Avaliação de Adequação Funcional:

Percepções e Tendências observadas na Indústria de Software

Lucas Lopes
Depto. de Ciências Exatas
UFPB
Rio Tinto - Paraíba - Brasil
lucas.fernando@dcx.ufpb.br

Juliana Saraiva
Depto. de Ciências Exatas
UFPB
Rio Tinto - Paraíba - Brasil
julianajags@dcx.ufpb.br

Yuska Aguiar
Depto. de Ciências Exatas
UFPB
Rio Tinto - Paraíba - Brasil
yuska@dcx.ufpb.br

RESUMO[1]

A demanda por softwares que sejam executados em diferentes contextos e usuários, traz à tona a importância da avaliação da adequação funcional. Esse atributo versa sobre o grau em que um produto de software fornece suas funções atendendo as necessidades declaradas e implícitas, quando usadas sob condições especificadas. A despeito de recomendações acerca do tema, a ISO SQuaRE 25000 preconiza o que precisa ser avaliado quando a adequação funcional é observada. Porém, a falta de pragmatismo pode ofuscar os insumos essenciais para se instanciar as recomendações em empresas de software. Portanto, neste trabalho relata-se os resultados do survey realizado através de entrevistas semi-estruturadas realizadas com 9 empresas de software, de 4 regiões do Brasil. O objetivo foi obter informações para compreender as percepções e tendências a respeito da avaliação da adequação funcional na indústria de software. Os resultados, embora não generalizáveis, subsidiaram as discussões iniciais sobre o processo de avaliação de qualidade, assim como ferramentas e métricas adotadas, assim como stakeholders envolvidos. Observou-se que o conceito deste atributo algumas vezes é desconhecido, e outras ignorado, tendo em vista a complexidade e o custo de realizar testes para medí-lo. A completude e a corretude, são as subcaracterísticas mais analisadas, em detrimento da adequação funcional propriamente dita.

CCS CONCEITOS

•Software and its engineering~Software creation and management~Software verification and validation~Software defect analysis~Software testing and debugging

PALAVRAS-CHAVES

Qualidade de Software, Adequação Funcional, ISO 25000, Indústria de Software

1 Contexto e Motivação

O uso de softwares para realizar atividades cotidianas tem crescido, sendo imprescindíveis para os negócios, entretenimento, saúde, educação e outras áreas. A pluralidade de contextos de uso de software e consequente heterogeneidade de perfis de usuários lança o desafio de garantia da qualidade desses produtos. Para que os usuários possam desempenhar suas atividades de forma adequada o software precisa passar por processos de avaliação de qualidade.

A garantia de qualidade pode ser compreendida como a definição de processos e padrões que devem levar a produtos de alta qualidade [1]. Sendo assim, esta qualidade não está apenas relacionada ao cumprimento dos requisitos funcionais, mas também no atendimento dos requisitos não funcionais - características implícitas do sistema, ao processo de desenvolvimento de software e aos padrões aplicados [2].

Nessa perspectiva, modelos de padronização de qualidade, como a ISO/IEC 25010 (SQuaRe - *Software Quality Requirements and Evaluation*), foram criados para definir quais características de qualidade devem ser levadas em consideração ao avaliar um produto de software [3]. Um dos atributos de qualidade definidos na referida norma é a **adequação funcional**, que estabelece o grau em que um produto fornece suas funções atendendo as necessidades declaradas e implícitas, quando usadas sob condições especificadas [3]. A adequação funcional

[1] Trabalho de conclusão de curso, sob orientação da professora Juliana Saraiva submetido ao Curso de Bacharelado em Sistemas de Informação do Centro de Ciências Aplicadas e Educação (CCAE) da Universidade Federal da Paraíba, como parte dos requisitos necessários para obtenção do grau de BACHAREL EM SISTEMAS DE INFORMAÇÃO.

possui três subcaracterísticas: a (i) *completude funcional*, grau em que um conjunto de funções cobre todas as tarefas especificadas e objetivos do usuário; (ii) *correção funcional*, que é o grau em que um produto ou sistema fornece os resultados corretos com o grau de precisão necessário e a (iii) *adequação funcional propriamente dita*, que é o grau em que as funções facilitam a realização de tarefas e objetivos especificados.

Num cenário onde os softwares são expostos a diferentes contextos de utilização, os usuários acabam possuindo necessidades de adaptações para a execução desses softwares. Aliado a isso, a quantidade de dispositivos de hardware e sistemas operacionais disponíveis no mercado, e consequente demanda de intercomunicação entre os softwares - através de desenvolvimento de software como serviço - também impõe desafios na garantia da qualidade do software [4]. Nesta conjuntura, a observância da adequação funcional no processo de desenvolvimento/avaliação de software torna-se imperativa, pois a conformidade do contexto de uso do software precisa ter tanta qualidade, quanto a completa e correta execução das funcionalidades do sistema.

A despeito da existência de padronizações relacionadas à qualidade de produto de software, como a Família ISO 25000, a aplicação prática da teoria preconizada por estes padrões pelas empresas de software precisa ser investigada, uma vez que a ISO recomenda o que deve ser feito e não como deve ser feito. Portanto, investigar, junto à indústria de software como a teoria tem sido adotada na prática de desenvolvimento/avaliação de software pode revelar insumos essenciais sobre como se instancia as recomendações normativas.

Portanto, o presente trabalho tem por objetivo investigar o estado-da-prática do gerenciamento da qualidade na indústria de software no contexto brasileiro, compreendendo o processo de avaliação da qualidade, pessoas envolvidas, ferramentas adotadas e mensuração, no que tange especificamente ao atributo de adequação funcional. Com isso, espera-se entender a percepção das empresas sobre os conceitos de adequação funcional e relatar as tendências de práticas cotidianas dessas no que se refere ao processo de avaliação deste atributo, considerando suas três subcaracterísticas.

É importante ressaltar que para esta pesquisa, **percepção** é entendida como sendo a forma como as empresas compreendem os conceitos acerca do tema: *adequação funcional e sua avaliação*. Por outro lado, **tendência** é analisada levando em consideração a maneira como as empresas avaliam a adequação funcional do software cotidianamente. Isto posto, as Questões de Pesquisa (QPs) que nortearam este trabalho foram: QP1: Quais as etapas do processo de avaliação de adequação funcional conduzido nas empresas de software? QP2: Quais os *stakeholders* envolvidos no processo de avaliação de adequação funcional? QP3: Quais ferramentas são utilizadas no processo de avaliação da adequação funcional? QP4: Quais métricas são adotadas como indicadores de qualidade de adequação funcional?

O artigo está organizado em 5 seções, incluindo esta seção: a Seção 2 apresenta a metodologia adotada no trabalho, enquanto a Seção 3 discute os resultados encontrados. A Seção 4 aponta os trabalhos relacionados com este, e por fim, as considerações finais são expostas na Seção 5.

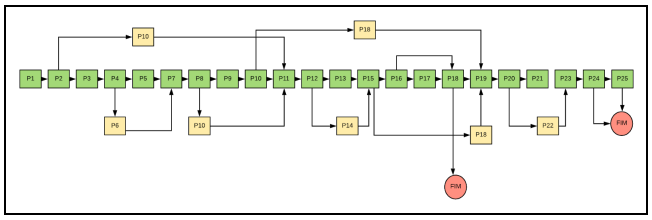
2 *Survey* sobre a prática da avaliação da Adequação Funcional na Indústria de Software Brasileira

S*urvey* é um método de pesquisa qualitativo cujo objetivo é coletar dados e informações a partir de características e opiniões de grupos de sujeitos observados [5, 6]. Levando em consideração que esta pesquisa tem o objetivo de compreender a percepção e a tendência sobre como empresas realizam avaliação da qualidade de software no que diz respeito à adequação funcional, este método foi adotado. A pesquisa foi realizada seguindo 6 etapas: Etapa 1: Levantamento dos sujeitos da pesquisa; Etapa 2: Recrutamento dos potenciais participantes; Etapa 3: Realização das entrevista; Etapa 4: Transcrição das entrevistas em planilhas eletrônicas; Etapa 5: Construção e Codificação do *corpus;* Etapa 6: Análise dos dados.

Adotou-se como instrumento base um roteiro de entrevistas semi-estruturada contendo 25 perguntas que compuseram as três principais linhas de investigação desta pesquisa, no tocante às subcaracterísticas de completude, corretude e adequação funcional propriamente dita. Esse roteiro foi construído por um dos autores deste trabalho e validado por especialistas em qualidade de software (docente e profissional da área). A entrevista sendo semi-estruturada, o entrevistador e o entrevistado tiveram uma flexibilidade nas perguntas/respostas que puderam ser modificadas, adaptadas, suprimidas ou adicionadas, a depender da interação entre os participantes durante sua realização [5].

O questionário norteador da entrevista semi-estruturada condicionou algumas delas a diferentes fluxos de respostas. Ou seja, a ordem das perguntas pôde ser alterada, dependendo da resposta obtida. A fim de guiar a entrevista para uma execução coerente, foi criado um fluxograma para que o entrevistador pudesse conduzi-la (Figura 1). As perguntas do questionário estão disponíveis

no Protocolo de Pesquisa e todas elas foram mapeadas junto às QPs, a fim de que as respostas das empresas subsidiassem as respostas a estas. Elas foram realizadas remotamente através de videoconferência - descrito no



Protocolo de Pesquisa[2]. O público-alvo das entrevistas foram funcionários da área de qualidade de empresas de software, de diferentes portes, contextos e escopos de projeto.

Figura 1: **Ilustração do fluxograma do roteiro de entrevista.**

A entrevista se deu através de 7 momentos: (i) Pedido de autorização para gravação; (ii) Entrega do Termo de Confidencialidade; (iii) Explicação do objetivo da pesquisa e esclarecimentos sobre termos técnicos usados expostos no glossário do protocolo; (iv) Levantamento do perfil do entrevistado - e-mail, função, tempo de trabalho na empresa, formação; (v) Realização das perguntas referentes ao processo de avaliação de qualidade de software; (vi) Agradecimento pelo disponibilidade e informes sobre a divulgação dos resultados e (vii) Espaço para o entrevistado indicar críticas, sugestões, informações adicionais, dentre outros.

2.1 Recrutamento das empresas participantes

Para tentar maximizar a disponibilidade das empresas em participar da pesquisa, abrangendo diferentes cenários - porte, localização, escopo de produto de software - foram convidadas 25 empresas: 12 do Nordeste, 2 do Norte, 2 do Centro-Oeste, 6 do Sudeste e 3 do Sul (Etapa 1). Essas empresas foram contactadas através de relações acadêmico-profissionais dos autores deste trabalho com a indústria de software no país. Esse recrutamento visou a tentativa de ter uma visão geral sobre a percepção e a tendência da avaliação da qualidade da adequação funcional em diferentes cenários de desenvolvimento de software no país.

Após este levantamento, foi enviado um e-mail para estas empresas contendo a apresentação e objetivo da pesquisa, o convite para participação da pesquisa e o Termo de Confidencialidade da equipe de pesquisa, que tratava do sigilo e privacidade das informações, além da anonimização dos dados fornecidos(Etapa 2). Com as empresas que se dispuseram a participar, foram realizadas as entrevistas (Etapa 3) que ocorreram remotamente através de videoconferência gravada com a permissão dos entrevistados. A duração das entrevistas variou entre 35 e 45 minutos.

2.2 Ferramentas adotadas no *Survey*

Todas as entrevistas aconteceram remotamente com suporte ferramental do Skype, Google Meet ou appear.in (Whereby), a depender da escolha do entrevistado. Elas foram gravadas, através do próprio Skype ou através da ferramenta da Apowersoft[3], um gravador de tela online que permite a gravação da tela do computador bem como a saída de áudio. De posse deste material gravado iniciou-se as transcrições de todas as entrevistas para uma planilha eletrônica.

Para auxiliar o processo de transcrição das entrevistas, foram utilizados o Voice Meeter[4], programa que permite alterar a saída e a entrada de áudio do computador por um alto falante e um microfone virtual. Desta forma, ao utilizar o sistema web https://dictation.io, houve a captação do áudio do alto falante, transcrevendo-o para o bloco de notas da própria aplicação web. Ao fim do processo de transcrição, realizou-se uma revisão do português corrigindo possíveis erros de concordância e ortografia resultantes desse procedimento automatizado.

2.3 Codificação e Análise das entrevistas

As últimas etapas foram as fases da codificação e análise do material persistido através das transcrições das entrevistas(Etapa 5 e 6). Essa avaliação se deu com o auxílio da ferramenta Iramuteq[5], que através da Interface R, realiza análises multidimensionais de textos advindos de questionários ou entrevistas, via investigação de *corpus*. O *corpus* desta pesquisa foi composto pelo texto transcrito das entrevistas realizadas, resultando em 30 páginas de documento. Em seguida, iniciou-se a codificação aberta desses dados(Etapa 5).

A codificação aberta, se deu pelo Método de Pesquisa da Teoria Fundamentada, é uma técnica de agrupamento e associação de discursos dos participantes das entrevistas, baseando-se nas terminologias e conceitos abordados na pesquisa e trazidas por eles durante a entrevista. Esses conceitos auxiliam a compreensão do

[2] Protocolo de Pesquisa.pdf

[3] https://www.apowersoft.com.br/gravador-de-tela-gratis

[4] https://www.vb-audio.com/Voicemeeter/

[5] http://www.iramuteq.org/

fenômeno ocorrido, que neste caso é a percepção e a tendência das empresas no que tange à avaliação de qualidade da adequação funcional.

As seguintes categorias de codificação foram criadas: (i) id_entrevistado, (ii) id_resposta_questionário, (iii) porte_empresa, (iv) regiao_empresa, (v) perfil_entrevistado e (vi) tipo_produto (interno x externo). As categorias no Iramuteq são denominadas de "variáveis". Ressalta-se que apesar deste estudo ter um caráter exploratório, sem necessidade de estratificação detalhada das informações, o algoritmo de correlação e similitude de discursos é melhor executado quando se tem um número maior de categorias (variáveis do Iramuteq). O processo de codificação, realizado na ferramenta, se deu associando um segmento do texto (parte dos trechos das entrevistas transcritas) às categorias às quais ele pertence.

A análise de dados foi alicerçada pelas saídas da ferramenta, subsidiando a discussão dos resultados. Destaca-se que a computação na ferramenta é realizada através da análise de similitude de palavras, frequência de palavras, agrupamento de discursos, classificação pelo Método de Reinert, Análise Fatorial de Correspondência, Dendrogramas, dentre outras abordagens disponíveis na ferramenta [8].

3 Avaliação da Adequação Funcional na Indústria de Software

Levando em conta que o objetivo deste trabalho é fazer um levantamento sobre a perceção e a tendência de avaliação de adequação funcional nas empresas de software, as subseções a seguir apresentam e discutem os resultados. O Quadro 1, sumariza o cenário das empresas que participaram das entrevistas. A primeira coluna traz a identificação da empresa (com anonimização) e a segunda coluna representa o porte da empresa. A classificação do porte segue o que é preconizado pelo SEBRAE, que se baseia no número de funcionários [9]. A terceira coluna informa o perfil profissional do participante da entrevista, enquanto a quarta coluna indica o escopo do produto de software. Por fim, a última coluna refere-se ao estado sede da empresa.

É possível observar que empresas de porte Pequeno, Médio e Grande foram contempladas nessa pesquisa, apesar da limitação da quantidade de empresas participantes do estudo. Este fato nos auxilia a ter uma visão que abrange três realidades distintas. Além disso, a diversidade de estados/regiões representados, assim como diferentes perfis profissionais dos entrevistados, também maximizam o aprimoramento da compreensão deste cenário. Nota-se que as pessoas escolhidas pelas empresas para participarem das entrevistas foram selecionadas de acordo com o seu envolvimento no processo de avaliação de qualidade de software, condizentes com o que se preconiza na literatura: Gerentes de Projeto e Qualidade, Analistas de Testes, Analista de Qualidade e Diretor de Tecnologia

Quadro 1. Visão geral das empresas participantes e dos entrevistados.

ID	Porte	Perfil	Escopo Produto	Sede
1	Médio	Gerente de Qualidade	Externo	Paraíba
2	Grande	Gerente de Projeto	Externo	Rio Grande do Norte
3	Médio	Diretor de Tecnologia	Interno	Santa Catarina
4	Pequeno	Diretor de Tecnologia	Externo	São Paulo
5	Grande	Analista de Teste	Externo	Pernambuco
6	Grande	Analista de Teste	Externo	Minas Gerais
7	Pequeno	Diretor de Tecnologia	Externo	Mato Grosso
8	Grande	Analista de Qualidade	Externo	Rio Grande do Sul
9	Grande	Gerente de Projeto	Externo	Rio Grande do Sul

Por outro lado, o escopo do produto de software é uma característica que pode influenciar no processo de avaliação de qualidade do software. Para este trabalho, os produtos comercializados foram classificados em (i) interno, quando desenvolvido para suprir demandas internas da empresa, normalmente menos complexos e (ii) externos, quando a solução é comercializada para outrem. Desta forma, os dois escopos são contemplados na amostra deste estudo.

Tendo em vista que muitas vezes a ISO preconiza normas que informa o que deve ser feito e não como deve ser feito, através dessa investigação, busca-se confrontar as definições de adequação funcional preconizadas pela ISO 25000, trazendo à tona possíveis incompreensões e desconhecimento conceituais sobre este atributo de qualidade de software. Além disso, será possível comparar o que é preconizado pela ISO com a percepção e aplicação nas empresas de software, ressaltando como as especificidades da equipe e escopo do projeto, podem determinar o estado da prática no processo de avaliação de qualidade do software.

3.1 Percepções sobre os conceitos de Adequação Funcional

A análise sobre a percepção de Adequação Funcional e dos conceitos relativos às suas subcaracterísticas foi observada avaliando os discursos dos entrevistados para as 4 perguntas: P1- Como avaliam adequação funcional (e se já ouviram falar)?, P2- O que compreendem por completude e como é avaliada na empresa?, P10 - O que

compreendem por corretude e como é avaliada na empresa?, e P18 - O que compreendem por adequação funcional propriamente dita e como é avaliada na empresa?.

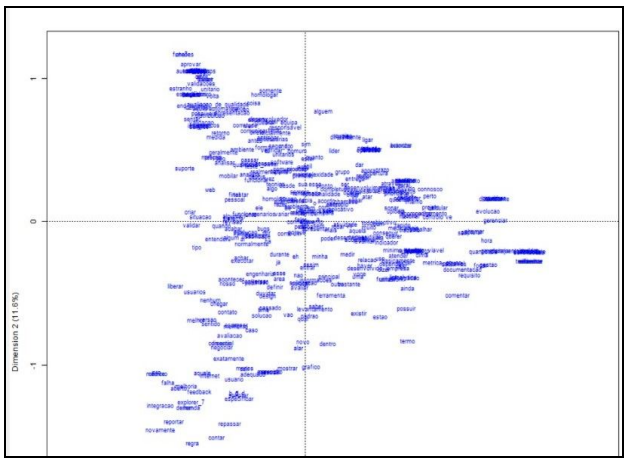


Figura 2: **Análise de similaridade no discurso das empresas.**

Dentre as representações visuais que o Iramuteq fornece após análise dos discursos, uma delas é a avaliação de similaridade textual, mostrada na Figura 2[6]. Ela é representada pela concentração das palavras no eixo central, havendo discrepância de discurso quando as palavras estão dispersas no gráfico. Ainda, quanto mais afastadas estão as palavras que compõem o *corpus*, maior distinção entre os discursos. E se estiverem em quadrantes opostos, além de diferentes, pode-se observar um comportamento de oposição de discursos, o que não acontece ao analisar as entrevistas.

A partir desta é possível observar como as palavras se agrupam quando os discursos de todas as empresas são sobrepostos. De maneira geral, independentemente das especificidades das empresas participantes (Quadro 1), de forma majoritária, o discurso acerca dos conceitos de avaliação da qualidade de software é convergente. A concentração do agrupamento das palavras no centro do gráfico ilustrado na Figura 2 indica que os discursos se distribuem e se comportam de forma semelhante, havendo inclusive sobreposição de palavras idênticas, revelando homogeneidade dos discursos.

Adicionalmente, o Iramuteq disponibiliza uma divisão de classes de palavras que é construída através da avaliação de categorias semânticas de partes do discurso. Ou seja, as classes, representadas por cores diferentes, são definidas de acordo com os significados das mensagens transmitidas através dos textos transcritos das entrevistas. Observa-se na Figura 2 que a ferramenta só forneceu uma única cor (azul) para a representação das classes. Assim sendo, o discurso entre as empresas é tão uniforme que não foi possível para a ferramenta computar significados distintos entre as falas dos entrevistados. Ressalta-se que quando se trata de *corpus* muito pequenos, há recomendações nos manuais do Iramuteq de se duplicar ou triplicar o texto no documento para não haver problemas de computação do algoritmo que faz análise de classes. Apesar de apenas 9 empresas terem participado do estudo, 30 páginas de *corpus* foram obtidas e esse procedimento de duplicação foi realizado. Entretanto, obteve-se o mesmo resultado, uma única classificação de classe decorrente da similaridade do discurso e não devido ao tamanho da amostra.

Após a identificação da similaridade nos discursos das empresas, independentemente do tópico abordado na entrevista (completude, corretude, stakeholders, ferramentas, métricas, etc), procurou-se compreender como elas percebiam os conceitos acerca de adequação funcional. A Figura 3 representa este entendimento, de forma resumida e ilustrativa em uma nuvem de palavras, com destaque em tamanho e centralização das palavras mais frequentes nos discursos.

No centro, a palavra "não" foi a mais mencionada durante as entrevistas, indicando o não (ou pouco) conhecimento sobre os conceitos de adequação funcional. Além da representação baseando-se na frequência em que a palavra aparece, a nuvem de palavras as agrupa de acordo com a proximidade em que elas aparecem no texto analisado. Desta forma, a palavra "não" está perto da palavra "teste" porque quando os entrevistados eram questionados sobre se realizavam a avaliação de adequação funcional, normalmente, a resposta foi "não" associada a um discurso que descrevia o tipo de teste realizado na empresa, mesmo sem estarem atrelados ao atributo de qualidade de software específico.

[6] Ressalta-se que por questão de limitação de espaço, as imagens não puderam ser expostas maiores, assim, todas elas estão disponíveis em um link externo para facilitar a compreensão.

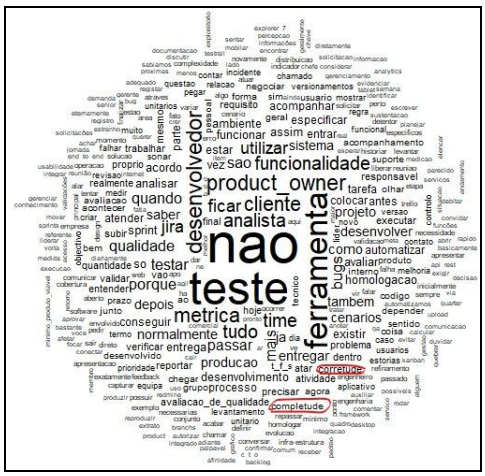

nao
teste
ferramenta
desenvolvedor
product_owner
funcionalidade
cliente
analista
time
metrica
tudo
jira
bugs
completude
correitude
homologacao
desenvolver
automatizar
cenarios
existir
problema
entregar
passar
producao
testar
qualidade
sistema
utilizar
ficar
quando
porque
depois
normalmente
conseguir
ambiente
acompanhar
requisito
negociar
versionamentos
precisar
agora
avaliacao
de_qualidade
levantamento
processo
desenvolvimento
chegar
reportar
verificar
entrega
usuarios
responsavel
colocar
antes
executar
projeto
tarefa
olhar
sprint
subir
atender
saber
acontecer
realmente
analisar
proprio
acordo
sonar
falhar
trabalhar
estar
vez
sao
simples
usuario
mostrar
forma
algo
pegar
unitarios
mesmo
parte
geral
especificar
funcionar
assim
entrar
acompanhamento
suporte
medicao
reuniao
parecido
servicos
etapa
trello
versao
controle
convidar
funcoes
necessidade
contato
validacao
rapido
basicamente
apresentar
produto
melhora
homologacao
incialmente
codigo
sempre
via
depender
sentido
comunicacao
upload
coisa
caso
evitar
usuarios
estorias
refinamento
possiveis
passado
engenharia
comentar
quadro
desktop
integracao
evolucao
homologar
definir
branch
conjunto
acabar
unitario
levantamento
completude
aplicativo
atividade
engenheiro
atar
dentro
dia
ja
hoje
tambem

Figura 4 apresenta a ilustração do agrupamento de palavras contidas nos discursos das empresas, quando foram questionadas sobre o processo de avaliação da qualidade.

Neste agrupamento de palavras emitido pelo Iramuteq, as palavras-chaves do discurso são representadas pelos nós-raízes do grafo, enquanto as palavras que os acompanham estão como ramos. Na Figura 5 as palavras "não", "teste" e "product_owner" foram as mais mencionadas pelos entrevistados, havendo a ramificação destas para as outras as quais foram associadas. Enfatiza-se que os trechos analisados para compreendermos o processo de avaliação foram advindos das respostas referentes às seguintes perguntas: P1: Considerando a definição de adequação funcional trazida pela ISO 25000, ela é avaliada na empresa? Se sim, como? P8: A completude do software é avaliada? Como? P9: A corretude do software é avaliada? Como? P17: A adequação funcional propriamente dita do software é avaliada? Como?

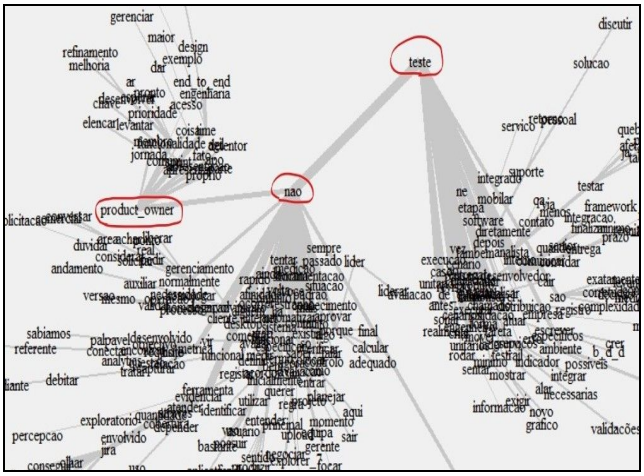


Figura 4: **Agrupamento de palavras - Discurso sobre definição do processo de avaliação.**

Infere-se, portanto, que o "não" apareceu quando eles foram questionados quanto à formalidade de se medir as características da adequação funcional e se elas eram de fato medidas. Como mencionado anteriormente, essa formalidade não existe e normalmente a corretude e completude são conferidas através de testes unitários, teste de validação ou teste funcional. E neste sentido a palavra "teste" aparece como nó-raiz conectando-se com a palavra "não". Os trechos a seguir ratificam esse entendimento:

Entrevistado A: "Dentro do processo de completude nós fazemos o design inicial e fatiamos as entregas, elas vão ser avaliadas em sua completude no sentido, se a entrega está ok e se aquele MVP está ok."

Entrevistado B: "[...] A gente colocou pessoas, pilotos para testar já estamos nessa fase entendeu? Então já foi utilizado por pessoas que vão utilizar no dia-a-dia."

Entrevistado C: "Isso. A gente não faz esse tipo de avaliação."

Entrevistado D: "[...] Liberou a versão 1 em cima dos requisitos que ele estimou no início. Ele (cliente) vai validar, vai fazer a homologação dele. Durante a homologação ele pode solicitar algumas melhorias, se ele solicitar algumas melhorias, ele vai criar as atividades, a gente vai desenvolver e vai passar pelo mesmo processo."

Neste processo, apesar de muitas vezes não formalizado ou pré-definido como único meio de avaliação da qualidade do software pela empresa, o papel do *Product Owner* (PO) parece ser crucial, estando presente na maioria dos discursos, sendo nó-raiz e representando como parte importante nesse processo (Figura 4). No tocante aos *stakeholders* (Figura 5) é possível reafirmar a importância do PO, além do testador, desenvolvedor, time de avaliação de qualidade e em menor intensidade, a participação do cliente no processo de avaliação. Trechos transcritos das entrevistas auxiliam na compreensão das pessoas envolvidas na análise de qualidade.

Entrevistado A: "[...] quem vai fazer a análise desse feedback é a minha PO (*product owner*)."

Entrevistado B: "[...] PO apresenta as features que ele levantou para ser desenvolvido naquelas sprint."

Entrevistado C: "Isso o time de QA, os desenvolvedores e a PO."

Entrevistado D: "Então, aqui tá na parte de corretude eu como PO, vejo que chegam muitas falhas, quando passa para mim." [...] Então algumas falhas, a gente fica olhando os logs e identificando os gaps e esse problemas funcionais e lançando fix imediatamente.

Há de se ressaltar que algumas vezes, os testes que deveriam verificar se todas as funcionalidades solicitadas estão disponíveis e corretas deveriam ser prévios aos testes de usabilidade. É necessário compreender que ao avaliar a adequação funcional, dispensa-se a avaliação da interação do usuário com o sistema [10]. Portanto, os testes funcionais de adequação propriamente dita deveriam ser voltados a examinar se o sistema, ao ser executado, faz com que todos os diferentes tipos de usuários atinjam seus objetivos [11, 12]. Este tipo de avaliação deveria ser realizada para tentar evitar que os usuários passem por experiências frustrantes de uso, que poderiam ser suprimidas pelas avaliações de adequação funcional antes de colocar o software para ser testado por usuários. Este cenário pode ser observado nos discursos abaixo:

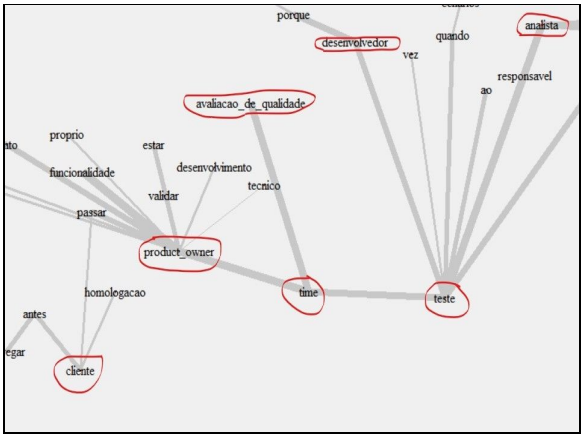


Figura 5: **Agrupamento de palavras - Discurso sobre *Stakeholders*.**

Entrevistado 2: "Passa sim, o desenvolvedor antes de entregar ele faz o teste unitário dele. [...] Quando está ok, o cliente é convidado, se senta para homologar com ele, remotamente e dá ok, aí a gente pede autorização para subir essa mudança em produção. Porque tanto a homologação quanto à produção são ambientes do cliente."
Entrevistado 8: "Temos usuários chaves, algumas vezes o próprio cliente, Design, membros do time técnico, time de engenharia e eu, PO ou PM."
Entrevistado 9: "A gente colocou pessoas, pilotos para testar já estamos nessa fase entendeu? Então já foi utilizado por pessoas que vão utilizar no dia-a-dia."

3.2.2 Ferramentas e Indicadores adotados na avaliação

Foi questionado aos entrevistados quais ferramentas, métricas e indicadores eles utilizam no processo de avaliação de qualidade. No que tange à ferramenta, o Jira[7] e o SonarQube foram as mais citadas nas entrevistas. O Jira dá suporte ao gerenciamento de projetos, agrupando tarefas e chamados de *help desk* sendo utilizado para a visualização de quais funcionalidades estão entregues (completude) e quais estão funcionando sem bugs (corretude). Na Figura 6 nota-se que as palavras "utilizar", "usar" e "jira" são as mais mencionadas pelos entrevistados, revelando a dinâmica de avaliação de qualidade dos softwares.

Entrevistado A: "A gente hoje está utilizando o sonar para a questão de cobertura de desenvolvimento de código."
Entrevistado B: "Sim, a gente usa o Jira."
Entrevistado C: "No Jira tem um extrato de tudo que eu fiz no dia, de todas as minhas atividades e a gente utiliza o jira mesmo."
Entrevistado D: "Não tem ferramenta que acompanha isso não."
Entrevistado R: "Nenhuma ferramenta, é exploratório, percepção, feeling."

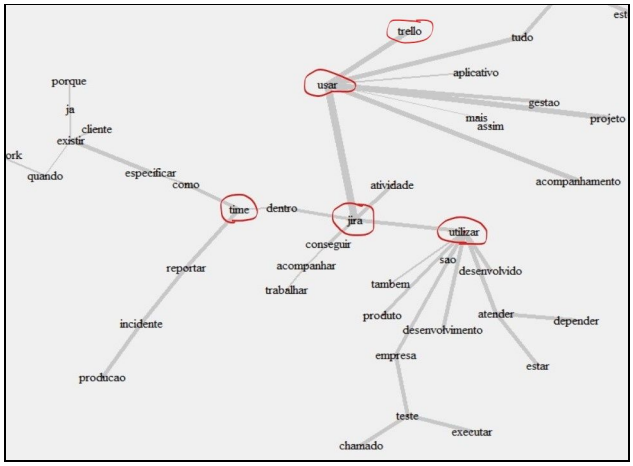


Figura 6: **Agrupamento de palavras - adoção de ferramentas.**

As empresas não adotam ferramentas específicas porque a avaliação da adequação se resume à análise de completude e corretude através de uso do sistema, enquanto ele ainda está em processo de testes e desenvolvimento, pela própria equipe de desenvolvimento ou mesmo por clientes "pilotos". Algumas desenvolvem ferramentas próprias, mas não entraram em detalhe. Assim, a avaliação baseia-se num *checklist*, onde é verificado qual a funcionalidade está implementada (completude) e se está funcionando corretamente (corretude). Adequação funcional propriamente dita é um atributo analisado implicitamente ou não avaliado. Outras ferramentas foram citadas, mas por apenas uma empresa, em casos mais específicos, como por exemplo:

Entrevistado A: *"Tem o **github** que mostra se o código está tudo ok. Cada pessoa tem que fizer responsável pela sua atividade e se der algum problema a gente tem que reportar."*
Entrevistado B: *"A gente usa o **trello** para fazer a gestão de tudo isso."*
Entrevistado C: *"Tem ferramentas **desenvolvidas por nós**, mas tem muitas ferramentas também que a gente utiliza, mas que não podemos falar pois são ferramentas específicas para as certificações."*

Sobre métricas e indicadores, os mesmos informaram que basicamente um "OK" ou "aprovação" da lista de funcionalidades entregues na *Sprint* é indicador de qualidade de software. Poucas empresas usam o SonarQube[8] para avaliar código-fonte, consequentemente, utilizam as métricas coletadas por ele. Na Figura 7 é possível observar que as palavras "não", "usar" e "métrica" destacam-se no discurso, indicando a não adoção de métricas e indicadores específicos no processo de avaliação de adequação. Alguns trechos consolidam os resultados trazidos pelo Iramuteq:

[7] https://www.atlassian.com/br/software/jira

[8] https://www.sonarqube.org/

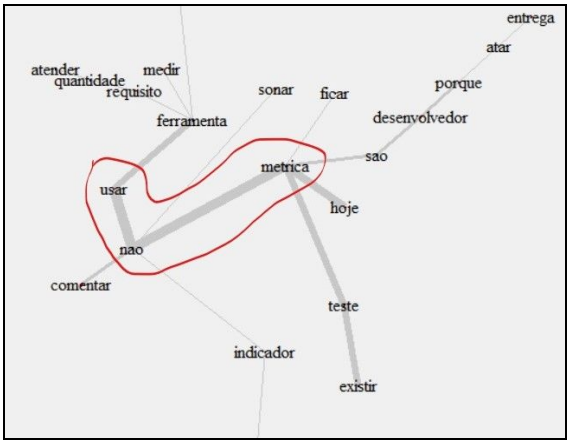


Figura 7: **Agrupamento de palavras - Adoção de métrica/indicadores para avaliação.**

Entrevistado A: *"Essa medição em termos de indicador não. O processo que acontece é que o desenvolvedor entrega para o analista de teste e esse analista faz uma homologação prévia antes de entrar em contato com o cliente para fazer a homologação concluinte."*
Entrevistado B: *"Não a gente não usa indicador, não."*
Entrevistado C: *"A gente não metrifica isso."*

4 Trabalhos Relacionados

Apesar de menos estudada, quando comparada com os outros atributos de qualidade de software preconizados pela ISO 25000, a adequação funcional vem sendo abordada em várias pesquisas ao longo dos últimos anos. Um dos trabalhos que lidam com o tema é o de Rodriguez e colaboradores (2016). Eles apresentaram um estudo de caso sobre a avaliação de adequação funcional de um sistema web denominado BitRRHH 1.0, dedicado à gerência de recursos humanos da empresa [10].

Neste trabalho, a completude, corretude e a adequação funcional propriamente dita foram avaliadas, levando os autores à conclusão de que a avaliação da adequação funcional deve ser realizada independentemente da tecnologia, da linguagem de programação e do produto de software analisado [10]. Durante a realização do estudo de caso eles tiveram que lidar com especificação de requisitos funcionais e não-funcionais, testes funcionais, código fonte e execução do software. Adicionalmente, eles observaram a relação entre a lógica do negócio, o produto do software e medições da adequação funcional.

No mesmo sentido, o trabalho de German-Sallo e coautores (2016) adota a avaliação da adequação funcional em um conjunto de laboratórios remotos, almejando encontrar parâmetros mensuráveis a fim de possuir uma ferramenta de análise da adequação funcional [13]. Este estudo avaliou a completude, a corretude e a adequação funcional propriamente dita, utilizando grupos de especialistas para avaliar os laboratórios em diferentes universidades. Foram definidas questões que possibilitaram medir e quantificar a adequação funcional dos laboratórios [13]. Os autores concluíram que todos os laboratórios avaliados, funcionavam corretamente e satisfizeram os parâmetros da adequação funcional.

Ali Idri et al. (2017) publicou um trabalho que estudou a qualidade externa de aplicações móveis, baseado no padrão de qualidade da ISO/IEC 25010, sendo elas, prontuários pessoais de monitoramento de gravidez [14]. O estudo levou em consideração a avaliação da qualidade externa como indicador para a qualidade de uso [ISO]. A adequação funcional foi avaliada através de um questionário que cobre as subcaracterísticas, a fim de identificar o grau em que as aplicações cumprem as características de qualidade externa [14]. Como conclusão, os autores verificaram que a adequação funcional e confiabilidade são as características com maior cobertura pelos aplicativos avaliados.

Um trabalho que também lida com este tema é o de Ajeng Puspaningrum e colaboradores (2017), que avaliou a adequação funcional de um Sistema de Informação Acadêmica (AIS) de algumas universidades, com uma abordagem orientada a objetivos. As subcaracterísticas da adequação funcional foram definidas, bem como todas as funções do AIS. As métricas e o método de medição foram definidos, levando em consideração a disponibilidade funcional, a completude, adequação funcional, enquanto subcaracterística, satisfação do usuário, a precisão, exatidão e consistência dos dados, adequação aos regulamentos acadêmicos e a adequação funcional ao procedimento [11].

Os resultados mostraram que devido a implantação do método de avaliação, o valor de qualidade aumentou de acordo com a realidade da instituição [11]. Levando os autores a concluir que o resultado da medição pode ser levado em consideração para refinar e melhorar a qualidade dos sistemas de informação acadêmica e que método pode ser implantado em qualquer universidade, sendo necessário ajustar os itens de medição.

O presente trabalho contribui na medida em que a análise proposta não é feita em uma aplicação, mas através das entrevistas, para levantar informações referentes ao processo de avaliação dos produtos na indústria, ou seja, com empresas e seus produtos de software. Buscou-se levantar informações do processo de avaliação de qualidade, com foco na adequação funcional, para entender como é avaliada e se suas subcaracterísticas são contempladas. O trabalho também traz informações referente às ferramentas utilizadas, métricas e indicadores de qualidade adotados pelas

empresas, bem como os stakeholders alocados no processo de avaliação.

5 Considerações Finais

Esta pesquisa buscou compreender o estado da prática do processo de avaliação de adequação funcional em empresas de software, através de um survey instrumentalizado por entrevistas semi-estruturadas. A adequação funcional é um atributo que demanda observação e avaliação constante pela equipe de desenvolvimento de software, tendo em vista o contexto atual de variação de dispositivos, disponibilidades de sistemas operacionais e usuários com demandas diferentes. Apesar de acreditarmos ser relevante os achados, compreendemos a limitação do trabalho quando apenas 9 empresas participaram do estudo. Entretanto, ressalta-se que buscou-se entrevistar empresas de diferentes portes e regiões do país, comercializando distintos produtos de software, a fim de minimizar este viés da pesquisa.

Adicionalmente, é notório que levando em conta que alguns entrevistados iniciaram a entrevista sem conhecer o conceito de adequação funcional adotado neste estudo, as perguntas poderiam ser mal compreendidas, atrapalhando a construção do conhecimento sobre o tópico estudado. Porém, antes de iniciar qualquer pergunta sobre o tema específico, o entrevistador explicou o conceito e definiu as características desse atributo de qualidade, inclusive dando exemplos práticos - no início e durante a realização da entrevista - para auxiliar a compreensão do entrevistado.

Por outro lado, apesar de saber das limitações de generalização dos resultados, foi possível depreender que as empresas avaliam a adequação funcional observando apenas a completude e a corretude dos softwares. Essas análises são feitas através da realização de testes unitários, funcionais e de validação. Em suas práticas cotidianas, apesar de não terem, de maneira majoritária, um processo formalizado para avaliar especificamente a adequação funcional, a completude é avaliada através de *checklist* de disponibilidade das funcionalidades previstas e a corretude é verificada através de testes unitários e funcionais.

Entretanto, a adequação funcional propriamente dita é subentendida como atingida quando a completude e a corretude são aprovadas pela equipe de teste e/ou com a participação do cliente/usuário. Neste sentido, os testes de usabilidade acabam dando suporte à avaliação dessa característica, quando na verdade, eles deveriam acontecer depois da avaliação da adequação funcional propriamente dita, a ser realizada pela equipe de software. Esse fato pode ocorrer porque normalmente as empresas possuem um produto 'chave' no negócio, assim o contexto de uso e adoção está caracterizado desde a entrega da primeira versão do sistema. É importante lembrar que realizar testes de usabilidade sem antes avaliar a adequação funcional propriamente dita pode fazer com que os usuários passem por experiências frustrantes de uso, que poderiam ser suprimidas pelas avaliações realizadas antes de colocar o software para ser testado por usuários.

No que tange à medição e suporte ferramental, a falta de um processo sistematizado para de avaliação de qualidade acaba levando às empresas a não adoção de métricas específicas e ou ferramentas de suporte. A mais adotada, o Jira, que na verdade visa auxiliar o trabalho de gerenciamento do projeto, é utilizada nas empresas como *dashboard* a respeito de quais funcionalidades foram entregues (completude) e quais estão com bugs (corretude). Esclarece-se que como algumas empresas adotam o SonarQube, as métricas coletadas por ele acabam sendo observadas.

Como trabalhos futuros, há uma demanda clara de se realizar mais entrevistas para confirmar ou refutar os resultados encontrados aqui. Ademais, os autores pretendem aplicar questionários com empresas a fim de maximizar a coleta de dados. Por fim, esclarece-se que um mapeamento sistemático da literatura foi finalizado a fim de compreendermos como a literatura reporta os conceitos sobre a avaliação da adequação funcional.

Referências

[1] Ian Sommerville. 2019. Engenharia de Software (10ª ed). Tradução Luis Cláudio Queiroz; Revisão técnica Fabio Levy Siqueira - São Paulo: Pearson, 2019.

[2] Roger Pressman. 2011. Engenharia de Software: Uma Abordagem Profissional (7ª ed). Tradução Ariovaldo Griesi; Revisão técnica Reginaldo Arakaki, Julio Arakaki, Renato Manzan de Andrade. Porto Alegre : AMGH, 2011.

[3] ISO / IEC 25000. 2014. Systems and Software Engineering — Systems and Software Quality Requirements and Evaluation (SQuaRE) — Guide to SQuaRE, 2014.

[4] Armando Fox e David Patterson. 2014. Engineering Software as a Service: An Agile Approach Using Cloud Computing. Obtido em http://www.saasbook.info/.

[5] Claes Wohlin, Per Runeson, Martin Hst, Magnus C. Ohlsson, Bjrn Regnell e Anders Wessln. 2012. Experimentation in Software Engineering. Springer Publishing Company, Incorporated.

[6] Barbara A. Kitchenham e Shari Lawrence Pfleeger. 2002. Principles of Survey Research part 2: designing a survey. SIGSOFT Softw. Eng. Notes 27, 1 (January 2002), 18–20. https://doi.org/10.1145/566493.566495

[7] Klaas-Jan Stol, Paul Ralph e Brian Fitzgerald. 2016. Grounded theory in software engineering research: A critical review and guidelines. In 2016 IEEE ACM 38th International Conference on Software Engineering (ICSE), pages 120–131. ISSN: 1558-1225.

[8] Maria Salviati. 2017. Manual do Aplicativo Iramuteq. Obtido em http://www.iramuteq.org/documentation/fichiers/manual-do-aplicativo-iramuteq-par-maria-elisabeth-salviati

[9] SEBRAE. 2013. Anuário do trabalho na micro e pequena empresa. Obtido em

https://www.sebrae.com.br/Sebrae/Portal%20Sebrae/Anexos/Anuario%20do%20Trabalho%20Na%20Micro%20e%20Pequena%20Empresa_2013.pdf
[10]Moisés Rodríguez , Jesús Ramón Oviedo e Mario Piattini.2016. Evaluation of software product functional suitability: a case study. Software Quality Professional, vol. 18, n. 3, p. 18, 2016.
[11] Ajeng Puspaningrum, Siti Rochimah e Rizky Akbar. 2017. Functional Suitability Measurement using Goal-Oriented Approach based on ISO/IEC 25010 for Academics Information System. Journal of Information Systems Engineering and Business Intelligence 3.2 (2017): 68-74.
[12]Functional Testing vs. Usability Testing. Qualitest. Obtido em https://www.qualitestgroup.com/white-papers/functional-testing-vs-usability-testing/#:~:text=A%20simple%20way%20to%20think,and%20successfully%20use%20the%20product
[13] Zoltan German Sallo, Horatiu Stefan Grif e Adrian Gligor. "Functional suitability analysis of remote laboratories in engineering education." Procedia Technology 22 (2016): 1169-1175.
[14] Ali Idri, Mariam Bachiri, José Luis Fernández-Aleman e Ambrosio Toval. 2017. ISO/IEC 25010 Based Evaluation of Free Mobile Personal Health Records for Pregnancy Monitoring, 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC), Turin, 2017, pp. 262-267, doi: 10.1109/COMPSAC.2017.159.