# 1. Background

DevOps has become a trending technology term and gained popularity in the software industry and academia. It refers to improving the performance of software development operations by incorporating the Software Development (Dev) team and IT Operations (Ops) team in one process (Hussain, Clear, & MacDonell, 2017). Security is one of the major concerns that limit the adoption of DevOps (Mohan & Othmane, 2016). In such case, the terms SecDevOps and DevSecOps have been created as security-oriented variants of DevOps. The common definition of these terms is the integration of security practices in DevOps processes to effectively reduce risks and address security issues, by promoting the collaboration amongst security team, development team and operations team (Zaydi & Nassereddine, 2019).

Some studies expound on DevSecOps with its perceptions, benefits, challenges, practices, tools, and metrics. Mohan and Othmane (2016) presented a paper which surveyed the literature from both academia and industry to identify SecDevOps. Wilde et al. (2016) reviewed the literature and conducted three case studies to discuss the security aspect in DevOps practices. Zaydi and Nassereddine (2019) presented five case studies of five Middle East and North Africa (MENA) organizations adopting DevSecOps in IT Service Management (ITSM) works, by conducting interviews and observation, and their results showed that DevSecOps practices really improved the performance of ITSM.

Some studies have proven that DevOps strategy can be adopted in globally distributed software projects. (Stray, Moe, & Aasheim, 2019). Gupta et al. (2019) presented their experience in a DevOps team distributed geographically across three countries (India, the USA, and Germany) that successfully established continuous delivery and short release cycles with agile scrum. Hussain et al. (2017) investigated online job advertisements and combined with interviews, to identify the required Knowledge, skills and capabilities for the DevOps roles in a New Zealand perspective, and they found that the global dimensions of DevOps roles were apparent in most job ads sometimes by explicit mention (16% job postings explicitly mentioned global aspects) but more often by implication. However, there are few studies related to DevSecOps in the GSE context. This has been one of the motivations to carry out this research for bridging the gap between DevSecOps and GSE.

# 2. Research Questions

We will review white and grey literature to identify recent research and practical trends of DevSecOps, aiming to: (a) observe, document and analyze the state of art of DevSecOps; and (b) investigate the application of DevSecOps in the Global Software Engineering (GSE) context. Regarding the research objectives, research questions were posed

➢ RQ1: What is the state of art of DevSecOps in the existing (white and grey) literature?

- Sub-question 1.1: What aspects of DevSecOps can be found in the existing (white and grey) literature?

- Sub-question 1.2: What themes do these aspects contain?

- Sub-question 1.3: How do the themes link to each other?

➢ RQ2. How does Global Software Engineering (GSE) relate to DevSecOps?

# 3. Research Methodology

A Systematic Literature Review (SLR) is a form of secondary study to identify, evaluate, analyze and interpret all of the possible existing literature relevant to particular research questions, and to synthesize these available researches in a fair manner (Kitchenham & Charters, 2007). However, a normal SLR mainly uses formally published literature (e.g., journal and conference papers) and is not quite adequate for this research, because we find that there are only a limited number of academic papers available relevant to this topic, after a quick pre-searching process. Therefore, a Multi-vocal Literature Review (MLR) needs to be conducted. MLR is a special form of SLR which does not only use formally published literature (called White Literature) but also includes the Grey Literature (any unpublished work such as technical reports, news, websites, blog posts, white papers, speeches, videos, etc) (Garousi, Felderer, & Mäntylä, 2019). Another important reason for conducting MLR is that the investigation of DevOps should contain both of the researcher-oriented and practitioner-oriented sources. Software engineering practitioners outside of academia constantly produce all kinds of grey literature based on their practical experience. Researchers should not ignore these valuable sources of knowledge and information during the process of literature review.

In this case, a protocol for a multi-vocal literature review of DevOps and its security has been developed, based on the guidelines for performing SLR in SE (Kitchenham & Charters, 2007) and guidelines for conducting MLR in SE (Garousi, Felderer, & Mäntylä, 2019).

# 4. Search Strategy

An exhaustive search strategy for the multi-vocal literature review is presented in this section, after having consulted the guidelines for both SLR and MLR in SE. Some subsections will describe separately, if there are differences between searching white literature and grey literature.

## 4.1. Source to be searched

### 4.1.1 Source to be searched for White Literature

- Automatic searching in well-known digital databases: ACM digital library, IEEE Xplore and Scopus.

- Searching in Google Scholar and the digital library of AUT University.

- Snowballing technique will be conducted on selected literature if necessary, so that more relevant studies can be included.

### 4.1.2    Source to be searched for Grey Literature

- General web searching engine i.e. Google.

- Source with high credibility, such as books, magazines, specialized databases, government reports, white papers, method creators and consultants' websites and case studies.

- Source with medium credibility, such as technical reports, news, Q/A sites (like StackOverflow), Wiki articles, blog posts, presentations and videos.

- Contacting individuals or organizations for un-published work or specialized databases of theses, via multiple methods, such as direct requests, emails and social media.

- Like snowballing in WL searching, backlinks can be navigated either forward or backward to find more relevant information.

## 4.2.  Search Strings

To address RQ1, Search String 1 = *"DevOps" AND ("security" OR "secure" OR "safe") OR "SecDevOps" OR "DevSecOps"*. After applying Search String 1 in all search sources, the results did not include any studies involving global DevSecOps. To address RQ2, we used an additional Search String 2 = *"DevOps" AND ("security" OR "secure" OR "safe") AND ("global software engineering" OR "global software development" OR "GSE" OR "GSD" OR "globally distributed software*") OR "SecDevOps" OR "DevSecOps"*. Search strings might vary according to the differences between databases, because of their acceptability of Boolean operators.

In addition, some limitations need to be preset on the searches. Strings will be searched within Metadata (title, abstract and keywords); the written language should be set to be English; the publication year should be between 2011 and 2021; and books (chapters), short papers, posters and abstracts would be excluded. All results sort by relevance so that searches can finish when the relevance is extremely weak. Based on the search results, the research questions and search terms would be refined; further, search strings would possibly be re-formulated. Meanwhile, a list of key papers will be generated to ensure the reliability and relevancy of search results (Jalali & Wohlin, 2010).

## 4.3. Study Selection

### 4.3.1. Study selection criteria

Study selection criteria are defined to ensure that selected studies provide direct evidence about the research questions, including both inclusion and exclusion criteria (Kitchenham & Charters, 2007). There are only selection criteria for White Literature being listed here, because in practice, inclusion and exclusion criteria for Grey Literature usually overlap and can be integrated with study quality assessment (Section 4.4) (Garousi, Felderer, & Mäntylä, 2019). Also, a large piece of GL has no accurate information so that selection should be made according to specific circumstance. Hence, the following inclusion and exclusion criteria apply to all WL, and have some reference value for selecting GL, although not entirely appropriate.

- Inclusion criteria:

  *a) The study fully or partially answers the research questions;*

  *b) The study is written in English;*

  *c) The study is published between 2012 and 2021;*

  *d) The study has a clearly stated methodology/research design;*

  *e) The study has credible source;*

- Exclusion criteria:

  *a) The study does not have a full-text;*

  *b) The study is external to the subject area of computer science and engineering;*

  *c) The study does not have a rigorous research method to prove the correctness of findings;*

  *d) Duplicate studies.*

### 4.3.2. Study selection process

- Selection process for White Literature:

  *a) Firstly, the defined search strings are applied to search full-text papers in digital databases, e.g. ACM digital library, IEEE Xplore and Scopus.*

  *b) Second, inclusion and exclusion criteria are applied to select papers quickly, based on their titles and abstracts.*

  *c) Duplicates will be removed.*

  *d) The full-texts of papers need to be read, if difficult to determine inclusion or exclusion based on titles and abstracts.*

*e) Snowballing technique can be performed based on the reference lists of included papers, so that more relevant papers can be found.*

*f) Finally, stop searching and selecting studies, when data exhaustion.*

- Selection process for Grey Literature:

*a) Firstly, web searching engine (Google) is used to search defined strings, then some relevant information (websites, news, white books, blogs, etc) will be found.*

*b) Second, all relevant information need be filtered and determined inclusion or exclusion, by browsing and reading them in detail.*

*c) Duplicates will be removed.*

*d) Backlinks can be navigated either forward or backward to search more relevant information.*

*e) Individuals or organizations of GL studies can be contacted for un-published work or specialized databases, if necessary.*

*f) Finally, stop searching and selecting studies, when theoretical saturation, effort bounded or evidence exhaustion.*

- After performing the selection process for WL and GL, all selected WL studies and GL studies will be combined together and ready for further processing.

## 4.4. Study Quality Assessment

Study quality assessment is a necessary procedure to provide more detailed inclusion/exclusion criteria; to determine the valid extent of sources; to assess importance of studies; and to minimize bias (Kitchenham & Charters, 2007). The table shows the study quality assessment checklist adapted from MLR guidelines of Garousi et al. (2019). Garousi et al. (2019) only presented quality assessment checklist of grey literature, here is an extension including both white and grey literature. The first 14 questions were grouped into 6 categories and would be answered YES/NO, so the criteria would be marked 0/1. "Literature Type" would be marked on a scale from 0 to 4. Out of 18 points (14+4), the borderline was set as 11 (60% of 18).

| Criteria | Questions |
|---|---|
| Authority of the producer (Measure = 0 or 1) | • Is the author or the publishing organization reputable? <br> • Has the author published other work in the field? <br> • Does the author have expertise in the area? |
| Methodology (Measure = 0 or 1) | • Does the work have a clearly stated aim? <br> • Does the work have a stated methodology? <br> • Does work have authoritative and contemporary references? <br> • Are any limits clearly stated? |
| Objectivity | • Does the work provide objective statements or credible findings? |

| | |
|---|---|
| (Measure = 0 or 1) | • Is there vested interest? E.g., a tool comparison by authors that are working for particular tool vendor.<br>• Is the conclusion supported by the data? |
| Publication Date<br>(Measure = 0 or 1) | • Does the work have a clearly stated date? |
| Novelty<br>(Measure = 0 or 1) | • Does the work have a novel idea or something unique?<br>• Does the work strengthen or refute a current position? |
| Impact<br>(Measure = 0 or 1) | • For WL, is the author's work cited often? / For GL, is the source viewed/shared/discussed often? |
| Literature Type<br>(Measure = 0 to 4) | • WL: peer-reviewed academic papers (Measure = 4).<br>• WL: PhD/Master thesis (Measure = 3).<br>• GL with high credibility, such as books, magazines, specialized databases, white papers, method creators and consultants' websites and case studies (Measure = 2).<br>• GL with medium credibility, such as technical reports, news, Q/A sites, blogs, presentations and videos (Measure = 1).<br>• GL with low credibility, like ideas/opinions/thoughts/commentaries without evidences (Measure = 0). |

## 4.5. Data extraction

Data extraction phase is to extract all relevant information from the selected papers. The following data should be extracted from each selected study:

• Paper information: paper id, title, authors, publication year, sources, etc.

• Key data items (e.g. for RQ1 and S1, data like definitions, terms, meanings etc. should be excerpted as key data items).

• Assessing paper by using exclusion/inclusion criteria

• Quality assessment score of the study.

• Context of study: study types.

• Qualitative data extraction

A data extraction form (available on final page) will be designed to accurately record all this information. Most contents of data extraction forms for MLR are similar to those in the SLR guidelines of Kitchenham and Charters (2007). Besides, Garousi et al. (2019) suggest that explicit traceable links between the extracted data and primary sources should be added in data extraction form, because it is helpful to deal with some GL sources without standardized structure.

Kitchenham & Charters (2007) recommend that data extraction process should be performed independently by at least two researchers (One is data extractor, another is data checker). For single researchers such as PhD students, it is necessary to use some checking techniques.

Moreover, during the process of data extraction, multiple publications of the same data should not be included because duplicate reports would cause heavy bias to influence the results. The right way is to use the most completed version (Kitchenham & Charters, 2007). For example, if having a PhD 1<sup>st</sup> or conference paper, its subsequent journal version would be more highly rated as most thoroughly reviewed (in theory depending on standing of the journal).

## 4.6. Data synthesis

Data synthesis is a procedure to collate and summarize the results of the included primary studies (Kitchenham & Charters, 2007). There are various data synthesis techniques, including: descriptive (narrative) synthesis, quantitative synthesis, qualitative synthesis, and the combination of quantitative and qualitative synthesis (Garousi, Felderer, & Mäntylä, 2019). Based on the type of primary studies (data), an appropriate data synthesis technique should be selected and used.

In this research, qualitative synthesis method will be mainly used, because there will be much qualitative data collected in the WL and qualitative and experience-based evidence is also very common in the GL. However, it may be changed as there will be some uncertainty until the data is actually collected and extracted. Therefore, the detailed activities of data synthesis will be discussed after data collection and extraction.

## 5. Evaluation, Validation and Amendment

The first draft of this review protocol will be presented to PhD supervisors (Dr Ramesh Lal & Assoc. Prof. Tony Clear) for evaluation and criticism. According to the SLR guidelines of Kitchenham & Charters (2007), the protocol can be approved to be validated if following conditions are checked to meet:

- The search strings are derived from the research questions.

- The study selection criteria and process for WL and GL are appropriate.

- The data extraction procedure is appropriate to address the research questions.

- The data synthesis procedure is appropriate to answer the research questions.

In addition, there may be some further amendments of this MLR protocol when executing the procedures in new situations. Some necessary changes can make up for deficiencies and improve the current version of this protocol. Each revision of the MLR protocol will be recorded timely and the protocol will be updated accordingly.

## 6. Dissemination

The final phase of an MLR is dissemination (Reporting the review). This review will be reported in a technical report or in a section of the PhD thesis.

# References

Garousi, V., Felderer, M., & Mäntylä, M. V. (2019), Guidelines for Including Grey Literature and Conducting Multivocal Literature Reviews in Software Engineering, *Information and Software Technology*, *106*, 101–121，doi: 10.1016/j.infsof.2018.09.006.

Gupta, R. K., Venkatachalapathy, M., & Jeberla, F. K. (2019), Challenges in Adopting Continuous Delivery and DevOps in a Globally Distributed Product Team: A Case Study of a Healthcare Organization, *2019 ACM/IEEE 14th International Conference on Global Software Engineering (ICGSE)*, Montreal, QC, Canada, 30-34, doi: 10.1109/ICGSE.2019.00020.

Hussain, W., Clear, T., & MacDonell, S. (2017), Emerging Trends for Global DevOps: A New Zealand Perspective, *2017 IEEE 12th International Conference on Global Software Engineering (ICGSE)*, Buenos Aires, 21-30, doi: 10.1109/ICGSE.2017.16.

Jalali, S., & Wohlin, C. (2010), Agile Practices in Global Software Engineering - A Systematic Map, *2010 International Conference on Global Software Engineering*, 45-54, doi: 10.1109/ICGSE.2010.14

Kitchenham, B., & Charters, S. (2007), Guidelines for Performing Systematic Literature Reviews in Software Engineering, doi: 10.1.1.117.471

Mohan, V., & Othmane, L. B. (2016), SecDevOps: Is It a Marketing Buzzword? - Mapping Research on Security in DevOps, *2016 11th International Conference on Availability, Reliability and Security (ARES),* Salzburg, 542-547, doi: 10.1109/ARES.2016.92.

Rütz, M. (2019), Devops: A Systematic Literature Review, *Seminar Paper*.

Souag, A., Mazo, R., Salinesi, C., and Comyn-Wattiau, I. (2016), Reusable knowledge in security requirements engineering: a systematic mapping study. *Requirements Engineering*, 21:251–283.

Stray, V., Moe, N. B., & Aasheim, A. (2019), Dependency Management in Large-Scale Agile: A Case Study of DevOps Teams, *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 7007-7016, ISBN: 978-0-9981331-2-6, https://hdl.handle.net/10125/60137.

Wilde, N., Eddy, B., Patel, K., Cooper, N., Gamboa, V., Mishra, B., & Shah, K. (2016), Security for DevOps Deployment Processes: Defenses, Risks, Research Directions. *International Journal of Software Engineering & Applications (IJSEA)*, 7(6), 1-16, doi: 10.5121/ijsea.2016.7601.

Zaydi, M., & Nassereddine, B. (2019). DevSecOps practices for an agile and secure it service management. *Journal of Management Information and Decision Sciences*, 22(4), 527-540.

# Data Extraction Form

| Phase1: Paper Selection | Your Response | Comments |
|---|---|---|
| **Paper Information** | | |
| Paper id | | ID of paper, including search No. and databases, e.g. S1_IEEE_01, S2_ACM_02. |
| Paper title | | Title of paper, short version suffices |
| Authors | | Authors' names of paper |
| Publication year | | Year of Publication |
| Key data items | | e.g. for RQ1, data like secure, security, SecDevOps, etc. should be excerpted as key data items |
| Date researcher analyzed this paper | | When researcher completed this form |
| **Exclusion/Inclusion Criteria** (Some criteria have been pre-performed during searching papers, e.g. language, years and subject area.) | | |
| Ex (a): Is the study external to the terms of Search 1 (DevOps security & SecDevOps/DevSecOps)? | | Paper needs to focus on DevOps and its security aspects (for RQ1) |
| Ex (b): Is the study external to the terms of Search 2 (DevOps & security & GSE/GSD)? | | Paper also needs to focus on DevOps security and GSE/GSD topic (for RQ1) |
| Ex (c): Is the study based on personal opinion? | | We reject papers without rigorous methodology or research design to prove the correctness of findings |
| Ex (d): Is this a repeated study? | | We only include the key study (most comprehensive) because repeated study would bias results |
| In (a): Is RQ addressed? | | Which research question is addressed by the paper |
| In (b): Is the study from acceptable source | | Include: conference and journal papers; Exclude: books (chapters), posters and abstracts. |
| Quality assessment score of study | | To assess and grade the quality |
| **Decision** | | |
| Decision status: {Accept/Reject/Waiting | | Define decision status. "Don't know" |

| | | |
|---|---|---|
| for Full paper/Don't Know} | | status will go to peer review and arbitration |
| Decision based on: {Title/Abstract/Introduction/ Conclusion/Methods/Whole Paper/Peer Review/Arbitration | | At what point did researcher make decision. |
| **Context of Study** | | |
| Type of papers | | Indicate type: solution proposals, philosophical papers, evaluation research, validation research, opinion papers, and personal experience papers. |
| **For Empirical Studies Add** | | |
| Type of empirical study methods | | Indicate type: experiments, survey, interviews, observation, case study, action research, focus groups, etc |
| Country/Location | | List countries involved in the study |
| **Phase2: Qualitative Data Extraction** (Go to Phase 2 if paper has passed all criteria in Phase 1 above.) | | |
| Understanding of DevSecOps (RQ1), including: definitions, challenges, practices, tools, and metrics. | | RQ1. What is the state of art of DevSecOps in the existing literature? (List as many as you find) |
| The relationship between GSE and DevSecOps. | | RQ2. How does Global Software Engineering relate to DevSecOps? (List as many as you find) |
| **Additional Data/Follow Up** | | |
| Other observations or useful quotes found in paper | | Record useful texts or exact quotes which can be used in our report |
| References found in paper/snowballing (to follow up) | | Include more relevant studies based on the references of selected papers. |