https://eventos.utfpr.edu.br//sicite/sicite2020

# *O uso de Crowdsourcing* em Revisões Sistemáticas da Literatura na Engenharia de Software.

# The use of Crowdsourcing in Systematic Literature Reviews in Software Engineering.

**Geovanne José Moro**
geovannemoro@alunos.utfpr.edu.br
Universidade Tecnológica Federal do Paraná, Cornélio Procópio, Paraná, Brasil

**Katia Romero Felizardo**
katiascannavino@utfpr.edu.br
Universidade Tecnológica Federal do Paraná, Cornélio Procópio, Paraná, Brasil

**Recebido:** 02 set. 2020.
**Aprovado:** 02 out. 2020.

**Direito autoral:** Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.

## RESUMO

Revisões Sistemáticas da Literatura (RSLs) estão sendo adotadas na área de Engenharia de Software (ES) para fornecer uma síntese de evidências. Entretanto, sua condução ainda exige muitas horas de dedicação dos pesquisadores envolvidos, sendo necessário explorar abordagens de condução mais rápidas e eficientes. Uma abordagem é o *Crowdsourcing* (CS), pois divide as tarefas entre um grande número de pesquisadores. **Objetivo**: identificar e resumir o corpo de conhecimento sobre CS para apoiar a condução de RSLs na ES. **Método**: inicialmente, foi realizado um Mapeamento Sistemático (MS) identificando os estudos envolvendo CS em RSLs na ES. Seguidamente, foi realizado uma coleta de dados com 39 pesquisadores de ES para identificar suas percepções ao realizar RSLs colaborativamente. **Resultados**: evidenciar agilização do processo de RSL; reduzir viés através de ampla participação; e expandir a experiência da equipe foram os benefícios mais potenciais associados ao uso de CS em RSL. Os principais desafios estavam associados ao controle de qualidade dos resultados. **Conclusões**: apesar dos desafios, o CS pode ser empregado com sucesso no contexto de RSL. Mais pesquisas empíricas são necessárias sobre como usar o CS na condução de RSL em ES e como minimizar os desafios identificados.

**PALAVRAS-CHAVE:** Crowdsourcing. RSL. Engenharia de Software.

## ABSTRACT

Systematic Literature Reviews (SLRs) have been adopted in Software Engineering (SE) to provide a summary of evidence. However, conducting an SLR remains time-consuming and expensive, especially with regard to the hours dedicated by researchers. It is necessary to explore faster and more efficient approaches to support SLR conduction. One approach is Crowdsourcing (CS), since it divides tasks between a large number of researchers. **Objective**: to identify and summarize the body of knowledge on CS to support the conduction of SLRs in SE. **Method**: initially, a Systematic Mapping (SM) was carried out, identifying studies involving CS in SLRs in SE. In the sequence, a survey was conducted with 39 SE researchers to identify their perception when performing SLR collaboratively. **Results**: speed up the SLR process; reduce bias through broad participation; and expand team's experience were the most potential benefits associated with the use of CS in SLR. The main challenges are associated with quality control of results. **Conclusions**: despite the challenges, CS can be successfully used in SLR context. More empirical researches are needed on how to use CS to conduct SLR in SE and how to minimize the identified challenges.

**KEYWORDS:** Crowdsourcing. SLR. Software Engineering.



4 EDUCAÇÃO DE QUALIDADE

## INTRODUÇÃO

A essência da Engenharia de Software Baseada em Evidências (ESBE) apoia-se na construção sólida do conhecimento através da coleta, interpretação e análise criteriosa de todos os dados empíricos disponíveis, atuais e relevantes, sobre um tópico ou fenômeno de interesse. Dessa forma, fornece conhecimento sobre quando, como e dentro de qual contexto tecnologias, processos, métodos ou ferramentas são apropriados para práticas da Engenharia de Software (ES) [1].

Nesse contexto, a Revisão Sistemática da Literatura (RSL) foi introduzida na ESBE em 2004, um dos métodos adotados para fornecer mecanismos para identificar e resumir evidências sobre o estado da arte de um tópico específico. Assim as RSLs ganharam importância substancial e foram aplicadas a vários tópicos de interesse tanto por pesquisadores como por profissionais [2].

Na última década, a maioria das pesquisas de ESBE enfatizaram vantagens e dificuldades do uso de RSLs [3-5] e apresentaram desafios e lições aprendidas [6]. Vários estudos constataram que a realização de uma RSL é trabalhosa e morosa, geralmente exigindo centenas de horas de contribuição dos revisores [7-9]. Essa quantidade de horas pode ser ainda maior devido ao crescimento exponencial da pesquisa científica nos últimos anos.

Dessa forma, pesquisadores de ES precisam explorar abordagens alternativas para conduzir RSL a custos menores. Abordagens para semi-automatizar as atividades do processo de RSL como, por exemplo, a busca por estudos candidatos, seleção de estudos relevantes, extração e sumarização de estudos, entre outros, têm sito propostas como um meio potencial de reduzir o tempo e a carga de trabalho [10-12]. Outra possibilidade é usar abordagens como o CS, foco deste projeto.

## FUNDAMENTAÇÃO TEÓRICA

**Estudos Secundários** – A ESBE foi proposta como um meio de aperfeiçoar a tomada de decisões relativas ao desenvolvimento e manutenção de software, por meio da integração de evidências de pesquisa com experiências práticas e valores humanos. O objetivo principal da ESBE é diminuir a lacuna existente entre a pesquisa e a prática na ES e, para tal, disponibiliza meios para construir um corpo de conhecimento sobre quando, como e quais processos, tecnologia, ferramentas, entre outros, são adequados para serem utilizados na ES [13]. Assim, pesquisadores da ESBE defendem o uso das RSLs para sumarizarem as melhores evidencias teóricas tanto para pesquisadores quanto para praticantes.

O processo de condução de um estudo secundário envolve três fases principais: planejamento, condução e publicação dos resultados [13]. Essas fases, bem como suas atividades, são conduzidas de modo iterativo. Durante o planejamento é definido o protocolo, que contém itens como as questões de pesquisa, a estratégia de busca, seleção e extração dos dados. Durante a condução os estudos são identificados, selecionados e os dados contidos nos estudos

incluídos são extraídos e sintetizados. A última fase do processo é a escrita dos resultados, que devem ser divulgados aos potenciais interessados [13].

**Crowdsourcing** – O termo CS foi cunhado pela primeira vez por Howe [14] em um artigo da revista *Wired* como: "O CS faz uso do mundo em rede. Nem sempre o trabalho é gratuito, mas custa muito menos do que pagar aos colaboradores tradicionais". De acordo com essa definição, o CS tem como pré-requisitos a força de trabalho em grande rede e indefinida juntamente com o formato de chamada aberta para as tarefas. Segundo Howe [14], o trabalho de CS pode ser feito por cooperação ou por indivíduos únicos.

A proliferação do uso do CS ocorreu em diversas áreas da Computação, sendo aplicado para executar o *design*, atividades de desenvolvimento e teste de software [15]. Analisando seu uso é possível identificar quatro características comuns que pertencem ao CS: acesso aberto na produção; flexibilidade na força de trabalho; livre arbítrio na participação; e benefícios mútuos entre as partes interessadas. Outros benefícios do CS destacam-se o fácil acesso a uma ampla gama de colaboradores, soluções diversas, menores taxas de mão de obra e menor tempo de execução das tarefas [15]. Desse modo o CS é uma forma emergente de terceirização, vista como um modelo *online* para produção distribuída e solução de problemas.

## O USO DO *CROWDSOUCING* EM RSL NA ES – UM MS

Uma pesquisa sistemática foi realizada em três fases para identificar a literatura atual sobre o uso de CS em RSL na ES.

**Fase 1: Planejamento –** Inicialmente o protocolo foi definido e as questões de pesquisa foram elencadas para esclarecer: "Q1 – Quando e onde foram publicados os estudos sobre o uso de CS na ES?"; "Q2 – Quais são os benefícios relacionados ao uso do CS em RSL na ES que foram relatados pelos autores?"; " Q3 – Quais são os desafios relatados pelos autores relacionados ao uso do CS em RSL na ES?"; e "Q4 – Quais são as atividades de RSL na ES relatadas pelos autores que podem ser apoiadas pelo CS?"

A *string de busca adotada foi: ("crowdsourcing"* **OR** *"citizen science"* ***OR*** *"widson of crowds"* **OR** *"collective action"* **OR** *"collective knowledge")* **AND** *("software engineering")*. Essa *string* foi aplicada na Scopus, uma das bases de dados mais utilizadas em Ciência da Computação [16].

Os critérios de seleção foram organizados em um critério de inclusão (CI) e dois critérios de exclusão (CE): **(CI1)** O estudo deve ser de ES e discutir o uso do CS no contexto de RSL; **(CE1)** O estudo não está no escopo de RSL, embora utilize CS em ES; e **(CE2)** O estudo está no contexto de RSL, no entanto, não usa CS no contexto de ES.

**Fase 2: Condução –** Três buscas: **–** Busca 1 (automática); **–** Busca 2 (manual); **–** Busca 3 (*Snowballing* para frente e para trás) foram conduzidas de forma complementares. A Busca 1 retornou 341 estudos. Foram considerados os estudos

publicados até setembro de 2019. Os critérios de seleção foram aplicados nos títulos, resumos e palavras-chave, resultando em 57 estudos. Na sequência, os critérios de seleção foram aplicados considerando os textos completos, resultando em um conjunto de zero estudos.

Definiu-se que a Busca 2 seria conduzida nas duas principais conferências de CS: (1) Conferencia da AAAI sobre Computação Humana e *Crowdsourcing* (HCOMP) no período de 2013/2018 e (2) *Crowdsourcing* e Computação Humana para Web – *World Wide Web Conference* (WWW) no período de 2012 até 2019. Com relação a Busca 2, no HCOMP, 255 estudos foram analisados e apenas um foi considerado relevante [17]. Dos 48 estudos encontrados na conferência da WWW, um foi considerado relevante [18]. As listas de publicações dos autores dos dois estudos foram analisadas individualmente, e mais um estudo foi encontrado [19].

Na sequência, os três estudos relevantes [17-19] selecionados para inclusão formaram o *seed set* para realizar iterações da Busca 3. Durante a primeira iteração, os três estudos que formaram o *seed set* inicial foram citados por 7 estudos candidatos à inclusão. Nenhuma dessas sete citações foram incluídas após aplicação dos mesmos critérios de seleção utilizados durante a busca automática. Os três artigos totalizaram 116 referências, dentre essas referências, uma foi incluída [20]. A iteração 2 começou com um estudo previamente selecionado na iteração 1 que foi citado por 14 estudos. Desses, um foi incluído [21]. Um total de 18 estudos compuseram as referências, mas nenhuma foi identificada como relevante.

Os achados das três buscas (Busca 1: 0 + Busca 2: 3 + Busca 3: 2), totalizaram 5 estudos [17-21]. No entanto, para fins de extração de dados, os estudos [17] [20] não foram considerados, pois eram versões mais antigas de outros estudos já considerados, [18] [21], respectivamente.

**Fase 3: Resultados –** A seguir são apresentados os resultados deste MS.

**Frequência de publicação (Q1):** a pesquisa de CS em RSL é recente: 2017 [21], 2018 [18], [19]. Analisando os meios de publicação, é possível afirmar que o tipo de publicação varia: um estudo foi publicado em conferência; um em serviço de impressão digital; e outro em periódico.

**Benefícios do CS em RSL (Q2):** os benefícios foram nomeados como B1 – B5 e são detalhados da seguinte forma: **B1:** eliminar a sobrecarga nas atividades de RSL; **B2:** reduzir o tempo de atividades de RSL – acelerar o processo de RSL; **B3:** minimizar o viés – um grupo maior tem maior probabilidade de reduzir o viés do que uma pequena equipe de pesquisadores; **B4:** obter especialistas no domínio da RSL – uma vasta gama de pesquisadores oferece uma enorme diversidade de experiências; e **B5:** aumentar a transparência, pois o processo e as atividades podem ser públicos, se desejado.

**Desafios do CS em RSL (Q3):** os desafios foram nomeados como D1 – D4 e são detalhados da seguinte forma: **D1:** selecionar colaboradores qualificados na multidão; **D2:** disponibilizar exemplos concretos na fase de treinamento para ajudar os colaboradores a entenderem o objetivo de uma atividade (fornecer um

rico conjunto de exemplos para os colaboradores); **D3:** garantir a qualidade das tarefas realizadas pela multidão; e **D4:** coordenar um processo de CS em RSL – sobrecarga adicional.

**Atividades RSLs apoiadas por CS (Q4):** dois estudos [18] [19] aplicaram o CS na atividade de seleção em 4 fases: treinamento, seleção de colaboradores, atividade de CS e agregação.

## O USO DO *CROWDSOURCING* EM RSL NA ES – UMA COLETA DE DADOS

Uma coleta de dados foi realizada para coletar a opinião de pesquisadores de ES sobre a adoção do CS em RSL [22]. O *coleta* compreendeu um questionário com 4 seções: 1) Perfil (7 perguntas); 2) Perspectiva técnica sobre CS para apoiar as atividades das RSL na ES (4 perguntas); 3) Percepção dos participantes sobre o uso do CS na RSL em ES (4 perguntas); e 4) Comentários finais (4 perguntas abertas). As opções possíveis de respostas para as questões das seções 1, 2 e 3 variaram de "concordo totalmente" a "discordo totalmente", em uma escala baseada no método *Likert Scale* [23], que é uma métrica usada em questionários como pesquisas de atitude). Uma versão *on-line* da pesquisa está disponível em: (https://forms.gle/UmKr4ryLqHTWe3Tz5).

Ao total 39 pesquisadores (9,9%), dos 391 convidados, responderam as questões. sendo: 20 (51,2%) docentes; 7 (17,9%) doutorandos; 4 (10,2%) mestrandos; 4 (10,2%) pesquisadores; e 2 parceiros da Indústria. 1 participante se classificou como Professor (2,5%) e outro como Pesquisador Sênior (2,5%).

Com relação a realização de RSLs, 1 participante (2,5%) realizou 20 RSLs, 3 (7,6%) realizaram 10, 4 (10,2%) entre 7 e 9, 15 (38,4%) entre 4 e 6 e o restante dos 16 (41,0%) participantes mencionaram entre 1 e 3 RSLs. O número total de RSLs realizadas pelos participantes foi de 182. Dessas 182 revisões, 127 (69,7%) foram publicadas em conferências ou periódicos. Outros 31 (79,4%) participantes têm experiência em tarefas de CS, incluindo testes de software, extração de dados, projeto, identificação de sinônimos específicos de domínio e computação para crianças.

Os participantes foram solicitados a indicar quais são os benefícios da aplicação do CS nas RSLs em ES. O benefício apontado pela maioria dos participantes (30 de 39 – 76,9%) foi reduzir o tempo de tarefas de RSL, acelerando o processo de condução da RSL. Já 58,9% (23 participantes) concordaram que o CS, se aplicado à RSL, pode reduzir o viés (um grupo maior tem maior probabilidade de reduzir o viés do que uma pequena equipe de pesquisadores).

Em relação aos principais desafios da aplicação do CS em RSL, a maior dificuldade destacada foi garantir a qualidade das tarefas realizadas pela multidão (30 – 76,9%); seguida pela seleção de colaboradores qualificados da multidão (26 – 66,6%); e coordenar um processo de CS em RSL – despesas gerais adicionais (22 – 56,4%).

A maioria dos participantes (30 – 76,9%) concordaram que a seleção de estudos e a extração de dados poderia ser apoiada pelo CS. A busca por estudos foi outra atividade que 28 participantes (71,7%) destacaram como possível a aplicação do CS.

## CONSIDERAÇÕES FINAIS

O principal objetivo deste estudo foi fornecer uma visão geral sobre CS para apoiar a condução de RSLs na ES. Até o momento, foram realizados um MS e uma coleta de dados com 39 pesquisadores da ES. Espera-se que os resultados deste estudo contribuam para o conhecimento existente sobre o uso do CS na RSL em ES. Em resumo, até a presente data, conclui-se que:

1. O conhecimento sobre o uso de CS no contexto de RSL na ES está concentrado em 5 estudos [17-21].

2. O CS no contexto de RSL é pouco explorado, poucos autores o investigaram, desta forma, não é claramente entendido pelos pesquisadores da RSL na ES.

3. Vários benefícios potenciais foram associados ao uso de CS na RSL, como acelerar o processo de RSL; reduzir o viés por meio de ampla participação; e expandir a gama de colaboradores proporciona uma enorme diversidade em suas experiências (diferentes especialistas no domínio da RSL).

4. Os principais desafios estão concentrados no controle de qualidade do processo para garantir a qualidade dos resultados. Aspectos diferentes podem afetar a qualidade, por exemplo: A descrição inadequada da tarefa pode resultar em uma tarefa entregue com baixa qualidade; A seleção inadequada de colaboradores para a tarefa; A correspondência inadequada entre colaborador e tarefa pode prejudicar a qualidade dos resultados. As tarefas precisam ser atribuídas aos colaboradores de acordo com sua experiência individual.

5. A falta de um processo de RSL baseado em multidões é outro desafio.

6. Não há plataformas para suportar especificamente tarefas de RSL.

Como trabalhos futuros destacam-se: implementar uma plataforma para suportar aplicativos de CS no contexto de RSL na ES; realizar estudos empíricos para demonstrar como o processo de RSL pode se beneficiar do CS; e avaliar a qualidade das tarefas de RSL baseadas na multidão.

## AGRADECIMENTOS

Os autores agradecem ao Programa de Bolsas de Iniciação Científica – PIBIC 2019/2020 (edital PROPPG – 02/2019), da UTFPR, pela bolsa de iniciação científica concedida ao aluno Geovanne José Moro, entre 08/2019 e 07/2020.

## REFERÊNCIAS

[1] KITCHENHAM, B. A.; BUDGEN, D.; BRERETON, P. **Evidence-based software engineering and systematic reviews**. CRC press, 2015.

[2] SKOGLUND, M.; RUNESON, P. Reference-based search strategies in systematic reviews. In: **13th International Conference on Evaluation and Assessment in Software Engineering (EASE) 13**. 2009. p. 1-10.

[3] RIAZ, M. et al. Experiences conducting systematic reviews from novices' perspective. In: **14th International Conference on Evaluation and Assessment in Software Engineering (EASE)**. 2010. p. 1-10.

[4] LAVALLEE, M.; ROBILLARD, P-N.; MIRSALARI, R. Performing systematic literature reviews with novices: An iterative approach. **IEEE Transactions on Education**, v. 57, n. 3, p. 175-181, 2013.

[5] KUHRMANN, M.; FERNÁNDEZ, D. M.; DANEVA, M. On the pragmatic design of literature studies in software engineering: an experience-based guideline. **Empirical software engineering**, v. 22, n. 6, p. 2852-2891, 2017.

[6] IMTIAZ, S. et al. A tertiary study: experiences of conducting systematic literature reviews in software engineering. In: **Proceedings of the 17th International Conference on Evaluation and Assessment in Software Engineering**. 2013. p. 177-182.

[7] MIAN, P. et al. A systematic review process for software engineering. In: **3rd Experimental Software Engineering Latin American Workshop (ESELAW)**. 2005. p. 1-10.

[8] BRERETON, P. et al. Lessons from applying the systematic literature review process within the software engineering domain. **Journal of systems and software**, v. 80, n. 4, p. 571-583, 2007.

[9] CARVER, J. C. et al. Identifying barriers to the systematic literature review process. In: **2013 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement**. IEEE, 2013. p. 203-212.

[10] MARSHALL, C.; KITCHENHAM, B. A.; BRERETON, P. Tool features to support systematic reviews in software engineering–a cross domain study. **e-Informatica Software Engineering Journal**, v. 12, n. 1, 2018.

[11] PULSIRI, N.; VATANANAN-THESENVITZ, R. Improving systematic literature review with automation and bibliometrics. In: **2018 Portland International Conference on Management of Engineering and Technology (PICMET)**. IEEE, 2018. p. 1-8.

[12] OSBORNE, F. et al. Reducing the effort for systematic reviews in software engineering. **Data Science**, v. 2, n. 1-2, p. 311-340, 2019.

[13] KITCHENHAM, B. A.; CHARTERS, S. Guidelines for performing systematic literature reviews in software engineering. 2007.

[14] HOWE, J. The rise of crowdsourcing. **Wired magazine**, v. 14, n. 6, p. 1-4, 2006.

[15] MAO, K. et al. A survey of the use of crowdsourcing in software engineering. **Journal of Systems and Software**, v. 126, p. 57-84, 2017.

[16] MAPLESDEN, D. et al. Performance analysis for object-oriented software: A systematic mapping. **IEEE Transactions on Software Engineering**, v. 41, n. 7, p. 691-710, 2015.

[17] KRIVOSHEEV, E. et al. Crowdsourcing paper screening in systematic literature reviews. **arXiv preprint arXiv:1709.05168**, 2017.

[18] KRIVOSHEEV, E.; CASATI, F.; BENATALLAH, B. Crowd-based multi-predicate screening of papers in literature reviews. In: **Proceedings of the 2018 World Wide Web Conference**. 2018. p. 55-64.

[19] RAMIREZ, J. et al. Crowdrev: a platform for crowd-based screening of literature reviews. **arXiv preprint arXiv:1805.12376**, 2018.

[20] WEISS, M. Crowdsourcing literature reviews in new domains. **Technology Innovation Management Review**, v. 6, n. 2, p. 5-14, 2016.

[21] WEISS, M.; ABUALHAOL, I.; AMIN, M. A leader-driven open collaboration platform for exploring new domains. In: **Proceedings of the 12th International Symposium on Open Collaboration**. 2016. p. 1-4.

[22] KITCHENHAM, B. A.; PFLEEGER, S. L. Personal opinion surveys. In: **Guide to advanced empirical software engineering**. Springer, London, 2008. p. 63-92.

[23] REISER, R. H. S. et al. Relation between polling and likert-scale approaches to eliciting membership degrees clarified by quantum computing. In: **2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)**. IEEE, 2013. p. 1-6.