TECHNICAL REPORT

ISO/IEC TR 9126-4

First edition
2004-04-01

Software engineering — Product quality —

Part 4: **Quality in use metrics**

Génie du logiciel — Qualité des produits —

Partie 4: Qualité en métrologie d'usage

Reference number
ISO/IEC TR 9126-4:2004(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

In exceptional circumstances, the joint technical committee may propose the publication of a Technical Report of one of the following types:

— type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;

— type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;

— type 3, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC TR 9126-4 which is a Technical Report of type 2, was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 7, *Software and system engineering.*

ISO/IEC TR 9126 consists of the following parts, under the general title *Software engineering — Product quality*:

— *Part 1: Quality model*

— *Part 2: External metrics*

— *Part 3: Internal metrics*

— *Part 4: Quality in use metrics*

Introduction

This Technical Report provides quality in use metrics for measuring attributes of quality in use defined in ISO/IEC 9126-1. The metrics listed in this Technical Report are not intended to be an exhaustive set. Developers, evaluators, quality managers and acquirers may select metrics from this Technical Report for defining requirements, evaluating software products, measuring quality aspects and other purposes. They may also modify the metrics or use metrics that are not included here. This report is applicable to any kind of software product, although each of the metrics is not always applicable to every kind of software product.

ISO/IEC 9126-1 defines terms for the software quality characteristics and how these characteristics are decomposed into subcharacteristics. ISO/IEC 9126-1, however, does not describe how any of these subcharacteristics could be measured. ISO/IEC 9126-2 defines external metrics, ISO/IEC 9126-3 defines internal metrics and ISO/IEC 9126-4 defines quality in use metrics, for measurement of the characteristics or subcharacteristics. Internal metrics measure the software itself, external metrics measure the behaviour of the computer-based system that includes the software, and quality in use metrics measure the effects of using the software in a specific context of use.

This Technical Report is intended to be used together with ISO/IEC 9126-1. It is strongly recommended to read ISO/IEC 14598-1 and ISO/IEC 9126-1, prior to using this Technical Report, particularly if the reader is not familiar with the use of software metrics for product specification and evaluation.

Software engineering — Product quality —

Part 4: Quality in use metrics

1 Scope

This Technical Report defines quality in use metrics for the characteristics defined in ISO/IEC 9126-1, and is intended to be used together with ISO/IEC 9126-1.

This Technical Report contains:

- an explanation of how to apply software quality metrics;
- a basic set of metrics for each characteristic;
- an example of how to apply metrics during the software product life cycle.

It includes as informative annexes a quality in use evaluation process and a reporting format.

This Technical Report does not assign ranges of values of these metrics to rated levels or to grades of compliance, because these values are defined for each software product or a part of the software product, by its nature, depending on such factors as category of the software, integrity level and users' needs. Some attributes may have a desirable range of values, which does not depend on specific user needs but depends on generic factors, i.e. human cognitive factors.

This Technical Report can be applied to any kind of software for any application. Users of this Technical Report can select or modify and apply metrics and measures from this Technical Report or may define application-specific metrics for their individual application domain. For example, the specific measurement of quality characteristics such as safety or security may be found in International Standards or Technical Reports provided by IEC 65 and ISO/IEC JTC1/SC 27.

Intended users of this Technical Report include:

- Acquirer (an individual or organization that acquires or procures a system, software product or software service from a supplier);
- Evaluator (an individual or organization that performs an evaluation. An evaluator may, for example, be a testing laboratory, the quality department of a software development organization, a government organization or user);
- Developer (an individual or organization that performs development activities, including requirements analysis, design and testing through acceptance during the software life cycle process);
- Maintainer (an individual or organization that performs maintenance activities);
- Supplier (an individual or organization that enters into a contract with the acquirer for the supply of a system, software product or software service under the terms of the contract) when validating software quality at qualification test;
- User (an individual or organization that uses the software product to perform a specific function) when evaluating quality of software product at acceptance test;

- Quality manager (an individual or organization that performs a systematic examination of the software product or software services) when evaluating software quality as part of quality assurance and quality control.

2 Conformance

There are no conformance requirements in this Technical Report.

NOTE General conformance requirements for metrics are in ISO/IEC 9126-1.

3 Normative References

ISO 8402, *Quality management and quality assurance — Vocabulary*

ISO/IEC 9126, *Software engineering — Product quality*

ISO/IEC 9126-1, *Software engineering — Product quality — Part 1: Quality model*

ISO/IEC TR 9126-2, *Software engineering — Product quality — Part 2: External metrics* [Technical Report]

ISO/IEC TR 9126-3, *Software engineering — Product quality — Part 3: Internal metrics* [Technical Report]

ISO 9241-11:1998, *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability*

ISO/IEC 14598-1, *Information technology — Software product evaluation — Part 1: General overview*

ISO/IEC 14598-3, *Software engineering — Product evaluation — Part 3: Process for developers*

ISO/IEC 14598-5:1998, *Information technology — Software product evaluation — Part 5: Process for evaluators*

ISO/IEC 12207:1995, *Information technology — Software life cycle processes*

ISO/IEC 14143-1, *Information technology — Software measurement — Functional size measurement — Part 1: Definition of concepts*

4 Terms and definitions

For the purposes of this Technical Report, the definitions contained in ISO/IEC 14598-1, ISO/IEC 9126-1 and the following apply. Some of the definitions from ISO/IEC 14598-1 and ISO/IEC 9126-1 are reproduced in Annex D.

4.1 context of use

the users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used

[ISO 9241-11]

4.2 goal

an intended outcome

[ISO 9241-11]

4.3 task

the activities required to achieve a goal

NOTE 1 These activities can be physical or cognitive.

NOTE 2 Job responsibilities can determine goals and tasks.

[ISO 9241-11]

5 Symbols and abbreviated terms

The following symbols and abbreviated terms are used in this Technical Report:

- SQA - Software Quality Assurance (Group)
- SLCP – Software Life Cycle Processes

6 Use of software quality metrics

These Technical Reports (ISO/IEC 9126-2, ISO/IEC 9126-3 and ISO/IEC 9126-4) provide a suggested set of quality metrics (external, internal and quality in use metrics) to be used with the ISO/IEC 9126-1 quality model. The user of these reports may modify the metrics defined, and/or may also use metrics not listed. When using a modified or a new metric not identified in these Technical Reports, the user should specify how the metrics relate to the ISO/IEC 9126-1 quality model or any other substitute quality model that is being used.

The user of these Technical Reports should select the quality characteristics and subcharacteristics to be evaluated from ISO/IEC 9126-1, identify the appropriate direct and indirect measures, identify the relevant metrics and then interpret the measurement result in objective manner. The user of these Technical Reports also may select product quality evaluation processes during the software life cycle from the ISO/IEC 14598 series of International Standards. These give methods for measurement, assessment and evaluation of software product quality. They are intended for use by developers, acquirers and independent evaluators, particularly those responsible for software product evaluation (see Figure 1).

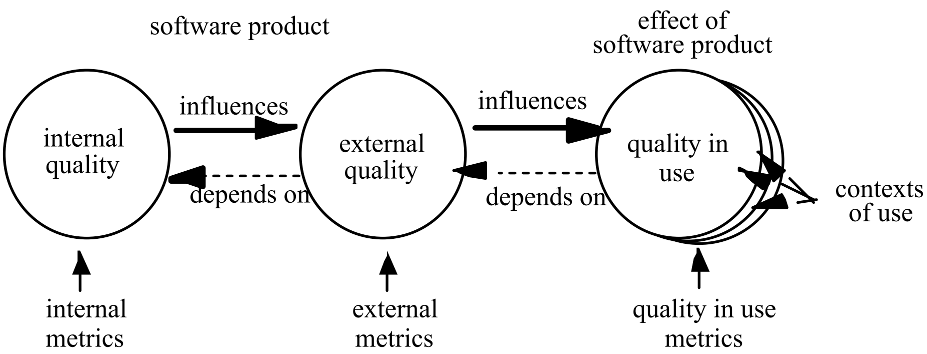

Figure 1 — Relationship between types of metrics

The internal metrics may be applied to a non-executable software product during its development stages (such as request for proposal, requirements definition, design specification or source code). Internal metrics provide the users with the ability to measure the quality of the intermediate deliverables and thereby predict the quality of the final product. This allows the user to identify quality issues and take corrective action as early as possible in the development life cycle.

The external metrics may be used to measure the quality of the software product by measuring the behaviour of the system of which it is a part. The external metrics can only be used during the testing stages of the life cycle process and during any operational stages. The measurement is performed when executing the software product in the system environment in which it is intended to operate.

The quality in use metrics measure whether a product meets the needs of specified users to achieve specified goals with effectiveness, productivity, safety and satisfaction in a specified context of use. This can be only achieved in a realistic system environment.

User quality needs can be specified as quality requirements by quality in use metrics, by external metrics, and sometimes by internal metrics. These requirements specified by metrics should be used as criteria when a product is evaluated.

It is recommended to use internal metrics having a relationship as strong as possible with the target external metrics, so that they can be used to predict the values of external metrics. However, it is often difficult to design a rigorous theoretical model that provides a strong relationship between internal metrics and external metrics. Therefore, a hypothetical model that may contain ambiguity may be designed and the extent of the relationship may be modelled statistically during the use of metrics.

Recommendations and requirements related to validity and reliability are given in ISO/IEC 9126-1:A.4. Additional detailed considerations when using metrics are given in Annex A of this Technical Report.

7 How to read and use the metrics tables

The metrics listed in clause 8 are categorised by the characteristics in ISO/IEC 9126-1. The following information is given for each metric in the table:

a) Purpose of the metric: This is expressed as the question to be answered by the application of the metric.

b) Method of application: Provides an outline of the application.

c) Measurement, formula and data element computations: Provides the measurement formula and explains the meanings of the used data elements.

NOTE In some situations more than one formula is proposed for a metric.

d) Interpretation of measured value: Provides the range and preferred values.

e) Metric scale type: Type of scale used by the metric. Scale types used are; Nominal scale, Ordinal scale, Interval scale, Ratio scale and Absolute scale.

NOTE: A more detailed explanation is given in Annex C.

f) Measure type: Types used are; Size type (e.g. Function size, Source size) , Time type (e.g. Elapsed time, User time) , Count type (e.g. Number of changes, Number of failures).

NOTE A more detailed explanation is given in Annex C.

g) Input to measurement: Source of data used in the measurement.

h) ISO/IEC 12207 SLCP Reference: Identifies software life cycle process(es) where the metric is applicable.

i) Target audience: Identifies the user(s) of the measurement results.

8 Metrics Tables

8.0 General

The metrics listed in this clause are not intended to be an exhaustive set and may not have been validated. They are listed by software quality characteristic.

Metrics, which may be applicable, are not limited to these listed here. Additional specific metrics for particular purposes are provided in other related documents, such as functional size measurement or precise time efficiency measurement.

NOTE It is recommended to refer a specific metric or measurement from specific standards, Technical Reports or guidelines Functional size measurement is defined in ISO/IEC 14143. An example of precise time efficiency measurement can be referred from ISO/IEC 14756.

Metrics should be validated before application in a specific environment (see Annex A).

NOTE This list of metrics is not finalized, and may be revised in future versions of this Technical Report. Readers of this Technical Report are invited to provide feedback.

The quality in use metrics in this clause measure the effectiveness, productivity, safety or satisfaction with which specified users achieve specified goals in a specified context of use. Quality in use depends not only on the software product, but also on the particular context in which the product is being used. The context of use is determined by user factors, task factors and physical and social environmental factors.

Quality in use is assessed by observing representative users carrying out representative tasks in a realistic context of use (see Annex E). The measures may be obtained by simulating a realistic working environment (for instance in a usability laboratory) or by observing operational use of the product. In order to specify or measure quality in use it is first necessary to identify each component of the intended context of use: the users, their goals, and the environment of use. The evaluation should be designed to match this context of use as closely as possible. It is also important that users are only given the type of help and assistance that would be available to them in the operational environment.

NOTE The term usability is sometimes used with a similar meaning to quality in use (but excluding safety) (e.g. in ISO 9241-11).

Some external usability metrics (ISO/IEC 9126-2) are tested in a similar way, but evaluate the use of particular product features during more general use of the product to achieve a typical task as part of a test of the quality in use.

Quality in use has four characteristics (effectiveness, productivity, safety and satisfaction) and no subcharacteristics.

8.1 Effectiveness metrics

Effectiveness metrics assess whether the tasks performed by users achieve specified goals with accuracy and completeness in a specified context of use. They do not take account of how the goals were achieved, only the extent to which they were achieved (see E.2.1.2).

Table 8.1 Effectiveness metrics

Metric Name	Purpose of the metrics	Method of application	Measurement, formula and data element computations	Interpretation of measured value	Metric scale type	Measure type	Input to measurement	12207 reference	Target audience
Task effectiveness	What proportion of the goals of the task is achieved correctly?	User test	M1 = $\|1-\Sigma A_i\|$ A_i= proportional value of each missing or incorrect component in the task output	0<= M1 <=1 The closer to 1.0 the better.	—	A= proportion	Operation (test) report User monitoring record	6.5 Validation 5.3 Qualification testing 5.4 Operation	User Human interface designer
NOTE Each potential missing or incomplete component is given a weight A_i based on the extent to which it detracts from the value of the output to the business or user. (If the sum of the weights exceed 1, the metric is normally set to 0, although this may indicate negative outcomes and potential safety issues.) (See for example G.3.1.1.) The scoring scheme is refined iteratively by applying it to a series of task outputs and adjusting the weights until the measures obtained are repeatable, reproducible and meaningful.									
Task completion	What proportion of the tasks is completed?	User test	X = A/B A = number of tasks completed B = total number of tasks attempted	0<= X <=1 The closer to 1.0 the better.	Ratio	A = Count B = Count X = Count/Count	Operation (test) report User monitoring record	6.5 Validation 5.3 Qualification testing 5.4 Operation	User Human interface designer
NOTE This metric can be measured for one user or a group of users. If tasks can be partially completed the Task effectiveness metric should be used..									
Error frequency	What is the frequency of errors?	User test	X = A/T A = number of errors made by the user T= time or number of tasks	0<= X The closer to 0 the better.	Absolute	A = Count	Operation (test) report User monitoring record	6.5 Validation 5.3 Qualification testing 5.4 Operation	User Human interface designer
NOTE This metric is only appropriate for making comparisons if errors have equal importance, or are weighted.									

8.2 Productivity metrics

Productivity metrics assess the resources that users consume in relation to the effectiveness achieved in a specified context of use. The most common resource is time to complete the task, although other relevant resources could include the user's effort, materials or the financial cost of usage.

Table 8.2 Productivity metrics

Metric Name	Purpose of the metrics	Method of application	Measurement, formula and data element computations	Interpretation of measured value	Metric scale type	Measure type	Input to measurement	12207 reference	Target audience
Task time	How long does it take to complete a task?	User test	X = Ta Ta = task time	0<= X The smaller the better.	Interval	T= Time	Operation (test) report User monitoring record	6.5 Validation 5.3 Qualification testing 5.4 Operation	User Human interface designer
Task efficiency	How efficient are the users?	User test	X = M1 / T M1 = task effectiveness T = task time	0<= X The larger the better.	—	T= Time X= proportion/ time	Operation (test) report User monitoring record	6.5 Validation 5.3 Qualification testing 5.4 Operation	User Human interface designer
NOTE 1 Task efficiency measures the proportion of the goal achieved for every unit of time. Efficiency increases with increasing effectiveness and reducing task time. It enables comparisons to be made, for example between fast error-prone interfaces and slow easy interfaces (see for example F.2.4.4).. NOTE 2 If Task completion has been measured, task efficiency can be measured as Task completion/task time. This measures the proportion of users who were successful for every unit of time. A high value indicates a high proportion of successful users in a small amount of time.									
Economic productivity	How cost-effective is the user?	User test	X = M1 / C M1 = task effectiveness C = total cost of the task	0<= X The larger the better.	—	C=value X= proportion/ value	Operation (test) report User monitoring record	6.5 Validation 5.3 Qualification testing 5.4 Operation	User Human interface designer

NOTE Costs could for example include the user's time, the time of others giving assistance, and the cost of computing resources, telephone calls, and materials

Metric Name	Purpose of the metrics	Method of application	Measurement, formula and data element computations	Interpretation of measured value	Metric scale type	Measure type	Input to measurement	12207 reference	Target audience
Productive proportion	What proportion of the time is the user performing productive actions?	User test	X = Ta / Tb Ta = productive time = task time - help time - error time - search time Tb = task time	0<= X <=1 The closer to 1.0 the better.	Absolute	Ta=Time Tb=Time X= Time/ Time	Operation (test) report User monitoring record	6.5 Validation 5.3 Qualification testing 5.4 Operation	User Human interface designer
NOTE This metric requires detailed analysis of a videotape of the interaction (see Macleod M, Bowden R, Bevan N and Curson I (1997) The MUSiC Performance Measurement method, Behaviour and Information Technology, 16, 279-293.)									
Relative user efficiency	How efficient is a user compared to an expert?	User test	Relative user efficiency X = A / B A = ordinary user's task efficiency B = expert user's task efficiency	0<= X <=1 The closer to 1.0 the better.	Absolute	X = proportion/ proportion	Operation (test) report User monitoring record	6.5 Validation 5.3 Qualification testing 5.4 Operation	User Human interface designer

NOTE The user and expert carry out the same task. If the expert was 100 % productive, and the user and expert had the same task effectiveness, this metric would give a similar value to the Productive proportion.

8.3 Safety metrics

Safety metrics assess the level of risk of harm to people, business, software, property or the environment in a specified context of use. It includes the health and safety of the both the user and those affected by use, as well as unintended physical or economic consequences.

Table 8.3 Safety metrics

Metric Name	Purpose of the metrics	Method of application	Measurement, formula and data element computations	Interpretation of measured value	Metric scale type	Measure type	Input to measurement	12207 reference	Target audience
User health and safety	What is the incidence of health problems among users of the product?	Usage statistics	X = 1-A / B A = number of users reporting RSI B = total number of users	0<= X <=1 The closer to 1 the better.	Absolute	A = count B = count X = count/ count	Usage monitoring record	5.4 Operation	User Human interface designer
NOTE Health problems can include Repetitive Strain Injury, fatigue, headaches, etc.									
Safety of people affected by use of the system	What is the incidence of hazard to people affected by use of the system?	Usage statistics	X = 1-A / B A = number of people put at hazard B = total number of people potentially affected by the system	0<= X <=1 The closer to 1 the better.	Absolute	A = count B = count X = count/ count	Usage monitoring record	5.3 Qualification testing 5.4 Operation	User Human interface designer Developer
NOTE An example of this metric is Patient Safety, where A = number of patients with incorrectly prescribed treatment and B = total number of patients									
Economic damage	What is the incidence of economic damage?	Usage statistics	X = 1-A/ B A = number of occurrences of economic damage B = total number of usage situations	0<= X <=1 The closer to 1 the better.	Absolute	A = count B = count X = count/ count	Usage monitoring record	5.4 Operation	User Human interface designer Developer

NOTE This can also be measured based on the number of occurrences of situations where there was a risk of economic damage