METRICS DEVELOPMENT FOR THE QUALIS OF SOFTWARE TECHNICAL PRODUCTION

Desenvolvimento de métricas para o Qualis de produção técnica de software

Marinho Jorge Scarpi

A B S T R A C T

Objective: To recommend metrics to qualify software production and to propose guidelines for the CAPES quadrennial evaluation of the Post-Graduation Programs of Medicine III about this issue. **Method**: Identification of the development process quality features, of the product attributes and of the software use, determined by Brazilian Association of Technical Standards (ABNT), International Organization Standardization (ISO) and International Electrotechnical (IEC), important in the perspective of the CAPES Medicine III Area correlate users, basing the creation proposal of metrics aiming to be used on four-year evaluation of Medicine III. **Results**: The in use software quality perception by the user results from the provided effectiveness, productivity, security and satisfaction that originate from its characteristics of functionality, reliability, usability, efficiency, maintainability and portability (in use metrics quality). This perception depends on the specific use scenario. **Conclusion**: The software metrics should be included in the intellectual production of the program, considering the system behavior measurements results obtained by users' performance evaluation through out the favorable responses punctuation sum for the six in use metrics quality (27 sub-items, 0 to 2 points each) and for quality perception proof (four items, 0 to 10 points each). It will be considered as very good (VG) 85 to 94 points; good (G) 75 to 84 points; regular (R) 65 to 74 points; weak (W) 55 to 64 points; poor (P) <55 points.

Key Words: Software. User-Computer Interface. Education, medical. Post-graduation. Medicine.

INTRODUCTION

Qualis is the set of CAPES determinations specific to each intellectual product, established by indicators, which stratifies the quality of the production in order to evaluate the stricto sensu postgraduate programs.

The demand for computerized procedural solutions generated by the various areas of knowledge, requires proper software development, involving engineering, internal and external quality of the product itself and the perception of quality during practice by the user. The engineer uses software programming languages, databases, tools, platforms, libraries, standards, procedures, media environments and software quality issues. These items can be classified as to its proposed use in: development (software construction), management activities (planning) and quality assurance activities[1].

The Brazilian Association of Technical Standards (ABNT), in partnership with International Organization Standardization (ISO) and the International Electrotechnical (IEC) has defined forms and approaches to understand the quality of a software system, establishing a model with quality components which includes the development process (ISO / IEC 14598-1 to 6); internal product quality attributes (ISO/IEC TR 9126-1 and 3: 2003, ISO / IEC 14598-1) and external product quality attributes (ISO/ IEC TR 9126-1 and 2: 2003, ISO/IEC 14598-1) and the quality perceived by the user (ISO/IEC 14598-1 and ISO/IEC TR 9126-1 and 4)[2]. ISO/IEC TR 9126 standards relate to the quality of the three model items and ISO 14598-1 relates to the evaluation criteria. The items related to product quality and to the effect produced by the software are reflected in the perception of use, appropriate skills related to the CAPES Medicine III, responsible for the implementation of quality metrics in use (Figure 1).

The objective of this study was to propose metrics in order to qualify the software production aimed at a quadrennial evaluation of the postgraduate courses in Medicine III/CAPES.

METHODS

This publication is the result of a study presented during the Fifth Meeting of Postgraduate Medicine III between December 8th and 9th, 2014, in São Paulo, Brazil. It was carried out starting from the principle that the quality assessment process is specific to the needs and expectation(s) of person(s), user(s) and area(s), not setting quality in an absolute way.
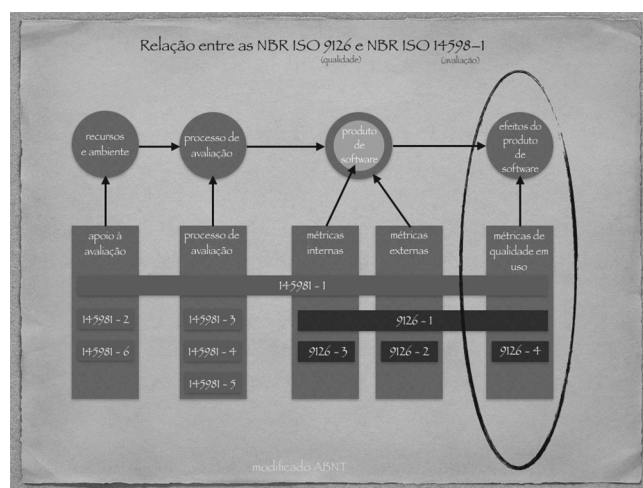

FIGURE 1 - *Model of software quality components (modified ABNT)*

In this sense, the choice of metrics took into account the user's perspective, the objective characteristics and known values, done in a particular use scenario.

The content of the standards established by the Brazilian Association of Technical Standards - ABNT, extracted the topics that could be considered quality items for the software in the perception of this specified user, related to Medicine III. Relevant standards were analyzed: ISO/IEC TR 9126, ISO/IEC 14598 and ISO/IEC 122072.

Use of outcome measures and software product quality characteristics, were quantified, guiding the evaluation metrics.

RESULTS

The range of specified goals for this particular context of use, called "quality model in use" (system behavior measurement), is evaluated according to four items listed in ISO/IEC 9126-13: efficiency, productivity, safety and satisfaction. Effectiveness, refers to the ability to achieve specified goals with accuracy and completeness. Productivity, is related to the use of appropriate amount of resources compared to the efficacy obtained. Security, is the fact of presenting acceptable levels of risk of injury to people, businesses, software, properties or environment. Satisfaction, is the ability to satisfy users. These items directly explain the perception of quality by this particular user, for each

of them bigger quantification for evaluation metrics: is given 0-10 points according to quality perception of the manifestation from the least to most favorable (total 40 points). This measurement should have sufficient accuracy, through rational observation, with a valid range, according to what is documented, i.e. after being clearly understood the definition of each of the four items.

In order to evaluate the product quality, the determinations of ISO 91 263 are used, which presents six quality features and 27 sub-features, called "quality attributes" (Figure 2).

For the purpose of establishing metrics, according to the user perception, each of these sub-features should be scored 0-2 points, being "zero" the failure to meet the requirements, "1" partial service and "2" complete satisfaction.

The sub-feature "compliance" appears in all the features, because it evaluates how the software meets the requirements of legislation and all kinds of standardization or normalization applicable to the context.

The characteristic "Functionality" (0-10 points) is related to the performance that satisfies the user in its context of use. Its five sub-features are:

- Adequacy, evaluates the set of features according to user needs;
- Accuracy, concerns the supply of accurate results and within what was agreed/requested;
- Interoperability, refers to the way the software interacts with other specified system(s);
- Security, evaluates the system's ability to protect users information and to provide it only (and always) to authorized people;
- Accordance.

The "Reliability" feature (0-8 points) demonstrates the performance level under the conditions prescribed. Its four sub-features are:

- Maturity, concerns the software's ability to prevent failures due to defects in the software;
- Fault Tolerance, it is related to the software's ability to maintain proper operation, even when there are defects in it or in its external interfaces;
- Recoverability, points out the ability to recover after a failure, restoring its performance levels and recovering its data;
- Accordance

The feature "Usability" (0-10 points) is the understanding of the software, the learning of its operation, its operation and the user's delight. In this case the dependence of the human-computer interface can be seen. Its five sub-features are:

- Intelligibility, how easily the user understands its features and evaluates if it can be used to meet his/her specific needs;
- Learnability, notes how easily the system can be learned;
- Operational, indicates how the product facilitates its operation, including how it tolerates operating errors;
- Attractiveness, concerns the characteristics that may attract a potential user to the system, which can range from the adequacy of information provided to the user until the visual refinements used in its graphical interface;
- Accordance.

The characteristic "Efficiency" (0-6 points) qualifies if the execution time and the resources involved are compatible with the level of performance of the software. Its three sub-features are:

- Behavior in Relation to Time Measures if the response time (or processing) are within specifications;
- Resource Usage measures both the consumed resources and the system ability to use the available resources;
- Accordance.

The characteristic "Maintainability" (0 to 10 points)

demonstrates how easy it is to modify in order to improve functionality, to correct defects, failures or errors. Its five sub-features are:

- Analyzability, demonstrates how easy it is to diagnose possible problems and to identify the causes of defects or failures;
- Modifiability is related to how easily the software behavior can be modified;
- Stability, measures the software ability to avoid side effects due to modifications;
- Testability expresses the ability to test the modified system, as well as the new features as those not directly affected by the change;
- Accordance.

The characteristic "Portability" (0-10 points) refers to the ability of the system to be transferred from one environment to another. It is considered as "environment", all the conversion factors such as different conditions of infrastructure (operating systems, versions of databases, etc.), different types and hardware resources (such as taking advantage of a greater number of processors or memory). Factors such as language or easy to create test environments should also be considered as portability features. Its five sub-features are:

- Adaptability, representing the software ability to adapt to different environments without the need of additional actions (settings);
- Ability to be Installed, identifies how easily the system can be installed in a new environment;
- Coexistence, measures how easily a software lives with others installed in the same environment;
- Ability to Replace represents the ability of the system to replace other specified system in a specific context of use and environment. This attribute interacts with both adaptability and with the ability to be installed;
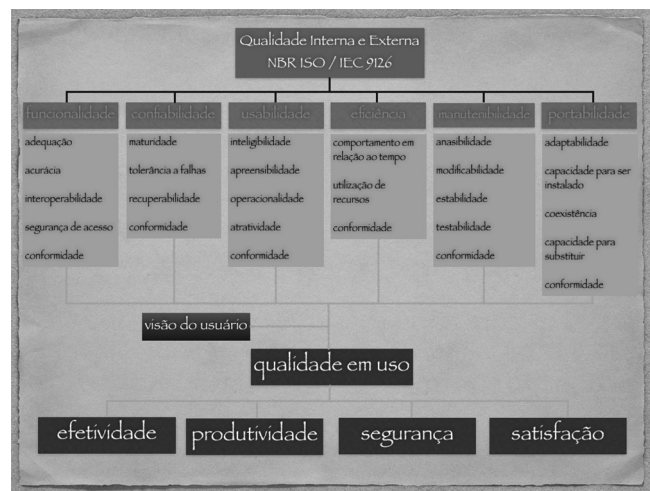- Accordance.

FIGURE 2 - *Items to be quantified as perceived by the user of the software*

DISCUSSION

The results of this study show that despite the high degree of subjectivity in the perception of quality in software use, explicit specification for this use and the topics of software product features allow quantitative analysis of qualitative data through scoring for intellectual production evaluation purposes in Medicine III CAPES.

The metrics proposal presented in this study is, in the opinion of the author, the nascent process of the evaluation in focus and susceptible to continuous improvement.

It is suggested that the submission of software for quality analysis, following the criteria of ABNT and the prospect of use, should be presented in the form of text containing the following topics:

I. Introduction: these should contain information about the subject matter, the target audience/user and details of the relevant ISO/IEC with deal with software quality (ISO/IEC 9126, 14 598 and 12 207);

II. Objective(s): to report its intention;

III. Method:

A. to display the developed software and the description of the user population that has been tested.

B. to develop description by activity according to ISO/IEC 12107, stating the following: 1) which procedures and methods will be used to implement the activities; 2) which tools and equipment will support the realization of the activities in order to simplify and automate the work; 3) what the appropriate profile of whom will perform the activities and what the required training in the procedures, methods and tools are so that they can carry out the activities as appropriate; 4) which metrics process can be used so that execution process could have its quality evaluated.

C. Brief description of the itens of the relevant engineering software[3] for the displayed items: programming languages, databases, tools, platforms, libraries, standards, procedures, media environment, software quality.

D. Comments on performance and productivity related to the presented software: software process, project planning and management, quality management, specification and requirements analysi s, system design, implementation and testing, delivery and maintenance.

E. Refer to the six features and the 27 sub-characteristics that make the quality attributes.

IV. Results: present the results in the quality study perception and the system behavior measures in the form of punctuation.

V. Discussion: discuss the results according to user needs and expectations and their perception of quality, against the specific context of use.

VI. Conclusion(s).

VII. References.

VIII. Statement of software.

The importance of software as intellectual production is evident in the Medicine III Area Document that highlights for Professional Master the technical production will be assessed according to importance/impact on area and consistency with technological lines of scientific expertise and area focus (it will be considered complete publishing articles in journals, processes and relevant protocols and impact in the area; patent registrations and technical production quality consistent with the proposal)"[4]. This emphasis is consistent with the highly interdisciplinary nature of the Professional Masters, offering creativity to software. Among the questions of evaluation, for "artistic production, technical, patents and other considered relevant productions", which include software, the weight of 40% is given within the discretion of "intellectual production" to which is assigned 35% of the weight of program evaluation. Added to this, is the evaluation criterion "Applicability of Produced Papers" with 20% of the item "Student Corpus and Study for Course Completion", due to market demand for software solutions.

CONCLUSION

Qualis metrics for software products should be included in the intellectual production of Area Document, with the highest percentage for the Professional Master. It will be considered very good (VG) from 85 to 94 points; good (G) from 75 to 84 points; regular (R) from 65 to 74 points; weak (W) from 55 to 64 points; Poor (P) <55.

R E S U M O

Objetivo: Recomendar métrica para qualificar a produção de software propondo diretrizes para a avaliação dos Programas de Pós-Graduação da Medicina III. **Método**: Identificação das características de qualidade para o processo de desenvolvimento, para os atributos do produto e para o uso de software, determinadas pela Associação Brasileira de Normas Técnicas (ABNT), International Organization Standardization (ISO) e International Electrotechnical (IEC), importantes na perspectiva dos usuários correlatos da Área Medicina III da CAPES, embasando a criação de proposta para métrica do tema, com vistas à avaliação quadrienal dos cursos de pós-graduação. **Resultados**: A percepção de qualidade em uso do software pelo usuário resulta da efetividade, produtividade, segurança e satisfação proporcionada, que têm origem nas suas características de funcionalidade, confiabilidade, usabilidade, eficiência, manutenibilidade e portabilidade (métricas de qualidade em uso). Tal percepção depende do cenário de uso específico. **Conclusão**: A métrica de software deve ser incluída na produção intelectual do Documento de Área do programa, ponderando os resultados nas medidas de comportamento do sistema em avaliação de desempenho por usuários, considerando a somatória da pontuação favorável para as seis métricas de qualidade em uso (27 sub-itens, de 0 a 2 pontos cada) e a comprovação da percepção de qualidade (quatro itens, de 0 a 10 pontos cada). Será considerado muito bom (MB) de 85 a 94 pontos; bom (B) de 75 a 84 pontos; Regular (R) de 65 a 74 pontos; fraco (F) de 55 a 64 pontos; deficiente (D) < 55.

Descritores: Software. Interface Usuário-Computador. Educação de Pós-Graduação em Medicina.

REFERENCES

1. Falbo RA. Engenharia de Software _ Notas de Aula. Disponível em: http://www.inf.ufes.br/~falbo/download/aulas/es-g/2005-1/Notas-DeAula.pdf. Acesso em 16/03/2015.

2. ABNT. ABNT Catálogo_ segurança, qualidade, padrão e confiança. Disponível em: http://www.abntcatalogo.com.br/. Acesso em:16/03/2015.

3. ABNT. ISO/IEC TR 9126-1_ Engenharia de software - qualidade de produto. Parte 1: Modelo de Qualidade, 2003.

4. CAPES. Documento de Área 2013 _ Medicina III. Disponível em: https://www.capes.gov.br/images/stories/download/avaliacaotrienal/Docs_de_area/Medicina_II_doc_area_e_comiss%C3%A3o_att08de-outubro.pdf. Acesso em: 17/03/2015.

Address for correspondence:
Marinho Jorge Scarpi
scarpi@terra.com.br