

Information & Management 38 (2001) 137-152



www.elsevier.com/locate/dsw

Quality metrics for intranet applications

Hareton K.N. Leung

Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong Received 19 August 1999; accepted 31 May 2000

Abstract

As the number of intranet application increases, software developers face a new software paradigm and possibly a new set of quality requirement. The work discussed here attempts to identify practical software metrics for intranet applications. The six software quality characteristics and 32 quality sub-characteristics of the Extended ISO model are used as a basis to identify the key quality characteristics of intranet applications. From the results of a user survey, three key quality characteristics are identified; they are reliability, functionality, and efficiency. Five sub-characteristics (availability, accuracy, security, suitability and time behaviour) are found to be the key attributes of intranet applications. Finally, a set of three metrics is developed. In order to verify their validity and applicability to intranet projects, an experiment was performed by computing these metrics in five intranet applications. The metric values were then compared with the results from a user satisfaction survey. The metric values and the survey results were closely correlated; a larger score in software metrics leads to a higher user satisfaction score. It is very likely that these software metrics can effectively reflect the quality of an intranet application. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Extended ISO model; Intranet applications; Quality metrics

1. Introduction

The intranet emerged in the early nineties as the application of internet technologies to internal business applications. When discussing the intranet, it is important to remember that the underlying technology is identical to that of the internet. Only in one way are the two different: private versus public usage [20]. More and more organizations from all sectors are attracted by the tremendous benefits of intranet applications: cost and time saving, increase of productivity, flexibility [17], open-architecture, consistency, reduced workload, timeliness, and removal of departmental boundaries. From a management perspective, successful intranet application implies higher return on investment, low risk, shorter payback time, more business opportunities, lower training cost, technology-enabled employees [16], and a shorter development life-cycle [3]. However, the use of the intranet may require dramatic changes in business processes and organizational culture [5].

Although many successful stories involving the use of the intranet have been reported, numerous intranet investors are disappointed and impatient as they wait for benefits that may never come. In the wave of use of the intranet, there is an increasing need for quality management. How do intranet applications differ from other applications? Can an intranet application be measured quantitatively? Can well-known software quality characteristics be applied?

Typically, metrics may be relevant to a specific application domain; a study indicated that different types of software applications have different quality profiles [15]. Hence, product quality metrics which are

relevant for non-intranet applications may not be suitable for intranet applications. Learnability is a typical example: when working with a standard web browser interface, users do not need to spend extra effort in learning about the user interface of a new intranet application. The value of this quality characteristic diminishes and thus, the corresponding quality metrics may not be applicable to an intranet application.

Unfortunately, there has been little study of quality measurement for intranet applications. Although many intranet purchasers are concerned about its quality, they do not apply quality metrics. For example, many organizations invest in many tools, such as a firewall, to improve security of intranet, but its value is rarely measured [7]. Some organizations have analyzed features [9,13] and cost-effectiveness [10] of the intranet without any apparent quantitative measurement. Use of quality metrics was rarely found; therefore our study attempts to identify practical product quality metrics for intranet applications.

A commonly accepted definition of quality is 'the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs' [12]. According to Deming, quality is customer-focused. Thus, the relative importance of software quality characteristics of intranet application should be determined for different types of users: endusers, end-user management, software developers and information system (IS) professionals.

In our study, the Extended ISO model was chosen as the reference point, due to its popularity and acceptance by the software industry [19]. Its software quality characteristics were used to identify key quality attributes of intranet application. The ranking of these characteristics were then used to determine a generic view of the importance of these quality factors. Their relevance was studied to identify the key quality characteristics.

Table 1 Six quality characteristics of ISO

Functionality Reliability Usability Efficiency Maintainability Portability

Shows the existence of a set of functions and their specified properties. The functions satisfy stated or implied needs. That capability of software which maintains its level of performance under given conditions for a given period of time. Attributes that determine the effort needed for use and the assessment of such use by a set of users.

Attributes that determine the effort needed for use and the assessment of such use by a set of users. The relationship between the level of performance of the software and the amount of resources used under stated conditions.

The effort needed to make specified modifications.

The ability of the software to be transformed from one environment to another.

2. Software quality and intranet application

2.1. Software quality model

McCall distinguished between two levels of quality attributes: quality factors and criteria. The former cannot be measured directly, while the latter can be measured subjectively [8]. Based on this, ISO further simplified the model, which is general and commonly accepted, being state-of-the-art in product quality specification. Software quality characteristics are gathered from practical situations. It uses the quality characteristics listed in Table 1 [21].

These quality characteristics can be further broken down into lower level quality sub-characteristics [22]. The Extended ISO model consists of 32 sub-characteristics, as shown in Fig. 1.

2.2. Relevance of ISO quality characteristics in intranet applications

Intranets may vary from a small-scale personal computer network to a corporate wide area network. Also, functionality of an intranet may range from a store of static documents to the enterprise knowledge architecture. There are several types of intranet as described:

2.2.1. Simple intraweb

An internal web server that publishes static pages in HTML format. The browser opens up a connection to its server, which, in turn, returns the page and closes the connection.

2.2.2. Interactive intraweb

Containing forms, fields, and buttons that take input from the user before opening a connection to the server for transmission of data and choices. The web server passes the information to a server program Extended ISO Model

Software Quality Reliability Functionality Efficiency Usability Maintainability Portability Suitability Maturity Time behaviour Understandability Analysability Adaptability Fault tolerance Resource behaviour Accuracy Learnability Changeability Installability Interoperability Recoverability Operability Stability Conformance Compliance Availability Explicitness Testability Replaceability Security Degradability Custom is a hility Manageability Traceability Luxurv Reusability Clarity Helpfulness

Fig. 1. Hierarchical view of the Extended ISO Model.

User-friendliness

for processing, which, in turn, passes a new page back to the browser for viewing. The interactive intraweb facilitates timely distribution of static information and allows end-users to obtain information from a variety of back end servers.

2.2.3. Distributed intraweb

Uses the distributed data sources throughout an organization. The simple action of opening a web page invokes a procedure on behalf of the user. Through the standard programming interface, some small programs can be incorporated into the pages sent to the browser. Planned actions can be automated. A richer user interface is provided and some tasks shifted from the server to the client.

2.2.4. Enterprise intranet

This is replete with true open client/server applications. Instead of downloading static web pages, a program called an *applet*, is loaded on the client and runs automatically. This applet opens its own communication session with the server. In addition, it can communicate with other servers to get access of a wide range of databases on the intranet. Eventually, analysis of the targeted data is presented to the user.

The nature of the Interactive intraweb, the Distributed intraweb and the Enterprise intranet are similar.

The major differences are in the extent of the use of applets, the level of automation of mission-critical business processes, the degree of utilization of the organizational structure and the network, and the capabilities of analytic power for management decision making. We grouped these three types of intranet when analyzing the relevance of the quality subcharacteristics. The results are summarized in Table 2; further details can be found in [14].

As intranet applications are typically open, conform to the standards, and are generally available, several quality sub-characteristics are implied. Therefore, they are less relevant to the quality of an intranet application. The value of measuring these sub-characteristics may be very small. For example, because intranet is generally suitable for a wide range of applications of various scales, suitability is less relevant. Moreover, some characteristics, such as interoperability and maturity, are generally not suitable for Simple intraweb, which is generally a static store of documents.

Since a Simple intraweb is a replacement for printing, minimal processing capabilities are needed. Compliance to HTML or similar methods guarantees interoperability, compliance, and reliability. Use of a browser interface facilitates the usability, by default, while the organization of documents and their appearance determine the popularity of the intranet. The

Table 2
Relevance of ISO quality characteristics to various types of intranet applications

Characteristics	Sub-characteristics	Relevance of sub-characteristic		
		Simple intraweb	Interactive intraweb, Distributed intraweb & Enterprise intranet	
Functionality	Suitability	\otimes^{a}	⊗	
	Accuracy	\otimes	*	
	Interoperability	\otimes	\otimes	
	Compliance	\otimes	\otimes	
	Security	\odot^{b}	♣ °	
	Traceability	\otimes	\odot	
Reliability	Maturity	\otimes	·	
·	Fault tolerance	\odot	\odot	
	Recoverability	\otimes	*	
	Availability	*	*	
	Degradability	\otimes	·	
Usability	Understandability	\otimes	*	
•	Learnability	\otimes	·	
	Operability	\otimes	*	
	Luxury	*	•	
	Clarity	\odot	*	
	Helpfulness	\otimes	*	
	Explicitness	\otimes	*	
	User-friendliness	\otimes	·	
	Customisability	\otimes	*	
Efficiency	Time behaviour	\odot	*	
•	Resource behaviour	\odot	· ⊙	
Maintainability	Analysability	\otimes	*	
•	Changeability	\odot	· ⊙	
	Stability	\otimes	*	
	Testability	\otimes	*	
	Manageability	\odot	*	
	Reusability	8	· ⊙	
Portability	Adaptability	8	⊗	
•	Installability	8	⊗	
	Conformance	8	8	
	Replaceability	⊗	⊗	

^a Not relevant.

property of openness solves the portability problems of intranet.

The other three types of intranet applications bear more processing and analytical power and perform more critical tasks. Therefore, more quality characteristics are relevant; they include: accuracy, recoverability, and speed. Also, as they may involve legacy systems and databases, interoperability is a matter of ensuring all form a harmony. Maintenance is usually the largest investment of all application development. As time is a critical factor for a successful

intranet application, maintainability becomes more relevant.

3. Methodology

The study involves four phases:

- 1. A survey of quality characteristics,
- 2. Development of metrics and their reduction,
- 3. Collection of data,
- 4. Survey of user satisfaction.

^b Relevant.

^c Highly relevant.

A survey was first performed to determine the users' view of the relative importance of quality characteristics in an intranet application. To increase the reliability of the survey, users with various background and different internet/intranet experiences were invited to participate in the survey. Respondents consisted of end-users, end-user management, developers, and IS professionals with or without intranet/internet project experience. The scope of the survey was single-project [1]. The method used included electronic mail and interview. Electronic mail was used because of its low cost, short transfer time, and convenience for respondents, while interview techniques provide more information and explanation from respondents and therefore can produce a more reliable result [11].

The questionnaire included a brief explanation of intranet applications and the Extended ISO quality model. Respondents were asked to rank the six quality characteristics of Extended ISO model in order of their perceived importance. Then, they were asked to rank the 32 quality sub-characteristics. Sub-characteristics that were less relevant for intranet application were to be ruled out.

Based on the survey findings, we then developed metrics for the most important and relevant quality sub-characteristics. By applying the goal-question-metric (GQM) paradigm, goals were identified, questions were formulated in quantifiable terms, and metrics were finally established [12]. The goal of this effort was to improve the quality of the highest ranked sub-characteristic of an intranet application. Questions focused on achieving the goals quantitatively. Areas of measurement that were practical and useful were identified. Finally, a set of product quality metrics, including those newly invented and extracted from the literature, were identified for further assessment.

The next phase involved evaluation of the developed product quality metrics. A series of interviews were conducted with prospective metrics users. This evaluation focused on the applicability and usability of the metrics; those that are not practical or infeasible were ruled out. The reduced set of product quality metrics was used for experiment in the next phase.

The experiment aimed at evaluation of the validity of the set of product quality metrics. These metrics were applied to a number of intranet applications. They covered document management, form-based interface, and replacement of pre-printed forms. The users of these applications included information consumers, information contributors, administrators and systems administrators. Data was collected from problem logs, problem incident forms, and project management data. The applications were then ranked by their metric values.

As user satisfaction is the ultimate validation of quality, we conducted a survey to measure satisfaction of these intranet applications. The applications were ranked based on user satisfaction. In our analysis, paired two-sample T-test [2] were applied to each pair of consecutive applications in their order of relative importance.

The rankings from the experiment and the user satisfaction survey were then examined for their correlation. The null hypothesis was: 'the ranking of a specific metric reflects the ranking of the corresponding quality sub-characteristic.'

Another null hypothesis was: 'the ranking of a specific metric reflects the ranking of the overall user satisfaction.' If this hypothesis held, the metric may be used to indicate the overall quality or a particular quality sub-characteristic of intranet application.

4. User survey of quality characteristics of the intranet application

Prior to the actual survey, a few respondents were invited to help in a pre-test of the questionnaire. They commented that the list was long and many subcharacteristics were similar. Then, a face-to-face survey was conducted so that additional explanation of the quality characteristics could be provided to respondents when they were completing the questionnaire.

The questionnaire consisted of seven questions:

- 1. Question 1 was for filtering respondents that had no experience of intranet or internet application.
- Question 2 identified whether the respondents had any hands-on intranet experience in their workplace.
- 3. Question 3 was used to verify the reply of Question 2.
- 4. The user type of the respondent was classified in Ouestion 4.

- 5. Question 5 attempted to ensure there were IS professionals having intranet/internet project experience in the set of respondents.
- 6. Question 6 asked respondents to rank the six quality characteristics in order of their importance.
- 7. Question 7 was the most difficult question, because it asked respondents to assign a ranking to 32 quality sub-characteristics without duplication.
- 8. Quality characteristics were listed in an alphabetical order to prevent their initial order from affecting respondent preference.

Fifty users from local companies of various industries, including education, finance, retail, marketing, and manufacturing participated in the survey. A total of 39 responses were received, giving a return rate of 78%. Three respondents had no intranet or internet experience. Another six responses were invalid, because duplicate rank was assigned to different quality characteristics in Question 6 or Question 7. Thus, the remaining 30 questionnaires formed the basis of the analysis.

Respondents consisted of all intended user types. 30% of them were end users and management, while the remaining 70% were IS professionals with various specialties, such as software/hardware supports personnel, consultants, programmers, systems analysts, and project managers. 27% of respondents had development or managerial experience in intranet or internet projects.

4.1. Relative importance of the quality characteristics

Respondents were asked to rank the relative importance of the six quality characteristics by assigning a

Table 3
Distribution of score assigned to each quality characteristic

Quality	Scor	Score					Total
characteristic	1	2	3	4	5	6	
Functionality	8	7	12	2	1	0	30
Reliability	11	12	5	1	1	0	30
Usability	4	0	2	12	6	6	30
Efficiency	7	7	6	6	4	0	30
Maintainability	0	2	3	7	12	6	30
Portability	0	2	2	2	6	18	30
Total	30	30	30	30	30	30	

number from one (most important) to six (least important) for each characteristic, without duplication. This number was used to calculate the mean score of each characteristic and determine the overall relative importance of the characteristics. The attribute with lowest score is the most important. Table 3 summarizes the responses; e.g. there were eight respondents who viewed functionality as the most important quality characteristic.

Table 4 gives the percentage of respondents that choose a characteristic as one of the top three most important characteristics. The reliability and functionality characteristics received over 90% of the votes.

In order to verify the ranking, a paired two-sample *t*-test was performed for each consecutive pair of quality characteristics. The null hypothesis was that the mean of the more important characteristic was smaller than or equal to the mean of the less important characteristic. The confidence level was 0.05. Thus, the probability of the type 1 error to reject the hypothesis is 5%. A one-tailed test was performed.

Results of the paired two-sample *t*-test of five pairs of consecutive quality characteristics are shown in

Table 4
Ranking of six quality characteristics

Quality characteristic	Mean	Maximum	Minimum	Standard deviation	% among top three important items	Ranking
Reliability	2.0	1	5	1.36	93.33	1
Functionality	2.4	1	5	1.05	90.00	2
Efficiency	2.8	1	5	1.10	66.67	3
Usability	4.1	1	6	1.22	20.00	4
Maintainability	4.6	2	6	0.98	16.67	5
Portability	5.2	2	6	1.50	13.33	6

Table 5 Paired two-sample *t*-test for the six quality characteristics

Null hypothesis (≥ more or equally important)	$P(T \le t)$ one-tail
Reliability≥functionality Functionality≥efficiency Efficiency≥usability Usability≥maintainability Maintainability>portability	0.0976 0.1133 0.0018 0.1352 0.0379

Table 5. Since the probability of $T \le t$ statistics of all pairs were less than the critical value of t (1.67), all null hypothesis hold. From the t-test analysis, the ranking of the six quality characteristics is that shown in Table 4, with reliability being the most important.

4.2. Relative importance of the quality subcharacteristics

Respondents were asked to rank the relative importance of the 32 quality sub-characteristics without duplication. The number was used to calculate the mean score of each sub-characteristic and determine the relative importance of the sub-characteristics. The one with the lowest score was the most important.

Table 6 gives the percentage of respondents that choose a characteristic as one of the top five most important sub-characteristics. The availability, accuracy, security, suitability and time behaviour sub-characteristics received over 30% of the votes.

Table 6 Ranking of 32 quality sub-characteristics

Sub-characteristic	Mean	Maximum	Minimum	Standard deviation	% among top five important items	Ranking
Availability	8.2	1	32	7.67	50.0	1
Accuracy	9.7	1	32	10.66	53.3	2
Security	11.5	1	32	7.64	30.0	3
Suitability	12.0	1	31	9.65	40.0	4
Time behaviour	12.0	1	30	9.25	43.3	5
Luxury	13.0	1	29	8.91	26.7	6
Clarity	13.4	5	31	6.94	10.0	7
User-friendliness	13.7	1	29	9.12	16.7	8
Fault tolerance	14.0	2	29	8.84	13.3	9
Recoverability	14.1	1	29	7.44	16.7	10
Operability	14.2	4	28	7.38	23.3	11
Compliance	14.3	3	26	6.98	16.7	12
Learnability	15.3	2	31	10.18	13.3	13
Analysability	16.4	3	30	9.08	16.7	14
Understandability	16.8	6	30	7.78	0.0	15
Manageability	17.1	2	28	7.76	10.0	16
Maturity	17.2	1	32	8.73	16.7	17
Adaptability	17.7	4	32	10.65	10.0	18
Explicitness	17.8	8	31	7.64	0.0	19
Stability	18.2	3	31	9.18	13.3	20
Resource behaviour	18.3	5	27	7.02	3.3	21
Reusability	18.7	2	31	8.94	16.7	22
Customisability	19.1	3	30	8.39	6.7	23
Helpfulness	19.3	5	32	8.31	3.3	24
Interoperability	19.4	4	30	6.09	6.7	25
Replaceability	19.6	1	28	8.99	13.3	26
Changeability	19.7	7	32	8.14	0.0	27
Installability	19.8	5	32	7.79	10.0	28
Conformance	19.8	5	31	7.25	10.0	29
Traceability	21.6	2	32	8.35	6.7	30
Degradability	22.7	1	32	9.56	13.3	31
Testability	23.6	2	31	8.77	6.7	32

Table 7
The top five quality sub-characteristics and their quality characteristics

Sub-characteristic	Relative importance among sub-characteristics	Corresponding quality characteristic	Relative importance among characteristics
Availability	1	Reliability	1
Accuracy	2	Functionality	2
Security	3	Functionality	2
Suitability	4	Functionality	2
Time behaviour	5	Efficiency	3

The large value of the standard deviation shows that the distribution of the scores of sub-characteristic is more scattered than that for the quality characteristics. Some respondents remarked that many sub-characteristics were very similar and that it was difficult to assign unique ranking.

As with the previous analysis, we carried out a paired two-sample *t*-test. The results confirm the ranking of Table 6. All five most-important sub-characteristics belong to the three most important quality characteristics. This is illustrated in Table 7.

5. Product quality metrics

A metric initiative must address the needs of its potential users. Practical metrics should be objective, cost effective and informative. A number of product quality metrics were defined by applying the goal-question-metric paradigm. It should be emphasized that we wanted to identify product quality metrics suitable for intranet applications. However, some of these metrics may also be applicable to non-intranet applications [18].

Because of its low relevance to the quality of an intranet application and the low usefulness of its

measures, suitability is not considered further. Therefore, we focus on the development of product quality metrics for the remaining four sub-characteristics.

We first identified the *user* and then the *levels of metric application*. As different users of product quality metrics may have different interests, it is essential to identify all of them [4]. Table 8 lists primary interests of the potential users.

In this study, the metric users were in the Computing Services Centre of a local university. The organization's mission was to create and maintain a computing environment for the pursuit of academic excellence with quality management; to provide quality professional services to students, staff, and administration: to ensure effective use of computing resources and to cooperate with other centres of higher education and professional bodies. The centre has provided numerous internet and intranet applications and services since its start in 1995. Numerous web, ftp, news, mail, and proxy servers have been set up for individual and departmental users. Staff may read public announcements and administrative notes posted onto web servers. Also, students can renew loans of library books. Currently, the staff-only intranet includes many administration and information

Table 8 Primary interests of metric users

Metric user	Primary interests
End-users	Quality and value of intranet application
End-user management	Overall control and improvement across projects in the business unit. Business value of intranet application
Software management	Overall control and improvement across specific intranet projects they manage
Software developers	Overall control and improvement across specific intranet project activities and work product in which they are involved
Software quality assurance	Cross-sections of primary interests above (depending on whether they are working at the business unit level or project level)

Table 9 Levels of metrics application

Level	Metric application
Company (or business unit)	Data across several projects may be lumped together to provide a view of attributes such as productivity, quality, and cycle-time across projects.
Project group	Data across several projects in the same product area may be lumped together to provide a view of the same attributes, but within the product group.
Project	Data within the project is tracked and analyzed both in-process and post mortem in order to plan and control the project, as well as to improve similar projects across the business unit.
Component	Data within a component, say a subsystem, of a project is tracked and analyzed for managing the development of that component, as well as improving its quality over time.

applications, such as the on-line phone book, message board system, and computer account applications.

Scope of measurement varies according to the need of the user, resources availability, company culture, awareness of quality within an organization, and organizational structure. It can be classified into four levels as shown in Table 9.

As one objective of this study was to identify product metrics for measuring the quality of inhouse developed intranet applications, we focused on the project group and the project levels of metric application.

5.1. Defining product quality metrics

Starting from the needs of metrics users, product metrics for the four short-listed sub-characteristics were developed. Table 10 summarizes them.

The number of metrics is overwhelming. Many scholars suggest the use of simple and fewer product metrics. Also, if the cost of obtaining the metric is too high, it will be hard to justify its use. To reduce the set, a series of interviews were conducted with software management and developers of intranet and internet applications. The interview focused on the following areas:

- 1. Scope of the quality measurement.
- 2. Importance and urgency of the metric.
- 3. Cost estimate for implementing the metrics.

Two project managers, three developers, and one application support staff member were invited to comment on each metric. All interviewees agreed that measurement should start from intranet applications developed by the centre to help it identify the best application. The metrics will be used to investigate the

quality of intranet application at the project level and to compare the quality of a number of intranet applications at the product group level.

5.2. Importance, urgency, and cost of metric

Interviewees were first asked to rate the importance and urgency of the 15 metrics on a three-point scale: low (0), neutral (1), and high (2). The results are shown in Table 11, with the mean shown in parentheses

Interviewees considered AVAIL, FR, and ERT to be useful metrics, because they reflect the important features of quality and influence the public image of the intranet application. Support staff commented that the availability is very important, because users notice when the system is down. NTF is also important, since illegal break-ins cause crisis and damage the reputation of the university. The need for AVAIL, FR, and SFR are most urgent, since service reliability is the primary concern of the centre.

As NRD indicates both product quality and software development process quality, it is important. However, as the centre adopts the 'simple is good' principle to develop intranet applications, most applications are small. Hence, the number of released defects is relatively lower than in large legacy applications. To meet ever-changing requirements, the rate of update of intranet application is high. Also, it is not easy to count the defects after release. As typing mistakes and incorrect URLs in web pages are the most common data error, DQ seems less useful.

After the implementation of an asynchronous transfer mode (ATM) network and use of powerful servers, response time was satisfactory and thus less urgent. Moreover, users are notified with appropriate mes-

Table 10 Quality metrics defined

Metric	Quality characteristic & sub-characteristic	Reference/Threshold
Availability AVAIL = $\frac{\sum Prime hours - \sum Outage hours}{\sum Prime hours}$	Reliability; availability	Normal application is above 0.95. Below 0.86 is unacceptable
Failure rate $FR = \frac{\sum \text{Number Of failues}}{\sum \text{Execution hours}}$	Reliability; availability	High rating indicates poor quality
Nomalized failure rate $NFR = \frac{FR}{\text{Application size}}$	Reliability; availability	High rating indicates poor quality
Number of released defects NRD = \sum No. of software defect	Functionality; accuracy	High rating indicates poor quality
Software failure rate SFR = $\frac{\sum No. \text{ of software failues}}{\sum Execution \text{ hours}}$	Functionality; accuracy	High rating indicates poor quality
Data quality $DQ = \sum_{All \text{ media type}} No. \text{ of defect}$	Functionality; accuracy	High rating indicates poor quality
Normalized number of released defects NNRD = $\frac{NRD}{Application size}$	Functionality; accuracy	High rating indicates poor quality
Normalized data quality NDQ = $\frac{DQ}{Application size}$	Functionality; accuracy	High rating indicates poor quality
Normalized software failure rate NSFR = $\frac{\text{SFR}}{\text{Application size}}$	Functionality; accuracy	High rating indicates poor quality
Trusted computer system evaluation criteria (TCSEC)	Functionality; security	Classification of grades indicates the security level, D is the least secure and A1 is the most secure
Number of threats found NTF = \sum No. of threat	Functionality; security	High rating indicates poor quality
End-user response time ERT=average response time for all end-user transactions	Efficiency; time behaviour	Below 250 ms is not acceptable
End-user response time of time-critical transaction. ERTTC= average response time for all time-critical end-user transactions	Efficiency; time behaviour	Below 250 ms is not acceptable
Largest end-user response time LERT=largest response time for all end-user transaction	Efficiency; time behaviour	Below 250 ms is not acceptable
System response time SRT=average response time for all transactions	Efficiency; time behaviour	High rating indicates poor quality

sages in case more processing time was needed. As the speed of operations was satisfactory, the metric for system response time was less important. However, interviewees agreed that measurement of time behaviour will be important later when computing resources become tight. Due to the satisfactory security implementation, measurement of security is less urgent.

Interviewees were also asked to estimate the cost of implementing each metric using a three-point scale:

Low cost: data is already available and a simple calculation is required.

Medium cost: additional effort is required for data collection and analysis, but collection and analysis are simple.

High cost: additional effort is required for data collection and analysis and also its collection and

analysis are complicated. Expensive tools or training is required. External assessors or advisors may be needed.

Due to the availability of data such as service outages and failures found in the user problem reports and system logs, the cost of collecting AVAIL, FR and SFR is low. Data analysis is also simple and requires little training. NP and NWP do not require much effort.

On the other hand, the remaining metrics, especially for speed and security, cost more to implement due to their complexity. For example, data is available for NTF from system logs and error reports, but it takes extra time to make any investigation.

From the analysis, AVAIL, FR and SFR are the most important, urgent, and cost-effective product quality metrics. The first two are associated with

Table 11 Importance and urgency of product quality metrics

Metric	Importance	Urgency	Cost
Availability (AVAIL)	High (2)	High (2)	Low
Failure rate (FR)	High (2)	High (2)	Low
Normalized failure rate (NFR)	Low (0)	Low (0)	Medium
Number of released defects (NRD)	High (1.7)	Low (0.3)	High
Software failure rate (SFR)	High (1.8)	High (1.7)	Low
Data quality (DQ)	Neutral (1)	Low (0.3)	High — tools may be required
Normalized number of released defects (NNRD)	Low (0)	Low (0)	High
Normalized data quality (NDQ)	Low (0)	Low (0)	High
Normalized software failure rate (NSFR)	Low (0.3)	Low (0)	Medium
Trusted computer system evaluation criteria (TCSEC)	Low (0)	Low (0)	High — external assessor is required
Number of threats found (NTF)	High (1.7)	Neutral (0.8)	Medium — extra effort is required for checking
End-user response time (ERT)	High (1.7)	Neutral (1)	High — tailor-made tools is required
End-user response time of time-critical neutral transaction (ERTTC)	Neutral (0.7)	Low (0.2)	High — tailor-made tools is required
Largest end-user response time (LERT)	Low (0)	Low (0)	High — tailor-made tools is required
System response time (SRT)	Neutral (0.8)	Low (0)	High — tailor-made tools is required

the availability characteristic while SFR relates to accuracy.

6. Evaluation of quality metrics

In order to verify the three product quality metrics, we applied them to five in-house intranet applications. Only applications that have been in use for at least six months were selected for our study. Generally speaking, they are popular applications that have been modified since their launching. All these intranet applications are protected through user authentication

and can be accessed via either the campus network or dial-up services. An overview of these applications is given in Table 12. Details are provided in Table 13.

6.1. Applying the product quality metrics

Data was collected from user problem logs, Help Desk monthly reports, system logs, and software maintenance logs over a 3-month period. Prime and outage-hours of the five applications were collected. Outage hours are the sum of all failures reported to the Help Desk or recorded in system logs. Lack of

Table 12 Description of five intranet applications

Application	Description
A	This application converts text-based documents into html format and posts them on the intranet. Administrators only need to upload a document into the application. Format enriching features, such as text formatting and hyper link creation, are available. Aging of documents is also supported. A handy and easy-to-use index page is provided for document retrieval.
В	This is one of the oldest applications. It is an information inquiry system. Users can specify very flexible search criteria. Data is transferred from the database of legacy systems within one-day.
C	This application is an information inquiry system. Users can inquire real-time information by category. Key word search features are also supported. Data is transferred from the database of legacy systems.
D	This application is a simple administrative system for staff. It provides electronic forms for a user to make requests for servicing. Upon completion, a notification is sent via electronic mail.
Е	This is one of the most frequently invoked applications. With this application, staff can enjoy sophisticated inquiry functions on sensitive data. To avoid unattended login session, a time-out feature is provided after 15 s or a user-defined idle time-out is exceeded. Data is transferred from the database of legacy systems within one-day.

Table 13
The five intranet applications

	A	В	C	D	E
Months of service	14	26	18	15	6
Users	Staff+students	Staff+students	Staff+students	Staff	Staff
Server involved	1 web server	1 Oracle web server	1 Oracle web server	1 Oracle web server	1 Oracle web server
Database server	None	Oracle database	Oracle database	Oracle database	Oracle database
Number of procedures	28	13	44	37	185
Service hours	24 h except daily, weekly and special maintenance	24 h except daily, weekly and special maintenance	24 h except daily, weekly and special maintenance	24 h except daily, weekly and special maintenance	Normal office hours except specia maintenance. Regular maintenance takes place in non-office hours.

availability due to a problem on the server side was also included in this figure.

Failure events are classified into the following categories [6]:

- 1. Software-related problem in a software application. Software maintenance can inject defects and causes failures. Moving a web site can produce invalid hyper links in the intranet applications. Other failures occur when transferring data from legacy systems. The time to solve this type of problem may involve a few hours to a week of effort.
- 2. *Hardware problem* with computing hardware but not network equipment. When a disk failure occurs, the time to repair may be a few days.
- 3. *Network communication problem*. This is reported as soon as it occurs. The time to repair may vary from several minutes to a day.
- 4. Server problem except hardware problem. The database server, web server, and host/clustered servers may fail. As host servers can recover by self-reboot, it only takes a few seconds or minutes to fix. Database servers need longer.
- Other unknown or miscellaneous problem. These
 often relate to the user account. For example, a
 user does not have sufficient disk space for some
 function.

The distribution of outage by failure type based on outage duration per failure type over total outage duration is shown in Table 14.

The major causes of outage are software and network communication problems. The software related outage is substantial for application A, since problems caused a long delay. Network communication problems occur mainly during normal office hours. As a result, they account for a major outage for application

Table 14 Distribution of service outage

Failure type	A (%)	B (%)	C (%)	D (%)	E (%)
Software problem	90.6	62.9	66.3	44.7	35.5
Hardware problem	0.9	7.5	7.3	6.0	0.0
Network problem	7.3	25.3	23.0	25.0	58.7
Server problem	0.4	2.3	1.1	1.2	0.6
Others	0.8	2.0	2.3	23.1	5.2
Total	100.0	100.0	100.0	100.0	100.0

Table 15 AVAIL, FR and SFR for applications A–E and ranking

Application	AVAIL (%)	Ranking based on AVAII		Ranking based on FR	SFR	Ranking based on SFR
A	83.37	5	1.22	5	1.00	5
В	95.19	2	0.20	2	0.11	3
C	94.71	4	0.42	3	0.26	4
D	95.13	2	0.58	4	0.08	2
E	95.63	1	0.07	1	0.04	1

E, whose prime-usage hours are the normal office hours.

Based on the prime and outage-hours, the metric results of AVAIL are shown in column 2 of Table 15. AVAIL of application E is highest, while that of application A is the worst. AVAIL of application A is less than the acceptable level (86%). This suggests that management should pay more effort to improve this application. Availability of application C is slightly below the standard (95%), while that of other applications are over 95%.

We can rank the applications from 1 to 5. But as the difference between AVAIL of B and D is very small, we assign them equal rank.

For the FR metric, the failures were collected as for the AVAIL metrics. Execution time is estimated by multiplying the number of invocations of an intranet application by the average duration of an invocation. Applications B and E are the most popular intranet applications with respect to the number of invocations. Average duration per invocation varies from 4 to 8 min. Application D is the least frequently invoked application and associates with the shorter average duration per invocation. Thus, it provides the smallest total execution. The distribution of failures by failure type based on count of failure per type divided by total count is shown in Table 16.

The software failures and network communication problems are substantial in all applications. The lower the failure rate of an application, the higher its rank. FR of application E is lowest, while application A has the highest FR, which is several times as large as the others.

Unlike FR, SFR only counts failures caused by software defects. The software failure rate of the five applications is shown in column 6 of Table 15. The

Table 16 Distribution of failures

Failure type	A (%)	B (%)	C (%)	D (%)	E (%)
Software problem	82.3	52.5	61.6	13.3	53.9
Hardware problem	0.7	4.9	4.1	6.7	0.0
Network problem	11.8	29.5	24.7	60.0	33.3
Server problem	3.3	8.2	6.9	16.7	5.1
Others	1.9	4.9	2.7	3.3	7.7
Total	100.0	100.0	100.0	100.0	100.0

lower the software failure rate, the higher its rank. Compared with other applications, application A has a notably higher SFR. When functions with defects are invoked frequently by users, the failure rate will be high. SFR of applications D and E are much lower than that of other applications.

6.2. Ranking from user satisfaction survey

We then carried out a user satisfaction survey for the five applications. The purpose of the survey was to conduct project-wise comparison on the quality of the five intranet applications. Four briefing sessions were organized to ensure that all respondents would interpret the questions in the same way. Each session started by describing the background and purpose of the survey, and then gave details of each question. Finally, the respondents were allowed to raise questions. Each session took 30 min, and the respondents were then asked to complete the questionnaire within 1 week. Twenty three people attended the sessions and all returned the questionnaire. Another 27 questionnaires were sent via electronic mail, and 12 questionnaire were returned. The respondents were all staff of the university. The questionnaire is available from the author.

Table 18 Combined results of experiment and user satisfaction survey

Application	Ranking from metric values (1: the best, 5: the worst)			Ranking from survey (1: the best, 5: the worst)		
	AVAIL	FR	SFR	Availability	Accuracy	Overall quality
A	5	5	5	5	5	5
В	2	2	3	2	3	2
C	4	3	4	4	4	4
D	2	4	2	3	2	3
E	1	1	1	1	1	1

Table 17
Mean ranking of application A–E from user satisfaction survey

Application	Mean rating					
	Availability	Accuracy	Overall quality			
A	4.86	4.83	4.77			
В	2.31	2.60	2.14			
C	3.23	3.94	3.80			
D	3.11	2.14	3.03			
E	1.49	1.49	1.26			

The questionnaire has four questions:

Question 1 was intended to verify that the respondent has used all five applications. In Questions 2 and 3, respondents were asked to rate the relative quality in term of availability and accuracy among the five applications. Question 4 collected the comparison on overall quality among all applications. Duplicate ranking was not allowed. All respondents had used all five applications within the previous 3 months.

Participants first ranked the relative quality of the applications on availability by assigning a number from one (the best) to five to each. The number was used to calculate the mean rank of each application. The one with the lowest score is the best. Application E is the best and application A is the worst. (Table 17)

Participants then ranked the relative quality of the five intranet applications on accuracy. Finally, participants ranked the relative overall quality of the five intranet applications.

T-test analysis was applied to check the ranking of the five applications according to each of the three characteristics. The ranking based on availability and overall quality are the same. The combined results from the experiment and the user satisfaction survey are summarized in Table 18.

As indicated by the paired two-sample *t*-test, results from the experiment and the user satisfaction survey are correlated at the 95% confidence interval. The availability metric, AVAIL, shows very similar results on availability and overall quality for the five intranet applications. Also, the threshold values of AVAIL are useful for metric users.

FR shows similar results in availability and overall quality, with a small difference (a swap between the third and the fourth applications in the ranking). Ranking generated from the accuracy metric, SFR, is the same as that from the user satisfaction survey for accuracy. Although it does not exactly match with the overall quality, it shows similar results with a small difference.

7. Conclusion

Based on a quality profile involving intranet application, 15 product quality metrics were found to be good indicators by using the GQM paradigm; their feasibility was examined by the target metricuser. Three product quality metrics were short-listed for evaluation and an experiment was performed on five live intranet applications. The metrics were shown to be valid by using a satisfaction survey. The evaluation shows a very encouraging result: all three metrics are consistent with the user's view of the specific quality sub-characteristic. Furthermore, the metric, AVAIL, indicates the overall satisfaction.

A weakness of our study is the relatively small sample size of the two user surveys. We could greatly improve the accuracy of our results if there had been more participants. However, we believe that the developed metrics are relevant to the targeted users. The users have agreed that the three product quality metrics are practical, useful and economical to use to measure the quality of intranet applications. With these metrics, the centre may start measuring the quality of their intranet applications and seek further improvement.

The product quality metrics are not only useful in intranet applications; they should also be applicable for internet applications. In fact, AVAIL, FR and SFR have been applied extensively in other applications.

References

- [1] V.R. Basili, D.H. Hutchens, Experimentation in software engineering, IEEE Transactions on Software Engineering 12 (7), 1986, pp. 733–743.
- [2] M.L. Berenson, D.M. Levine (Ed), Business Statistics, A First Course, Prentice Hall, Englewood Cliffs, NJ, 1998, pp. 361–406.
- [3] I. Campbell, The Intranet: Slashing the cost of business, International Data Corporation, http://home. netscape.com/ comprod/announce/idc/summary/html, 1998.
- [4] M.K. Daskalantonakis, A practical view of software measurement and implementation experiences within Motorola, IEEE Transactions on Software Engineering 18 (11), 1992, pp. 998–1010.
- [5] J. Desborough, Intranet Web Development, New Riders Publishing, 1996.
- [6] P. Dixit, M.A. Vouk, D.L. Bitzer, Reliability and availability of a wide area network-based education system, in: Software Reliability Engineering - ISSRE'96, The IEEE Computer Society Press, 1996, pp. 213–219.
- [7] S. Fuller, K. Pagan, Intranet Firewalls: Planning & Implementing Your Network Security System, Library of Congress Cataloging-in Publication Data, 1997, pp. 57–98, pp. 127, pp. 133–135.
- [8] M.L. Hines, A.A. Goerner, Software Quality: Attributes and modalities, in: Software Quality Management III Vol. 2. Measuring and Maintaining Quality, BCS, 1995, pp.137–146.
- [9] C. Holtham, M. D'Cruz, A. Tiwari, The application of intranet and business groupware technologies to support collaborative learning with face-to-face students, in: The Digital University: Reinventing the Academy, Springer, London, 1998, pp. 267–280.
- [10] B.D. Ingraham, Multimedia and Higher Education: Some Lessons from CAMILLE, Empowering Teachers and Learners through Technology, SEDA Paper 90, Staff and Educational Development Association, 1995, pp. 119–125.
- [11] S.H. Kan, Metrics and Models in Software Quality Engineering, Addison Wesley, Reading, MA, 1995, pp. 1–12, pp. 273– 294.
- [12] H. Kempter, H. Frank, Measurement based quality control of real-time system specification, in: Software Quality Management III Vol. 2 Measuring and Maintaining Quality, BCS, 1995, pp. 159–180.
- [13] J. Kirakowski, Quality of use of multimedia learning systems: practical considerations, in: The Digital University: Reinventing the Academy, Springer, London, 1998, pp. 233–244.
- [14] R.W.K. Lam, Practical Product Quality Metrics for Intranet Applications, MSc dissertation, Hong Kong Polytechnic University, 1999.
- [15] T. Musson, E. Dodman, An investigation of quality profile for different type of software, in: Software Quality Management III Vol. 2 Measuring and Maintaining Quality, BCS, 1995, pp. 119–135.
- [16] Process Software Corporation, Intranet A Guide to Intraprise-Wide Computing, http://www.process.com/news/ intrawp.htm, 1998.
- [17] A. Schneider, C. Davis, Intranet Architecture Integrating

- Information Design with Business Planning, http://www.intrack.com/intranet/iarchi.shtml, 1997.
- [18] N.F. Schneidewind, Methodology for validating software metrics, IEEE Transactions on Software Engineering 18 (5), 1992, pp. 410–422.
- [19] E.P.W.M. Van Veenendaal, J.J.M. Trienekens, Testing based on Users' Quality needs, Reliability, Quality and Safety of Software-Intensive Systems, Chapman & Hall, London, 1997, pp. 243–255.
- [20] D.E. Welch, Intranets The New Engine of Corporate Productivity, http://www.welch.com/wpapers/intranet.html, 1996.
- [21] M. Xenors, D. Christodoulakis, Software quality: the user's point of view, in: Software Quality and Productivity — Theory, Practice, Education and Training, Chapman & Hall, London, 1995, pp. 266–272.

[22] R.H.J. Zeist, P.R.H. Hendriks, Specifying software quality with the extended ISO model, in: Software Quality Management IV — Improving Quality, BCS, 1996, pp. 145–160.



Hareton K.N. Leung