Data Evaluation Practice for Central Management of Energy in Industrial Complexes

Seung-hwan Ju[1], Hee-suk Seo[2 1]
[1]Electrical Standards Center, Korea Testing Laboratory
[2]Dept. of Computer Engineering, Koreatech

Article Info
Page Number: 634 – 643
Publication Issue:
Vol. 71 No. 3 (2022)

Article History
Article Received: *12 January 2022*
Revised: *25 February 2022*
Accepted: *20 April 2022*
Publication: *09 June 2022*

Abstract. The Industrial Complex Corporation seeks ways to save energy and provide additional services by collecting energy data from industrial complexes. FEMS and CEMS are energy management systems for factories and complexes and are established for central data management. EMS gets data from energy meters and various mobile sensors. This study defines a standard-based data quality evaluation method to operate an information system that collects, stores, and analyzes energy data through the consumer's measurement infrastructure. This study is an example of applying the standard data quality evaluation method in the demonstration stage of the establishment of the information system for the industrial complex.

Keywords: data quality, industrial complex, energy platform demonstration, information assurance

1. Introduction

The Industrial Complex Corporation establishes an industrial complex energy management system (Kachi et al., 2013). Based on it, the network connects scattered energy demand and supply resources such as factories, industrial and living infrastructure, and distributed power sources (Lee et al., 2016). This connection aims to spread new energy industry businesses such as factory energy efficiency improvement and achieve energy efficiency and self-sufficiency in industrial complexes (Kosmala et al., 2016). Customers operate the FEMS (Factory Energy Management System) that measures and communicates energy data generated from production facilities and utilities (Khan et al., 2016). The Industrial Complex Energy Management System (CEMS) is an information system that collects the factory energy management data of consumers (Zhou et al., 2017).

[1]+ Corresponding author. Tel.: +82-41-560-1495; fax: +82-41-560-1462
E-mail address: histone@koreatech.ac.kr.

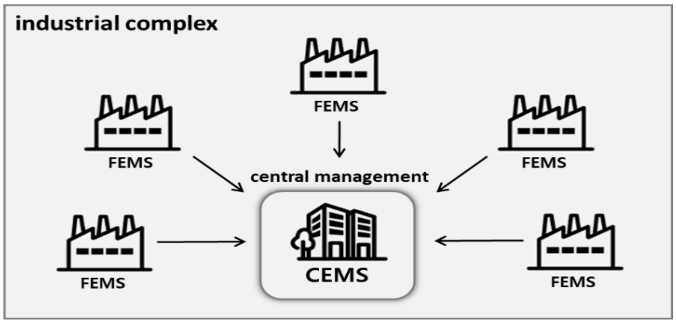

Fig. 1: FEMS and CEMS structure

Consumers access CEMS through web/mobile and manage their own energy such as facility control through monitoring, analysis, and prediction of energy data (Pipino et al., 2002). For such an information system, interoperability of communication and data is essential (Song et al., 2002). The industrial complex, the subject of this study, adopted IEC 62541 OPC UA as the communication standard for communication interoperability (Ju et al., 2022). The demonstration target defines the data model and operates the CEMS system. This study is a method to evaluate the data acquisition quality for information service of industrial complex (Park et at., 2003). To test the data acquisition, it was necessary to verify the error precision and reliability of the meter (Eppler et al., 2000). The data model supports energy measurement and integrated management in factory operations and energy platforms.

2. Related Standards

2.1. ISO 8000 (Data Quality)

It expands management activities based on high-quality data-based decision-making in quality management (ISO 9000). In 2005, ISO 8000 was proposed as a quality standard for the catalog management system for NATO military data quality management (De Guio et al., 2014).

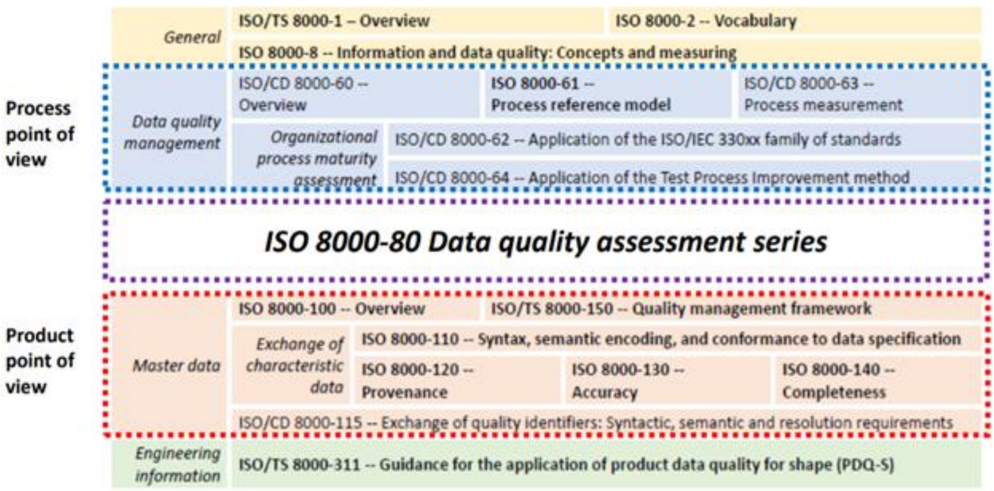

Fig. 2: ISO 8000-80 series for data quality assessment

2.2. ISO/IEC 2502n Series

ISO/IEC 25024 defines a data quality measurement method for quantitatively measuring data quality. The object of measurement in this standard is an object (model, architecture, data file, etc.) created/managed during the data life cycle and expressed in various formats (Corrales, et al., 2018). ISO/IEC 25024 manages data quality characteristics into 15 categories and defines detailed metrics for each (Cichy et al., 2019). For industrial complex data quality, we use data quality as a measurement index for data that has been processed, not from the point of view of the data design and collection stage.

- ISO/IEC 25020(Measurement reference model and guide)
- ISO/IEC 25021(Quality measure elements)
- ISO/IEC 25022(Measurement of quality in use)
- ISO/IEC 25023(Measurement of system and software product quality)
- ISO/IEC 25024(Measurement of data quality)

The ISO/IEC 25024 defines the data quality measurement sheet as shown in the following table:

Table 1: Classifications on ISO/IEC 25024 data quality

Classification	Measurement
Accuracy	• Syntactic data accuracy • Semantic data accuracy • Data accuracy assurance • Risk of data set inaccuracy • Data model accuracy • Metadata accuracy • Data accuracy range
Completeness	• Record completeness • Attribute completeness • Data file completeness • Data values completeness • Empty records in a data file • Metadata completeness • Conceptual data model completeness • Conceptual data model attributes completeness
Consistency	• Referential integrity • Data format consistency • Risk of data consistency • Architecture consistency • Data values consistency coverage • Semantic consistency
Credibility	• Values credibility • Source credibility • Data dictionary credibility • Data model credibility
Currentness	• Update frequency • Timeliness of update • Update item requisition
Accessibility	• User accessibility • Device accessibility • Data format accessibility
Compliance	• Regulatory compliance of value and/or format • Regulatory compliance due to technology
Confidentiality	• Encryption usage • Non vulnerability

Efficiency	• Efficient data item format • Usable efficiency • Data format efficiency • Data processing efficiency • Risk of wasted space • Space occupied by records duplication • Time delay of data update
Precision	• Precision of data values • Precision of data format
Traceability	• Traceability of data values • Users access traceability • Data values traceability
Understandability	• Symbols understandability • Semantic understandability • Master data understandability • Data values understandability • Data model understandability • Data representation understandability • Linked master data understandability
Availability	• Data availability ratio • Probability of data available • Architecture elements availability
Portability	• Data portability ratio • Prospective data portability • Architecture elements portability
Recoverability	• Data recoverability ratio • Periodical backup • Architecture recoverability

The above table is the data quality index of ISO/IEC 25024, and this standard provides the measurement function of the measured value.
The method for selecting reliability evaluation criteria in ISO/IEC 25024 is easy to access to authorized users who need data for a set period of time, provides a tracking audit function when accessing data, and maintains the reliability of data and established rules. Measures to evaluate the degree of compliance and correct update of data are selected. (i.e. availability, accessibility, traceability, credibility, conformability, realism, etc.)

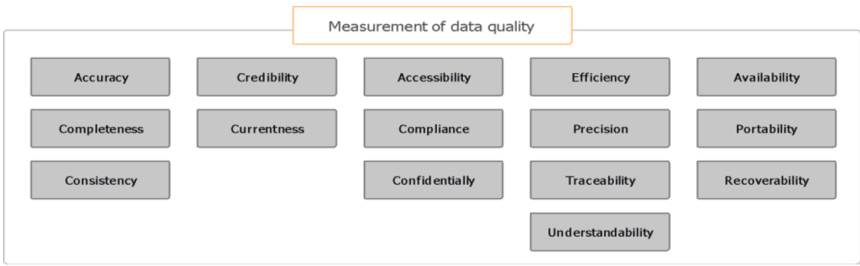

Fig. 3: Categories of ISO/IEC 25024 data quality

3. Data Communication Structure

3.1. Data Model

The data model is divided into workplace information, measuring instrument information, and energy measurement information. This demonstration uses power energy usage information as a data model. The data set contains the following information:

Table 2: meta-data model

Properties	Description
CEMS.Site.TS	Business site information
CEMS.Meta.TS	Equipment meta information
CEMS.Opertaion.TS	Equipment installation and operation information
CEMS.Communication.T	Communication information
CEMS.Power.TS	Power energy instantaneous information
CEMS.PeakPower.TS	Instantaneous information on the energy peak
AMS.Common.TS	Common information
AMS.Asset.TS	Equipment common information
AMS.Sensor_Common.TS	Sensor common information
AMS.AssetEval.TS	Equipment handling information
AMS.MO.TS	Equipment Motor Information
AMS.MO_Sensor.TS	Motor-separated sensor information
AMS.MO_BuiltIn_Sensor.TS	Motor-integrated sensor information
AMS.MO_Sensor_Raw.TS	Measuring data on motor-separated sensor
AMS.MO_BuiltIn_Sensor_Raw.TS	Measuring data on Motor-integrated sensor
CEMS.Calculation.TS	Processing data
CEMS.CodeProperties.TS	Code management for logic

Table 3: Power data attribute between CEMS and FEMS

Properties	Description	Properties	Description
Kwh	active power	V{a,b,c}	phase voltage
Kvarh	reactive power	I{a,b,c}	phase current
Hz	frequency	P{a,b,c}	phase active power
Vh3{a,b,c}	phase voltage 3-harmonic	D{a,b,c}	phase angle
Ih3{a,b,c}	phase current 3-harmonic	Pf{a,b,c}	phase power factor
Ih5{a,b,c}	phase current 5-harmonic	Q{a,b,c}	phase reactive power
Ih7{a,b,c}	phase current 5-harmonic	status	communication status
Ih11{a,b,c}	phase current 11-harmonic		

By defining data items as shown in Table 1 and Table 2 and selecting OPC UA as the protocol, CEMS can provide Completeness, Compliance, Accuracy, and Currentness of ISO/IEC 25024. By defining data items as shown in Table 1 and Table 2 and selecting OPC UA as the protocol, CEMS can provide completeness, compliance, and accuracy of ISO/IEC 25024. The Accuracy of ISO/IEC 25024 consists of syntax and semantics, data set, and data range accuracy. To test this, the ratio of data items with correct values and the ratio of outlier data values are measured. The compliance with ISO/IEC 25024 indicates conformance to a standard of data or format, and measures the proportion of data items that conform to a standard, convention or regulation for testing. CEMS is compliant because it defines data items, units, and representation methods. The completeness of ISO/IEC 25024 consists of record integrity and data value integrity, and is confirmed by measuring the ratio of non-null data items among data items and the ratio of data values that satisfy expected values in data items.

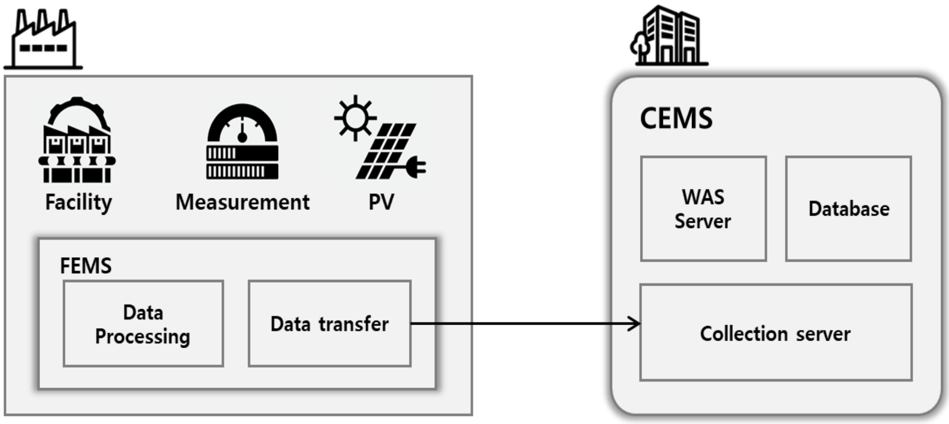

Fig. 4: FEMS and CEMS structure

The Currentness of ISO/IEC 25024 is confirmed by the ratio of information items with update requests according to the update cycle and conditions for data update requests.

Data quality is ensured by demonstrating a common data model for optimal plant operation. This includes Timestamp, instrument type, sampling rate, measurement data, units, etc. The common data model for CEMS integrated management through energy platform includes Timestamp, instrument type, weight, verification site, sampling rate, data, unit, status, etc. In particular, for the effectiveness of CEMS, factory operation information (personnel, contract, facility, inventory, etc.) is reflected. The data model can also be extended to monitor the condition of equipment using environmental data (vibration, noise).

3.2. Data Communication

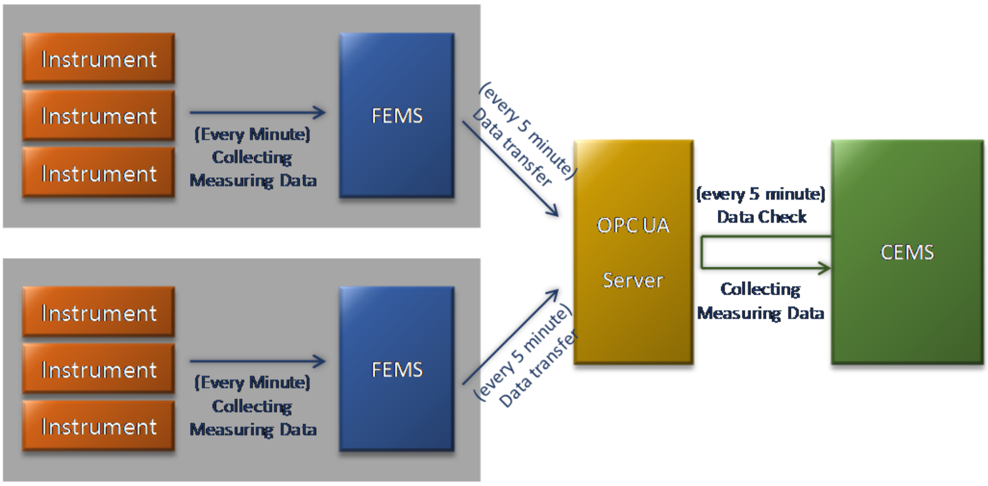

Fig. 5: Linkage flow between FEMS and CEMS

Every minute, FEMS collects data on equipment, instruments, and distributed power. FEMS is in charge of the OPC UA server, which registers FEMS data every 5 minutes. As an OPC UA client, CEMS requests and receives data from OPC UA server of FEMS once every 5 minutes (Ju et al., 2022). CEMS collects data sets based on FEMS's metering data according to rules. For example, if there are 1,440 weighing data per minute, CEMS needs to collect 288 data sets.

4. Data evaluation practice

This study evaluates the data quality between the FEMS and CEMS of the demonstration site. This study selects evaluation indicators at the level supported by CEMS data among the data evaluation items (Bansal et al., 2015) of ISO/IEC 25024 (Ge et al., 2017): Consumers and meters under the following conditions were not included in the test subject. Consumer and meter in the state of communication inability between consumer and industrial complex CEMS. A meter that does not utilize a weighing dataset for the purpose of peak measurement. Customers and meters with system configuration changes such as additional meter installation work

The demonstration is for the Banwol-sihwa Industrial Complex in Gyeonggi-do, Korea. The CEMS data acquisition consistency test consists of 231 instruments for 13 customers.

- Accuracy: The FEMS data includes all data from all CEMS.
- Completeness: The CEMS data is complete in the data set structure without omissions.
- Currentness: The CEMS acquires data according to the request period.
- e.g. Since the demonstration case acquires data at 5-minute intervals, CEMS needs to acquire 288 data sets per day.
- Data creation date: 2021-12-15
- Data subject to test: 13 consumers, 247 meters

Table 4: Data Completeness between FEMS and CEMS

Site code	number of meters (A)	Number of FEMS Data	Number of Data in OPC server (B)	Number of CEMS data (C)	Completeness (C/B)
bwsh-03	22	31,672	6,333	6,304	99.54%
bwsh-04	21	27,653	5,530	5,147	93.07%
bwsh-07	7	10,072	2,012	2,003	99.55%
bwsh-09	20	26,878	5,370	2,308	42.98%
bwsh-12	23	32,320	6,463	6,378	98.68%
bwsh-17	21	30,170	6,027	5,995	99.47%
bwsh-18	11	15,815	3,159	3,147	99.62%
bwsh-20	19	27,350	5,470	1,055	19.29%
bwsh-21	27	38,835	7,750	7,653	98.75%
bwsh-25	7	10,053	2,008	689	34.31%
bwsh-31	21	30,239	6,047	5,950	98.40%
bwsh-34	16	31,037	6,194	5,835	94.20%
bwsh-35	16	34,559	6,911	6,557	94.88%
Summary	231	346,653	69,274	59,021	85.20%

Table 5: Data Accuracy between FEMS and CEMS

Site code	number of meters (A)	Number of measurement data stored in FEMS	The number of CEMS data does not match the FEMS data
bwsh-03	22	132,405	1
bwsh-04	21	108,087	2
bwsh-07	7	42,105	-
bwsh-09	20	48,468	2
bwsh-12	23	133,938	-
bwsh-17	21	125,895	-
bwsh-18	11	66,087	-
bwsh-20	19	113,232	-
bwsh-21	27	160,734	-
bwsh-25	7	14,469	-
bwsh-31	21	124,950	-
bwsh-34	16	123,564	-
bwsh-35	16	137,697	-
Summary	231	346,653	5

Table 6: Data Currentness between FEMS and CEMS

Site code	number of meters (A)	Number of data in CEMS (C)	Number of data that fit the collection rules (D)	Currentness (C/D)
bwsh-03	22	31,672	6,270	99.01%
bwsh-04	21	27,653	5,123	92.64%
bwsh-07	7	10,072	1,991	98.96%
bwsh-09	20	26,878	2,278	42.42%
bwsh-12	23	32,320	6,340	98.10%
bwsh-17	21	30,170	5,964	98.95%
bwsh-18	11	15,815	3,132	99.15%
bwsh-20	19	27,350	1,021	18.67%
bwsh-21	27	38,835	7,621	98.34%
bwsh-25	7	10,053	679	33.81%
bwsh-31	21	30,239	5,921	97.92%
bwsh-34	16	31,037	5,804	93.70%
bwsh-35	16	34,559	6,519	94.33%
Summary	231	346,653	59,021	85.20%

5. Conclusion

To establish CEMS for centrally managing energy data of industrial complexes, information systems require sufficient data quality for services. The study is a method for applying the ISO/IEC 25024 data quality standard to the demonstration stage. This study derives items that can be tested during ISO/IEC 25024 based on data communication rules of FEMS and CEMS. The Industrial Complex Corporation utilizes CEMS to provide additional services such as demand response, electricity transaction, and ESS connection using energy usage information. Furthermore, energy efficiency and self-sufficiency at the level of industrial complexes can be expected through the connection of distributed power sources.

Through this study, CEMS has a quality management method for weighing data. Data quality provides a cause of data acquisition errors in CEMS and records of errors in weighing data such as replacement and testing of meters. CEMS may have data requirements such as data traceability collected by recording CEMS system time at the time of data acquisition of CEMS.

This data consistency test verified the grammatical quality of data from the CEMS point of view. This includes the data structure, data format, data unit, and number of data collected according to the collection cycle. The next study will involve both semantic and data syntactic quality verification (Nikolaidis, 2019). The researchers will keep searching on the procedure to perform semantic relevance of data, linkage with metadata, and validation of data values.

6. Acknowledgments

This paper was supported by the Education and Research promotion program in 2022.

This result was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE)(2021RIS-004)

7. References

1. Bansal, S. K., & Kagemann, S. (2015). Integrating big data: A semantic extract-transform-load framework. *Computer*, 48(3), 42-50.
2. Cichy, C., & Rass, S. (2019). An overview of data quality frameworks. *IEEE Access*, 7, 24634-24648.
3. Corrales, D. C., Ledezma, A., & Corrales, J. C. (2018). From theory to practice: A data quality framework for classification tasks. *Symmetry*, 10(7), 248.
4. De Guio, F., & Cms Collaboration. (2014, June). The CMS data quality monitoring software: experience and future prospects. In Journal of Physics: Conference Series (Vol. 513, No. 3, p. 032024). IOP Publishing.
5. Eppler, M. J., & Wittig, D. (2000). Conceptualizing Information Quality: A Review of Information Quality Frameworks from the Last Ten Years. *IQ*, 20(0), 0.
6. Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data mining and analytics in the process industry: The role of machine learning. *Ieee Access*, 5, 20590-20616.
7. Khan, A., & Turowski, K. (2016). A survey of current challenges in manufacturing industry and preparation for industry 4.0. In Proceedings of the First International Scientific Conference "*Intelligent Information Technologies for Industry"(IITI'16)* (pp. 15-26). Springer, Cham.
8. Kachi, M., Yoshimoto, Y., Makita, H., Nozue, N., Shida, Y., Kitagami, S., & Sawamoto, J. (2013). FEMS: Factory Energy Management System based on production information. Recent Advances in Energy and Environment Integrated Systems.
9. Kosmala, M., Wiggins, A., Swanson, A., & Simmons, B. (2016). Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*, 14(10), 551-560.
10. Lee, D., & Cheng, C. C. (2016). Energy savings by energy management systems: A review. Renewable and Sustainable Energy Reviews, 56, 760-777.
11. Nikolaidis, Y., Pilavachi, P. A., & Chletsis, A. (2009). Economic evaluation of energy saving measures in a common type of Greek building. Applied energy, 86(12), 2550-2559.
12. Park, J., John Park, A. S. D., & Mackay, S. (2003). Practical data acquisition for instrumentation and control systems. Newnes.
13. Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. Communications of the ACM, 45(4), 211-218.
14. S.H. Ju, S.H. Song, H. S. Seo, (2022). System to verify interoperability of smart industrial complex, International Journal of Mechanical Engineering, 7(1), 180-185
15. Zhou, Z., Gong, J., He, Y., & Zhang, Y. (2017). Software defined machine-to-machine communication for smart energy management. IEEE Communications Magazine, 55(10), 52-60.