Systematic approach for migrating to a structured reporting context based on free-text reports and the DICOM Structured Reporting guidelines

Eduardo Beckhauser[1]**, Vinicius Andreoli Petrolini**[1]**, Alexandre Savaris**[1]**, Aldo von Wangenheim**[1]**, Dirk Krechel**[2]

[1]Brazilian Institute for Digital Convergence – INCoD
Florianópolis, Santa Catarina, Brazil

[2]Hochschule RheinMain
Wiesbaden - Germany

Abstract. *Even with the many advantages of adopting a structured reporting system, there is little convention on how to disseminate this routine into the report environment. This work proposes a systematic approach to migrate a system routine from free-text reports to structured reports, focusing on the DICOM Structured Reporting guidelines. We evaluated this proposal by creating a reporting module in the context of a telemedicine system, and performing case study covering ultrasonography reports. Using the AdEQUATE model, the evaluation showed a high user perception from the system, directly reflecting the quality of our proposal. The results are a set of defined premises and steps that turns a telemedicine system into a complete structured reporting environment.*

1. Introduction

Medical reports have always been a great mediator between doctors and patients. It's an essential component for many tests, capturing critical elements of the study and recording information for future use [Homorodean et al. 2012]. In this scenario, badly written and incomplete reports lead to a series of problems, ranging from misinterpretation to wrong diagnosis, affecting the final interpretation and the patients' life directly [Pool and Goergen 2010] [Homorodean et al. 2012].

The use of free-text reports contributes to this scenario because the physicians have the liberty to write these documents in a natural way, without the necessity to follow more strict procedures to provide their medical opinion [Homorodean et al. 2012]. Furthermore, it's known that physicians are still reluctant to use structured reporting approaches, despite the increasing interest for structured reporting systems [Pool and Goergen 2010] [Bosmans et al. 2011].

One main reason for this to happen is the little standardization of how to design a system that can be considered a structured reporting provider. In other words, it's still difficult to measure technological and scientific needs for a system to achieve such status.

To ensure that report systems provide more reliable and organized reports, techniques and approaches that aim to conserve information have to be used [Hussein et al. 2004]. Additionally, when a system uses guidelines to create and manipulate these kind of documents, the doctors can use those rules to create a report routine that eases their work and increases reliability in the provided reports [Barcellos et al. 2011].

Finally, it is most likely that physicians that emit reports look at images and findings in a systematic way, contributing for the adoption of a structure reporting standard [Bosmans et al. 2015].

One of the most common standards that proposes a set of guidelines to structure report documents is the DICOM Structured Reporting (DICOM SR) [Clunie 2000], defining a set of rules to standardize how a report can be created, manipulated and how it's data has to be organized in order to give a complete report, without information dubiety.

Homorodean [Homorodean et al. 2012] defined foundations for structured reporting in echocardiography by creating a consistent set of templates using the coded concepts of a previously developed ontology and the DICOM standard rules. Faggioni [Faggioni et al. 2017] assessed the opinion on structured reporting and its usage by radiologist members of the Italian Society of Medical and Interventionist Radiology (SIRM). The assessment was performed using an online survey composed by ten questions about the physician's interest in structured reporting. Towbin [Towbin and Hawkins 2017] created a web-based application that calculates length discrepancies using a calculator, which was evaluated by two radiologists using two sets of ten radiographs to measure the difference between the *pen and paper* method and the calculator. Other authors proposed the creation of a structured reporting system focusing on a specific context of use. These authors focused their works on specific medical fields, without concerning themselves with the systematization of the migrating process to a SR system.

They focused their work on specific steps to achieve a structured reporting scenario for a medical specialty, without investigating a general approach to the problem.

There is no consensus on what steps to take to create a fully functional structured reporting system, independent of medical specialties, that could be extended to every medical domain. In order to facilitate the migration from text-based reports to structured reports, this work proposes a systematic approach that covers every step that a system has to complete to achieve a structured reporting provider status, starting from a set of text-based reports and the physicians' expertise. Our proposal focuses on a routine that covers the subject in every aspect of how to perform this migration. Finally, this proposal was developed to be able to be applied in any telemedicine environment.

To evaluate the proposal, we conducted the construction of a telemedicine module in the context of the Santa Catarina State Integrated Telemedicine and Telehealth System (STT/SC) following each step proposed in the methodology. In this context, the STT/SC offers a wide range of services since 2010 [Beckhauser et al. 2016]. STT/SC is composed by a set of information systems, services and modules that provide access to healthcare and healthcare education in the Santa Catarina state, Brazil, responsible for the maintenance of over 7 million examinations in several exam modalities [Beckhauser et al. 2016].

It is known that a telemedicine system must adhere to high quality standards in order to increase productivity and to offer the best possible user experience [Sánchez-Pi and Molina 2010]. To evaluate our structured reporting system, we conducted a case study evaluation based on the AdEQUATE model (questionnAire for Evaluation of QUAlity in TElemedicine systems), an evaluation model based on the ISO/IEC 25010 standard built accordingly to the specificities of the telemedicine systems. We focused this evaluation on obstetric ultrasonography because it is a specialty that leads

directly with findings and patient' images, so a well built system is necessary to attend its needs.

This study is organized as follows: Section 2 conceptualizes DICOM, DICOM Structured Reporting and the AdEQUATE model; Section 3 shows the main goals of our proposal; Section 4 describes the proposal, followed by Section 5 where we discuss the system created and the evaluation. In Section 6 we discuss the results, present our conclusions and suggest future directions for the topic. Section 7 describes the threats to the validity of our work.

2. Background

DICOM: Created in 1983 by NEMA (National Electrical Manufactures Association) and by ACR (American College of Radiology), the Digital Imaging and Communications in Medicine is a standard for exchanging information between various nodes in a hospital environment [Clunie 2000]. It allows the storage and transference of medical data between medical units transparently, rising interoperability and promoting a series of protocols to be adopted by those devices [Pianykh 2009].

DICOM Structured Reporting: DICOM SR is a standard used for structured reporting proposed by the Association of Electrical Equipment and Medical Imaging Manufacturers (NEMA) [Association et al. 2015]. It brings image and textual information together, linking report measurements with other data types [Barcellos et al. 2011] [Clunie 2000], unifying medical images and textual information, creating a bond between medical data and measurements [Barcellos et al. 2011].

The AdEQUATE model: developed as a customization of the ISO/IEC 25010 standard, its main goal is to evaluate the quality of telemedicine systems based on the perception of its end users [Alves et al. 2016]. The model decomposes quality concepts by revising the characteristics from the ISO/IEC 25010 standard and attributing these concepts based on their relevance for telemedicine systems [von Wangenheim et al. 2013]. The result is a set of new metrics customized to a generic Telemedicine and Telehealth application derived in a systematic manner and molded in the shape of a questionnaire [Alves et al. 2016]. This questionnaire aims to collect data on the perception of end users about the quality of telemedicine systems being evaluated.

3. Objectives

Our main objective is to propose and evaluate a systematic approach for migrating text-based reports to a structured reporting environment, with the following specific objectives:

(1) Formalize every step for a system to achieve a structured reporting status into a defined methodology;
(2) Integrate the methodology proposed on (1) with the STT/SC reporting routine;
(3) Create a structured reporting module that covers every step proposed on (1) and (2);
(4) Evaluate the structured reporting system to ensure its high external quality, supporting and justifying the initial proposal.
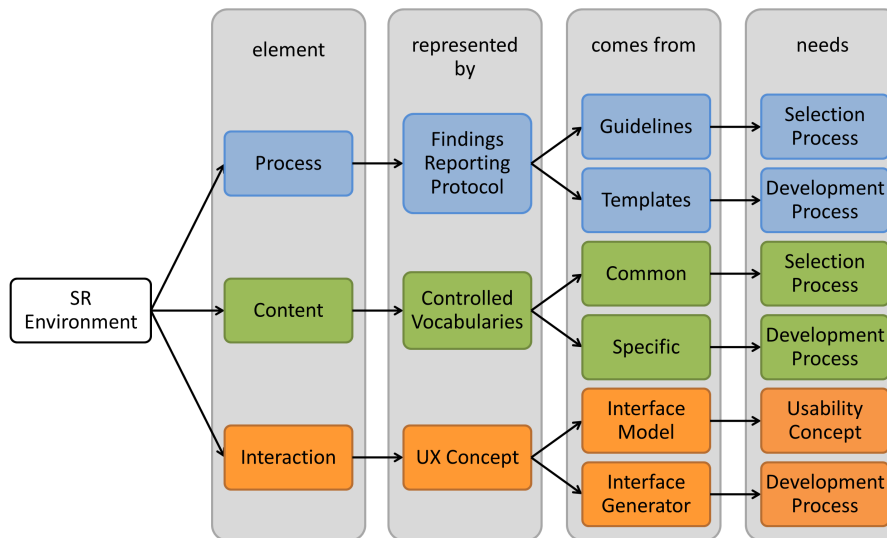
Figure 1. The systematic approach proposal.

4. Methodology Proposal

In order to cover every step to complete our proposal, we based our work on the DICOM Structured Reporting Guidelines and on the state-of-the-art related works, which served as complementary source of necessities and information showed by the authors that were not necessarily covered by the DICOM SR standard.

In this section, we discuss the proposal and associate each step to the structured reporting system built inside the STT/SC in this case study. The approach overview can be seen at Fig. 1. It represents a linear flowchart, where each color represents an individual flow of our approach. Each column represents a different stage that defines what to achieve in each step, defining how we represent and create each concept.

A report system can be considered a structured reporting environment if it is built following three main directions:

- A *finding reporting process*: the first phase of our proposal consists on the *process* needed to create structured documents. When a physician writes a report, it is natural that he looks at images and findings in a systematic way [Bosmans et al. 2015], following some *reporting protocol*. To support this scenario, the DICOM Structured Reporting standard offers a set of *guidelines* that can be followed to create structured documents, applying constraints to the report's creation:
    - item types that a report can use, like *containers*, *texts*, *numbers*, reference to *images*, etc.
    - item relationships, which gives meaning to two associated nodes, like *contains*, *has properties*, etc.
    - rules for the generation and filling of the structured document and how it can be stored and recovered;
    - rules for the document construction, like patient's and physician's information, report versioning, etc.
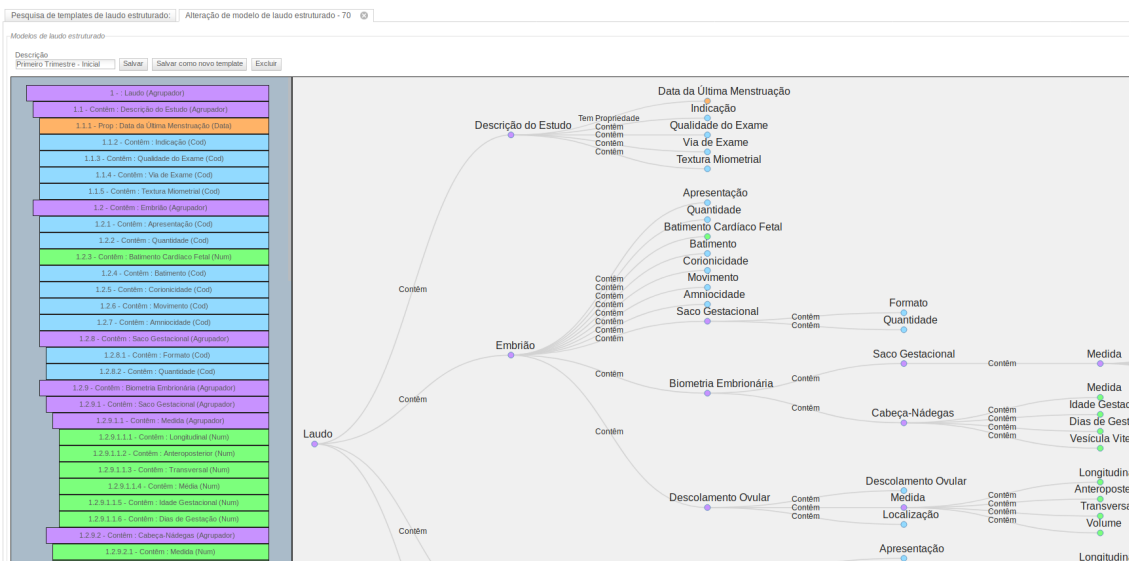
Figure 2. Template construction module

In addiction to that, a structured reporting system has to possess a well-defined process to construct structured documents following these guidelines, so the physicians have the ability to create new templates. To achieve this goal, it is vital that a structured reporting system has a *template development process*, which guides the user through the report construction and applies rules for the user to create items with meaning to the report in general. A template development process gives transparency and flexibility for the user to create several types of reports that will be used depending on the case, while following the standard constraints. The Fig. 2 shows how we conducted the development of a template generator module inside the STT/SC. Both sides of the shows two different ways to create the report templates. While the list is easier to understand, the tree structured gives an overview of the entire document. Each item type has a specific colour and shows the main informations for the user to understand the template, like the *relationship* from one item to another, its *type*, a *unique identifier* associated with every item, etc. The user can create new items by right-clicking an item and choosing the new item's type, concept name and context group according to the DICOM SR standard. When created, both list and tree update, keeping consistency in the template construction.

- A *formally defined content*: the report's content is the basis for a structured document, giving meaning to the report and flexibility in the use, eliminating the dubiety inherent to the free-text reports.

  Controlled terminology are collections of terms, organized according to a methodology that can be used to specify codes and relationships between nodes, making it easy to recover and process information from the report documents due to its uniqueness and immutability. There are two main ways to obtain controlled terminology:

  - *selection process* of already defined terms: where the structured reporting developers collect, in association with physicians, disseminated vocables already presented in controlled terminology, like RadLex, SNOMED,

UCUM, ONTONEO, etc.

    – *vocabulary development process*: by creating local vocabulary due to some impeding imposed by the other vocabularies, like language or context coverage.

Fig. 2 shows how this step was covered by the creation of a local vocabulary for the obstetric ultrasonography specialty. This vocabulary was specifically developed to serve as a translated version of the vocables found in other english vocabularies. As we can see in Fig. 2, the created vocables were not translated back to the English language to maintain the original vocabulary language of the created local vocabulary.

- A *well-defined interaction model*: the last phase of our proposal refers to how the report is showed for the user to fill its inputs. The user's experience influence the final product perception [Garrett 2010]. Every step in the structured reporting system process has the responsibility to secure a high user perception. Two main subjects were covered:

    – *development process of an interface generator*: the structured report templates have all the information and constraints needed to give meaning to the report, but the DICOM SR standard do not restrict how this information is presented to the user [Clunie 2000]. This information has to be generated in an easy and procedural way to be filled, leading to the creation of a tool that can read the structured template, interpret its nodes and constraints and create the appropriate interface to be presented to the user.

    – *usability concept of the interface model*: a user interface can be seen as the closest layer between a user and the machine [Dufresne 1994], but it is worthless to have a interface generator if the system that doesn't follow any usability concepts. In order to be natural to the user and to have a high usability concept, it is important that the reporting module follows the system visual identity, focusing on colour, shape and component disposition. Furthermore, some improvements could be done in order to make the interface generated more appealing, like displaying two items per row, create the same kind of components for similar item types, etc.

To increase the usability of the interface generated, our structured reporting module followed the STT/SC system visual identity, with the same set of components used in the entire system. The interface generator handled similar items with the same kind of components, sizes and shapes.The Fig. 3 shows how the structured reporting tree showed in Fig. 2 was transformed into inputs that could be more easily managed and filled by the physicians (a) and how the process of reporting generated the final version of the report document (b).

5. Evaluation of the SR system

In order to evaluate external quality, we adopted the case study research design [Yin 2017].

Figure 3. STT/SC's visual identity applied to the created structured reporting system. a) the report to be filled by the physician. b) the final report

5.1. Case study

5.1.1. Case study definition

The objective of this case study is to evaluate the external quality of the STT/SC' structured report system in order to verify the system in terms of the end user's point of view, analyzing the perceived ease of use by healthcare professionals that work with examination reports in terms of the minimal quality needed for the construction of report templates and in terms of the minimal quality needed for its content filling and visualization.

5.1.2. Case study execution

During the execution phase, we applied the AdEQUATE questionnaire, collecting data from the end-users of the structured reporting system utilization. Users that are familiar with the obstetric ultrasonography reports were invited to participate in the evaluation. We obtained 10 valid responses from three hospitals which are part of the STT/SC context. Every respondent was a physician in the obstetric ultrasonography area, with years of experience, having issued several reports over the career, being part of the STT/SC for more that 3 years. All respondents were experienced users and we considered that their opinion is highly consolidated. The questionnaire data were used to calculate the distribution of each response for each evaluation item. The overall quality of the structured reporting system, as well as its stronger and weaker points, were inferred using statistical considerations for Likert scale as defined in [Boone and Boone 2012]. The acquired data were analyzed and the questionnaire items were grouped by its subcharacteristics, generating two metrics (M1 and M2) that were calculated for every subcharacteristic (SC). During the analysis of the discrete data we used descriptive statistical techniques, adopting the median as a representative for the middle of the distribution and quartiles for grouping

[Boone and Boone 2012].

Figure 5 shows a percentage distribution of each possible answer grouped by the ISO/IEC 25010 sub-characteristic from the AdEQUATE model. As the chart shows, the majority of the sub-characteristics have been positively evaluated, with more than 50% of the answers being either *totally agree* or *agree*. In fact, only 4 sub-characteristics had a percentage less or equal 50%: *acessibility*, *maturity*, *fault tolerance* and *recoverability*. The best sub-characteristics according to the evaluation were: *flexibility*, *functional Correctness*, *functional appropriateness* and *autenticity* with a score of 100% being *totally agree* or *agree*. The *accessibility* sub-characteristic was the only one with more than 50% of the answers being either *non-applicable* or *don't know*. This result suggests that the users thought that this item was irrelevant to the test and in the structured reporting system context. The points system chart, shown in Fig. 4 presents scores calculated similarly to the SUS score [Finstad 2010], using a weight system to differentiate each type of answer. The main goal of this chart is to facilitate the system quality comprehension. The score is calculated as follows in 1:

$$S_n = \left(\sum_{i=1}^{10} W(\tilde{R}_i) \right) \times 10030 \tag{1}$$

The equation $\tilde{R}_i$ is the median of responses given by the users for the sub-characteristic n by respondent i. Each possible response value is based on a weight function $W(\tilde{R}_i)$, whose values are gained from a [0, 3] Likert scale.

The response options *don't know*, *non applicable* and *don't understand* are ignored. The last component in (1) is S_n, and its value ranges from 0 (worst quality degree – every one responded *totally disagree*) to 100 (best quality degree possible – every one responded *totally agree*).

Analyzing the Fig. 4, we can observe discrepancies between this score compared with the one showed in Fig. 5. While the overall order is maintained, some SCs have been plotted with different scores. This happens because the first chart shows the real distribution responses, while the second one considers the difference from every response option. The SC *time behaviour*, for example, shows a positive score in Fig. 4, however, it achieved a lower score in Fig. 5. Some other SCs, like *accessibility*, *fault tolerance*, *recoverability* and *maturity* scored less than 50 points as showed in Fig. 4 [1].

6. Discussion and Conclusions

Based on the results acquired during the case study, we perceive a positive quality perception from the STT/SC system by its end users regarding the use of the structured reporting platform, which obtained a good overall quality, supported by the positive responses obtained in Fig. 5. The majority of the highest *scores* were related directly to the use of the structured reporting module, like *flexibility*, *functional correctness* and *functional appropriateness*. This shows that the system is more prominent to have a high quality when it's developed according to a defined methodology that covers every step to it's completeness. Using the DICOM SR guidelines made sub-characteristics like *comfort*, *trust* and

---

[1]Further data regarding this evaluation are available at https://tinyurl.com/cbms2019systematicApproach
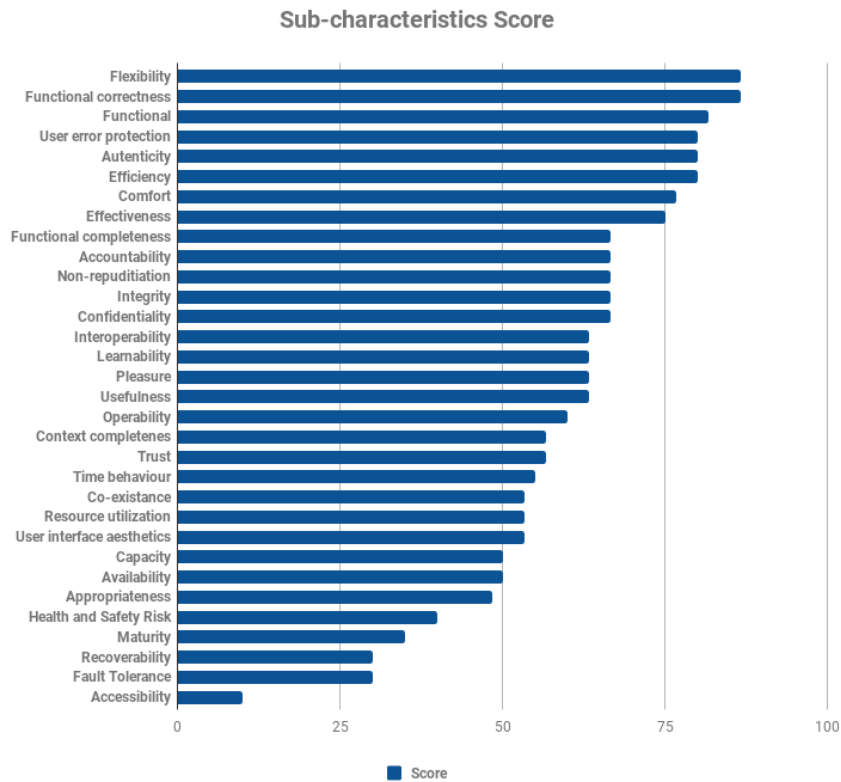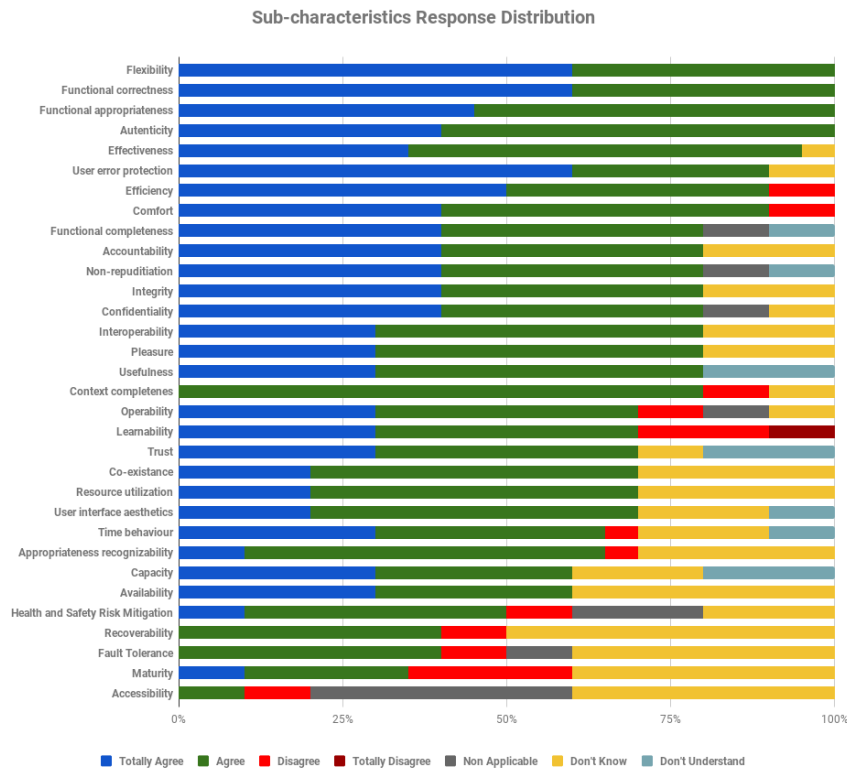
Figure 4. M1 for ISO/IEC 25010 quality subcharacteristics.

learnability achieve higher scores, because the standard brings a specific order on how the tasks should be done, favoring the system learnability and raising the user's system confidence.

Other characteristics, like *pleasure* and *non-repudiation* achieved a high score, with more than 75% of the answers being either *totally agree* or *agree*. This reflects the use of a *well-defined interaction model*, in which the system is presented to the user in a understandable and enjoyable way. Even with good results in the evaluation, as previously shown, some sub-characteristics had bad results, like *accessibility*, *maturity* and *fault tolerance*. While some of them are expected for a newly created structured reporting module, like *maturity*, it shows that some improvements can be done to maximize user experience and to minimize system failures.

As a general aspect, we conclude that a system that follows a systematic approach is more prominent to achieve a high user perception when developed correctly. The *finding reporting process* routine brings flexibility and context coverage in the use of the structured reporting module. The *defined content* brings interoperability and organizes the report data, removing dubiety inherent to free-text reports. Finally, the use of *usability concepts* and a *defined interface generator development* bring comfort and pleasure in the system use. Besides that, the system has the potential to bring similar results for every template used to issue structured reports, turning the system into a structured reporting environment.

Figure 5. M2 for ISO/IEC 25010 quality subcharacteristics.

Further work is yet to be done in this field. Due to the fact that it is a work focused in obstetrical ultrasonography reports, the system can be expanded to other medical modalities to create a more complete system. Furthermore, the systematic approach can be enhanced when other specialties are adhered to the system, improving the system quality as a whole.

7. Threats to Validity

To mitigate problems with internal validity, specially regarding selection bias, the execution of the case study was performed by a group of health professionals that issue reports daily. One possible threat to a construction bias is the use of the AdEQUATE evaluation model. This model was not been extensively validated by a large population in the context of other systems; however, we believe that this threat is minimized because the adopted questionnaire had its items systematically derived from the ISO/IEC 25010 standard, TAM, and SUS, which are largely used methods. The small number of performed interviews may also present a threat related to external validity. We had to select only physicians that were identified as potential users (n=10). Despite the small sample, it represents a homogeneous group in terms of experience in reporting exams in the context of the STT/SC.

8. Acknowledgments

Foundation - FAPESC and the Coordination for Higher Level Graduates Improvement - CAPES.

References

Alves, J. M., Savaris, A., von Wangenheim, C. G., and von Wangenheim, A. (2016). Software quality evaluation of the laboratory information system used in the santa catarina state integrated telemedicine and telehealth system. In *Computer-Based Medical Systems (CBMS), 2016 IEEE 29th International Symposium on*, pages 76–81. IEEE.

Association, N. E. M. et al. (2015). Nema-the association of electrical equipment and medical imaging manufacturers.

Barcellos, C. L., von Wangenheim, A., and Andrade, R. (2011). A reliable approach for applying dicom structured reporting in a large-scale telemedicine network. In *Computer-Based Medical Systems (CBMS), 2011 24th International Symposium on*, pages 1–6. IEEE.

Beckhauser, E., Petrolini, V. A., Savaris, A., Alves, J. M., and von Wangenheim, A. (2016). Are single-board computers an option for a low-cost multimodal telemedicine platform?: First tests in the context of santa catarina state integrated telemedicine and telehealth system. In *Computer-Based Medical Systems (CBMS), 2016 IEEE 29th International Symposium on*, pages 163–168. IEEE.

Boone, H. N. and Boone, D. A. (2012). Analyzing likert data. *Journal of extension*, 50(2):1–5.

Bosmans, J. M., Neri, E., Ratib, O., and Kahn Jr, C. E. (2015). Structured reporting: a fusion reactor hungry for fuel. *Insights into imaging*, 6(1):129–132.

Bosmans, J. M., Weyler, J. J., De Schepper, A. M., and Parizel, P. M. (2011). The radiology report as seen by radiologists and referring clinicians: results of the cover and rover surveys. *Radiology*, 259(1):184–195.

Clunie, D. A. (2000). *DICOM structured reporting*. PixelMed Publishing.

Dufresne, A. (1994). Deuxièmes journées francophones" hypermédias et apprentissage". *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation*, 1(1):125–129.

Faggioni, L., Coppola, F., Ferrari, R., Neri, E., and Regge, D. (2017). Usage of structured reporting in radiological practice: results from an italian online survey. *European radiology*, 27(5):1934–1943.

Finstad, K. (2010). The usability metric for user experience. *Interacting with Computers*, 22(5):323–327.

Garrett, J. J. (2010). *Elements of user experience, the: user-centered design for the web and beyond*. Pearson Education.

Homorodean, C., Olinic, M., and Olinic, D. (2012). Development of a methodology for structured reporting of information in echocardiography. *Medical ultrasonography*, 14(1):29.

Hussein, R., Engelmann, U., Schroeter, A., and Meinzer, H.-P. (2004). Dicom structured reporting: Part 2. problems and challenges in implementation for pacs workstations. *Radiographics*, 24(3):897–909.

Pianykh, O. S. (2009). *Digital imaging and communications in medicine (DICOM): a practical introduction and survival guide*. Springer Science & Business Media.

Pool, F. and Goergen, S. (2010). Quality of the written radiology report: a review of the literature. *Journal of the American College of Radiology*, 7(8):634–643.

Sánchez-Pi, N. and Molina, J. (2010). Adaptation of an evaluation system for e-health environments. *Knowledge-Based and Intelligent Information and Engineering Systems*, pages 357–364.

Towbin, A. J. and Hawkins, C. M. (2017). Use of a web-based calculator and a structured report generator to improve efficiency, accuracy, and consistency of radiology reporting. *Journal of Digital Imaging*, pages 1–5.

von Wangenheim, C. G., von Wangenheim, A., McCaffery, F., Hauck, J. C. R., and Buglione, L. (2013). Tailoring software process capability/maturity models for the health domain. *Health and Technology*, 3(1):11–28.

Yin, R. K. (2017). *Case study research and applications: Design and methods*. Sage publications.