

Análise quali-quantitativa sobre a influência da diversidade na produtividade de equipes ágeis: um estudo na indústria

Taianne Valerie Alves Motta¹, Gláucia Braga e Silva¹

¹Instituto de Ciências Exatas e Tecnológicas - Universidade Federal de Viçosa (UFV)
- Florestal – MG – Brazil

{taianne.mota, glaucia}@ufv.br

Resumo. *Este artigo investiga a influência da diversidade na produtividade de equipes que utilizam metodologias ágeis e também na qualidade dos artefatos produzidos. Para isso, foi realizada uma pesquisa quali-quantitativa, no contexto da indústria, envolvendo quatro times com diferentes perfis de diversidade. O estudo utilizou dois métodos de coleta de dados: um questionário organizado em três dimensões e métricas aplicadas a dados de uma ferramenta de apoio ao desenvolvimento. Os resultados mostraram que os times mais diversos apresentaram melhores indicadores de produtividade e qualidade em seus produtos.*

Abstract. *This paper investigates the influence of diversity on the productivity of teams that use agile methodologies and also on their artifacts quality. A quali-quantitative research was applied with four teams from a software company with different diversity profiles. Data was collected from two methods: a questionnaire organized in three dimensions and some metrics applied to data from a software development tool. The results showed that the most diverse teams presented better productivity and quality indicators in their products.*

1. Introdução

Com a cultura ágil, amplia-se a satisfação dos clientes, uma vez que se aumenta as entregas de valor e se reduz as entregas de produtos finais mal sucedidos. A busca pela qualidade do produto tem dependência direta com a qualidade do processo envolvido e com o ambiente de trabalho que precisa ser um espaço livre, acolhedor e inclusivo, de forma a melhorar a produtividade e explorar o potencial criativo da equipe. Isso vai de encontro ao que defende o primeiro valor do Manifesto Ágil, definido por Kent Beck e outros dezoito programadores [Beck et al. 2001], segundo o qual indivíduos e a interação entre eles são mais importantes do que processos e ferramentas. A Engenharia de Software Ágil evidencia a importância de times diversos dentro do desenvolvimento de software, pois possui um conjunto de métodos e práticas baseados nos valores e princípios expressos no Manifesto Ágil [Silveira and Prikladnicki 2019], como o planejamento adaptativo e o desenvolvimento com melhoria contínua. Times diversos são múltiplos e plurais, compostos por pessoas com várias características que se diferenciam e acrescentam, nas resoluções de problemas, gerando um conjunto mais amplo de habilidades, necessidades e vivências.

A diversidade no contexto da Engenharia de Software (ES) tem sido objeto de estudo de vários trabalhos na literatura [Silveira and Prikladnicki 2019, Albusays et al. 2021, Rodríguez-Pérez et al. 2021], com um foco maior em questões de

gênero. No entanto, pouco se conhece sobre o impacto de outras diversidades no contexto do desenvolvimento de software [Rodríguez-Pérez et al. 2021]. Além disso, existe também uma lacuna no que compete ao impacto da diversidade no contexto das metodologias ágeis [Silveira and Prikładnicki 2019].

Dessa forma, com o intuito de contribuir com novos estudos sobre a influência da diversidade, em especial, no contexto da Engenharia de Software Ágil, este trabalho será guiado pela investigação das seguintes questões de pesquisa:

- QP1: Existe alguma influência na produtividade em função do índice de diversidade da equipe?
- QP2: Existe alguma influência na qualidade do produto em função do índice de diversidade da equipe?
- QP3: Existe alguma influência da diversidade na presença de *Soft Skills* na equipe?

Este trabalho propõe um estudo quali-quantitativo sobre a influência da diversidade em times ágeis, por meio de uma pesquisa com desenvolvedores e da aplicação de métricas de produtividade e de qualidade de software no contexto de projetos reais da indústria. O estudo será conduzido em uma empresa de TI e contará com a participação de quatro equipes com diferentes perfis de diversidade, as quais adotam metodologias ágeis em seus desenvolvimentos. Para medir o grau de diversidade das equipes avaliadas, o estudo irá contar com o índice de heterogeneidade de Blau, que é a medida mais comumente utilizada na literatura, para diversidade variedade.

O restante deste artigo está organizado de acordo com as seções: seção 2, que apresenta conceitos sobre Metodologias Ágeis e Diversidade; seção 3, com os trabalhos relacionados; seção 4, em que se abordam materiais e métodos; seção 5, com os resultados obtidos; seção 6, com as ameaças a validade; e a seção 7, que traz as considerações finais.

2. Metodologias Ágeis e Diversidade

As metodologias ágeis surgiram para alinhar os valores e princípios do Manifesto Ágil [Beck et al. 2001], que sugere quatro conceitos fundamentais, sendo eles:

1. “Indivíduos e interação entre eles mais que processos e ferramentas.”
2. “Software em funcionamento mais que documentação abrangente.”
3. “Colaboração do cliente mais que negociação de contratos.”
4. “Responder a mudanças mais que seguir um plano.”

O primeiro conceito das Metodologias Ágeis demonstra o entendimento de que pessoas são as responsáveis por realizar o desenvolvimento de software e, portanto, a qualidade da interação entre elas pode resolver problemas crônicos de comunicação. A exemplo de que a comunicação é necessária para uma implementação bem-sucedida, e a criação de um entendimento compartilhado entre as partes interessadas constitui um fator importante para o sucesso do projeto [Diegmann and Rosenkranz 2017].

Na Engenharia de Software, a habilidade de comunicação é considerada uma *soft skill* imprescindível em um profissional da área. Essa se torna especialmente importante no contexto de algumas práticas do desenvolvimento ágil, com destaque para as atividades de requisitos, programação em pares (*Pair Programming*) e revisão em pares

(*Pair Review*). No que diz respeito à atividade de levantamento de requisitos, a habilidade de comunicação em conjunto com um time mais diverso aumenta a qualidade das especificações, pois o time consegue entender e perceber as dificuldades e necessidades dos usuários mais de perto, [Barbosa et al. 2021].

Quando se tem um foco em pessoas e não em processos ou algoritmos, o ambiente de trabalho fica mais humanizado criando um contexto integrativo necessário para que a diversidade se torne o passo essencial, abrangendo e incluindo todas as pessoas. Por isso que ao se considerar indivíduos e interações, o desempenho do trabalho deve ser feito em conjunto, abarcando as diferentes ideias do time, as pluralidades e diversidades, sejam elas culturais, físicas ou de gênero.

No contexto da Engenharia de Software, muito se fala sobre a importância da diversidade. Estudos mostram que equipes de software mais diversas podem aproveitar de informações mais amplas e possuem experiências e habilidades mais aprimoradas para resolução de problemas, tornando-se, portanto, mais eficazes [Vasilescu et al. 2015]. No entanto, em geral, as discussões têm se afunilado muito para questões de gênero, uma vez que o ambiente de TI é majoritariamente masculino. Entretanto, se faz necessário perceber o problema de maneira ampla, abordando outros aspectos sociais e outras características de diversidade que são igualmente importantes no contexto dos times [Rodríguez-Pérez et al. 2021].

3. Trabalhos Relacionados

González-Morales et al. [Gonzalez-Morales et al. 2011] apresentaram experiências de cinco anos que enfatizam a necessidade crítica de combinar, na Ciência da Computação, as *hards skills* com as *soft skills*. Em colaboração com empresas, o estudo utilizou uma aprendizagem baseada em projetos, com o intuito de ajudar alunos a adquirir as novas habilidades necessárias.

No que compete à avaliação da produtividade de equipes ágeis, Melo et al. [Melo et al. 2011] acompanharam duas equipes para saber quais os conceitos e fatores de produtividade são utilizados em seus projetos, chegando à conclusão de que alguns princípios tradicionais de produtividade ainda impactam os times de desenvolvimento, mesmo com a adoção das práticas ágeis. Através de revisões literárias, alguns fatores de produtividade, como reuso e características de software, foram levantados. O trabalho também destacou a experiência e a capacidade de cumprir prazos como fatores pessoais de cada membro da equipe, e como fatores de projeto, comunicação. Em seu estudo Hernández et al. [Hernández et al. 2019], faz uma revisão de literatura sobre métricas que avaliam a produtividade no contexto de desenvolvimento ágil de software e discute sobre como as medidas são utilizadas majoritariamente de forma quantitativa.

O presente artigo pretende contribuir, ampliando a base de estudos sobre a relevância da diversidade em equipes que utilizam metodologias ágeis, no que compete aos possíveis ganhos em termos de produtividade e qualidade do produto final. Ademais, essa pesquisa busca mostrar a importância de *softs skills*, como a habilidade de comunicação, para se melhorar a qualidade do processo e do produto final, trazendo para um contexto real os estudos de [Hernández et al. 2019].

4. Materiais e Métodos

Este estudo foi guiado por dois métodos de coleta de dados: um formulário para análise quali-quantitativa; e a extração de dados de uma ferramenta de apoio ao desenvolvimento, para aplicação de métricas de produtividade e qualidade. Os dados foram coletados no contexto de uma empresa de soluções tecnológicas e envolveu times que aplicassem Scrum¹ em seus projetos de software. Os times foram escolhidos arbitrariamente. Ao todo, quatro times participaram da pesquisa, nomeados aqui como: “Time A”, “Time B”, “Time C” e “Time D”.

Para se calcular o índice de diversidade dos times, foi utilizada a métrica conhecida como Índice de Blau [Collins 1979]. O índice analisa a medida da diversidade de espécies em um ecossistema com valores de zero (população homogênea) a um máximo indicado pela quantidade de categorias. Neste estudo, foram consideradas as categorias: identidade de gênero, raça, Pessoas com Deficiência (PCD), orientação sexual e Estado.

Os resultados dependem do cálculo: $1 - (\sum_{x=0}^N p_i^2)$

Onde p_i é a porcentagem de indivíduos dentro da categoria e N é o número de categorias. O valor máximo de cada índice segue o seguinte padrão: 2 categorias com 35% = 0,65, 3 categorias com 30% = 0,70; 4 categorias com 25% = 0,75; 5 categorias com 20% = 0,8; E assim, sucessivamente.

4.1. Pesquisa com desenvolvedores

Para conhecer o perfil de diversidade dos times e também sua percepção quanto às *soft skills* e às medidas de produtividade e de qualidade do produto em seus projetos, foi elaborado um questionário, organizado em três dimensões:

1. **Perfil de Diversidade**, com perguntas sobre a identidade do respondente, incluindo sua identificação de gênero, orientação sexual, idade, raça, origem e se é pessoa com deficiência (PCD). Além de conter perguntas sobre sua posição/cargo dentro da empresa.
2. **Perfil Técnico**, com o objetivo de perceber como o respondente enxergava o contexto do seu time, incluindo prazos de *Sprints*, quantidades de bugs/débitos técnicos, retrabalhos, entre outros.
3. **Soft Skills**, com o objetivo de avaliar a qualidade das comunicações no contexto dos times ágeis e como o respondente considera sua percepção do time, realizando boas leituras emocionais e de impedimentos.

Os formulários foram elaborados no Google Forms² e continham um total de 28 perguntas quali-quantitativas, sendo apenas uma aberta. As perguntas foram distribuídas em seções, sendo uma para cada dimensão: 11 perguntas, na seção de “Perfil de Diversidade”, 10 perguntas, na seção de “Perfil Técnico”, e 7 na seção de “Soft Skill”. As perguntas foram elaboradas com base nos trabalhos de Melo et al. [Melo et al. 2011], Silveira e Prikladnicki [Silveira and Prikladnicki 2018] e Kitchenham et al. [Kitchenham et al. 2009]. A maioria das perguntas emprega a escala de Likert utilizando 5 pontos, para resultados médios. A escolha dessa escala considerou o estudo de

¹<https://www.scrum.org/>

²<https://www.google.com/intl/pt-BR/forms/about/>

K. Vieira e M. Dalmoro [Kelmara M. Vieira 2008], que mostra a importância da mesma quando pondera escalas com formatos iguais, contínuo de direção e com a opção de neutro, necessária para a confortabilidade dos respondentes. O mesmo formulário foi aplicado para os diferentes times participantes de forma individual, com a finalidade de melhor organizar as respostas e cruzá-las com as métricas de cada time separadamente. A aplicação foi realizada remotamente, com cada respondente utilizando o próprio computador. O prazo para as respostas desses formulários foi de duas a três semanas e, no total, 20 indivíduos participaram.

4.2. Métricas de Produtividade e de Qualidade

Para avaliar a produtividade e a qualidade dos artefatos produzidos, foram utilizadas métricas que levaram em consideração os contextos dos times participantes e seus respectivos ambientes [Cesar 2017]. As métricas foram, em sua maioria, coletadas individualmente pelos times no ambiente da plataforma Microsoft Azure, o Visual Studio (VSTS)³. Nesse ambiente, é possível desenvolver e monitorar os produtos elaborados, em nuvem, localmente e, posteriormente, colocá-los em produção. Cada time escolheu um integrante para a retirada das métricas na ferramenta. Para isso, foram utilizadas consultas na plataforma, que selecionavam os dados separadamente, dentro dos seus contextos.

A produtividade de cada time foi avaliada usando-se três métricas aplicadas sobre os dados do VSTS:

- a. MP1 - Média de features por *Sprint*, calculada pela relação do número de trabalhos realizados (*features*) pelo número de *Sprints*;
- b. MP2 - Média de *user stories* por *Sprint*, calculada pela relação do número de *user stories* pelo número de *Sprints*; e
- c. MP3 - Percentual de tarefas cumpridas no prazo, calculada pela relação de tarefas que respeitaram as datas de entrega (*target dates*) dentre todas as tarefas previstas.

Para analisar a qualidade, outras duas métricas foram aplicadas no contexto dos times participantes e também de acordo com as possibilidades do VSTS:

- a. MQ1- Média de bugs por *Sprint*, calculada pela relação do número de bugs pelo número de *Sprints*;
- b. MQ2 - Percentual de débitos técnicos, calculada pela relação do número de débitos técnicos registrados por features.

5. Análise e Discussão dos Resultados

Esta seção apresenta os resultados com a coleta e análise de dados, de forma a responder às três questões de pesquisa endereçadas neste trabalho. As análises foram realizadas cruzando-se, sempre que possível, os dados do formulário aplicado aos desenvolvedores com os resultados das métricas de produtividade (MPs) e de qualidade (MQs).

Participaram do estudo, quatro times compostos por pessoas com diferentes níveis de diversidade, sendo seis pessoas no Time A (5 homens e 1 mulher); três pessoas (1 mulher e 2 homens) no Time B, três no time C (todos homens); e oito pessoas no time D (2 mulheres e 6 homens). Para se conhecer o perfil de diversidade de cada time, foram identificadas as categorias de diversidade presentes e contabilizados os respectivos indivíduos em cada uma delas, segundo as respostas da primeira dimensão do formulário (Tabela 1).

³<https://azure.microsoft.com/>

Tabela 1. Indivíduos por Categoria de Diversidade

Quantidade	A (6)	B (3)	C (3)	D (8)
Gênero	2	2	1	2
Raça	2	2	2	3
Orientação	3	2	2	3
PCD	1	1	1	2
Estado	5	2	2	4

Para cada categoria, calculou-se o índice de Blau correspondente (Figura 1). Importante destacar que quando um time apresenta índice de Blau igual a 0.0 em uma determinada categoria, indica ausência de diversidade para aquela categoria.

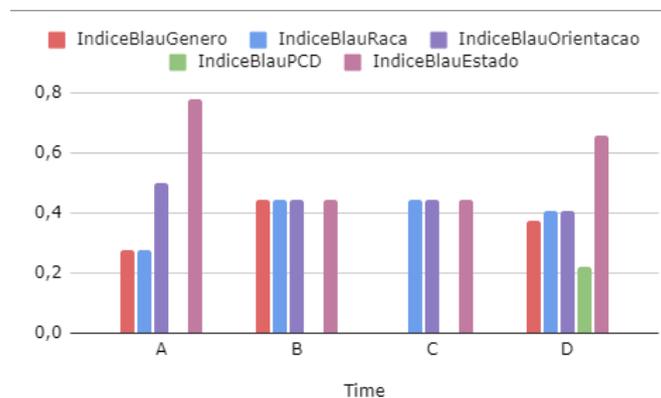


Figura 1. Índices de Blau calculados para cada time e por categorias.

Dentre os resultados mostrados no gráfico da Figura 1, é possível observar que todos os times participantes se mostram relativamente diversos, o que pode ser decorrente de políticas inclusivas e de diversidade da empresa envolvida. Os times com maiores índices para a categoria Estados, foram os times A e D. Os times B e C possuem médias parecidas, embora B contenha diversidades de gênero e C não.

Os times que podem ser considerados em relação à média como mais diversos são os Times B e D, possuindo níveis de diversidade próximos ao valor máximo. Além disso, o time D foi o único com diversidade para PCD, diferencial importante quando se considera a construção de sistemas mais inclusivos.

5.1. QP1: A produtividade varia em função do índice de diversidade da equipe?

Procurando responder à essa questão de pesquisa, foi construído um gráfico que une os resultados das três métricas de produtividades (MPs), dentro do contexto dos 4 times considerados, conforme mostra a Figura 2. Importante destacar que as 3 métricas analisam produtividade em função das *Sprints*, sendo maiores resultados um melhor cenário.

Ressalta-se que o Time A não disponibilizou os cadastros das features do seu time, não sendo considerado na avaliação da MP1 e MP3, uma vez que as datas de entregas são utilizadas apenas nas features cadastradas dentro do VSTS.

ser decorrente do fato de muitos não terem uma visão muito clara do que é a diversidade em si, valorizando mais as habilidades técnicas, ainda que trabalhem em um ambiente diverso. Já os respondentes que reconhecem essa relação, fizeram associações de como discussões sob diferentes visões podem mitigar riscos e encontrar soluções diversas para os problemas do desenvolvimento.

5.2. QP2: Existe alguma influência na qualidade do produto em função do índice de diversidade da equipe?

Para responder a essa questão de pesquisa, os resultados das duas métricas de qualidade propostas estão apresentados na Figura 4.

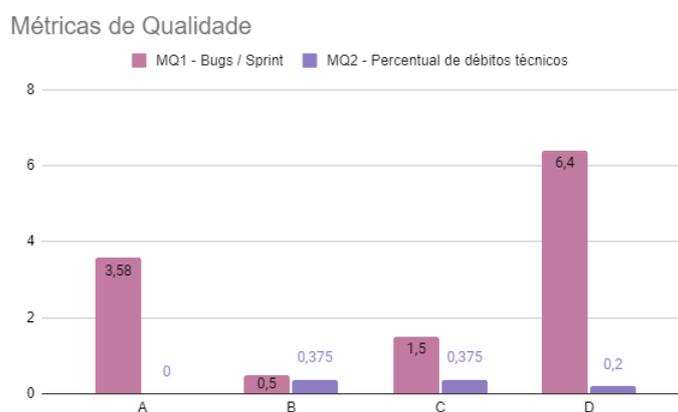


Figura 4. Resultados das métricas de qualidade para os 4 times.

Através da MQ1, foi possível observar que o número de bugs por *Sprint* dos times A e D se destaca em relação aos demais. O time B obteve a menor média de bugs por *Sprints* relatada, sendo 0.5 a cada ciclo da *Sprint*. Já o Time C, alcançou uma média de 1.4 bugs por *Sprint*. Em relação aos débitos técnicos (MQ2), todas as equipes alcançaram valores baixos, e o time A não obteve nenhum débito técnico. Nota-se que o time com a maior qualidade de produto é o time B, sendo um dos times mais diversos analisados no estudo. O time tem diversidade de gênero, orientação, cor/raça e estado.

Curiosamente, na visão dos participantes, segundo as respostas do formulário, a taxa de retrabalho obteve nota 4 (escala Likert de 0 a 5) em todos os cenários, mostrando que todos os times consideram ter baixa ou nenhuma taxa de retrabalho. Quando se cruza essas respostas com os resultados das métricas de qualidade (Figura 4), observa-se que os integrantes dos times A e D parecer ter pouco conhecimento sobre a realidade de seus times, no que compete aos bugs identificados.

5.3. QP3: Existe alguma influência da diversidade na presença de *Soft Skills* na equipe?

Para avaliar a presença de *soft skills* na equipe e buscar encontrar uma relação com a diversidade, a terceira dimensão do formulário continha perguntas sobre aspectos como a capacidade de comunicação dos membros e a qualidade das interações.

As respostas mostram que todos os times consideraram ter boa interação com seus colegas de trabalho e saber compartilhar suas dificuldades e impedimentos nas reuniões

diárias de forma honesta. Ainda com bons resultados, na pergunta sobre as condições desfavoráveis ou desafiadoras, a maioria considera que consegue manter seu bom desempenho. Os times B e C consideram que conseguem perceber e ajudar se algum membro possui algum impedimento, com média 4, já os times A e D obtiveram uma média 5. Dentre as respostas sobre quais *soft skills* cada pessoa considera ter, destacaram-se: “Facilidade de trabalhar em equipe”, “Flexibilidade” e “Aprendizagem contínua”. Por fim, o time C, que apresentou o menor grau de diversidade (formado apenas por homens), também apresentou a menor confiança em relação às *soft skills* avaliadas.

A avaliação positiva dos times em relação à presença das *soft skills* pode ter alguma relação com o perfil de diversidade, uma vez que todos eles apresentaram em maior ou menor grau algum índice de diversidade em alguma das categorias consideradas. No entanto, alguns dos resultados obtidos para as métricas de produtividade (QP1) e qualidade (QP2) mostraram-se divergentes das percepções dos times nas dimensões **Perfil de Diversidade** e **Perfil Técnico** do formulário, indicando, por exemplo, uma falta de transparência no processo e até mesmo uma subutilização e/ou divulgação dos relatórios gerados pela ferramenta de apoio.

6. Ameaças à Validade

Dentre as ameaças à validade, destacam-se: o tamanho da amostra de dados utilizada e a participação de times de uma única empresa; limitações nas comparações entre os times, devido à ausência de times pouco/nada diversos, que pode ser explicado pelo estudo ter sido realizado no contexto de uma empresa com políticas inclusivas; e falta de uma análise estatística para se avaliar correlações entre a diversidade dos times e os índices de produtividade e qualidade.

7. Considerações Finais

A falta de times diversos no contexto de ES é uma realidade que deve ser alterada. O objetivo do estudo foi coletar e analisar dados de 4 times com diferentes níveis de diversidade, aplicando métricas no contexto real desses times e um formulário. Após as análises desenvolvidas nas questões de pesquisa do presente estudo, foi possível perceber, na QP1, que a produtividade variou em função do índice de diversidade da equipe, uma vez que os times que mais se destacaram foram também os mais diversos. De forma similar, em relação à QP2, os resultados mostraram que o time com maior destaque, em termos de qualidade, tinha o maior grau de diversidade. Com relação a QP3, observou-se resultados positivos e similares para todos os times, o que pode ser explicado pelos bons índices de diversidade de todos. Contudo, observou-se também divergências entre as percepções dos times e alguns dos resultados para as métricas de produtividade e qualidade. Para trabalhos futuros, é necessário que haja uma correlação entre o índice de diversidade e os indicadores de produtividade e qualidade. Além disso, é importante considerar nas análises, as características do projeto e a complexidade do produto em desenvolvimento.

Referências

Albusays, K., Bjorn, P., Dabbish, L., Ford, D., Murphy-Hill, E., Serebrenik, A., and Storey, M. (2021). The diversity crisis in software development. *IEEE Software*, 38(02):19–25.

- Barbosa, M., Brilhante, I., Andrade, A., Conceição, J. L., Gomes, G., Conte, T., and Gadelha, B. (2021). Diversidade de gênero elicitación de requisitos: Uso do gender-mag como estratégia de identificação de requisitos de usabilidade. In *Workshop On Requirements Engineering*, volume 24.
- Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., et al. (2001). Manifesto for agile software development.
- Cesar, E. (2017). *Fatores de influência na produtividade de desenvolvedores de organizações de software*. PhD thesis, Instituto de Computação / UFAM.
- Collins, R. (1979). Inequality and heterogeneity: A primitive theory of social structure. by peter m. blau. new york: Free press, 1977. *Social Forces*, 58(2):677–683.
- Diegmann, P. and Rosenkranz, C. (2017). Team diversity and performance-how agile practices and psychological safety interact. In *International Conference on Information Systems (ICIS)*.
- Gonzalez-Morales, D., Moreno, L., and Roda-García, J.-L. (2011). Teaching “soft” skills in software engineering. In *2011 IEEE Global Engineering Education Conference (EDUCON)*, pages 630 – 637.
- Hernández, G., Martinez, , Jimenez, R., and Jimenez, F. (2019). Productivity metrics for an agile software development team: A systematic review. *TecnoLógicas*, 22:63 – 81.
- Kelmara M. Vieira, M. D. (2008). Dilemas na construção de escalas tipo likert: o número de itens e a disposição influenciam nos resultados? v. 6 n. 3 (2013): *Edição Especial Epistemologia e Métodos de Pesquisa em Administração e Contabilidade*.
- Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., and Linkman, S. (2009). Systematic literature reviews in software engineering – a systematic literature review. *Information and software technology*, 51(1):7–15.
- Melo, C., Cruzes, D. S., Kon, F., and Conradi, R. (2011). Agile team perceptions of productivity factors. In *2011 Agile Conference*, pages 57–66.
- Rodríguez-Pérez, G., Nadri, R., and Nagappan, M. (2021). Perceived diversity in software engineering: a systematic literature review. *Empirical Software Engineering*, 26(5):1–38.
- Silveira, K. K. and Prikladnicki, R. (2018). Perceptions on diversity in brazilian agile software development teams: A survey. In *Proceedings of the 1st International Workshop on Gender Equality in Software Engineering*, GE '18, page 37–40, New York, NY, USA. Association for Computing Machinery.
- Silveira, K. K. and Prikladnicki, R. (2019). A systematic mapping study of diversity in software engineering: a perspective from the agile methodologies. In *Proceedings of the 12th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)*, CHASE '19, pages 7–10. IEEE, IEEE.
- Vasilescu, B., Posnett, D., Ray, B., van den Brand, M. G., Serebrenik, A., Devanbu, P., and Filkov, V. (2015). Gender and tenure diversity in github teams. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, page 3789–3798, New York, NY, USA. Association for Computing Machinery.