Assistant for the Evaluation of Software Product Quality Characteristics Proposed by ISO/IEC 25010 Based on GQM-Defined Metrics

Julieta Calabrese, Rocío Muñoz, Ariel Pasini[1], Silvia Esponda, Marcos Boracchia, Patricia Pesado

Computer Science Research Institute LIDI (III-LIDI)*
School of Computer Science – National University of La Plata.
50 y 120, La Plata, Buenos Aires, Argentina
*Partner Center of the Scientific Research Agency of the Province of Buenos Aires (CICPBA)

{jcalabrese,rmunoz,apasini,sesponda,marcosb,ppesado}@lidi.info.unlp.edu.ar

Abstract. An assistant to evaluate the characteristics, proposed by ISO/IEC 25010, of a software product using GQM (Goal, Question, Metric) is presented. A set of questions was defined whose combined answers allow obtaining a logical metric applicable to the characteristics proposed by ISO/IEC 25010. For this work, the characteristic of Security was used as case study, the corresponding metrics were defined, and the results obtained when applying them to three case studies are presented.

Keywords: Quality, Software product, GQM, ISO/IEC 25000

1 Introduction

The number of software developing companies has increased significantly together with the increase in the demand for products from the sector. For this type of companies, software quality has a key role, in particular as a differentiating element for competitiveness and corporate image, and because the monetary losses these companies can suffer as a consequence of software quality issues are significant. In this context, the activities related to software quality and its evaluation are becoming ever so important [1].

An organization can be interested in evaluating its product as a differentiator from its competitors by ensuring delivery times and lower failure rate in the product after

[1] Corresponding author

its implementation to production; through establishing agreements in the service sector by defining quality parameters that the product must meet before delivery; by detecting software product flaws and remove them before delivery; by evaluating and controlling the performance of the software product developed, ensuring that it will be capable of yielding the results taking into account given time and resource restrictions; by ensuring that the software product developed complies with the necessary levels for security characteristics (*Confidentiality*, *Integrity*, *Authenticity*, *Non-Repudiation*, etc.); and so forth.

In this sense, the ISO 25000 family, known as SQuaRE (Software Product Quality Requirements and Evaluation) is born as a response to these needs. Its objective is creating a common framework to assess the quality of the software product. It is a replacement of the previous ISO/IEC 9126 and ISO/IEC 14598, Quality models and metric generation [2-4].

In this article, we propose a software product evaluation assistant based on the metrics defined in ISO/IEC 25010 using the GQM approach [5].

Section 2 briefly describes the ISO/IEC 25000 family and the approach proposed in GQM. Then, Section 3 describes the model used to evaluate the characteristics proposed by ISO/IEC 25010 using the GQM approach, in particular, the characteristic of ***Security***. After that, three case studies are presented, where the evaluation model is applied and the results obtained are discussed. Finally, conclusions are presented.

2 Quality Models and Metric Generation

2.1 The ISO/IEC 25000 Family.

Quality management is required in organizations due to the significance it has on various fronts: on the product level, establishing the quality achieved and the characteristics present in the products; on the level of the organization, establishing a procedural framework that allows improvement; as well as on the level of the processes.

To organize and unify all standards related to software product quality, ISO/IEC published in 2005 the document ISO/IEC 25000:2005 - SQuaRE (Software Product Quality Requirements and Evaluation), also known as the ISO 25000 family. Within ISO/IEC 25000, ISO/IEC 25010 - *System and software quality models* and ISO/IEC 25040 - *Evaluation process*, described below, stand out.

ISO/IEC 25010 - System and software quality models.

It replaces ISO/IEC 9126-1:2001. It adds new internal and external characteristics, grouping them under the name of software product quality. The main change made is the addition of the ***Compatibility*** characteristic, which is related to the possibility of exchanging information between systems, and the ***Security*** characteristic, which is related to the concepts of confidentiality and access to information [6].

Each of these software product quality characteristics is subdivided into sub-characteristics that define them in more detail, as shown in Figure 1.

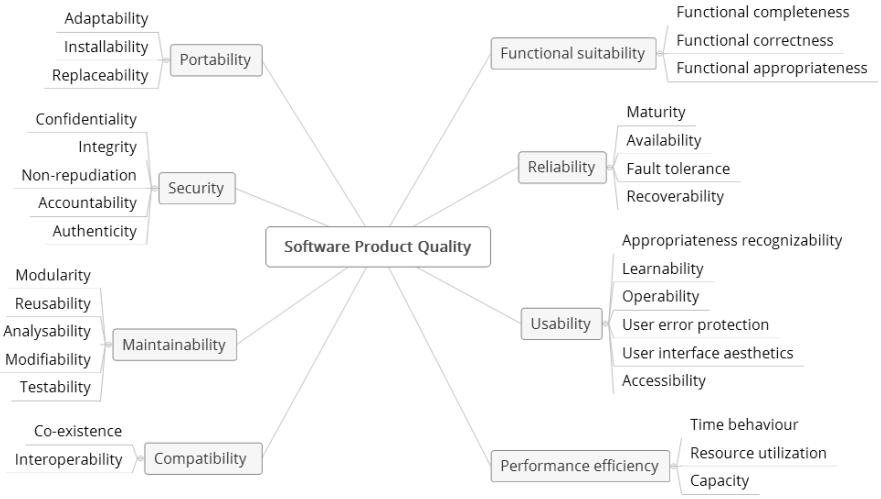

Figure 1- Software product quality characteristics

ISO/IEC 25040 - Evaluation process.

It replaces ISO/IEC 14598-1:1999. The new version defines 13 processes in five stages:

1) Establishing evaluation requirements: a. Establishing the purpose of the evaluation. b. Obtaining product quality requirements. c. Identifying the parts of the product that should be evaluated. d. Defining the strictness of the evaluation.

2) Specifying the evaluation: a. Selecting the evaluation modules. b. Defining decision criteria for metrics. c. Defining decision criteria for the evaluation.

3) Designing the evaluation: a. Planning evaluation activities.

4) Performing the evaluation: a. Carrying out the measurements. b. Applying decision criteria for metrics. c. Applying decision criteria for the evaluation.

5) Finishing the evaluation: a. Reviewing the results of the evaluation. b. Creating the report for the evaluation. c. Reviewing the quality of the evaluation and obtaining feedback. d. Treating evaluation data [7].

2.2 GQM. (Goal, Question, Metric)

GQM (Goal, Question, Metric) is a method that uses a metric to measure certain goal in a given way. The measurement model has three levels:

- Conceptual Level (Goal): a goal is defined for an object, which can be a product, a process or a resource, in relation to several quality models, from several points of view and relative to a specific environment.
- Operational Level (Question): a set of questions is refined based on the goal and aimed at checking if the goal is met. These questions are intended to characterize the object being measured (product, process or resource) in relation to a given quality issue and establishing its quality from that point of view.
- Quantitative Level (Metric): a set of metrics, which can be objective or subjective, is linked to every question, so that each question can be answered quantitatively.

A GQM model is developed by identifying a set of quality and/or productivity goals, at a corporate, division or project level. From those goals and based on models of the object being measured, questions are created to define those goals as thoroughly as possible. The next step is specifying the measures that should be taken to answer the questions and follow up in relation to how the products and processes meet those goals. Once the measures are specified, information compilation procedures should be developed, including validation and analysis procedures [5].

3 Quality Characteristics Evaluation Model

The model developed here consists in defining a set of questions based on the GQM approach that will then show, through logical connectors, to which extent the goals proposed are met.

As case study, we considered the characteristic ***Security***, which includes the following sub-characteristics: *Confidentiality*, *Integrity*, *Non-Repudiation*, *Responsibility* and *Authenticity*.

CONFIDENTIALITY: It evaluates the ability to protect against non-authorized access to data and information, be this accidental or deliberate.

INTEGRITY: It evaluates the ability of the system or computer to prevent non-authorized accesses or modifications to computer data or programs.

NON-REPUDIATION: It evaluates the ability to prove the actions or events that have taken place, so that such actions or events cannot be repudiated later on.

RESPONSIBILITY: It evaluates the ability to unequivocally track the actions carried out by an entity.

AUTHENTICITY: It evaluates the ability to prove the identity of an individual or a resource.

3.1 Questionnaire

Based on the characteristics described above, 33 true/false questions were defined. Table 1

Table 1. Questionnaire for the ***Security*** characteristic

ID	QUESTION
Q1	Is it a requirement for the password to be at least 8 characters long?
Q2	Is it a requirement for the password to include both upper- and lower-case characters?
Q3	Is it a requirement for the password to include both numbers and letters?
Q4	Is it a requirement for the password to include special characters?
Q5	Does the system use secure connection through HTTPS?
Q6	Are database data encrypted?
Q7	Does the system allow access to functionalities to which no permissions have been granted?
Q8	Does the system allow access to the database by any individual?
Q9	Does the system allow access to application server code by any individual?
Q10	Is the physical server accessible to any individual?
Q11	Is the remote server accessible to any individual?
Q12	Does the system redirect to non-secure sites?
Q13	Does the system request registration confirmation via e-mail when a new user registers?
Q14	Does the system allow any individual to modify the database?
Q15	Does the system allow any individual to modify the application server code?
Q16	Does the system allow SQL injections?
Q17	Does the system keep a history of actions performed?
Q18	Does the system have data encryption algorithms?
Q19	Does the system have a cryptographic method, such as digital signature?
Q20	Does the system request confirmation when an action is performed?
Q21	Is the system protected with SSL certificates?
Q22	Does the system issue a warning when accessing from an unknown location?
Q23	Does the system send an e-mail report of the operations done?
Q24	Does the system keep a record of date and time of logins to the system?
Q25	Does the system record the type of browser and operating system used to enter the site?
Q26	Does the system record the IP address from which the site is accessed?
Q27	Does the system check identity through a digital certificate?
Q28	Does the system have two-step verification?
Q29	Is a second-level key required to enter the system?
Q30	Does the system check identity through biometric data?
Q31	Does the system check identity through a code card?
Q32	Does the system check identity through credentials?
Q33	Does the system check identity through a digital signature?

3.2 Evaluation Criteria (EC) Description

To achieve our objective, the answers to the questions were combined through logic, establishing a score for each EC.

Table 2. Evaluation Criteria (EC) Description

ID	*Name*	*Description*	*Equation*	*Points*
C-1	Secure connections	A connection is considered to be secure if it uses HTTPS and there is no redirectioning to non-secure sites.	Q5 & ~Q12 = T	1
C-2	Access control	No unauthorized access to functionalities, the database, application code, and physical or remote servers should be allowed.	if Q7 | Q8 | Q9 | Q10 | Q11 = F	1
C-3	Data encryption	Database data should be encrypted.	Q6 = T	1
C-4	Low-level password	A password is considered to be low level if it is less than 8 characters long, does not include upper and lower case, does not include letters and numbers and does not include special characters.	Q1 | Q2 | Q3 | Q4 = F	0
	Mid-level password	A password is considered to be mid-level if it is at least 8 characters long or it includes upper and lower case or letters and numbers or special characters.	Q1 | Q2 | Q3 | Q4 = T	0.5
	High-level password	A password is considered to be high level if it is at least 8 characters long and it includes upper and lower case, letters and numbers and special characters.	Q1 & Q2 & Q3 & Q4 = T	1
I-5	Access prevention	Unauthorized access to functionalities, the database and application code should be prevented; SQL injections should not be allowed.	Q7 | Q8 | Q9 | Q16 = F	1
I-6	Modification prevention	Unauthorized database data modification and application code should be prevented.	Q14 | Q15 = F	1
I-7	Data confirmation	Registration should be confirmed via e-mail.	Q13 = T	1
NR-8	Operations carried out	A history of actions should be available, or they should be sent by e-mail.	Q17 | Q23 = T	1

NR-9	Encryption method	There should be a data encryption algorithm or a cryptographic method, such as digital signature, or protection through SSL certificates.	Q18 | Q19 | Q21 = T	1
NR-10	Action confirmation	A confirmation should be requested when performing a given action.	Q20 = T	1
NR-11	Location registration	A notification should be issued if the system was accessed from an unknown location.	Q22 = T	1
R-12	Action and data records	A history of actions should be available, or a record showing system access date and time or the IP address from which the system was accesses and the type of browser and operating system used,	Q17 | Q24 | Q25 | Q26 = T	1
R-13	Location control	A notification should be issued if the system is accessed from an unknown location.	Q22 = T	1
A-14	Identity verification	The system should check identity through any of the following methods: biometric data, code card, credentials, digital signature or digital certificate.	Q27 | Q30 | Q31 | Q32 | Q33 = T	1
A-15	Additional verification	Two-step verification should be used, or a second level key should be required to enter the system, or a registration confirmation via e-mail should be used.	Q28 | Q29 | Q13 = T	1

3.3 Metrics for Each Sub-Characteristic.

EC were combined to define the metrics that meet the goals of each sub-characteristic. For each of them, a name, a purpose, an application method, input values and equation used were defined.

Confidentiality.
Metric: *Confidentiality*
Purpose: *How efficient is the system when protecting against non-authorized access to data and information, be this accidental or deliberate?*
Application method: *Answering EC questions corresponding to the sub-characteristic "Confidentiality" and calculating the score obtained by adding the score for each EC that meets the expected goal. "Total score" is the maximum score that can be obtained.*
Inputs: *A = Score obtained. B = Total score.*
Equation: $X = A/B$

Observations: The EC to be used are: C-1, C-2, C-3 and C-4.

Integrity.
Metric: *Integrity*
Purpose: How capable is the system to prevent non-authorized accesses or modifications to computer data or programs?
Application method: *Answering EC questions corresponding to the sub-characteristic "Integrity" and calculating the score obtained by adding the score for each EC that meets the expected goal. "Total score" is the maximum score that can be obtained.*
Inputs: *A = Score obtained. B = Total score.*
Equation: *X = A/B*
Observations: The EC to be used are: I-5, I-6 and I-7.

Non-Repudiation.
Metric: *Non-Repudiation*
Purpose: *How capable is the system to prove the actions or events that have taken place, so that such actions or events cannot be repudiated later on?*
Application method*: Answering EC questions corresponding to the sub-characteristic "Non-Repudiation" and calculating the score obtained by adding the score for each EC that meets the expected goal. "Total score" is the maximum score that can be obtained.*
Inputs*: A = Score obtained. B = Total score.*
Equation: *X = A/B*
Observations: The EC to be used are: NR-8, NR-9, NR-10 and NR-11.

Responsibility.
Metric: *Responsibility*
Purpose: *How capable is the system to unequivocally track the actions carried out by an entity?*
Application method*: Answering EC questions corresponding to the sub-characteristic "Responsibility" and calculating the score obtained by adding the score for each EC that meets the expected goal. "Total score" is the maximum score that can be obtained.*
Inputs: *A = Score obtained. B = Total score.*
Equation: *X = A/B*
Observations: The EC to be used are: R-12 and R-13.

Authenticity.
Metric: *Authenticity*
Purpose: *How capable is the system to prove the identity of an individual or a resource?*
Application method*: Answering EC questions corresponding to the sub-characteristic "Authenticity" and calculating the score obtained by adding the score for each EC*

that meets the expected goal. "Total score" is the maximum score that can be obtained.
Inputs: *A = Score obtained. B = Total score.*
Equation: *X = A/B*
Observations: The EC to be used are: A14 and A15.

The equations used for each sub-characteristic are listed in Table 3.

Table 3. Equations for Each Sub-Characteristic.

METRIC	EQUATION
CONFIDENTIALITY	(C1+C2+C3+C4)/4
INTEGRITY	(I5+I6+I7)/3
NON-REPUDIATION	(NR8+NR9+NR10+NR11)/4
RESPONSIBILITY	(R12+R13)/2
AUTHENTICITY	(A14+A15)/2

4 Case Studies

The evaluation process was carried out, following the structure presented in ISO/IEC 25040, for three web applications to evaluate their ***Security*** characteristic.

Case a) It has been in production for approximately 18 months, it has more than 3,200 users, and an average of 500 daily logins.
Case b) It has been in production for approximately 30 months, it has more than 160 users, and an average of 75 daily logins.
Case c) It is in a testing stage, it has ten users with a minimal login frequency by the users that are testing the application.

4.1 Establishing evaluation requirements

The purpose of the evaluation is to measure the security of three web systems by analyzing different aspects of such systems. The characteristic of "***Security***," as defined in ISO/IEC 25010, was selected.
Two of the systems to be evaluated are in their final version and are currently being used by different users. The remaining system is running in a test version and is being used by different people responsible for carrying out tests.

4.2 Specifying the evaluation

The metrics used for each sub-characteristic are those defined in Section 3. Acceptance criteria for these sub-characteristics are:

Not acceptable: 0 <= X < 40

Minimally acceptable: 40 <= X < 60

Target range: 60 <= X < 90

Exceeds requirements: 90 <= X <=100

The result shall be considered as acceptable if all sub-characteristics are rated as Minimally acceptable, Target range or Exceeds requirements.

4.3 Designing the evaluation

To carry out the evaluation, three developers from the web systems to be analyzed (one developer from each system) were asked to use T/F to answer the questions listed in Section 3.1. They were asked to use an Excel spreadsheet that was provided to them, which was set so that values A and B corresponding to each metric were automatically calculated.

4.4 Carrying out the evaluation

The evaluation was carried out as planned and the following results were obtained:

Case a) *Confidentiality* 88%, *Integrity* 67%, *Non-Repudiation* 50%, *Responsibility* 50% and *Authenticity* 0%

Case b) *Confidentiality* 75%, *Integrity* 100%, *Non-Repudiation* 75%, *Responsibility* 50% and *Authenticity* 50%

Case c) *Confidentiality* 0%, *Integrity* 33%, *Non-Repudiation* 50%, *Responsibility* 0% and *Authenticity* 0%

4.5 Completing the evaluation

Case study a) has the sub-characteristics *Confidentiality* and *Integrity* within the acceptable range, *Non-Repudiation* and *Responsibility* are minimally acceptable, and *Authenticity* is unacceptable.

Case study b) exceeds expectation in relation to *Integrity*, has acceptable *Confidentiality* and *Non-Repudiation* characteristics, and *Responsibility* and *Authenticity* that are minimally acceptable.

Case study c) is minimally acceptable as regards *Non-Repudiation*, and unacceptable for *Responsibility*, *Authenticity*, *Integrity* and *Confidentiality*.

Figure 2 shows a comparison of the sub-characteristics evaluated in each of the cases.

Characteristic Analysis – *Security*

Case study a) does not meet evaluation requirements since its *Authenticity* sub-characteristic is in an unacceptable range: the system does not check identity through any valid method and it does not have a two-step verification process, a second-level key or registration confirmation via e-mail.

Case study b) is considered to meet evaluation requirements because all of its sub-characteristics are within an acceptance range.

Case study c) does not meet evaluation requirements because only its *Non-Repudiation* sub-characteristic is in an acceptable range. In this case, the *Authenticity* sub-characteristic is unacceptable for the same reasons as in case study a). As regards *Responsibility*, it is considered to be unacceptable because the system does not keep a history of actions or an access record, nor does it issue a notification when it is accessed from an unknown location. In relation to *Confidentiality*, the system does not have secure connections or access control, and there is no encryption of database data or criteria set for creating secure passwords. Finally, as regards *Integrity*, even if the system uses a procedure to prevent non-authorized modifications to the database and system code, this is not enough to achieve an acceptance level because it does not have any procedures to prevent unauthorized access to functionalities and there are no data confirmations sent via e-mail.

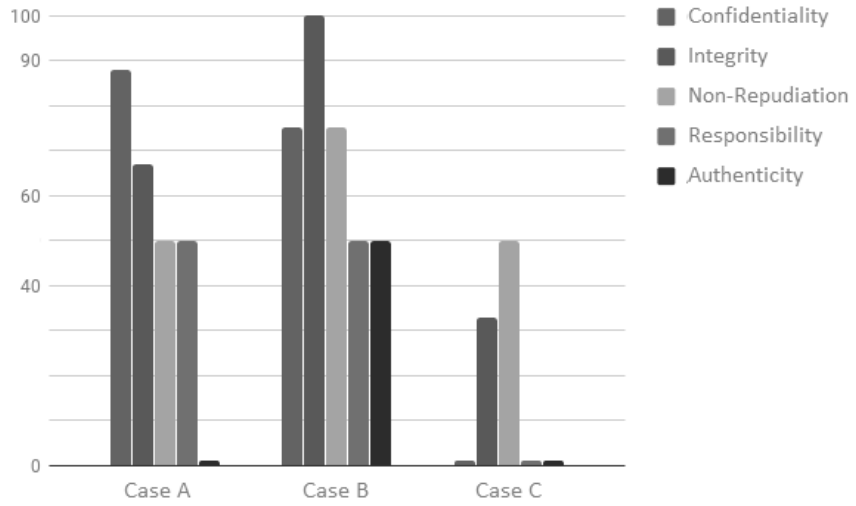

Figure 2 Sub-characteristics in each case study

5 Conclusions

The ISO/IEC 25010 standard offers a quality model to evaluate a set of characteristics applicable to a software product. We presented an evaluation model applicable to characteristics and sub-characteristics that is based on the GQM approach, which starts from a specific goal and then creates questions related to that goal. The answers to these questions are then combined to obtain the metric in question. Questions were created for the sub-characteristics of the ***Security*** characteristic, and a set of rules was generated to evaluate the answers to those questions. Then, these answers were combined to produce the metrics corresponding to each sub-characteristic and, therefore, for the characteristic as a whole.

Three web systems were evaluated, and only one successfully passed the test. The evaluation of the other systems was useful to detect shortcomings in them.

In the future, we plan to expand the model by creating questions and evaluation criteria for the other characteristics included in the ISO/IEC 25010 standard.

References

1. S. Esponda, P. Pesado, "Ambiente para la ayuda a la mejora de procesos en las PyMEs", 2013. Master dissertation, School of Computer Science – National University of La Plata.
2. ISO, "ISO/IEC 25000:2014 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Guide to SQuaREtle," 2014.
3. IRAM and ISO, "IRAM-NM-ISO IEC 9126-1 Information technology. Software engineering. Product quality. Part 1 - Quality model.," 2009.
4. IRAM;ISO, "IRAM-ISO-IEC 14598-1 Information technology. Software engineering. Software product evaluation. Part 1: General overview," 2006.
5. V. R. Basili, G. Caldiera, and H. D. Rombach, "The goal question metric approach" Encyclopedia of software engineering, 1994, vol. 2, no 1994, p. 528-532.
6. ISO, "ISO/IEC 25010:2011 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models," 2011.
7. ISO, "ISO/IEC 25040:2011 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Evaluation process," 2011.