sensor data faults with respect to data quality characteristics such as accuracy, consistency, timeliness, completeness, and validity, and modifies the sensor data by applying data cleansing or IoT-specific rules to improve data quality in the processing layer of the IoT system. As for IoT data analytics algorithms, a data aggregation scheme [23] is proposed for highly uncertain raw IoT sensor data collected in fog cloud servers to remove uncertainties such as noise, outliers and missing values in an unsupervised fashion. A cross validation method [24] is also proposed to seek a validating crowd to ratify the contributing crowd and reshape sensor data for data quality improvement.

However, DAQUA-MASS [11] proposes a process-centric approach, i.e., a process to improve the data quality of sensor data in IoT or WSN environments, which is based on ISO 8000-61 [6]. However, this process-centric approach only considers the improvement of the quality of the sensor data in the acquisition layer where sensor data are collected, which makes it only usable for the initial steps of the lifecycle of the IoT data. Therefore, for a more complete approach, all data included in SCP operations must be considered, such as the metadata, external data, and business data, which are used across the entire IoT data lifecycle, i.e., not only in the acquisition layer but also in the processing layer related to the data storage, analysis, etc., and in the utilization layer where the post-processed data are exploited by the services or applications [9].

Bearing in mind the importance of data quality management in IoT environments, and the growing importance of IoT in future economy, we felt motivated to propose a process-centric framework of the IoT data quality management (DQM). This proposed framework is aimed to improve the quality of the IoT data used in SCP operations in every IoT application layers, i.e., during the acquisition, processing, and utilization activities. The framework considers the PRM for the DQM (referred as DQM-PRM) as defined in ISO 8000-61 and extends it by integrating SCP operation processes.

For the sake of readers' understanding, we will first introduce and summarize not only the characteristics of the IoT and the IoT data required for SCP operations but also the types of IoT data currently defined. Second, an IoT case study is provided to illustrate how some common SCP operations works regarding the equipment installed and operated in a plant. In addition, it determines the scope of the SCP operations for DQM.

The processes required for the SCP operations are then defined to reflect the characteristics of the IoT and IoT data described previously. A process-to-process relationship diagram is presented to show the execution of the processes in conjunction through data flows among them. In addition, the structure of the processes and data flows in the technology stack of the SCP operations is also described. Finally, a new IoT DQM-PRM is proposed by integrating the SCP operation processes into the DQM-PRM defined in ISO 8000-61. One of the proposed processes and an example scenario are provided for a better understanding of the IoT DQM-PRM.

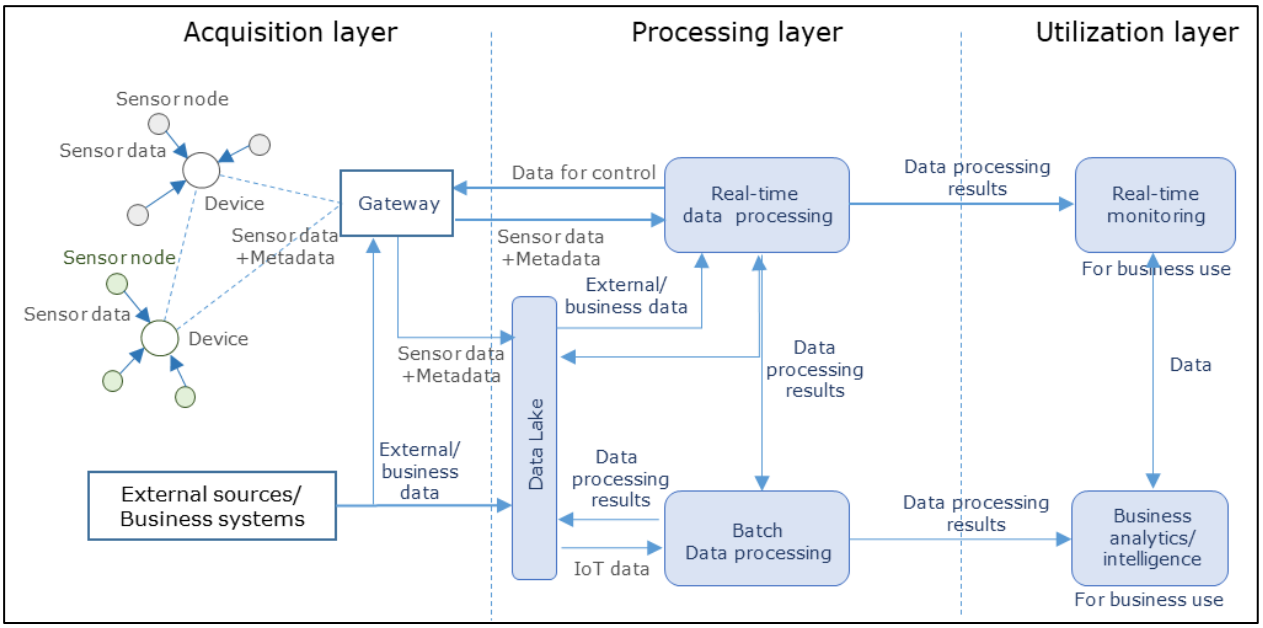

FIGURE 1. IoT data flows in the IoT system layers.

The rest of this paper is organized as follows: Section II presents characteristics of the IoT and IoT data, Section III introduces the types of IoT Data, Section IV defines the processes of the SCP operations and their relationships, Section V describes the DQM-PRM of the SCP operations, and finally Section VI presents the conclusions and future works.

II. CHARACTERISTICS OF IOT AND IOT DATA

This section presents the main characteristics of the IoT and IoT data that affect the quality of the IoT data. According to the ITU-T Y.2060 report [2], the IoT presents fundamental characteristics of interconnectivity with global information and communication infrastructure, things-related services such as privacy protection and semantic consistency between physical things and their associated virtual things, and heterogeneity of devices similar to systems based on different hardware platforms and networks. Furthermore, the states of the IoT change dynamically, and the number of devices and their data volume vary in an enormous scale. Therefore, in conjunction with these IoT characteristics, the IoT should satisfy the high-level requirements that will eventually affect not only the performance of the SCP operations but also the quality of the IoT data. These requirements are summarized in Table 1.

In connection with the IoT characteristics, Karkouch *et al.* [7] defined characteristics of IoT data as 1) uncertain, erroneous, and noisy because of the numerous factors endangering the data quality, 2) voluminous and distributed because the devices are situated in various locations, 3) smoothly varying and continuous, 4) correlated temporally or spatially, 5) periodic in patterns, and 6) having a Markovian behavior in which a sensor value depends on the sensor value generated at the previous timestamp.

III. TYPES OF IOT DATA

SCPs can be connected in large complex networks covering the IoT data lifecycle throughout acquisition, processing, and utilization layers [9]. Under the IoT layers as shown in Figure 1, the IoT data used or generated from connected products consist of various data types as follows:

- Sensor data that are generated by sensors, digitalized in a computer-readable format, and include timestamps. The sensor data are typically used in real time for monitoring and modified for further data analysis and utilization.
- Observed metadata that characterize the behavior of the sensor data such as the information about the rules for data value changes, and the range of data value changes within a certain time period (sometimes referred to as ''dynamic metadata'').
- Device metadata that characterize the device or sensor itself such as the device model, sensor model, manufacturer, precision, unit of measure, value type, maximum and minimum value of sensor (sometimes referred to as ''static metadata'').
- Business data used for a business purpose such as operation, maintenance, and service.
- External data such as weather, traffic, commodity and energy prices, social media, and geo-mapping that informs product capabilities.
- Technical metadata that contain data standards (for example, the rule of table name or field name) and data structures of physical data storage (for example, table name, field name, and field data type). This metadata may be considered part of business data.

The embedded sensors continuously measure the characteristics of a physical object in the acquisition layer. The device transfers the sensor data, observed metadata and device metadata to the processing layer through

TABLE 1. High-level requirements for the IoT.

Characteristics of IoT	High-level requirements
Inter-connectivity	• Identification-based connectivity that supports the connectivity between a thing and the IoT based on the thing's identifier.
Things-related services	• Autonomic service provision by capturing, communicating, and processing automatically the data of things based on rules. • Location-based capability that communications and services depend on the location information of things
Heterogeneity	• Interoperability ensured among heterogeneous and distributed systems. • Autonomic networking to adapt to large numbers and various types of devices, different application domains, and communication environments.
Dynamic and scalable	• Rapid application development that enables instant generation, composition, or configurations for the seamless integration of things with applications. • Scalable big data management of real-time and historic data of things. • Manageability of the entire operation process of IoT applications by the relevant parties.
Security	• Security that ensures confidentiality, authenticity, and integrity of both data and services. • Privacy protection during data transmission, aggregation, storage, and processing of private information. • High quality and highly secure human body related services based on laws and regulations

the connectivity. In the processing layer, the sensor data are analyzed by real time, and the analysis results are fed back to the device in the acquisition layer to control the physical object or applied to real-time monitoring for business use in the utilization layer. The sensor data are also modified for further batch data analysis in the processing layer. The analysis results are applied to the real-time data analysis in the processing layer or to business use such as business analytics/intelligence and services in the utilization layer. The external and business data coming from external sources and business systems can be used in all layers in combination with the sensor data when necessary.

As the global level quality of the IoT data is influenced by the quality of each type of IoT data, it is not desirable to improve only the quality of sensor data. A process-centric approach can sustain high data quality by improving the quality of the IoT data that, as already said, consist of various data types. In the following sections, the scope of SCP operations and corresponding processes that are subject to IoT data quality management are proposed, and the DQM processes that can improve the IoT data by integrating the SCP operation processes with the DQM-PRM defined in ISO 8000-61.

IV. PROCESSES FOR SCP OPERATIONS

IoT use cases vary widely depending on the applications such as transportation, home, offices, factories, process plants, and healthcare [12]. For a better understanding of SCP operations, let us drive our explanation through an example of a power production plan using gas turbines. In this case, the plant is difficult to maintain with this equipment due to a lack of relevant expertise on equipment manufactured by other companies. Therefore, the manufacturer who has the necessary expertise and information is the most qualified to manage the equipment in the plant. For effective and efficient management of the equipment, the manufacturer can deploy an IoT system with the product cloud [3]. The sensors attached to the equipment form a local mesh network, and the gateway of this network is connected via the Internet to the product cloud that stores IoT data, processes them, and provides service responses.

The example case comprises the monitoring and control by remotely suited operators or systems, remote or on-site maintenance by repairers, and service parts order by procurement staff. While managed remotely, appropriate access protocols such as authentication and authorization should be operated to ensure security and privacy. Confidentiality and/or availability may also be required for security and privacy when a critical remote management operation is performed.

A typical process for SCP operations that monitors, controls, and maintains SCPs can be depicted in Figure 1. The SCPs are connected to users, product cloud, external sources, business systems, or other SCPs when necessary through the networks. The connected products collect raw sensor data such as location, condition, and use from the device of the SCPs, external data such as weather, traffic, and geo-mapping from external sources, or business data such as service history and warranty status from business systems through the networks. The raw data are later aggregated to the data lake in the product cloud which is used for applications in the processing and utilization layers. These data are analyzed by using various analytics, and the results are used for monitoring and control, services/ maintenance, performance improvement, or other business purposes. The scope of the SCP operations for the IoT data quality management is limited exclusively to the processes that are performed during the internal operations of SCPs in the product lifecycle, laying out of the scope of the processes that generate and manage in external sources, other SCPs, or business systems.

A. PROPOSED PROCESSES FOR SPECIFIC SCP OPERATIONS

For our investigations we decide to gather a set of processes that could be required for SCP operations. We depict this set from the best practices for product lifecycle management (PLM) [13]. The processes in this set include:

- Equipment Monitoring and Maintenance Management (EMMM);
- Interactive Service Diagnostics and Responses (ISDR);
- Service Order Management and Field Service (SOFS);
- Warranty Compliance Management (WCM);
- Remote Software Update Management (RSUM);

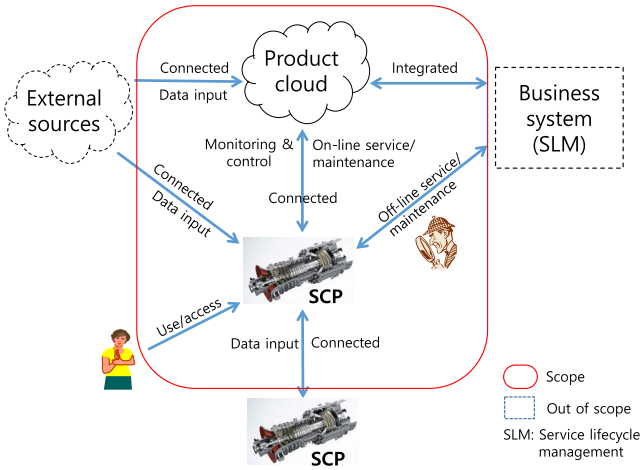

FIGURE 2. Scope of processes for SCP operations.

- Connected Product Failure Analysis (CPFA);
- Smart Connected Product Enablement (SCPE).

Herein, the process titles capitalize the first letter of the words to distinguish them from normal words. These processes reflect the characteristics of IoT and IoT data and the high-level requirements described in Section 2. Here, only the processes required for the operation phase during the SCP lifecycle are considered. The processes such as connected product usage analysis and connected causal forecasting are excluded from the proposed processes as they are considered part of the business systems for product design improvements and service parts forecasting to determine inventory levels, respectively.

The following subsections introduces the proposed seven processes. Each proposed process is described by the following elements: 1) a descriptive title, 2) the purpose which describes the goal of performing the process, 3) the outcomes which express the observable data or results expected from the successful performance of the process, and 4) the activities which are actions that can achieve the outcomes. This conforms to the process description format of ISO 8000-61 [6].

1) EQUIPMENT MONITORING AND MAINTENANCE MANAGEMENT (EMMM)

a: PURPOSE

The purpose of EMMM is to continually monitor and capture the equipment conditions and event information across the installed base, pre-emptively detect and diagnose early indications of potential issues, and deliver service recommendations to minimize product failures and downtime.

b: OUTCOMES

– The physical characteristics of the connected products are tracked and diagnosed in real time and remotely by location, performance, utilization, and condition to pre-emptively identify potential connected product issues and approaching preventive maintenance events.

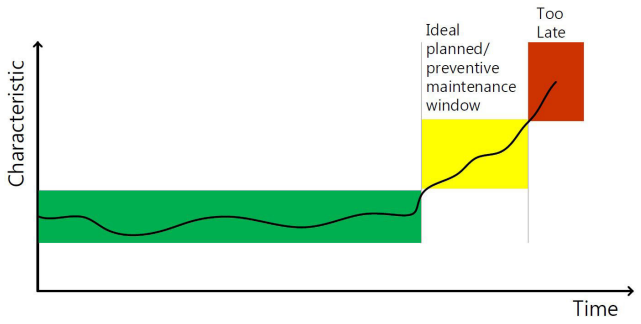

FIGURE 3. Three stages of the EHM [14].

Example: The battery lifetime of devices should be tracked to prolong the communication between connected products and the IoT.

– Notifications are issued to initiate service responses that minimize the downtime and avoid potential product failures.
– Service/maintenance responses are initiated based on the thresholds and trends that indicate a potential failure of a connected product and rules that drive service/maintenance decisions.
– The asset and fleet history data such as product as-maintained configuration, utilization, and service/maintenance history are updated based on the connected product data and service events performed across the service network.
– The asset and fleet history data are analyzed for contract compliance and optimization, deferred and approaching service/maintenance, and product replacement opportunities.

c: ACTIVITIES

– Real-time equipment health monitoring (EHM): Provide connected product monitoring to pre-emptively identify potential product issues and approaching preventive maintenance events. Issue notifications to initiate service responses that minimize the downtime and avoid potential product failures.
Note: This is a real-time model-based activity. As shown Figure 3, the EHM monitors steady state physical characteristics of the equipment and provides an alarm when these characteristics deviate from the pre-defined ranges. The data faults may then be quickly investigated, and the problem corrected before it becomes serious and the equipment shuts down.
– Condition-based maintenance (CBM): Monitor connected product conditions to identify the thresholds and trends that indicate potential product failures. Automatically initiate service responses to minimize the downtime and avoid potential failures.
Note: The CBM is a maintenance method that monitors the actual condition of the connected product to decide if maintenance is required. The CBM dictates that main-

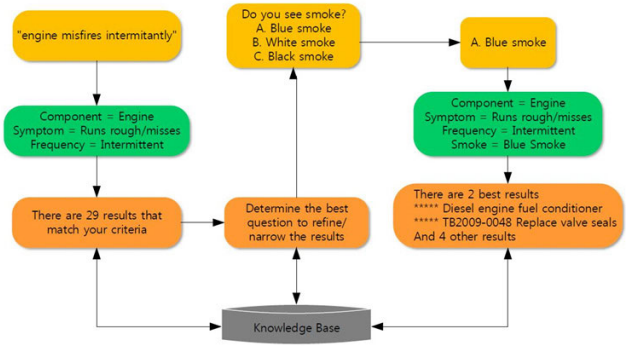

FIGURE 4. Example of a knowledge-based interactive diagnosis process.

tenance should only be performed when certain physical characteristics show signs of decreasing performance or upcoming failure. Compared with planned or preventive maintenance, this increases the time between maintenances as the maintenance is done only when required.

– Connected asset and fleet management: Automatically update product as-maintained configuration, utilization, and service history based on connected product data and service events performed across the service network. It analyzes the asset and fleet history for contract compliance and optimization, deferred and approaching service requirements and product upsell and replacement opportunities.

2) INTERACTIVE SERVICE DIAGNOSTICS AND RESPONSES (ISDR)

a: PURPOSE

The purpose of ISDR is to examine the causes of reduced product performance or failure by interactive diagnostics leveraging real-time and trending product information and initiate proactive service responses. Figure 4 shows an example of a knowledge-based interactive diagnosis process. In this process, when a connected product or an operator inputs anomalous symptom in the engine, the diagnostic function searches the knowledge base for all possible causes and suggests the most probable causes and service responses through additional questions.

b: OUTCOMES

– The procedure and knowledge base for the interactive service diagnostics and responses are provided in the system.
– The product issues are identified and diagnosed interactively with connected products using real-time product information at the contact center, field service, depot operations, and via self-service
– Based on the diagnostic results, the best service responses for the troubleshooting are initiated.
Note: Service response examples: Perform remote services, conduct additional remote diagnostics, dispatch a service technician, provide customer ''self-service'' specific actions.

c: ACTIVITIES

– Connected interactive diagnostics: Interact with connected products to diagnose the product issues following the procedure provided by the knowledge base for service diagnostics and responses.
– Connected service responses: Initiate proactive service responses corresponding to the diagnostic results to eliminate unnecessary service calls, avoid product failures and downtime and improve customer satisfaction.

3) SERVICE ORDER MANAGEMENT AND FIELD SERVICE (SOFS)

a: PURPOSE

The purpose of SOFS is to perform service order events across all service network resources based on part and resource availability to maximize the service level agreement (SLA) compliance and customer satisfaction.

b: OUTCOMES

– The best service response is determined through the diagnosis of issues related to the service events notified pre-emptively.
– An order for the service response is issued to correct issues based on the SLA compliance and resource availability across the service networks.
– The service is performed on-site and/or remotely:
  • Technicians are dispatched on-site and correct the issues;
  • The remote service is performed for machine adjustments, software updates, self-tests, etc.
– The service information and completion status are recorded after completion of the service.

c: ACTIVITIES

– Automated service execution: Automatically trigger service events based on preemptive connected product alerts, diagnose issues, determine the best service response, and dispatch technicians based on the SLA entitlements and resource availability to correct issues before product failures
– Connected remote service: Interact in real time with the connected product to perform remote service activities including machine adjustments, software updates, and self-tests to avoid downtime and eliminate the need for on-site service calls.

4) WARRANTY COMPLIANCE MANAGEMENT (WCM)

a: PURPOSE

The purpose of WCM is to define the warranty and contract policies for connected products based on predicted reliability and/or contracts with customers and comply with the warranty and contract policies based on actual reliability and usage.

b: OUTCOMES

- The warranty and contract policies for connected products are defined based on the predicted reliability and/or warranty/service contracts with customers.
  Note: Manufacturer warranty policies are pre-defined in the system. These policies comprise at least the following parameters:
  - Duration of the warranty coverage;
  - Usage (mile/km/hour) allowed under the warranty;
  - Warranty start date – sale date, delivery date, and first used date;
  - Covered components/parts;
  - Reimbursement rate – 50% on parts, 100% on labor, 50% of travel, etc.;
  - Applicability – model, make, year, etc.;
  - Effective dates.
- Connected product usage is monitored, recorded and compared with the warranty policies to identify potential warranty compliance issues.
- The compliance issues identified are automatically notified to operators to avoid potential product failures and warranty issues.

c: ACTIVITIES

- Connected product usage monitoring: Continually monitor the connected product usage to identify potential warranty compliance issues.
- Warranty compliance issues notification: Automatically notify the operator when compliance issues occur to avoid potential product failures and warranty issues.

5) REMOTE SOFTWARE UPDATE MANAGEMENT (RSUM)

a: PURPOSE

The purpose of RSUM is to remotely manage the as-maintained configuration of connected products, systems, and assets and automatically update software systematically.

b: OUTCOMES

- The service bill of materials (BOM) and configuration are managed remotely with the controlled access of relevant views which capture and trace the updates of the field product, systems, and assets.
- The software updates and security patches are automatically installed remotely.
- The packages of software instructions are distributed remotely to field products, systems, and assets.

c: ACTIVITIES

- Connected remote product configuration management: Remotely identify and manage the as-maintained configuration of the fielded products, systems, and assets.
- Connected remote software update: Automatically install software updates and security patches and easily distribute packages of software instructions remotely.

6) CONNECTED PRODUCT FAILURE ANALYSIS (CPFA)

a: PURPOSE

The purpose of CPFA is to conduct a systematic analysis of the connected product failure data and provide the analysis results for performance improvements.

b: OUTCOMES

- The smart connected product failures and objective data captured directly from a system or device are identified and recorded.
  Note: The relevant objective data captured include incident log data such as operating conditions, environmental conditions, system state at the time of failure, and time of failure.
- Through analysis tools or methods, real-world failure rates and reliability metrics are generated for performance tracking and comparison with expectations or contractual requirements.
- The analysis results are provided to improve the root cause analysis and corrective actions, product quality, reliability and safety, preventive maintenance, and service.

c: ACTIVITIES

- Connected product failure data collection: Collect field failure data through remote connectivity with sensor-equipped products or systems.
- Failure data analysis: Analyze the collected field failure data to improve the root cause analysis and corrective actions, product quality, reliability, preventive maintenance, and service.

7) SMART CONNECTED PRODUCT ENABLEMENT (SCPE)

a: PURPOSE

The purpose of SCPE is to enable the organization to securely connect and efficiently create values from smart and connected products through end-to-end IoT solutions.

b: OUTCOMES

- Identification-based connectivity: The IoT supports that the connectivity between a device and the IoT is established with the device identifier. This includes that possible heterogeneous identifiers of the different devices are processed in a unified manner.
- Location-based capability: Device-related communications and services depend on the location information of the devices and/or users. It senses and tracks the location information automatically.
  Note: Location-based communications and services may be constrained by laws and regulations and should comply with security requirements.
- Interoperability: Interoperability is ensured among heterogeneous and distributed systems for provision and consumption of a variety of information and services. This ensures the identification of devices or real-time

operating systems which support simple and secure paths for data exchange between the devices and product cloud, remote software upgrade, etc.
- Autonomic networking: Autonomic networking (including self-management, self-configuring, self-healing, self-optimizing and self-protecting techniques and/or mechanisms) is supported in the networking control functions of the IoT in order to adapt to different application domains, different communication environments, and large numbers and types of devices. Communication protocols are identified to facilitate real-time data exchange between the device and the IoT.
- Autonomic services provisioning: The services are provided by a complex event processing that collects, communicates, and processes automatically the real-time data generated from multiple devices based on the rules and analysis configured by operators or customized by subscribers.
  Note 1: Autonomic services may depend on the techniques of automatic data fusion and data mining.
  Note 2: Complex event processing includes data generation from devices, collection of the data, queuing of events, transformation of events (data analysis or stream processing), preparation for a long-term storage of events, and multi-format presentation of the analyzed events for reporting or further action [15].
  Note 3: Platforms are available for real-time data processing such as Apache Hadoop and Azure which can ingest, process and store millions of events per second.
- Scalability: A big-data database system is supported to enable aggregation, normalization, and management of real-time and historical product data sent in the form of streams. Big-data storage and data manipulation technologies that leverage a collection of horizontally coupled resources are provided to achieve a nearly linear scalability in performance.
  Note: The data storage automatically provides partitions, normalization, backups, fast reads, and writes and ensures the high availability, distribution transactions, and consistency of data.
- Rules/analytics engine: The rules/analytics engine is supported to establish rules, business logic, and big data analytical capabilities that populate the algorithms involved in product operation and reveal new product insights.
  Note 1: This engine may use machine learning on data to facilitate data analytics.

c: ACTIVITIES
- Flexible IoT product connectivity: Provide means to leverage proved agent technology and connectivity services to easily and flexibly connect to any wired or wireless asset via third-party device clouds, direct network connections, open APIs or edge devices.
- Scalable IoT data management: Establish a highly scalable system for complex event processing and data storage to aggregate and manage large volumes of unstructured, time series, and transactional data from people, systems, and things
- Simplified IoT data analytics: Establish rules, business logics, and algorithms that analyze and correlate unstructured, time-series and transactional data to improve sensor data quality, optimize business processes and discover new opportunities and insights that answer key business questions.

B. RELATIONSHIPS BETWEEN THE PROCESSES OF SCP OPERATIONS

The processes for SCP operations have a relationship to each other in the product cloud. The relationship reveals data flowing among the processes, and additionally, explains how the processes are implemented to provide services for SCP operations. The processed data evolve into information as it progresses through the processes. The output information from one process is used as an input to the other processes. The information includes improvement figures, specific facts and notifications.

The relationships between the processes are illustrated in Figure 5. The process of SCPE performs the processing of data and provides functional support of agent technology, connectivity services, a highly scalable system for complex event processing and data storage, and simplified IoT data analytics to identify sensor data faults and improve the sensor data quality in the product cloud. The remaining processes utilize the processed data to deliver high-performance SCP operations. The EMMM performs monitoring connected product conditions, issues a notification when potential product issues and preventive maintenance events are pre-emptively identified, and initiate service responses for maintenance when the thresholds and trends that indicate potential product failures are identified. When product issues such as reduced product performance or failure are found in the EMMM, the ISDR begin to diagnose the monitored product issues and initiate service responses to improve product performance or avoid failure. The service responses are executed online by a remote service or offline by operators/technicians across all service network resources by the SOFS. Before executing service responses, warranty compliance issues are checked from the WCM. After completing the service, the results are sent back to the EMMM, which automatically updates product as-maintained configuration, utilization, and service history with the activity of connected asset and fleet management.

The CPFA receives failure data from the EMMM and the ISDR. This process analyzes the failure data and applies the analysis results to the improvement of the root cause analysis and corrective actions, product quality, reliability, preventive maintenance, and services. The WCM monitors the connected product usage based on the warranty and contract policies, receives analyzed asset and fleet history information from the EMMM and identifies the potential warranty compliance issues of the connected product usage.

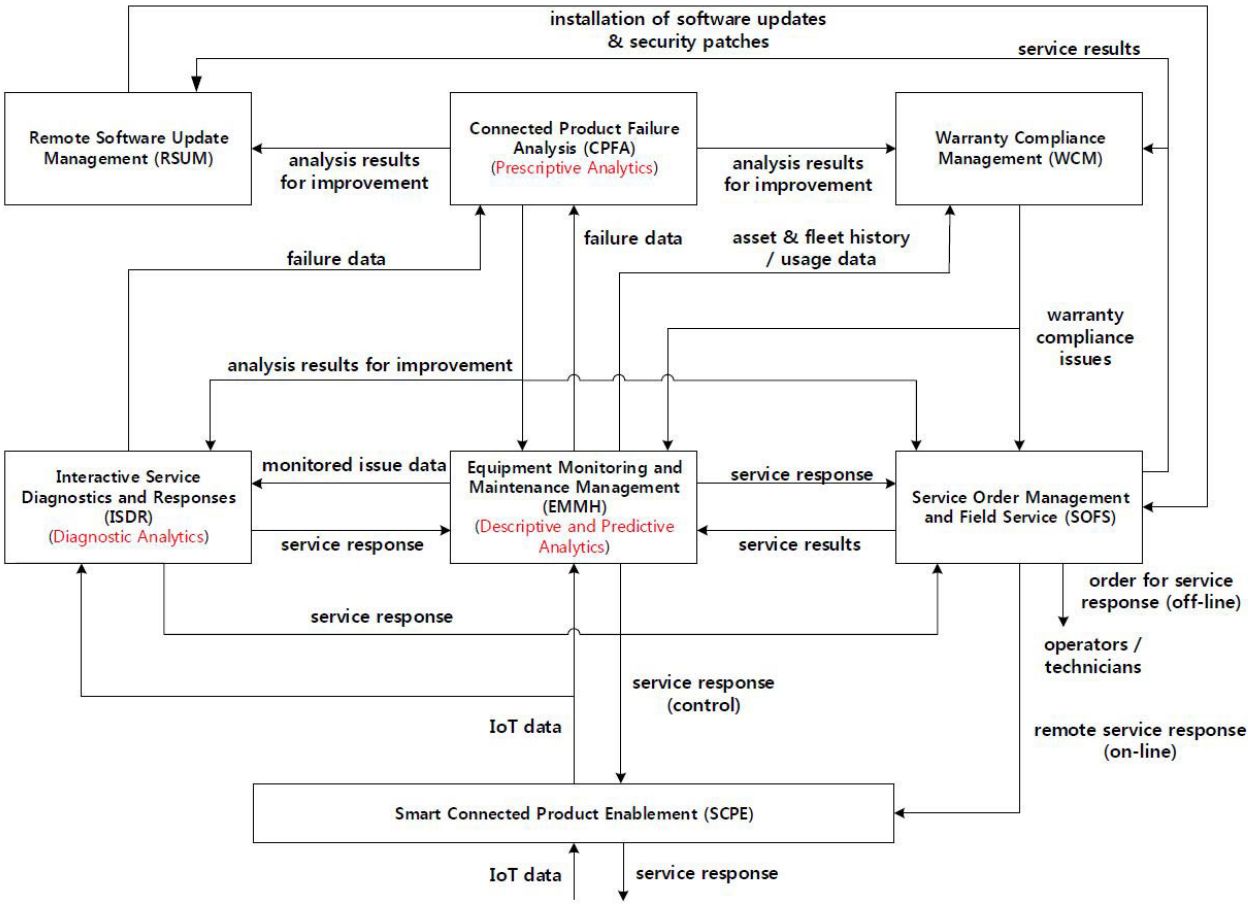

FIGURE 5. Relationships between the processes of the SCP Operations.

This process also notifies the warranty compliance issues to the EMMM and SOFS to avoid potential product failures and warranty issues. The RSUM executes the installation of the software update and security patches remotely or by offline through the SOFS and is fed back service results.

The data are analyzed by descriptive, diagnostic, predictive, and prescriptive analytics [4]. All the analytics are essential to improve not only the quality of sensor data but also the utilization of the SCP operations. The descriptive analytics is a type of analysis which provides performance insights, for instance, capturing connected product conditions, operations, and environments. The diagnostic analytics investigates the causes of poor performance or failure of a connected product or answers the question what is happening. The predictive analytics predict the future trends or answers the questions what will happen. Therefore, the predictive analytics detects patterns that indicate impending events in the SCP operations. Finally, the prescriptive analytics identify the methods to improve possible outcomes or correct problems. The data analytics is supported with rules, business logics, and algorithms by the activities of simplified IoT data analytics in the SCPE. The tasks of descriptive and predictive analytics are performed by the EMMM, and those of diagnostic and prescriptive analytics by the ISDR and the CPFA, respectively.

C. THE TECHNOLOGY STACK OF SCP OPERATIONS

The technology stack of the SCP consists of physical, smart, and connectivity components [3]. The physical components of the SCP can be electrical and mechanical, and the smart components comprise the sensors, microprocessors, data storage, controls, software, an embedded operating system, and enhanced user interface. The connection of the SCP is achieved through ports, antennae, and protocols that enable wired or wireless connections between the product and product cloud. The connectivity could be one to one, one to many, and many to many. The product cloud consists of a big-data database system, an analytics engine, an application development and executing environment, and software applications that manage the monitoring, control, optimization, and the autonomous operations of the product functions. The technology stack of the SCP comprises the embedded product hardware and software, connectivity protocols, product cloud, as well as the components for identity and security tools, a gateway for information from external sources, and integration capability with enterprise business systems [3].

Figure 6 illustrates the SCP operation processes and data flows in the technology stack of the SCP that consists of the acquisition layer, processing, and utilization layers:

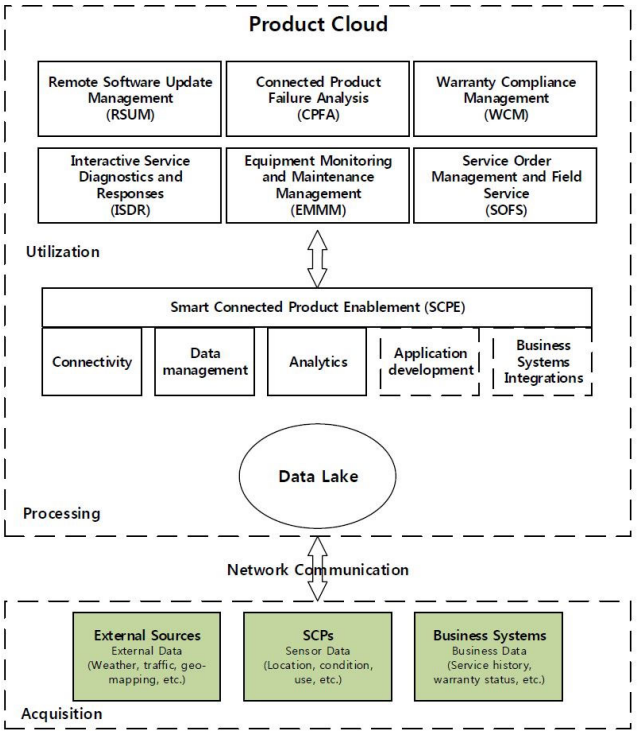

FIGURE 6. Technology stack of the SCP operations.

- In the acquisition layer, sensor data are collected from sensors, external data from external sources and business data from business systems such as SLM, PLM, or ERP.
- In the processing layer, the sensor data are transferred to the data lake that aggregate raw data in multiple formats through a network communication and the activity of the flexible IoT product connectivity in the SCPE. The collected sensor data and related business and external data are managed in the data lake for real-time or batch data processing by the activity of scalable IoT data management in the SCPE and analyzed by the activity of simplified IoT data analytics in the SCPE. In the utilization layer, the analyzed data are used for the SCP operation processes such as monitoring, diagnosis, services/maintenance, warranty compliance and further analysis.

In order to efficiently support the SCP operations, the SCPE requires the activity of rapid IoT application development that leverages an IoT platform with a model-based development environment and the activity of seamless IoT business systems integrations that rapidly integrates business systems and external data sources. However, herein these activities are not considered as they are related to the service quality but not to data quality.

V. DATA QUALITY MANAGEMENT FOR SCP OPERATIONS

The processes for SCP operations presented in the previous section focus on increasing the utilization of SCPs or sustaining the quality of sensor data by monitoring and controlling the stream sensor data. For example, the EMMM monitors

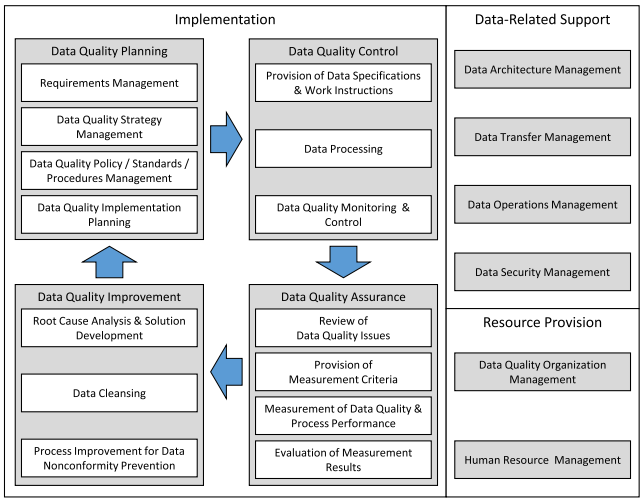

FIGURE 7. Detailed structure of the DQM-PRM in ISO 8000-61 [6].

specific sensor data generated by SCPs, identifies potential product issues, approaching preventive maintenance events and potential product failures in advance, and maintains the sensor data in the normal range through service responses. However, ordinary SCP operations do not include any process to solve data nonconformities that occur other than the sensor data within the IoT data. For example, if data nonconformities exist in the information of connected asset and fleet such as the service BOM (Bill of Materials) or product as-maintained configuration information provided by the business system, on-line/off-line maintenance or software upgrade is not performed properly and causes serious problems. To prevent this, DQM processes are required to improve the quality of IoT data, including not only sensor data but also the other structured or semi-structured data among the IoT data.

Therefore, in this section, we present our main proposal, an IoT DQM-PRM that accommodates both the SCP operation processes to improve the quality of stream sensor data and the DQM processes to improve the quality of the other structured or semi-structured data among the IoT data. First, the DQM-PRM defined in ISO 8000-61 is introduced. Second, the IoT DQM-PRM is proposed by integrating the SCP operation processes with the DQM-PRM through the sensor fault management procedure. Finally, one of the proposed processes and its work products are provided to show how the outcomes and activities are integrated and what data types are interrelated in the process. An example scenario is also provided to explain how the processes of the IoT DQM-PRM contribute to IoT data improvement.

A. PROCESS REFERENCE MODEL FOR DATA QUALITY MANAGEMENT IN ISO 8000-61

The DQM-PRM defined in ISO 8000-61[6], as a part of the DQM series, specifies the processes required for the efficient and effective management of data quality. This model is used as a reference to assess and improve data quality management at the organizational level for digital data sets that are

structured or semi-structured. The model requires fundamental principles of process-centric approach, continuous improvement by process, and involvement of people to improve the data quality. As shown in Figure 7, the structure addresses the following processes:

- Implementation to improve the data quality;
- Data-Related Support to supply data-related information and technology to Implementation;
- Resource Provision to supply resources and human training to both Implementation and Data-Related Support.

The Implementation process is performed based on the plan-do-check-act cycle [18], i.e., Data Quality Planning (plan), Data Quality Control (do), Data Quality Assurance (check), and Data Quality Improvement (act). The previous process or sub-process in the plan-do-check-act cycle is performed as a base for the establishment of the subsequent process. However, the sub-processes of the data-related support and resource provision can be used individually when required. In the Implementation process, measurements and improvements of the data quality and process performance should be considered without disruption of the normal functioning of the organizational business process.

In order to extend the DQM-PRM for the SCP operations, it must integrate their specific processes. The following subsection provides the concepts of this integration.

B. INTEGRATION OF THE SCP OPERATION PROCESSES WITH DQM-PRM

According to ISO 8000-61, the Data Processing process in DQM-PRM is an integral part of business processes of any organization that uses this model and this core process, is performed by end users across the organization. Therefore, the Data Processing process is specific to the business processes of the organization and usually varies with the domain where ISO 8000-61 is applied. In this sense, the proposed processes for the SCP operations in Section IV.A should be integrated as part of the Data Processing process. However, these proposed SCP operation processes are performed to monitor abnormal symptoms of connected products, identify sensor data faults and their causes, and remove the causes to improve the data quality as well as the utilization of the connected products.

In general, sensor data faults can be managed by the following procedure:

Step 1: Identify target profiles to manage sensor data faults.
Resultant data: physical characteristics of the connected products to be monitored based on planned or preventive maintenance.

Step 2: Monitor the conditions of the physical characteristics of the connected products in operation and return symptoms by detecting abnormal conditions at runtime.
Resultant data: 1) connected product conditions and symptoms that pre-emptively identify the potential product issues and approaching preventive maintenance events, 2) connected product conditions and symptoms to identify thresholds and trends that indicate potential product failures or performance degradation, 3) connected product conditions and symptoms that cause reduced performance or failure.

Step 3: Identifies product issues or the types of sensor data faults occurring by the returned symptoms.
Resultant data: product issues or the types of sensor data faults that cause reduced product performance, failure, potential failure, or potential performance degradation.

Step 4: Determines the causes that lead to the identified sensor data faults.
Resultant data: causes identified by the diagnosis procedure and knowledge base.

Step 5: Performs service responses to remove the determined causes. If the causes are removed successfully, Go to Step 1 to modify the information of the targets. Otherwise, the sensor data faults occur in other targets, and go to Step 3 to identify the faults.
Resultant data: service responses such as planned maintenance, preventive maintenance, condition-based maintenance, fault data corrections, and software updates.

This procedure has been adapted from reference [16] and modified to suit the SCP operations by using terms specific to the SCP operations and adding resultant data to every step. In this procedure, Step 1 to Step 3 monitor the connected product conditions and identify product issues or types of sensor data faults. Therefore, these steps can be performed by the EMMM for the SCP operations. Step 4 performs a diagnosis to identify the causes of product issues or sensor data faults, and therefore, corresponds to the ISDR. As Step 5 performs the service responses to remove the causes of sensor data faults, this step corresponds to the SOFS in the SCP operations. This relationship is illustrated in Figure 8.

From the viewpoint of data quality management, the procedure to manage sensor data faults is like the DQM-PRM processes as it identifies data nonconformity that encompasses sensor data faults, analyzes the corresponding root causes, and improves the data quality by removing the root causes. Therefore, as shown in Figure 7, the EMMM which identifies the sensor data faults can be included in the Data Quality Monitoring and Control of the DQM-PRM. The ISDR that determine the causes of the sensor data faults can be included in the Root Cause Analysis and Solution Development of the DQM-PRM. Moreover, the SOFS that removes the causes of the sensor data faults can be included in the Process Improvement for Data Nonconformity Prevention of the DQM-PRM that removes the root causes of the data nonconformity by improving the corresponding process.

Furthermore, the CPFA is considered as part of the evaluation of measurement results of the DQM-PRM that quantitatively analyzes the measurement results of the data

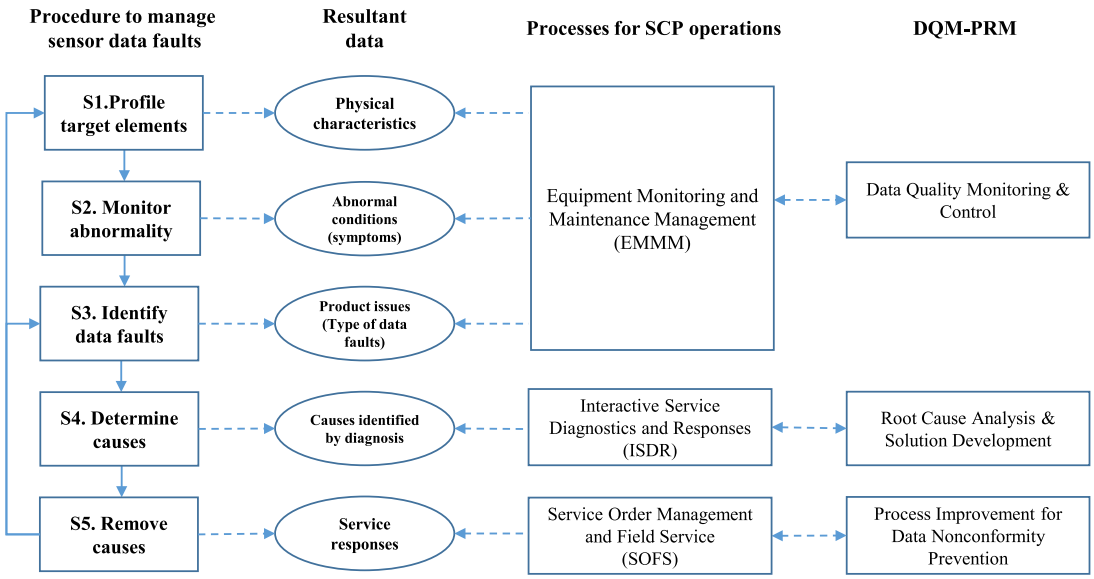

FIGURE 8. Mapping of the SCP operation processes to the DQM-PRM.

quality and process performance. The remote software update management can be included in the data operations management of the DQM-PRM as it comprises a software update task. The WCM is included in the data quality monitoring and control of the DQM-PRM as it continually monitors the connected product usage to identify potential warranty compliance issues. On the other hand, the SCPE is used to enable the organization, securely connect, and efficiently create value from smart connected products. Its function is similar to that of the data operations management of the DQM-PRM due to the database connectivity, data exchange, etc., but also different due to the flexible IoT product connectivity, scalable IoT data management, and IoT data analytics. Accordingly, the SCPE is included as an individual process in the data-related support of the DQM-PRM.

The entire view of the IoT DQM-PRM integrated with the processes for the SCP operations is represented in Figure 9. The name of each SCP operation process is prefixed with the IoT symbol to distinguish them from the DQM-PRM process.

C. AN EXAMPLE PROCESS OF IOT DQM-PRM

In this section, the Data Quality Monitoring and Control (DQMM) in the IoT DQM-PRM is presented as an example to show how SCP operation processes are integrated with DQM, especially in outcomes and activities. Other processes are no longer described in detail as they are integrated in the similar way. The DQMM also conforms to the process description format of ISO 8000-61 [6]. Work products required for and generated from the DQMM are also provided to show input and output data to and from the process. Herein, outcomes, activities and work products related to the SCP operation processes are prefixed with the IoT symbol, which implies they should be further considered in the SCP operations environment.

1) DATA QUALITY MONITORING AND CONTROL(DQMM)

a: PURPOSE

The purpose of the DQMM is, by following applicable work instructions, to identify and respond when the data processing fails to deliver data that meet the requirements in the corresponding data specifications.

b: OUTCOMES

① The risks are identified and quantified against the data specifications covering the corresponding impacts on the organization, systems, or other stakeholders.
IoT Based on the preventive maintenance plans, the risks are identified and quantified against the data specifications.

② The priorities are identified according to the monitoring and control of risks.
IoT The warranty and contract policies for connected products are prioritized for monitoring and control of risks.

③ Records are kept for comparing performance with planned results for processes monitored the with respect to identified risks.
Note: The comparison of performance may occur at intervals or continuously.

④ End users are notified when planned results are not achieved for processes so that they can follow data specifications and work instructions more effectively in implementing and maintaining the processes.

⑤ The data nonconformity is identified, classified, and corrected.
IoT The connected products are tracked and diagnosed in real time and remotely by location, performance, utilization, and condition to pre-emptively identify potential connected product issues and approaching preventive maintenance events.

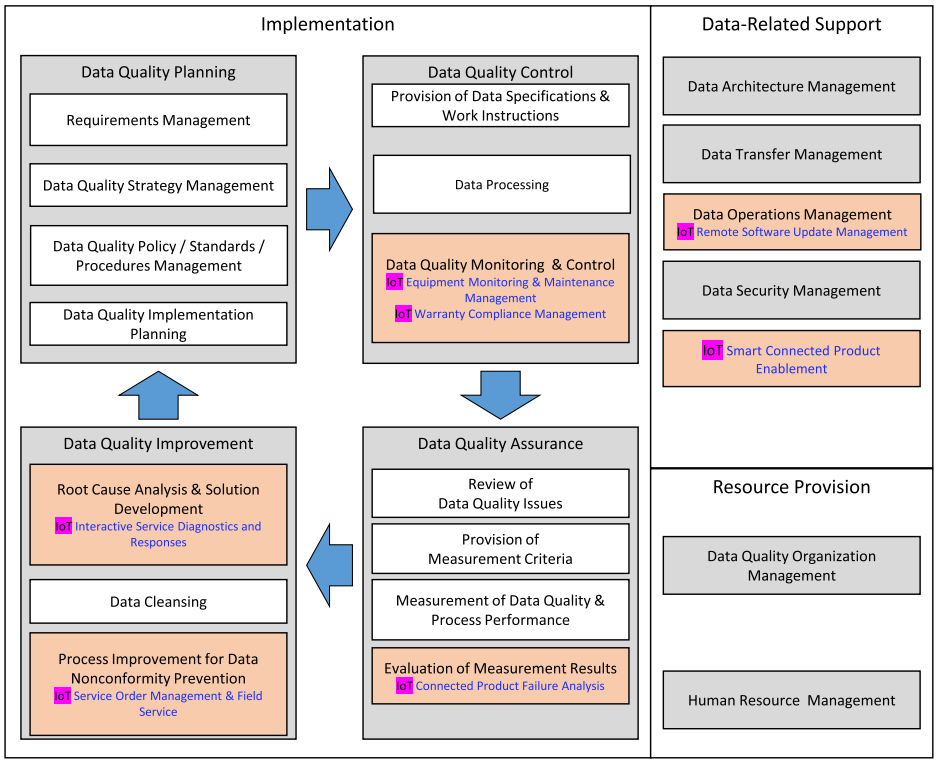

FIGURE 9. Mapping of the SCP operation processes to the DQM-PRM.

Example: The battery lifetime of the devices should be tracked to prolong the communication between the connected products and IoT.

Note: When product issues are identified during the data tracking, an alert system is usually automatically triggered to warn operators or systems against the product issues.

IoT The service/maintenance responses are initiated based on the thresholds and trends that indicate a potential failure of a connected product and rules that drive service/maintenance decisions.

IoT The connected product usage is monitored, recorded, and compared with the warranty policies to identify potential warranty compliance issues.

⑥ The records of the actions taken to address data nonconformity are collected.

IoT The asset and fleet history data such as product as-maintained configuration, utilization and service/maintenance history are updated based on connected product data and service events performed across the service network. The asset and fleet history data are analyzed for contract compliance and optimization, deferred and approaching service/maintenance needs, and product replacement opportunities.

⑦ Stakeholders are notified of actions taken to address data nonconformity.

IoT The initiation of service responses is notified to minimize downtime and avoid potential product failures.

IoT The identified compliance issues are automatically notified to avoid potential product failures and warranty issues.

Note: In IoT environments, end users include customers of IoT products, operators monitoring and controlling IoT products, experts analyzing data and diagnosing IoT products, stakeholders using IoT data for business systems, and IoT solution providers such as the device, network, platform, and application providers.

⑧ Guidelines, rules, and procedures are refined and applied to prevent recurrence of data nonconformity.

c: ACTIVITIES

– Data quality risk assessment: Identify risks throughout the data lifecycle, analyze the impact if each risk was to occur and determine the risk priorities to establish the basis for monitoring and controlling the processes and data.

– Monitoring and controlling of the processes: Monitor and measure process performance according to the identified risk priorities. Monitoring and measuring occur at intervals or continuously and in accordance with applicable work instructions. If planned results have not been achieved during the data processing, the end

users respond by updating and maintaining processes to ensure future conformity of the data.
- Monitoring and controlling of the data: Monitor and measure conformity of data to the applicable specification according to the identified risk priorities. Monitoring and measuring take place at intervals or continuously and in accordance with applicable work instructions. If data nonconformity is found, correct the data when viable and distribute to stakeholders a record of the viability and degree of success for each corrective action.
  IoT Real-time equipment health monitoring: See Section 4.2.1.
  IoT Condition-based maintenance: See Section 4.2.1
  IoT Connected asset and fleet management: See Section 4.2.1.
  IoT Warranty compliance management: This process is not detailed to the activity level as the content is relatively small. See Section 4.2.4.
- Prevention of the data nonconformity recurrence: Act to prevent recurrence of similar data nonconformity by refining and applying guidelines, rules, and procedures.

2) WORK PRODUCTS FOR DATA QUALITY MONITORING AND CONTROL

Table 2 summarizes the work products, consisting of input data required for and output data produced from DQMC process execution. In this table, the outcome refers to the one related to input and output data. The work products include both the sensor-related data marked with the IoT symbol and the other data among the IoT data. This implies that by performing this process, it is possible to simultaneously improve the quality of sensor data and the other data among the IoT data. These work products are used as indicators to measure the extent to which the process purpose is achieved.

TABLE 2. Work products generated by the DQMC.

Work products			
Input data		Output data	
Name	Outcome	Name	Outcome
Data specification.	①	Identified and quantified risks	①
Work instructions	①		
IoT preventive maintenance plans	①	Prioritized risks for monitoring and control.	①②
IoT Warranty and contract policies	②	IoT Prioritized risks (physical characteristics, warranty and contract policies) for monitoring and control	①②
Data of processes monitored	③	Records for comparing process performance with respect to identified risks.	③
		Notification to end users	④
Data nonconformities	⑤⑥	Records of actions taken for data nonconformities.	⑤⑥
IoT Device identifiers	①②⑤⑥⑦	IoT Data tracked and diagnosed by location, performance, utilization, and condition	⑤
IoT Device location information.	①②⑤⑥⑦	IoT Data of connected product usage	⑤
IoT Physical characteristics of devices to be monitored	①②⑤⑥⑦	IoT Service/maintenance responses initiated based on thresholds and trends	⑤
		IoT Asset and fleet history data updated	⑥
		Notification of actions taken for data nonconformities to stakeholders.	⑦
		IoT Notification of service response initiations	⑦
		IoT Notification of compliance issues to warranty and contract policies	⑦
		Guidelines, rules and procedures refined to prevent recurrence of data nonconformities	⑧

D. AN EXAMPLE SCENARIO TO IMPROVE IOT DATA QUALITY

As an example, to which the IoT DQM-PRM process is applied, it is assumed that a gas turbine has been installed in a power plant and physical characteristic data are being collected through the SCPE. The sensor data are monitored by the EMMM, where they are processed to improve the data quality and analyzed for further utilization by the support of the data analytics function in the SCPE. Among the various physical characteristic data monitored, an exit temperature of the combustor suddenly begins to rise and exceeds its threshold. Therefore, the sensor data fault is diagnosed by the ISDR and it is found that there is something wrong with a component part of the combustor. As a result, condition-based maintenance becomes necessary and corresponding service response is initiated by the SOFS. At this time, sensor data are excluded from normal data and separated into abnormal data by the Data Processing or SCPE. However, the similar phenomenon re-occurs after a while. Therefore, stopping the operation of the gas turbine, detailed analysis is performed by the Root Cause Analysis and Solution Development. It is found that the root cause is the old version of the software of the sensor-equipped device. Accordingly, an on-line service response is initiated to upgrade the software by the RSUM.

However, if the service BOM (or the product as-maintained configuration) of the gas turbine has not been updated even when the device was replaced with a new device in previous

maintenance services, then the old version of the software will be provided again, and the sensor data fault occurs repeatedly. Therefore, it will be necessary to update the service BOM information by the activity Connected asset and fleet management in the EMMM, trace data transfer records of the wrong data by the Data Transfer Management, and correct all related erroneous service information and sensor data that have been processed previously by the Data Cleansing. If the service BOM has not been updated due to the problem of the updating process, the process should also be improved by the Process Improvement for Data Nonconformity Prevention.

As can be seen in this example scenario, there may exist data nonconformities not only in sensor data but also in other data types. Therefore, in order to maintain the IoT data composed of various data types with high quality, it is necessary to continuously execute the IoT DQM-PRM processes proposed in this study.

VI. CONCLUSION

In this research, seven processes required for SCP operations have been proposed, which can be used for sensor data management in the product cloud. Furthermore, the IoT DQM-PRM that integrate these processes with the DQM-PRM defined in ISO 8000-61 have been proposed for the first time. Like the ISO 9000 certification model for product quality improvement [18], the process assessment model for IT process improvement [21] and the CMMI (Capability Maturity Model Integration) model for software quality improvement [27], the IoT DQM-PRM can be used to improve the quality of IoT data that include both real-time stream sensor data and other structured and semi-structured data and assess the DQM maturity of the organization in IoT environments.

The integrated processes as shown in Figure 9 are not enough to represent a complete IoT DQM-PRM. Other processes of the DQM-PRM also require tailoring to suit the SCP operations environment. However, their tailoring content is smaller than those of the integrated processes and is not discussed further in this study.

The proposal presented here is part of a long-term research project. We are working in various areas in order to apply in the future the IoT DQM-PRM to a real-life, industrial context. Main current research areas are the following: First, as the SCPE was added as an individual process to the Data-Related Support in the DQM-PRM, the DQM maturity model defined in ISO 8000-62 should be modified accordingly. The process is expected to present the same maturity level as the Data Processing and Data Security Management in the DQM-PRM as it is primarily required for collecting, analyzing, and managing IoT data. Further, in order to assess the IoT DQM-PRM for SCP operations, it is also necessary to derive work products for each process. The IoT DQM-PRM model will be validated by applying the development results to various IoT use cases.

REFERENCES

[1] M. Hung, "Leading the IoT-gartner insights on how to lead in a connected world," Gartner, Stamford, CT, USA, Tech. Rep., 2017. [Online]. Available: https://www.gartner.com/en/publications/iot-business

[2] *Next Generation Networks-Frameworks and Functional Architecture Models: Overview of the Internet of Things*, Standard ITU-T Y.2060, Jun. 2012.

[3] M. E. Porter and J. E. Heppelmann, "How smart, connected products are transforming competition," *Harvard Bus. Rev.*, vol. 92, no. 11, pp. 64–88, Nov. 2014. [Online]. Available: https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition

[4] M. E. Porter and J. E. Heppelmann, "How smart, connected products are transforming companies," *Harvard Bus. Rev.*, vol. 93, no. 10, pp. 96–114, Oct. 2015. [Online]. Available: https://hbr.org/2015/10/how-smart-connected-products-are-transforming-companies.

[5] S. Kim and C. Lee, "The process reference model for the data quality management process assessment," *J. Soc. e-Bus. Stud.*, vol. 18, no. 4, pp. 83–105, Nov. 2013.

[6] *Data Quality—Part 61: Data Quality: Process Reference Model*, Standard ISO 8000-61:2016, 2016.

[7] H. Karkouch, H. Mousannif, H. A. Moatassime, and T. Noel, "Data quality in Internet of Things: A state-of-the-art survey," *J. Netw. Comput. Appl.*, vol. 73, pp. 57–81, Sep. 2016.

[8] G. Jesus, A. Casimiro, and A. Oliveira, "A survey on data quality for dependable monitoring in wireless sensor networks," *Sensors*, vol. 17, no. 9, p. E2010, Sep. 2017. doi: 10.3390/s17092010.

[9] C. C. G. Rodríguez and S. Servigne, "Managing sensor data uncertainty: A data quality approach," *Int. J. Agricult. Environ. Inf. Syst.*, vol. 4, no. 1, pp. 35–54, Jan. 2013.

[10] Y. Qin, Q. Z. Sheng, N. J. G. Falkner, S. Dustdar, H. Wang, and A. V. Vasilakos, "When things matter: A survey on data-centric Internet of Things," *J. Netw. Comput. Appl.*, vol. 64, pp137-153, Apr. 2016.

[11] R. Perez-Castillo, A. G. Carretero, I. Caballero, M. Rodriguez, M. Piattini, A. Mate, S. Kim, and D. Lee, "DAQUA-MASS: An ISO 8000-61 based data quality management methodology for sensor data," *Sensors*, vol. 18, no. 9, p. 3105, Sep. 2018.

[12] *Information Technology—Internet of Things (IOT)—IOT Use Cases*, Standard ISO/IECTR22417, 2017.

[13] *PTC Value Roadmap Version 7.2 Discrete Manufacturing: Best Practices*, PTC, Boston, MA, USA, 2015.

[14] Honeywell. *Equipment Health Monitoring (EHM) Systems*. [Online]. Available: http:www.stevenengineering.com

[15] S. Sudhi R, P. Youngchoon, *Building an Effective IoT Ecosystem for Your Business*. Cham, Switzerland: Springer, 2017.

[16] C. W. Park and S. D. Kim, "Practical methods for managing faults in IoT computing," *J. Internet Comput. Services*, vol. 16, no. 5, pp. 75–86, Oct. 2015.

[17] *Data Quality—Part 62: Data Quality Management: Organizational Process Maturity Assessment: Application of Standards Relating Process Assessment*, Standard ISO 8000-62:2018, 2018.

[18] *Quality Management Systems—Requirements*, Standard ISO 9001, 2015.

[19] *Geographic Information—Data Quality*, Standard ISO 19157, 2013.

[20] *Software Engineering—Software Product Quality Requirements and Evaluation (SQuaRE)—Data Quality Model*, Standard ISO 25012, 2008.

[21] *Information Technology—Process Assessment—Process Measurement Framework for Assessment of Process Capability*, Standard ISO 33020, 2015.

[22] E. Al-Masri and Y. Bai, "A service-oriented approach for assessing the quality of data for the Internet of things," in *Proc. IEEE Int. Conf. Service-Oriented Syst. Eng. (SOSE)*, Apr. 2019, pp. 9–97.

[23] S. Sanyal and P. Zhang, "Improving quality of data: IoT data aggregation using device to device communications," *IEEE Access*, vol. 6, pp. 67830–67840, 2018.

[24] T. Luo, J. Huang, S. S. Kanhere, J. Zhang, and S. K. Das, "Improving IoT data quality in mobile crowd sensing: A cross validation approach," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5651–5664, Jun. 2019.

[25] T. Banerjee and A. Sheth, "IoT quality control for data and application needs," *IEEE Intell. Syst.*, vol. 32, no. 2, pp. 68–73, Mar./Apr. 2017.

[26] A. G. Carretero, F. Gualo, I. Caballero, and M. Piattini, "MAMD 2.0: Environment for data quality processes implantation based on ISO 8000-6X and ISO/IEC 33000," *Comput. Standards Interfaces* vol. 54, pp. 139–151, Nov. 2017.

[27] *CMMI Model 2.0*, CMMI Institute, Pittsburgh, PA, USA, 2018.

SUNHO KIM received the B.S. degree from Seoul National University, in 1979, and the M.S. and Ph.D. degrees from Pennsylvania State University, USA, in 1989, all in industrial engineering.

He was a Researcher with the Agency for Defense Development, from 1979 to 1994, and a Senior Researcher with the Korea Institute of Machinery and Metals, from 1989 to 1991. He was the Chair of the Korea University-Industry Collaboration Council, from 2014 to 2015. Since 1992, he has been a Professor with the Department of Industrial and Management Engineering, Myongji University, South Korea. His research interests include data quality management and data analytics for data quality.

Dr. Kim is currently a member of ISO TC184/SC4 and a Project Leader of the ISO 8000-60 series development project. He received the Jeongheon Scholarship Grand Prize from KIIE, in 2009, and the Korean deputy prime minister's commendation in the field of industry-academia cooperation, in 2016.

RICARDO PÉREZ DEL CASTILLO received the Ph.D. degree in computer science from the University of Castilla-La Mancha, Spain, where he is currently with the Information Systems and Technologies Institute. His research interests include architecture-driven modernization, model-driven development, business process archaeology, and enterprise architecture.

ISMAEL CABALLERO received the M.Sc. and Ph.D. degrees in computer science from the University of Castilla-La Mancha, Spain, in 2004, where he has been an Associate Professor with the Information Systems and Technologies Department, since 2001, and he was appointed as a Training Head of the spinoff DQTeam SL, in 2017.

He has been researching on data quality management and data governance, since 1998, coauthoring several conference and journal articles. He teaches data quality management and data governance foundations in many universities and companies. He holds CISA certification by ISACA, since 2016, and CDO-1 certification by UALR-MIT, since 2017.

He is currently a member of ISO TC184/SC4 and a Project Editor of several parts of ISO 8000-60 series development project as well as a Project Editor of ISO 8000-62.

JIMWOO LEE received the B.S. degree in business administration from Korea University, in 1986, and the M.S. degree in management information system from Temple University, USA, in 1990.

He was a Planner on IT strategies with POSCO, from 1990 to 1996. Since 1996, he has been a Consultant on business and data management with 2e Consulting. He consistently participated in various consulting projects for government and major leading companies, such as Samsung, Hyundai, and Korea Telecom, and he is currently a Vice President of 2e Consulting. He is participating in the international government project, Data Quality for Internet of Things, researching new technology in the IoT environment, and developing data quality management processes for this.

CHANGSOO LEE received the B.S., M.S., and Ph.D. degrees in industrial engineering from Seoul National University, in 1987, 1989, and 1994, respectively. From 2002 to 2003, he was a Guest Researcher with the National Institute of Standards and Technology, USA. Since 1992, he has been a Professor with the Department of Industrial and Management Engineering, Gangneung-Wonju National University, South Korea. His research interest includes data analytics. He is a Project Leader of ISO 8000-66 (data quality management: assessment indicators for data processing in manufacturing operations).

DOWNGWOO LEE received the B.S. degree in applied mathematics from Pukyong National University, in 2000.

From 2008 to 2012, he was a Consultant on data quality and metadata management with GTOne, primarily in the financial sector. He has conducted various data analysis projects as a Data Analyst, from 2013 to 2017. Since 2017, he has been developing analysis-based solutions for data quality assessment in big data and IoT environment. He is participating in the international government project, Data Quality for Internet of Things, researching new technology in the IoT environment, and developing software for this.

SANGYUB LEE received the bachelor's degree in information statistics from Korea University, in 2007.

He analyzed market data in Nielsen, international market research company, from 2007 to 2013, and was a Team Manager for Online Research Development, until 2015. He joined several projects for big data analysis in AreumICT, data analysis consulting company, from 2016 to 2017. Since 2017, he has been with the Research and Development Center, GTOne, data and application governance software company. He is participating in the international government project, Data Quality for Internet of Things, researching new technology in the IoT environment, and developing software for this.

ALEJANDRO MATE received the degree in computer science engineering, the M.Sc. degree in computer science technology, and the Ph.D. degree from the University of Alicante, in 2009, 2010, and 2013, respectively.

He held a postdoctoral researcher position in Italy, from 2014 to 2016, and worked for Lucentia Lab as a Business Intelligence and Big Data Architect as part of a Torres Quevedo grant, from 2016 to 2017. He has been an Associate Professor with the University of Alicante, since 2019. He has collaborated with several research groups across the globe, most notably, including the requirements engineering group led by John Mylopoulos at the University of Trento, Italy, and the software engineering group led by Eric Yu at the University of Toronto, Canada. His research has been mainly focused on BI and analytics, ranging from the definition of strategic plans and key performance indicators to the extraction of insights by means of dashboards and algorithms. He has published over 50 articles related to BI and analytics. Most of these articles are published in high impact international conferences (e.g. ER, CAiSE, and RE) and JCR journals (*Information Systems, Future Generations*, and *Information & Software Technology*). Nevertheless, his career has not been limited to the research field. In the professional department, he has developed analytic systems and software for several national and international projects. Among these projects, we can find European Research Council grants (Lucretius) and large-scope national projects from private initiatives (LPS-Bigger). The algorithms developed granted him the Best Demonstration Award at the IBM conference CASCON in Canada. He is currently working on several projects related to the Internet of Things (IoT) combining real-time analytics, machine learning, and artificial intelligence.