About the framework of quality evaluation using SQuaRE

Tsuyoshi Nakajima
Shibaura Institute of Technology
Tokyo, Japan
tsnaka@shibaura-it.ac.jp

***Abstract*—Of the five core divisions of the SQuaRE series, the revision study of the quality evaluation division has started. As systems and software have become more complicated and diversified in recent years, it is important to have a mechanism to ensure the quality evaluation of them by a whole society. In order to do so, this revision aims to provide a framework for systematic quality evaluation using quality models and quality measures in the SQuaRE series. Furthermore, the framework will expand its scope of application to include four purposes: comparison, qualification to standard, conformity checking to requirements, and suitability evaluation in the market. This paper organizes and reports on the idea.**

Keywords- quality evaluation, framework, quality measurement, quality rating

I. INTRODUCTION

ISO/IEC 25000 (SQuaRE) series is a set of international standards for quality requirements and evaluation for a wide range of systems and software [1]. Fifteen standards have been published since 2006 to cover all the five core divisions. Among them, the quality evaluation division (ISO/IEC 2504n) is currently studying its revision in the study group.

In the development of a system or software, there are no absolute criteria for its quality to be achieved[2]. This is because the quality goal is a relative one needed in the context of social demands and competition with others. Actual products need to fulfill different types of quality, such as performance efficiency and security, and so difficult trade-offs are often required to achieve those qualities in a balance.

On the other hand, if a system or software operates as an important service in the society or works with the other systems already operating in the society, the quality evaluation may exceed the responsibility of one company because of the significance of its impact on society[3] [4] .

Nevertheless, many of the quality evaluations for systems and software, even if it drives the society, are carried out based on the empirical knowledge of development organizations for testing and ambiguous pass/fail criteria. Although it may be managed by an experienced development organization with high quality awareness, a small organization with little experience in quality control hardly can perform quality evaluation adequately for itself. This may lead to a project failure.

To overcome this situation, it is important to have a mechanism to ensure the quality evaluation of a system and software by a whole society, including objective comparison techniques for its quality with other products in the industry, and third parties' reliable quality certification.

This paper proposes a quality evaluation framework as a basis to support such a mechanism, including systematic quality evaluation in three stages: quality measurement, quality rating, and quality evaluation, using quality models and quality measures in the SQuaRE series. Furthermore, the framework will expand its scope of application to deal with four purposes: comparison, qualification to standard, conformity checking to requirements, and suitability evaluation in the market. This paper organizes and reports on the idea.

II. REVISION OF QUALITY EVALUATION DIVISION

A. *Quality evaluation division in SQuaRE*

SQuaRE divides the 25000 series into SQuaRE divides the 25000 series into five core divisions and one extension division. Figure 1 shows how quality evaluation division relating to the other division in SQuaRE. At the center are the

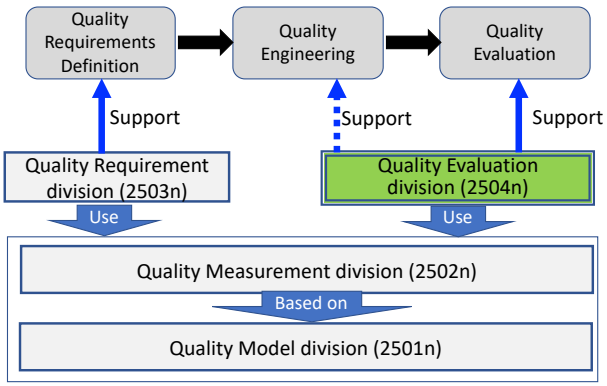

Figure 1 Quality evaluation division in SQuaRE

quality model division in 2501n and the quality measurement division in 2502n. The quality model division currently has four quality models: quality in use and product quality (ISO/IEC 25010), IT service quality (ISO/IEC TS 25011), and data quality (ISO/IEC 25012), each of which defines characteristics and sub-characteristics. The quality measurement division defines the quality measures to measure the quality (sub)characteristics defined in each quality model. The quality evaluation division provides a framework and usage guides to support the quality evaluation process (partly, validation and verification activities in quality engineering) using the pairs of quality model and quality measurement.

B. Revision of quality evaluation division

The SQuaRE Future Direction Study Group (2018-2019) in ISO/IEC JTC1/SC7 WG6 concludes has improved the ISO / IEC 2504n quality evaluation department as follows[5].

[Problems]

There is a great demand from the industry for methods and techniques to support how to plan inspections and testing on quality, and concrete assessment based on their results. In addition, due to the development of 2502n and the revision of 25030, the quality requirements and their measurement have been clarified, so modifications aligning with these will also be necessary.

[Proposed solution]

- Ensuring consistency with 2502n and 25030R
- Improving the concept of evaluation modules (EVs) (and encouraging industries to provide ANNEXs)
- Guidelines for the following activities:
  - Organizing quality testing including inspections, aligning with **29119** (WG 26)
  - Comprehensive quality evaluation (e.g., for judgment of delivery) based on measurement results
    - How to devise a set of quality measure suitable for evaluation
    - Concept of evaluation (analysis of testing results, etc) and rating
  - Selecting the right quality characteristics from some evaluation goal
  - Choosing an appropriate evaluation module for the characteristics or to make a new evaluation module.

Based on this policy, we are currently studying a study group within WG6 for revision.

III. FRAMEWORK OF QUALITY EVALUATION

A. Quality evaluation and its types

Quality evaluation is a systematic examination of the extent to which an entity is capable of fulfilling stated and implied needs. This is a set of activities to perform objective confirmation work on the target to produce quality evaluation results.

However, there exist many cases in which the quality evaluation based on the current ISO/IEC 25040 and 25041, cannot be applicable because it requires a full set of quality requirements for the target. Therefore, we decided to expand its scope so that it can deal with the cases without fixed requirements. We categorize the assumed situations that require quality evaluation and reconstructed the quality evaluation framework so that such cases can be included.

We have identified the following four types of quality evaluation:

(T1) Quality evaluation for comparison

This is a type of quality evaluation:
- for the purpose of obtaining information for product/component/data selection
- to finds out:
  - whether the candidate entities meet the setting criteria,
  - which entities are better overall, and
  - what are the strengths and weaknesses of each entity,
- by comparing multiple entities based on quality requirements or general guidelines (including industry standards)
- mainly performed by the acquirer or development organization (integrator).

(T2) Quality evaluation for qualification to quality standard

This is a type of quality evaluation:
- for the purpose of obtaining quality certification or information for quality improvement of the target entity
- based on the standards in the industrial domain
- to find out:
  - whether the entity meets the setting standards (mostly minimum set) and,
  - what are its strengths and weaknesses from the quality perspective
- mainly performed by the development organization (quality assurance) or an independent evaluation and certification organization.

(T3) Quality evaluation for conformity checking to requirements

This is a type of quality evaluation:
- for the purpose of confirming the satisfaction of the contract items
- based on the agreed requirements
- to confirm whether the product satisfies the quality requirements
- mainly performed by:
  - the ordering party (at its acceptance) or
  - the development organization (quality assurance)
    (at the time of final inspection before delivery).

(T4) Quality evaluation for suitability to the market

This is a type of quality evaluation:
- for the purpose of obtaining information for management decisions
- on in-house products / services

- based on the requirements derived from the assumed stakeholder needs
- to check how well the product meets the needs of the assumed market

- mainly performed by the development organization (quality assurance).

The revision of ISO/IEC 25040 needs to provide the framework which covers these four types of quality evaluation.

B. Relationships between Quality models/measures and quality valuation

Quality evaluation can be performed objectively, quantitatively, and comprehensively by using SQuaRE's quality model and quality measures. As shown in TABLE I, the SQuaRE series defines a pair of quality model and quality measures for each entity type to be evaluated.

TABLE I Quality models for entity types and related documents

Entity	Quality model	Standard for model	Standard for measures
Information system	Quality in use	ISO/IEC 25010	ISO/IEC 25022
ICT product	Product quality	ISO/IEC 25010	ISO/IEC 25023
IT service	IT service quality	ISO/IEC TS 25011	ISO/IEC TS 25025
Data	Data quality	ISO/IEC 25012	ISO/IEC 25024

As shown in Figure 2, the Quality model classifies the quality that the target entity should have as quality characteristics and subcharacteristics. Quality property is an attribute (information needs) of the target entity which you want to measure. Quality measure measures a specific quality property. In general, one quality property corresponds to multiple quality measures.

The quality evaluator must first select a quality model for quality evaluation based on the type of the target entity. Furthermore, all the quality characteristics and subcharacteristics defined in the model are examined for the target to identify the important quality characteristics and sub-characteristics for the quality evaluation. For ICT products, important quality (sub)characteristics differ depending on the category of its system type. For example, user interface aesthetics is important for consumer products while operational operability and user error prevention are required more for mission-critical systems.

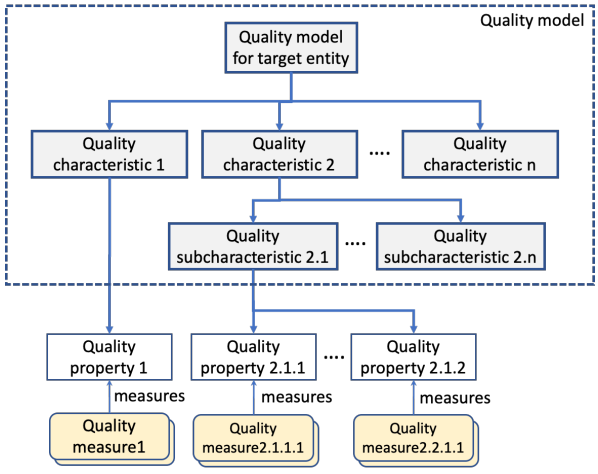

Figure 2 Relationship between quality model and measures

Furthermore, even if the same quality characteristics/ sub-characteristics are selected, the quality properties to be evaluated differ depending on the target category. For example, even with the same time behavior, it is necessary to evaluate "efficiency in processing requests" (throughput) on the server and "quickness of response to input events" (response time) on the client. Even with the same learning ability, it is "the degree of operation without a manual" for consumer products, and "a lot of support for learning operations" for mission-critical systems. These are measured by different quality measures.

C. Flow of quality evaluation

Use of the SQuaRE quality model and quality measures, quality evaluation can be systematically performed according to the flow shown in Figure 3, including three activities: quality measurement, quality rating, and quality evaluation. These activities are described below.

(1) Quality measurement

Quality measurement is an activity to measure the quality property of the target entity using a quality measure under a certain measurement condition. Since a quality measure defines the measurement method, whose concrete tasks includes testing or inspection (including static analysis) on the target entity, detailed measurement conditions (how many subjects to use, what to do with the parameters of the execution environment, etc.) at that time must be given.

ISO/IEC 2502n provides quality measures corresponding to each quality model defined in ISO/IEC 2501n. If no suitable quality measures for the target entity are found, a new quality measure that meets the requirements specified in ISO/IEC 25020 can be used.

(2) Quality rating

Quality rating is an activity to classify the values of quality measures into preset rating levels. The rating level gives the meaning of the degree of quality to the measured value. An example of rating levels that can be used to determine pass/fail.

Figure 4 shows an example of a typical pre-determined rating level, in which four rating levels are set for the measurement scale of quality measurements. Acceptable is a level that has an unconditionally acceptable value as the lower limit with the target value as the upper limit. When scoring, the level of Acceptable can be divided into multiple levels. In the case of Type 3, the allowable value uses the value determined by agreement with the customer. Boundary is not unconditional but acceptable by setting conditions. Not-acceptable is a level that means unconditional fail. Excessive quality (optionally defined) represents a level above the target value, which means unnecessarily high quality. In Figure 4,

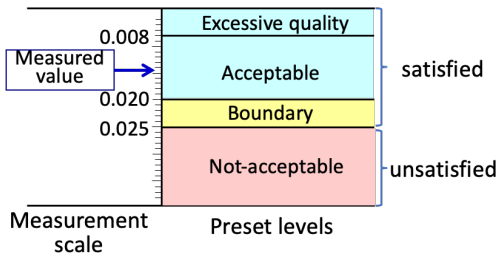

Figure 4 Typical preset rating levels

the lower limit values of Excessive quality, Acceptable, and Boundary are respectively 0.008, 0.02, and 0.025.

(3) Quality evaluation

Quality evaluation is an activity to determine pass/fail or score regarding quality (sub)characteristics using the quality rating results for each quality property. The quality evaluation is the integration of the quality rating results for the lower quality properties, or the integration of the quality evaluation results for the quality subcharacteristics immediately below.

In case of determining pass/fail, the following calculation can be used.

$$E(s) = \min\left(L(\mathbb{p})\right)$$
$$E(c) = \min\left(E(\mathbb{s})\right)$$
$$E_{overall} = \min\left(E(\mathbb{c})\right)$$

where

s : a quality subcharacteristic

c : a quality characteristic

$\mathbb{p}$: a vector of all quality properties under s

$\mathbb{s}$: a vector of all the selected subcharacteristics under c

$\mathbb{c}$: a vector of all the selected characteristics

$E(s\ or\ c)$: evaluation result of s *or* c

$L(\mathbb{p})$: a vector of rating result for each element of $\mathbb{p}$

$E_{overall}$: total quality evaluation result for the target

The evaluation value $E(s)$ for the quality subcharacteristic s is the minimum value among the rating results of $\mathbb{p}$, and the evaluation value $E(c)$ for the quality characteristic c is the minimum value among the evaluation result of $\mathbb{s}$. The quality evaluation value $E_{overall}$ of the target are calculated as the minimum value among the evaluation results of $\mathbb{c}$.

With this calculation method, the lowest level of rating results for all quality properties goes to the overall result. If all are boundary or higher, it passes (conditionally), and if there is even one Not-acceptable, the whole fails.

In case of obtaining a score for quality of the target entity, it is reasonable to assign a value to the quality level. This is because a quality measure basically has a value ranging from 0 to 1 but cannot be used as a universal scale. To interpret the levels into some score value, it is recommended that Boundary should be 0, Acceptable should be further divided into multiple levels and integer values of 1 or more should be assigned to the levels, and Not-acceptable should be assigned values of -1 or less.

From the quality rating results, the quality evaluation score can be calculated for each stage by the following calculation. Here, w_i, w_j, and w_k are weighting coefficients for each stage. It is good to normalize each (adding up to 1 for all).

$$E(s) = \sum_i w_i * Score(L(p_i))$$
$$E(c) = \sum_j w_j * E(s_j)$$
$$E_{overall} = \sum_k w_k * E(c_k)$$

where $Score(l)$ is a function to map from quality level to score.

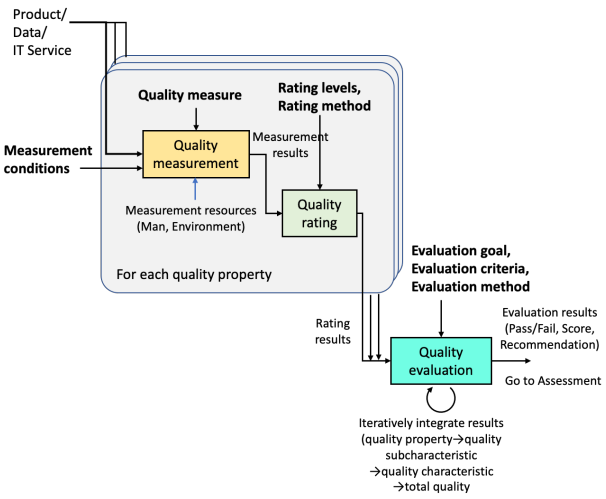

Figure 3 Flow of quality evaluation_

IV. DIFFRENCES IN FOUR TYPES OF QUALITY EVALUATION

Quality evaluation is classified into four types according to its purpose, which is useful to clarify variation points on how to conduct quality evaluation activities. We identified these points as below:

a) Model: Which quality model should be applied, Quality in use, Product quality, Data quality, or IT service quality (TABLE III "Model to use")

b) Charas & measures: Which quality characteristics and sub-characteristics are important, which quality properties and quality measures should be used.

c) Criteria: Quality criteria to set quality levels, including weights for scoring for quality properties, quality subcharacteristics and quality characteristics.

TABLE II Variation points of quality evaluation and their factors

Variation point	QE type	Entity type & category	Industrial domain
a) Model	V	V	
b) Chars & measures		V	V
c) Criteria		V	V
d) Output	V		

TABLE III Four types of quality evaluation and their differences

Type		Number of targets	Model to use			Expected results				Source of quality criteria
			QiU	PQ	DQ	Pass/ Fail	Candidate selection	Score	Strength and weakness	
T1	QE for comparison	Multiple	M	M	M		V	V	V	Quality requirements or general guidelines (including industry setting standards)
T2	QE for determining to achieve quality standard	Single	S	M	M	V			V	Standards in the industrial domain
T3	QE for conformity to requirements	Single	S	M	M	V			V	Agreed requirements
T4	QE for suitability to the market	Single	M	S	S			V	V	Requirements derived from the assumed stakeholder needs (established during development)

M: major, S: supplementary

a) Output: What kind of output are needed (TABLE III "Expected results")

TABLE II shows variation points of quality evaluation and their factors. The factors to determine these variation points includes the type of the quality evaluation, entity type of the quality evaluation, and industrial domain of the target entity.

TABLE III shows each type has its own pattern of quality model selection to use and expected results. Only T1 can have multiple targets while the others cannot. Concerning model to use, T1 can use all types of quality models, T2 and T3 mainly use product quality (PQ) and data quality (DQ) model, but supplementarily use quality in use (QiU) model. Contrary to that, T4 mainly use QiU model. Expected results have four possibilities: pass / fail, candidate selection, score, and strength and weakness. Which results are needed also depends on the evaluation types, as shown in this table. Source of quality criteria describes what kinds of information source should be used to determine the quality criteria.

V. CONCLUSION

This paper proposes a quality evaluation framework, which will be hopefully used for the revision of the quality evaluation division (ISO/IEC 2504n). The framework would serve as a basis to support systematic quality evaluation, which consist of three activities: quality measurement, quality rating, and quality evaluation, using quality models and quality measures in the SQuaRE series. Furthermore, the framework will expand its scope of application to include four purposes: comparison, qualification to standard, conformity checking to requirements, and suitability evaluation in the market. This paper organizes and reports on the idea.

We hope to get many feedbacks from SQuaRE users on the ideas in this paper to successfully complete the revision of the quality evaluation division.

REFERENCES

[1] Esaki, Kazuhiro, Motoei Azuma, and Toshihiro Komiyama. "Introduction of quality requirement and evaluation based on ISO/IEC square series of standard." International Conference on Trustworthy Computing and Services. Springer, Berlin, Heidelberg, 2012.

[2] Cavano, Joseph P., and James A. McCall. "A framework for the measurement of software quality." *Proceedings of the software quality assurance workshop on Functional and performance issues*. 1978.

[3] Nakajima, Tsuyoshi, and Toshihiro Komiyama. "Applying Quality Requirements Framework to an IoT System and its Evaluation." *International Journal on Advances in Internet Technology Volume 12, Number 1 & 2, 2019* (2019).

[4] Vargas, Iohan Gonçalves, Thiago Gottardi, and Rosana Teresinha Vaccare Braga. "Approaches for integration in system of systems: a systematic review." *2016 IEEE/ACM 4th International Workshop on Software Engineering for Systems-of-Systems (SESoS)*. IEEE, 2016.

[5] Nakajima, Tsuyoshi. "Study group report on square future direction." CEUR Workshop Proceedings. Vol. 2545. CEUR-WS, 2020.