ISO-BASED MODELS TO MEASURE SOFTWARE PRODUCT QUALITY

Alain Abran[1], Rafa E. Al-Qutaish[1],
Jean-Marc Desharnais[1] and Naji Habra[2]
[1]École de Technologie Supérieure, 1100 Notre-Dame Ouest, Montréal, Québec H3C1K3, Canada
[2]Institut d'Informatique, Universitaire de Namur, Rue Grandgagnage 21, 5000 Namur, Belgium
alain.abran@etsmtl.ca, rafa@rafa-elayyan.net,
jean-marc.desharnais@etsmtl.ca, nha@info.fundp.ac.be

ABSTRACT
*The ISO is developing a new ISO 25000 series on **S**oftware Product **Qua**lity **R**equirements and **E**valuation (SQuaRE) to improve the interpretation and use of quality measures for software products. This chapter explains how the ISO 19539 Measurement Information Model can be used to implement the ISO 9126 models for software product quality. It also identifies some of the harmonization issues arising as a result of the addition of new documents like ISO 25020 and ISO 25021, in particular with respect to previously published measurement standards for software engineering.*

KEYWORDS: ISO 25020, ISO 25021, ISO 9126, ISO 15939, ISO 25000-SQuaRE, Software Measurement, Software Product Quality.

1. INTRODUCTION

In 1991, the ISO published its first international consensus on terminology for the quality characteristics for software product evaluation (ISO 9126:1991) [1]. During the period 2001 to 2004, the ISO published an expanded version, containing the ISO quality models and a consensus on inventories of proposed measures for these models. This version of the ISO 9126 series consists of one International Standard (IS) [2] and three Technical Reports (TR) [3-5]:

- ISO IS 9126-1: Quality Models
- ISO TR 9126-2: External Metrics[1]
- ISO TR 9126-3: Internal Metrics
- ISO TR 9126-4: Quality in Use Metrics

The ISO has recognized a need for further enhancement of ISO 9126, primarily as a result of advances in the information technologies (IT field) and changes in the IT environment. Consequently, the ISO is now working on the next generation of software product quality standards, which will be referred to as Software Product Quality Requirements and Evaluation (SQuaRE – ISO 25000). Once completed, this series of standards will replace the current ISO 9126 and ISO 14598 series of standards. The SQuaRE series is made up of five divisions:

- Quality management division (ISO 2500n)
- Quality model division (ISO 2501n)

[1] The term 'metrics' used in ISO 9126 is replaced by 'measures' in the new series of standards, in accordance with ISO 15939.

- Quality measurement division (ISO 2502n)
- Quality requirements division (ISO 2503n)
- Quality evaluation division (ISO 2504n)

One of the main objectives of the SQuaRE series and the current ISO 9126 series (which also constitutes the difference between them) is the coordination and harmonization of its contents with the ISO standard on software measurement process – ISO 15939 [6, 7]. In particular, the quality measurement division (ISO 2502n) will consist of the following five documents that are scheduled to replace the current four-part ISO 9126 series:

- ISO 25020: Measurement Reference Model and Guide [6]
- ISO 25021: Quality Measure Elements [7]
- ISO 25022: Measurement of Internal Quality
- ISO 25023: Measurement of External Quality
- ISO 25024: Measurement of Quality in Use

This chapter presents some of the harmonization issues arising with respect to previously published measurement standards for software engineering, including ISO 15939, ISO 14143-1 and ISO 19761, and proposes ways to address them using the measurement information model of ISO 15939 on software measurement process.

This chapter is organized as follows: Section 2 presents the ISO Measurement Information Model adopted in ISO 15939. Section 3 presents our solution for alignment of the ISO models of software product quality with the measurement information model of ISO 15939. Section 4 presents next the outstanding harmonization issues in terminology and coverage in ISO FDIS 25020 and ISO 25021. Finally, a discussion and a summary are presented in section 5.

2. ISO 15939 MEASUREMENT INFORMATION MODEL

The ISO has produced an information model within ISO 15939 (2002) (see Figure 1) to help in determining what must be specified during measurement planning, performance and evaluation.

Figure 1 shows that a specific measurement method is used to collect a base measure for a specific attribute. The values of two or more base measures can then be used in a computational formula (by means of a measurement function) to produce and construct a specific derived measure. These derived measures are in turn used in an analysis model to arrive at an indicator, which is a value, and to interpret the indicator's value in order to explain the relationship between it and the information needed, doing so in the language of the measurement user, to produce an Information Product for the user's Information Needs.

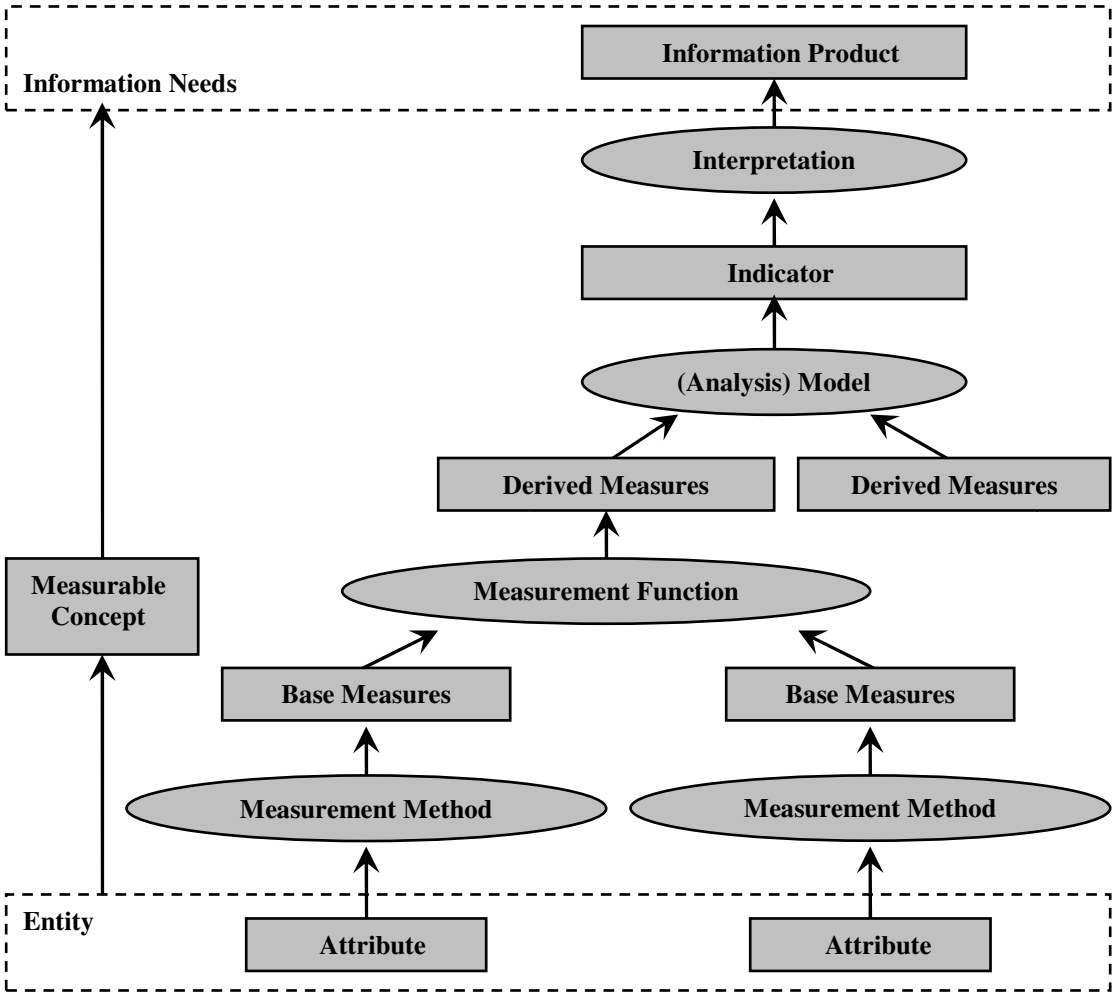

Figure 1: Measurement Information Model from ISO 15939 (2002)[2]

There already exists a very mature measurement terminology which is well documented in the ISO International Vocabulary of Basic and General Terms in Metrology (ISO VIM) [8]. This terminology is widely accepted and used in most fields of science, and has been adopted in ISO 15939 [9] as the agreed-upon measurement terminology for software and system engineering-related ISO standards.

3. MAPPING THE QUALITY MODEL TO THE MEASUREMENT INFORMATION MODEL

The following two expressions come from the ISO standard on software measurement process, ISO 15939 [9], which is itself based on the definitions in the ISO International Vocabulary of Basic and General Terms in Metrology (VIM 1993) [8]:

- Base measure: a measure defined in terms of an attribute and the method for quantifying it. A base measure is functionally independent of other measures.
- Derived measure: a measure defined as a function of two or more values of base measures. A transformation of a base measure using a mathematical function can also be considered as a derived measure.

[2] We added the arrows to the ISO 15939 Measurement Information Model to point up the dataflow relationships. Ovals represent activities, and rectangles represent the input and output of an activity.

In practice, the data collection associated with a property of an object (or concept), and quantification of that property, occurs at the base measure level, at which time a measurement unit is assigned based on the rules of the measurement method used for the quantification.

At the derived measure level, the base measures have been already collected and are being assembled according to the rules of combination (e.g. a computational formula) defined within each derived measure. A derived measured is therefore the product of a set of properly combined measurement units (through a measurement function). This combination is then labelled to represent an attribute (of a quality characteristic or subcharacteristic) of a software product.

Table 1 shows examples of base measures used in the measure definitions documented in ISO 9126-2, -3 and -4 (see Table A1 in Appendix A for the complete list of base measures). Table 1 shows the name of each base measure and the unit of measurement that is assigned to its value. These base measures can be used to calculate each of the derived measures (akin to metrics) in ISO 9126-4.

Table 1: Examples of Base Measures in ISO 9126-4.

Quality in Use Base Measures		
	Measure Name	**Unit of Measurement**
1	Task Effectiveness	(a given weight)
2	Total Number of Tasks	Task (number of)
3	Task Time	Minute
4	Cost of the Task	Dollar
5	Help Time	Second
6	Error Time	Second
7	Search Time	Second
8	Number of Users	User (number of)
9	Total Number of People Potentially Affected by the System	Person (number of)
10	Total Number of Usage Situations	Situation (number of)

Each of these base measures must be collected individually. They can be used at least once, or multiple times, to obtain the derived measure required to quantify the software properties specified in the ISO 9126 quality model. Table 2 provides an example of where some base measures are used throughout ISO 9126-2. For instance, the base measure ‘number of inaccurate computations encountered by users’ is used only once in ‘external functionality – accuracy measures’, while the base measure ‘number of items requiring compliance’ can be used in 6 subcharacteristics of external quality (ISO 9126-2). The construction of derived measures is based on a computational formula consisting of two or more base measures (see Table A2 in Appendix A for the complete cross-reference of base measure usage).

Table 2: Examples of the Use of Base Measures in ISO 9126-2.

	Measure Name	External																								
		Functionality					Reliability				Usability					Efficiency			Maintainability					Portability				
		F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5
1	Number of functions	✓									✓	✓	✓											✓			✓	
2	Operation time		✓	✓	✓		✓		✓			✓	✓			✓	✓			✓	✓	✓		✓		✓		
3	Number of inaccurate computations encountered by users		✓																									
4	Number of data formats			✓																								
5	Number of illegal operations				✓																							
6	Number of items requiring compliance					✓				✓					✓			✓					✓					✓
7	Number of interfaces requiring compliance					✓																						
8	Number of faults						✓																	✓	✓			

Such lists of base measures and usage cross-references are currently missing from ISO 9126 and would be helpful to those designing programs for implementing measurement of the quality of software products using ISO 9126 quality models and related measures. In particular, these lists can help in:

- Identifying, selecting and collecting a base measure (once), and then using this base measure to evaluate a number of derived measures;
- Knowledge of which base measures are required to evaluate specific software quality attributes (characteristics and subcharacteristics).

Below, we present a mapping of both the measures and the quality models in ISO 9126 to the measurement information model described in ISO 15939.

First, we refer to the bottom section of Figure 2 by the term ‘Data Collection’ (e.g. the measurement methods and the base measures), the middle section by the term ‘Data Preparation’, using agreed-upon mathematical formulas and related labels (e.g. measurement functions and derived measures), and the top section by the term the ‘Data Analysis’ (e.g. analysis model, indicator and interpretation).

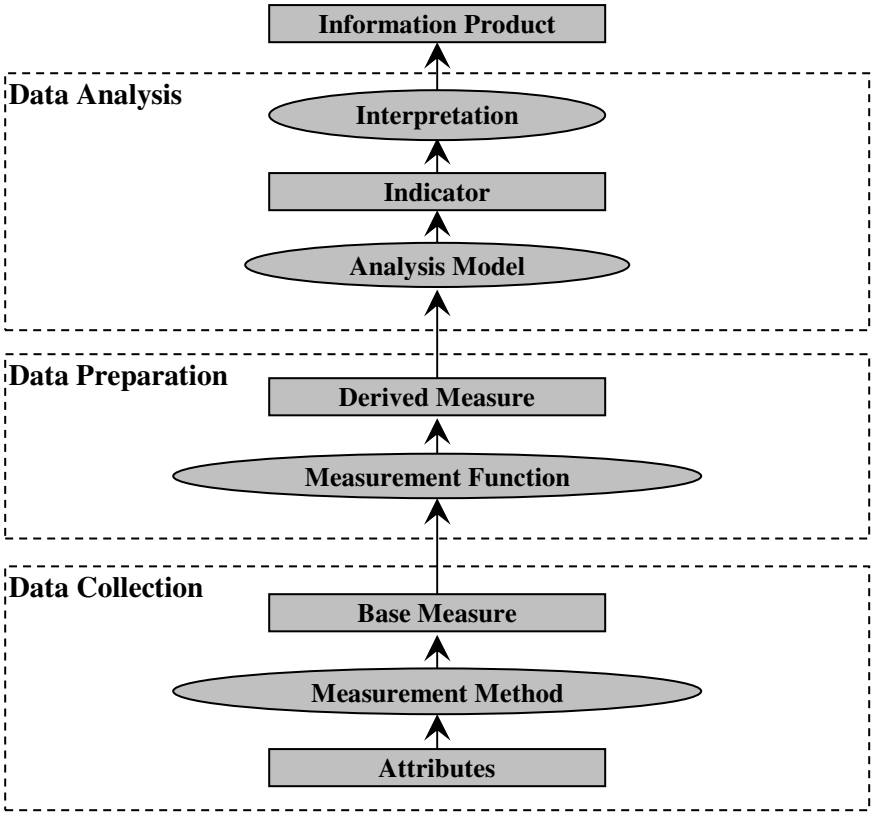

Figure 2: ISO 15939 (2002) Measurement Information Model – three different sections.

As we have already discussed both data collection and data preparation, we now focus on the 'Data Analysis' section. It is in the 'Analysis Model' part of the ISO 15939 measurement information model that the ISO 9126 models of software product quality are to be used. Figures 3, 4 and 5 present these generic models of ISO 9126 [2].

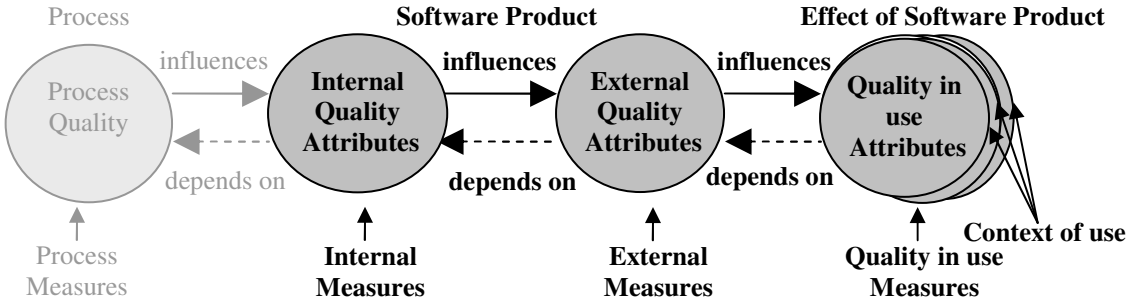

Figure 3: Quality in the life cycle – ISO 9126-1 [2].

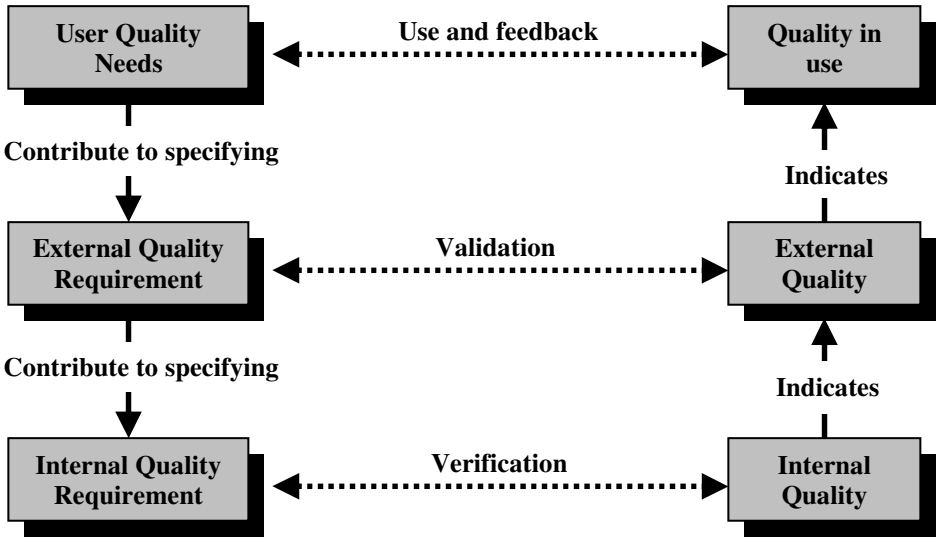

Figure 4: Quality in the software life cycle – ISO 9126-1 [2].

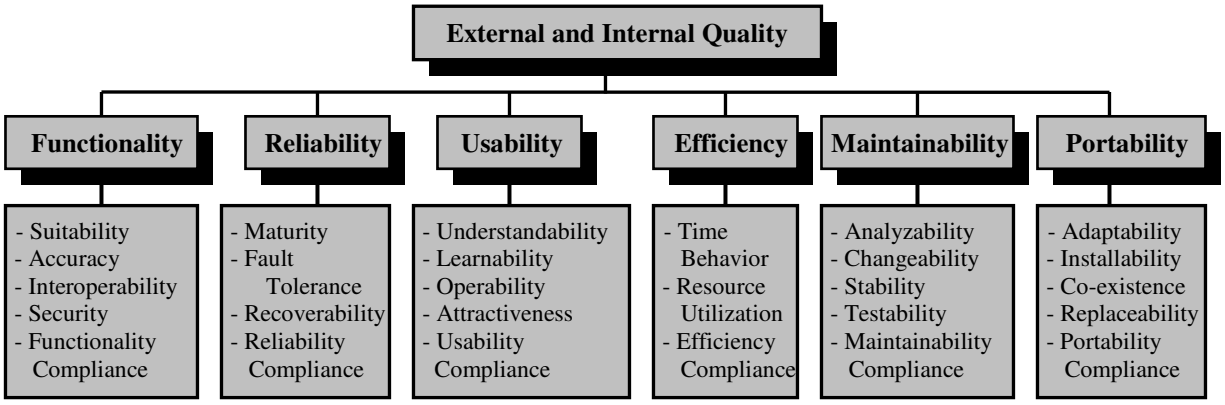

Figure 5: Quality model for External and Internal Quality – ISO 9126-1 [2].

These generic ISO models (Figures 3, 4 and 5) are to be instantiated in any particular context of measuring the quality of a specific software product. This is usually performed in a four-step process:

1. Identification of quality-related requirements, that is, the selection of the parts of the ISO quality models that are relevant to a particular context of quality evaluation (Figure 5);
2. Identification of the context of interpretation (Figure 6), that is:
   - the selection of reference values, such values being either generic or specific threshold values, or
   - the determination of targets specified for a particular context;
3. Use of the derived measures from the data preparation phase to fill out the instantiated quality model determined in 1 (Figure 5);
4. Comparison of the results of step 3 with either the set of reference values or targets determined in step 2 (Figure 6).

This process is summarized in Figure 6, and a number of examples are presented in the side box.

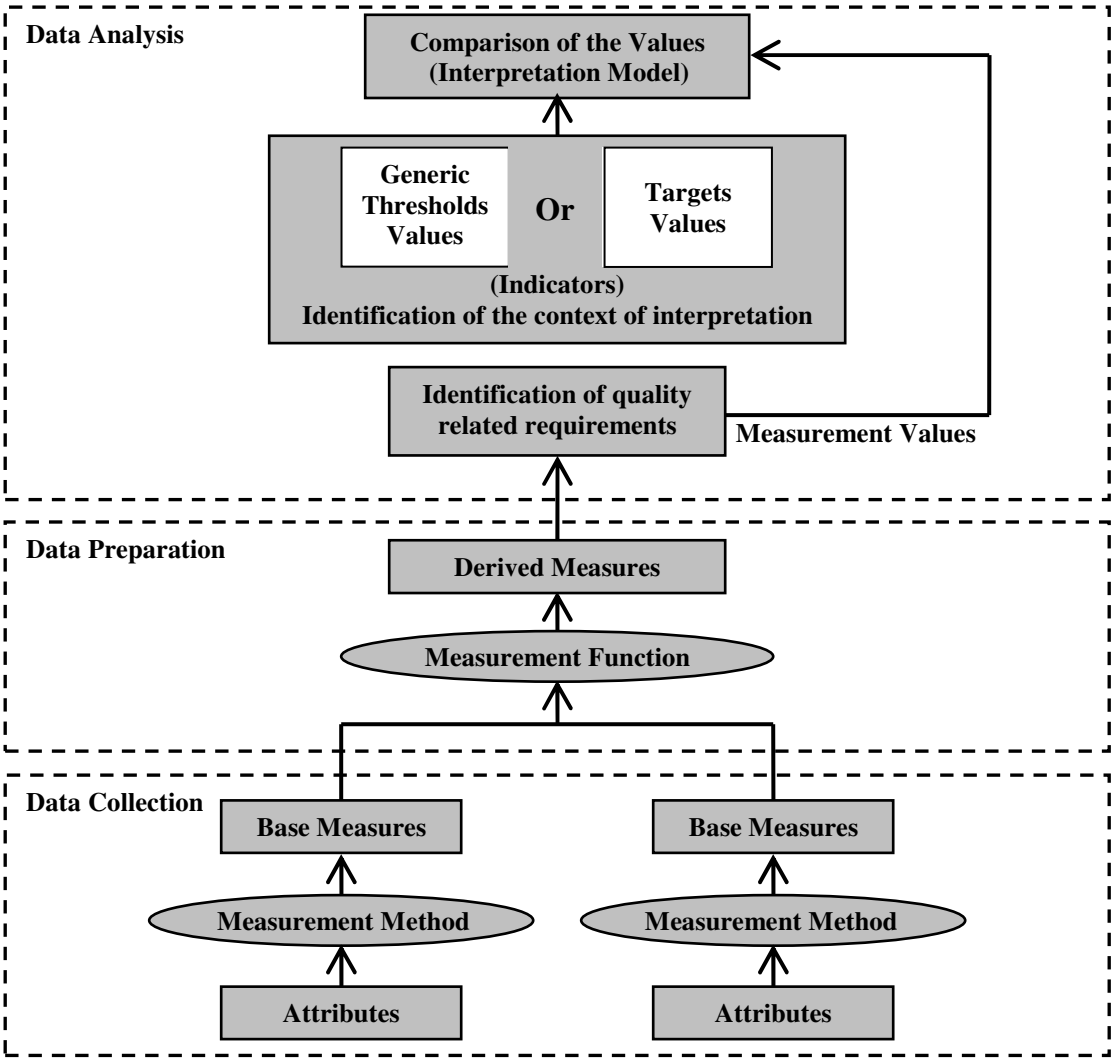

Figure 6: Mapping to the Measurement Information Model.

Examples

The examples presented here illustrate the process described in Figure 6. We include some of the ISO 9126 base measures, and show the way in which they are combined to construct a derived measure using a computational formula (measurement function):

Example 1:

Data Collection:

- Base Measure 1 (B1): Number of inaccurate computations encountered by users.
- Base Measure 2 (B2): Operation time.

Data Preparation:

- Derived Measure: B1 / B2
- Name of Derived Measure: Computational accuracy.

Data Analysis

- Quality group name: External quality measures.
- Characteristic: Functionality. - Subcharacteristic: Accuracy.
- Comparison of values obtained with the indicators (generic thresholds and/or targets).

Example 2:

Data Collection:

- Base Measure 1 (B1): Number of detected failures.
- Base Measure 2 (B2): Number of performed test cases.

Data preparation:
- Derived Measure: B1 / B2
- Name of Derived Measure: Failure density against test cases.

Data Analysis:
- Quality group name: External quality measures.
- Characteristic: Reliability. - Subcharacteristic: Maturity.
- Comparison of values obtained with the indicators (generic thresholds and/or targets).

Example 3:

Data Collection:
- Base Measure 1 (B1): Task time.
- Base Measure 2 (B2): Help time.
- Base Measure 3 (B3): Error time.
- Base Measure 4 (B4): Search time.

Data Preparation:
- Derived Measure: (B1-B2-B3-B4) / B1
- Name of Derived Measure: Productive proportion.

Data Analysis:
- Quality group name: Quality in Use measures.
- Characteristic: Productivity.
- Comparison of values obtained with the indicators (generic thresholds and/or targets).

Example 4:

Data Collection:
- Base Measure 1 (B1): Number of errors made by user.
- Base Measure 2 (B2): Number of tasks.

Data Preparation:
- Derived Measure: B1 / B2
- Name of Derived Measure: Error frequency.

Data Analysis:
- Quality group name: Quality in Use measures.
- Characteristic: Effectiveness.
- Comparison of values obtained with the indicators (generic thresholds and/or targets).

4. OUTSTANDING HARMONIZATION ISSUES

4.1 Terminology

The ISO 9126 working group has come up with three new expressions which they introduced in ISO TR 25021, namely: *quality measure element categories*, *quality measure elements* and *quality measures* [6, 7]. The introduction of these new terms raises the following concern: either the proper mapping to the set of classic metrology terms has not yet been completed or there are concepts and related terms missing in the metrology vocabulary. The latter would be surprising, since metrology is a rather mature domain of knowledge based on centuries of expertise in the field of measurement and related international standardization. In this section, we revisit the new documents, ISO FDIS 25020 and ISO 25021, with a view to recommending a proper mapping of concepts to the related metrology terms, as well as to ISO 15939.

In ISO 25021, it is claimed that a quality measure element is either a base measure or a derived measure [6, 7]; however, the consensual metrology terms are then ignored in favor of locally defined WG6 measures, thus bypassing the ISO and SC7 harmonization requirements on measurement terminology.

Quality measure elements are described as an input for the measurement of the software quality measures of external quality, internal quality and quality in use [6, 7]. Figure 7 shows

the relationship between the quality measure elements and the software quality measures, and between the software quality measures and the quality characteristics and subcharacteristics. In metrology, these would correspond to base measures and derived measures respectively. It can be observed that these measures, in particular the derived measures, are defined specifically to measure the subcharacteristics of internal and external quality or the characteristics of quality in use. None of these is directly related to the top level of software quality (which is itself broken down into three models, then into 16 characteristics and further into a large number of subcharacteristics). Therefore, the expression 'software quality measures', which was selected in ISO 25021, is at a level of abstraction that does not represent the proper mapping of the measures to the concept being measured.

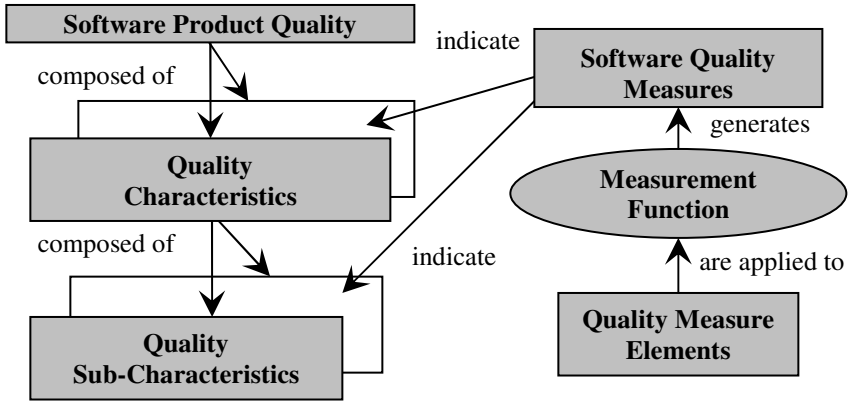

Figure 7: Quality Measure Elements Concept in the Software Product Quality Measurement Reference Model [6, 7]

Now, there are 15 categories of quality measure elements [7] in ISO FDIS 25021 (see Table 3)[3].

Table 3: The 15 Quality Measure Elements Categories

Data Size	Number of I/O Events	Number of Test Cases
Number of Data Items	Number of Requirements	Number of Trials
Number of Failures	Number of Restarts	Number of User Operations
Number of Faults	Number of System Operations	Product Size
Number of Functions	Number of Tasks	Time Duration

It can be observed in Table 3 that a number of the quantities have a label starting with 'number of'. However, these do not use a reference scale typical of measures in the sciences or in engineering, but are rather counts of entities. For any of these proposed counts, such as Number of Functions, no specific method is proposed for identifying this number in a consistent manner across measurers and organizations; for instance, the definition of the word 'function' could differ from one individual to another within a single organization, and more so across organizations. Therefore, to say in ISO TR 25021 that such numbers are obtained by

[3] It is to be noted that in ISO DIS 25021 there are no specific quality measure elements proposed within the Number of User Operations or Number of System Operations categories.

an ‘objective’ method is an overstatement, since they must be obtained mostly on the basis of the judgment of the person performing the count.

Of the 15 categories, only ‘time’ comes from a classic base measure, in that it uses the international standard unit of the second (or a multiple or submultiple of it) as its reference scale. There are also measuring instruments to ensure that time measurements are indeed obtained in an objective manner.

It can also be observed that, of the 15 categories in Table 3, at most four are directly related to the quality of software: number of faults, number of failures, number of restarts and number of trials. None of the other 11 measures is directly or indirectly related to the quality of software. In fact, they are strictly independent of it, as they are often used for normalization purposes, for instance.

For the Product Size category, ISO TR 25021 lists only one way to measure this dimension, that is, in non-comment lines of code. There are also other ways to measure product size, such as function points, modules, classes and visual structures. Furthermore, there are various methods for counting lines of code and for measuring function points. Therefore, this quality measure element category could be further split into different quality measure elements (base measures). Moreover, the ISO has specified mandatory requirements for functional size measurement methods (ISO 14143-1) [10], and has recognized four of these as ISO standards which meet these requirements, such as COSMIC-FFP [11]. None of these existing ISO software engineering standards, which are referenced in ISO 90003 [12], has been mentioned or referenced in ISO TR 25021. Also, the various methods available to obtain those numbers have their strengths and weaknesses, from a measurement perspective, in terms of repeatability, reproducibility, software domains of applicability and accuracy.

In summary, from our point of view, it is not necessary to introduce new terms such as *quality measure element categories*, *quality measure elements* or *quality measures*: the terminology and concepts currently in ISO VIM [8] and ISO 15939 [9] are sufficient.

4.2 Limited coverage of the ISO quality models and corresponding measures

ISO TR 9126, parts 2 to 4, presents the ISO inventory of measures for the full coverage of the ISO software product quality models (internal quality, external quality and quality in use) for measuring any of their quality characteristics and subcharacteristics. The full sets of base measures in these three parts of ISO 9126 is presented in Appendix A, and includes 82 base measures.

Of these 82 base measures, only 57 are included in ISO 25021; this means that the coverage in this new ISO document is limited, and the reasons for this are not obvious. In addition, out of the 197 measures in ISO 9126, only 51 were selected for ISO 25021 as quality measures. The content coverage of this subset of quality measures (derived measures) is limited, and no specific criteria are provided on how they were selected. Some generic information is provided in this standard to suggest that they were derived from a questionnaire-based survey; however, no information is provided about the criteria for selection, the size and representativeness of the sample in the countries where the data were

collected, nor about the representativeness of this sample outside these countries. Another claim, that "they represent a default kernel of quality measures which are proven to be used in common practice" [7], is not supported by documented evidence, nor is there a discussion of their generalizability outside their data collection context.

Appendix B (Tables B1, B2 and B3) presents a detailed analysis of the coverage of the quality measures in ISO 25021, together with their corresponding availability in ISO 9126. Table B1 specifically illustrates that 34 measures for the external quality of a software product were selected in ISO 25021 out of an inventory of 112 in the corresponding ISO 9126-2, while 78 measures were excluded, again without a documented rationale.

Table B2 provides similar information on measures for the internal quality of a software product, as selected for ISO 25021; out of 70 measures, only 15 were selected, and they cover only 4 of the 6 quality characteristics of the ISO model of internal quality, and only 9 of 27 subcharacteristics; again, there is no documented rationale for excluding any characteristic or subcharacteristic. The same is true of Table B3 for the Quality in Use quality measures:

- Included: only 2 measures of the 15 already available in ISO 9126-4;
- Excluded: 2 quality-in-use characteristics, which are 'safety' and 'satisfaction';
- Also excluded: any quality measure elements related to the Number of User Operations and Number of System Operations.

4.3 Redundancy issues

Some additional information included in ISO 25021 has already been covered in ISO 9126 documents, and will be included in the ISO 25000 series; for instance, information about scale types is covered through rephrasing information contained in other documents, once again increasing synchronization and harmonization right away and over the long term.

The situation is similar for commentaries on the measures of internal software quality, external software quality and software quality in use, as well as for commentaries on the software measurement methods. This is contrary to the ISO practice of avoiding duplication, redundancy or the rephrasing of information across ISO documents, and increases the possibility of inconsistencies arising across documents, all of which could lead to significant effort over the long term in maintaining synchronization of documents covering similar subsets of information.

These examples point to configuration management issues over the long term which will represent additional cost to the purchasers of these ISO documents, since they will be required to pay twice for the same information, which is, in fact, a subset of the full inventory. This could lead to some confusion for standards users as to which of these documents is most valuable to a standard purchaser, and under what circumstances.

We have illustrated in this section how the issue of ambiguity and redundancy in ISO FDIS 25020 and ISO 25021 with respect to the use of the new terms *quality measure elements categories*, *quality measure elements* and *quality measures*, which can be avoided through the use of the corresponding metrology concepts and terms.

5. SUMMARY & DISCUSSION

Within their mandate to upgrade its set of technical reports on the measurement of the quality of software products (ISO 9126), the ISO has come up with a new structure for upgrading the current series of ISO 9126 documents for the measurement of the quality of software products. This new structure is referred to as Software Product Quality Requirements and Evaluation, or SQuaRE. In this chapter, we have presented practitioners with an alignment of the ISO models of software product quality with the measurement information model of ISO 15939, and explained how to use them for data collection, data preparation and data analysis. Some examples have also been provided.

In addition, some issues have been raised in this chapter concerning three new concepts proposed in SQuaRE; that is, *quality measure element categories*, *quality measure elements* and *quality measures* The following is a summary of the harmonization issues identified:

A) Terminology in ISO 25021:

- what is referred to as a 'quality measure element' corresponds to the classic concept of the 'base measure' in ISO 15939;
- what is referred to as 'software quality measure':
  - corresponds to the classic concept of the 'derived measure' in ISO 15939;
  - is not at the proper level of abstraction for the concept being measured when mapped to the hierarchy of concepts for software product quality adopted by the ISO.
- In both ISO FDIS 25020 and ISO 25021, the 'measurement method' is defined as "a logical sequence of operations, described generically, used in quantifying an attribute with respect to a specific scale." But, in ISO 25021, it is used in the 'set of quality measure elements' to represent the type of the measurement method (objective or subjective). In contrast, a new field called 'detail' is used to represent the measurement method.

B) Harmonization with the Information Model of ISO 15939:

- Unless the terminology is harmonized with ISO International Vocabulary of Basic and General Terms in Metrology, then it is challenging to align the older versions of ISO 9126 and ISO 14598, and it will be even more challenging with the ISO 25000 updates.
- Should the proposed terminology harmonization be accepted, it will become easier to map each of these ISO 9126 and ISO 14598 series to the ISO 15939 Information Model.

C) Description harmonization:

- A large number of the base measures proposed in ISO 25021 are counts of entities rather than measures per se with the required metrological characteristics, such as unit, scale, dimension, measurement method, measurement procedures, etc.
- In ISO 25021, in some instances, like Product Size for example, there is no reference to other existing ISO standards for software size, such as ISO 19761, etc.
- There are a number of claims that the proposed base measures are objective, while they are obviously derived from a manual process without precisely documented measurement procedures, thereby leaving much to the measurer's judgment.

D) Coverage harmonization in ISO 25021:

- The documented set of base measures represents only a limited subset of the base measures within ISO 9126, parts 2 to 4; the rationale for inclusion or exclusion is not documented.

- The set of base measures does not allow coverage of the full spectrum of quality characteristics and subcharacteristics in ISO 9126, parts 2 to 4; again, the rationale for inclusion or exclusion is not documented.

These concerns can be summarized as follows:

- *Quality measure element categories* and *quality measure elements*: non alignment with the classical terminology on measurement is puzzling.
- *Quality measures*: inconsistency in the terminology used, and ambiguity about which level of the ISO 9126 multi-level standard is being applied.

From the above analysis, we propose the following recommendations:

- Ensure that the terminology on software product quality measurement is fully aligned with the classical measurement terminology in the sciences and in engineering;
- Provide full coverage of the base measures for all three ISO models of software quality;
- Provide improved documentation of the base measures using the criteria from metrology;
- Provide clear mapping and traceability of the new ISO 25000 documents to the ISO 15939 Information Model.

We have analyzed some of weaknesses of the new terms, and have proposed ways to address them by using the ISO 15939 measurement information model on software measurement process. In summary, the previously defined terms such as 'base measure' and 'derived measure', as well as the proper mapping to the Measurement Information Model in well-developed standards like ISO 15939, and the International Vocabulary of Basic and General Terms in Metrology (ISO VIM) are more useful than the weakly defined terms that have recently been introduced.

REFERENCES

[1] ISO, *ISO/IEC IS 9126: Software Product Evaluation - Quality Characteristics and Guidelines for Their Use*. Geneva, Switzerland: International Organization for Standardization, 1991.

[2] ISO, *ISO/IEC 9126-1: Software Engineering - Product Quality - Part 1: Quality Model*. Geneva, Switzerland: International Organization for Standardization, 2001.

[3] ISO, *ISO/IEC TR 9126-2: Software Engineering - Product Quality - Part 2: External Metrics*. Geneva, Switzerland: International Organization for Standardization, 2003.

[4] ISO, *ISO/IEC TR 9126-3: Software Engineering - Product Quality - Part 3: Internal Metrics*. Geneva, Switzerland: International Organization for Standardization, 2003.

[5] ISO, *ISO/IEC TR 9126-4: Software Engineering - Product Quality - Part 4: Quality in Use Metrics*. Geneva, Switzerland: International Organization for Standardization, 2004.

[6] ISO, *ISO/IEC FDIS 25020: Software Engineering - Software Product Quality Requirements and Evaluation (SQuaRE) - Measurement Reference Model and Guide*. Geneva, Switzerland: International Organization for Standardization, 2007.

[7] ISO, *ISO/IEC TR 25021: Software Engineering - Software Product Quality Requirements and Evaluation (SQuaRE) - Quality Measure Elements*. Geneva, Switzerland: International Organization for Standardization, 2007.

[8] ISO, *International Vocabulary of Basic and General Terms in Metrology (VIM)*. Geneva, Switzerland: International Organization for Standardization, 1993.

[9] ISO, *ISO/IEC IS 15939: Software Engineering - Software Measurement Process*. Geneva, Switzerland: International Organization for Standardization, 2002.

[10] ISO/IEC, *ISO/IEC 14143-1: Software Measurement - Functional Size Measurement - Part 1: Definition of Concepts*. Geneva, Switzerland: International Organization for Standardization, 1998.

[11] ISO/IEC, *ISO/IEC 19761: Software Engineering - COSMIC-FFP - A Functional Size Measurement Method*. Geneva, Switzerland: International Organization for Standardization, 2003.

[12] ISO/IEC, *ISO/IEC 90003: Software Engineering - Guidelines for the Application of ISO 9001:2000 to Computer Software*. Geneva, Switzerland: International Organization for Standardization, 2004.

APPENDIX A

Table A1: The List of ISO 9126 Base Measures and their Measurement Units.

External Base Measures		
Measure Name		**Unit of Measurement**
1	Number of Functions	Function (number of)
2	Operation Time	Minute
3	Number of Inaccurate Computations Encountered by Users	Case (number of)
4	Total Number of Data Formats	Format (number of)
5	Number of Illegal Operations	Operation (number of)
6	Number of Items Requiring Compliance	Item (number of)
7	Number of Interfaces Requiring Compliance	Interface (number of)
8	Number of Faults	Fault (number of)
9	Number of Failures	Failure (number of)
10	Product Size	Byte
11	Number of Test Cases	Case (number of)
12	Number of Breakdowns	Breakdown (number of)
13	Time to Repair	Minute
14	Down Time	Minute
15	Number of Restarts	Restart (number of)
16	Number of Restoration Required	Restoration (number of)
17	Number of Tutorials	Tutorial (number of)
18	Number of I/O Data Items	Item (number of)
19	Ease of Function Learning	Minute
20	Number of Tasks	Task (number of)
21	Help Frequency	Access (number of)
22	Error Correction	Minute
23	Number of Screens or Forms	Screens (number of)
24	Number of User Errors or Changes	Error (number of)
25	Number of Attempts to Customize	Attempt (number of)
26	Total Number of Usability Compliance Items Specified	Item (number of)
27	Response Time	Second or Millisecond
28	Number of Evaluations	Evaluation (number of)
29	Turnaround Time	Second or Millisecond
30	Task Time	Minute
31	Number of I/O Related Errors	Error (number of)
32	User Waiting Time of I/O Device Utilization	Second or Millisecond
33	Number of Memory Related Errors	Error (number of)
34	Number of Transmission Related Errors	Error (number of)
35	Transmission Capacity	Byte
36	Number of Revised Versions	Version (number of)

37	Number of Resolved Failures	Failure (number of)
38	Porting User Friendliness	Minute

Internal Base Measures		
Measure Name		**Unit of Measurement**
1	Number of Functions	Function (number of)
2	Number of Data Items	Item (number of)
3	Number of Data Formats	Formats (number of)
4	Number of Interface Protocols	Protocol (number of)
5	Number of Access Types	Access-Type (number of)
6	Number of Access Controllability Requirements	Requirement (number of)
7	Number of Instances of Data Corruption	Instance (number of)
8	Number of Compliance Items	Item (number of)
9	Number of Interface Requiring Compliance	Interface (number of)
10	Number of Faults	Fault (number of)
11	Number of Test Cases	Test-Case (number of)
12	Number of Restoration	Requirement (number of)
13	Number of Input Items Which Could Check for Valid Data	Item (number of)
14	Number of Operations	Operation (number of)
15	Number of Messages Implemented	Message (number of)
16	Number of Interface Elements	Element (number of)
17	Response Time	Second or Millisecond
18	Turnaround Time	Second or Millisecond
19	I/O Utilization (Number of Buffers)	Buffer (number of)
20	Memory Utilization	Byte
21	Number of Lines of Code Directly Related to System Calls	Line (number of)
22	Number of I/O Related Errors	Error (number of)
23	Number of Memory Related Errors	Error (number of)
24	Number of Items Required to be Logged	Item (number of)
25	Number of Modifications Made	Modification (number of)
26	Number of Variables	Variable (number of)
27	Number of Diagnostic Functions Required	Function (number of)
28	Number of Entities	Entity (number of)
29	Number of Built-in Test Function Required	Function (number of)
30	Number of Test Dependencies on Other System	Dependency (number of)
31	Number of Diagnostic Checkpoints	Checkpoint (number of)
32	Number of Data Structures	Data-Structure (number of)
33	Total Number of Setup Operations	Operation (number of)
34	Number of Installation Steps	Step (number of)

Quality in Use Base Measures		
Measure Name		**Unit of Measurement**
1	Task Effectiveness	(a given weight)
2	Total Number of Tasks	Task (number of)
3	Task Time	Minute

4	Cost of the Task	Dollar
5	Help Time	Second
6	Error Time	Second
7	Search Time	Second
8	Number of Users	User (number of)
9	Total Number of People Potentially Affected by the System	Person (number of)
10	Total Number of Usage Situations	Situation (number of)

Table A2: The Cross-Reference Table of ISO 9126 Base Measure Usage.

	Measure Name	Unit	External																							Internal																					Quality in use			
			Functionality					Reliability				Usability					Efficiency			Maintainability					Portability					Functionality					Reliability				Usability					Efficiency			Maintainability					Portability					Q1	Q2	Q3	Q4
			F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5	F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5				
1	Number of Functions	Function	✓									✓	✓	✓											✓			✓		✓	✓					✓			✓	✓	✓							✓														
2	Operation Time	Minute		✓	✓	✓		✓		✓			✓	✓			✓	✓			✓	✓	✓		✓		✓																																							
3	Number of Inaccurate Computations Encountered by Users	Case		✓																																																											
4	Number of Data Formats	Format			✓																										✓																																		
5	Number of Illegal Operations	Operation				✓																													✓																														
6	Number of Items Requiring Compliance	Item					✓				✓					✓			✓					✓					✓					✓			✓					✓					✓					✓					✓					✓				
7	Number of Interfaces Requiring Compliance	Interface					✓																										✓																																
8	Number of Faults	Fault						✓																	✓	✓																																									
9	Number of Failures	Failure						✓	✓									✓		✓		✓					✓																																											
10	Product Size	Byte						✓																																																											
11	Number of Test Cases	Case				✓		✓	✓														✓												✓																																
12	Number of Breakdowns	Breakdown							✓	✓																																																								
13	Time to Repair	Minute								✓																																																								
14	Down Time									✓																																																							
15	Number of Restarts	Restart								✓																																																								
16	Number of Restoration	Restoration								✓																												✓																												
17	Number of Tutorials	Tutorial										✓																																																					
18	Number of I/O	Item										✓																																																					

Measure Name		Unit	External																											Internal																											Quality in use			
			Functionality					Reliability				Usability					Efficiency			Maintainability					Portability					Functionality					Reliability				Usability					Efficiency			Maintainability					Portability					Q1	Q2	Q3	Q4
			F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5	F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5				
	Data Items																																																											
19	Ease of Function Learning	Minute											✓																																															
20	Number of Tasks	Task											✓				✓																											✓													✓			
21	Help Frequency	Freq.											✓																																															
22	Error Correction	Minute												✓																																														
23	Number of Screens/Forms	Screen												✓																																														
24	Number of User Errors/Changes	Error												✓																																										✓			
25	Number of Attempts to Customize	Attempt												✓																																														
26	Number of Corrected Faults	Fault																																		✓																										
27	Response Time	Second															✓																											✓																
28	Number of Evaluations	Evaluatio n															✓	✓																																										
29	Turnaround Time	Second															✓																											✓																
30	Task Time	Minute															✓																																									✓		
31	Number of I/O Related Errors	Error												✓				✓																											✓															
32	User Waiting Time of I/O Device Utilization	Second																✓																																										
33	Number of Memory Related Errors	Error																✓																											✓															
34	Number of Transmission Related Error	Error																✓																																										
35	Transmission Capacity	Byte																✓																																										

Measure Name		Unit	External																											Internal																											Quality in use			
			Functionality					Reliability				Usability					Efficiency			Maintainability					Portability					Functionality					Reliability				Usability					Efficiency			Maintainability					Portability					Q1	Q2	Q3	Q4
			F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5	F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5				
36	Number of Revised Versions	Version																			✓																																							
37	Number of Resolved Failures	Failure						✓													✓	✓	✓																																					
38	Porting User Friendliness	Minute																							✓																																			
39	Number of Data Items	Item																		✓					✓			✓			✓		✓																					✓						
40	Number of Interface Protocols	Protocol																														✓																												
41	Number of Access Types	Type																															✓																											
42	Number of Controllability Requirements	Req.																															✓																											
43	Number of Instances of Data Corruption	Instance				✓																											✓																											
44	Number of Items Which Could Check for Valid Data	Item																																							✓																			
45	Number of Operations	Operatio n												✓																											✓																			
46	Number of Messages Implemented	Message																																							✓																			
47	Number of Interface Elements	Element													✓																										✓	✓																		
48	I/O Utilization (Number of Buffers)	Buffer																																											✓														
49	Number of Line of Code Directly	Line																																											✓														

	Measure Name	Unit	External																									Internal																									Quality in use			
			Functionality					Reliability				Usability					Efficiency			Maintainability					Portability					Functionality					Reliability				Usability					Efficiency			Maintainability					Portability					Q1	Q2	Q3	Q4
			F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5	F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5				
	Related to System Calls																																																								
50	Number of Items required to be Logged	Item																																																								
51	Number of Modifications Made	Modification																																																								
52	Number of Variables	Variable																																																								
53	Number of Diagnostic Functions Required	Function																																																								
54	Number of Entities	Entity																																																								
55	Number of Built-in Test Functions Required	Function																																																								
56	Number of Test Dependencies on Other Systems	Dependency																																																								
57	Number of Checkpoints	Checkpoint																																																								
58	Number of data Structures	Data-Structure																																																								
59	Number of Setup Operations	Operation																							✓																																		
60	Number of Installation Steps	Step																																																								
61	Task Effectiveness	(a Given Weight)																																																							✓	✓		
62	Cost of the Task	Dollar																																																								
63	Help Time	Second																																																								✓		

	Measure Name	Unit	External																							Internal																					Quality in use			
			Functionality					Reliability				Usability					Efficiency			Maintainability					Portability					Functionality					Reliability				Usability					Efficiency			Maintainability					Portability					Q1	Q2	Q3	Q4
			F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5	F1	F2	F3	F4	F5	R1	R2	R3	R4	U1	U2	U3	U4	U5	E1	E2	E3	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5				
64	Error Time	Second																		✓	✓																																					✓		
65	Search Time	Second																																																						✓		
66	Number of Users	User																																																					✓	
67	Number of People Potentially Affected by the System	Person																																																						✓	✓
68	Number of Usage Situations	Situation																																																					✓	

Legend of the Quality in Use characteristics

Q1	Effectiveness
Q2	Productivity
Q3	Safety
Q4	Satisfaction

Legend of the External and Internal sub-characteristics

F1	Suitability	E1	Time Behavior
F2	Accuracy	E2	Resource Utilization
F3	Interoperability	E3	Efficiency Compliance
F4	Security	M1	Analyzability
F5	Functionality Compliance	M2	Changeability
R1	Maturity (Hardware/Software/Data)	M3	Stability
R2	Fault Tolerance	M4	Testability
R3	Recoverability (Data/Process/Technology)	M5	Maintainability Compliance
R4	Reliability Compliance	P1	Adaptability
U1	Understandability	P2	Instability
U2	Learnability	P3	Co-existence
U3	Operability	P4	Replaceability
U4	Attractiveness	P5	Portability Compliance
U5	Usability Compliance		

Appendix B

Table B1: External Quality Measures in ISO 9126-2 and ISO 25021.

Quality Characteristics	Quality Subcharacteristics		Measure Names	ISO DTR 25021	ISO 9126-2
Functionality	Accuracy	1	Computational accuracy	√	√
		2	Precision	√	√
		3	Accuracy relative to expectations		√
	Interoperability	4	Data exchangeability (data format-based)	√	√
		5	Data exchangeability (user's success, attempt-based)		√
	Security	6	Access controllability	√	√
		7	Access auditability		√
		8	Data corruption prevention		√
	Suitability	9	Funcional implementation completeness	√	√
		10	Functional adequacy	√	√
		11	Functional implementation coverage	√	√
		12	Functional specification stability (volatility)		√
	Functionality Compliance	13	Functional compliance		√
		14	Interface standard compliance		√
Reliability	Maturity	15	Failure density against test cases	√	√
		16	Failure resolution	√	√
		17	Fault removal	√	√
		18	Mean time between failures (MTBF)	√	√
		19	Test maturity	√	√
		20	Estimated latent fault density	√	√
		21	Fault density	√	√
		22	Test coverage (specified operation scenario testing coverage)		√
	Recoverability	23	Restartability	√	√
		24	Availability		√
		25	Mean down time		√
		26	Mean recovery time		√
		27	Restorability		√
		28	Restore effectiveness		√
	Fault Tolerance	29	Breakdown avoidance		√
		30	Failure avoidance		√
		31	Incorrect operation avoidance		√
	Reliability Compliance	32	Reliability compliance		√
Usability	Learnability	33	Effectiveness of the user documentation and/or help system	√	√
		34	Help accessibility	√	√
		35	Ease of function learning		√
		36	Ease of learning to perform a task in use		√
		37	Effectiveness of user documentation and/or help system in use		√
		38	Help frequency		√

Quality Characteristics	Quality Subcharacteristics		Measure Names	ISO DTR 25021	ISO 9126-2
	Operability	39	Physical accessibility	√	√
		40	Operational consistency in use		√
		41	Error correction		√
		42	Error correction in use		√
		43	Default value availability in use		√
		44	Message understandability in use		√
		45	Self-explanatory error messages		√
		46	Operational error recoverability in use		√
		47	Time between human error operations in use		√
		48	Undoability (user error correction)		√
		49	Customizability		√
		50	Operation procedure reduction		√
	Understandability	51	Completeness of description	√	√
		52	Function understandability	√	√
		53	Understandable input and output	√	√
		54	Demonstration accessibility		√
		55	Demonstration accessibility in use		√
		56	Demonstration effectiveness		√
		57	Evident functions		√
	Attractiveness	58	Attractive interaction		√
		59	Interface appearance customizability		√
	Usability Compliance	60	Usability compliance		√
Efficiency	Resource Utilization	61	I/O loading limits	√	√
		62	Maximum memory utilization	√	√
		63	Maximum transmission utilization	√	√
		64	Mean occurrence of transmission error	√	√
		65	I/O device utilization		√
		66	I/O-related errors		√
		67	Mean I/O fulfillment ratio		√
		68	User waiting time of I/O device utilization		√
		69	Mean occurrence of memory errors		√
		70	Ratio of memory error/time		√
		71	Media device utilization balancing		√
		72	Mean transmission error per time		√
		73	Transmission capacity utilization		√
	Time Behavior	74	Response time (mean time to respond)	√	√
		75	Throughput (mean amount of throughput)	√	√
		76	Turnaround time (mean time for turnaround)	√	√
		77	Response time		√
		78	Response time (worst-case response time ratio)		√
		79	Throughput		√
		80	Throughput (worst-case throughput time ratio)		√

Quality Characteristics	Quality Subcharacteristics		Measure Names	ISO DTR 25021	ISO 9126-2
		81	Turnaround time		√
		82	Turnaround time (worst-case turnaround time ratio)		√
		83	Waiting time		√
	Efficiency Compliance	84	Efficiency compliance		√
Maintainability	Analyzability	85	Audit trail capability	√	√
		86	Diagnostic function support		√
		87	Failure analysis capability		√
		88	Failure analysis efficiency		√
		89	Status monitoring capability		√
	Changeability	90	Software change control capability	√	√
		91	Change cycle efficiency		√
		92	Change implementation elapsed time		√
		93	Modification complexity		√
		94	Parameterized modifiability		√
	Stability	95	Change success ratio		√
		96	Modification impact localization (emerging failure after change)		√
	Testability	97	Availability of built-in test function		√
		98	Retest efficiency		√
		99	Test restartability		√
	Maintainability Compliance	100	Maintainability compliance		√
Portability	Adaptability	101	Adaptability of data structures	√	√
		102	Hardware environmental adaptability (adaptability to hardware devices and network facilities)	√	√
		103	System software environmental adaptability (adaptability to OS, network software and cooperated application software)	√	√
		104	Organizational environment adaptability (organization‘s adaptability to its infrastructure)		√
		105	Porting user-friendliness		√
	Installability	106	Ease of installation	√	√
		107	Ease of setup retry		√
	Coexistance Replaceability	108	Availability coexistence		√
	Replaceability	109	Continued use of data		√
		110	Function inclusiveness		√
		111	User support functional consistency		√
	Portability Compliance	112	Portability Compliance		√

Table B2: Internal Quality Measures in ISO 9126-3 and ISO 25021.

Quality Characteristics	Quality Subcharacteristics		Measure Names	ISO DTR 25021	ISO 9126-3
Functionality	Accuracy	1	Computational accuracy	√	√
		2	Precision	√	√
	Interoperability	3	Data exchangeability (data format-based)	√	√
		4	Interface consistency (protocol)		√
	Security	5	Access controllability	√	√
		6	Access auditability		√
		7	Data corruption prevention		√
		8	Data encryption		√
	Suitability	9	Funcional implementation completeness	√	√
		10	Functional adequacy	√	√
		11	Functional implementation coverage	√	√
		12	Functional specification stability (volatility)		√
	Functionality Compliance	13	Functional compliance		√
		14	Intersystem standard compliance		√
Reliability	Maturity	15	Fault removal	√	√
		16	Fault detection		√
		17	Test adequacy		√
	Recoverability	18	Restorability		√
		19	Restoration effictiveness		√
	Fault Tolerance	20	Failure avoidance		√
		21	Incorrect operation avoidance		√
	Reliability Compliance	22	Reliability compliance		√
Usability	Learnability	23	Completeness of user documentation and/or help facility	√	√
	Operability	24	Physical accessibility	√	√
		25	Input validity checking		√
		26	User operation cancellability		√
		27	User operation undoability		√
		28	Customizability		√
		29	Operation status monitoring capability		√
		30	Operational consistency		√
		31	Message clarity		√
		32	Interface element clarity		√
		33	Operational error recoverability		√
	Understandability	34	Completeness of description	√	√
		35	Function understandability	√	√
		36	Demonstration capability		√
		37	Evident functions		√
	Attractiveness	38	Attractive interaction		√
		39	User interface appearance customizability		√
	Usability Compliance	40	Usability compliance		√
Efficiency	Resource Utilization	41	I/O utilization		√
		42	I/O utilization message density		√

Quality Characteristics	Quality Subcharacteristics		Measure Names	ISO DTR 25021	ISO 9126-3
		43	Memory utilization		√
		44	Memory utilization message density		√
		45	Transmission utilization		√
	Time Behavior	46	Response time		√
		47	Throughput time		√
		48	Turnaround time		√
	Efficiency Compliance	49	Efficiency compliance		√
Maintainability	Analyzability	50	Activity recording		√
		51	Readiness of diagnostic function		√
	Changeability	52	Change recordability		√
	Stability	53	Change impact		√
		54	Modification impact localization		√
	Testability	55	Completeness of built-in test function		√
		56	Autonomy of testability		√
		57	Test progress observability		√
	Maintainability Compliance	58	Maintainability compliance		√
Portability	Adaptability	59	Adaptability of data structures	√	√
		60	Hardware environmental adaptability (adaptability to hardware devices and network facilities)	√	√
		61	System software environmental adaptability (adaptability to OS, network software and cooperated application software)	√	√
		62	Organizational environment adaptability		√
		63	Porting user-friendliness		√
	Installability	64	Ease of setup retry		√
		65	Installation effort		√
		66	Installation flexibility		√
	Co-existence	67	Availability of coexistence		√
	Replaceability	68	Continued use of data		√
		69	Functional inclusiveness		√
	Portability Compliance	70	Portability compliance		√

Table B3: Quality in Use Measures in ISO/IEC TR 9126-4 and ISO/IEC TR 25021.

Quality Characteristics		Measure Names	ISO DTR 25021	ISO 9126-4
Effectiveness	1	Task effectiveness		√
	2	Task completion	√	√
	3	Error frequency		√
Productivity	4	Task time	√	√
	5	Task efficiency		√
	6	Economic productivity		√
	7	Productive proportion		√
	8	Relative user efficiency		√
Safety	9	User health and safety		√
	10	Safety of people affected by use of the system		√
	11	Economic damage		√
	12	Software damage		√
Satisfaction	13	Satisfaction scale		√
	14	Satisfaction questionnaire		√
	15	Discretionary usage		√