

An Exploratory Study for Identification, Classification, and Removal of Waste in Lean Software Development: A Multivocal Literature Review Protocol with Preliminary Results

Muhammad Ilyas*, Izzah Manzoor†, Rafi Ullah‡

Abstract

Lean Software Development (LSD) is an emerging area adopted by the software industry from manufacturing industry. LSD provides value to the customer by delivering the product with only value-adding activities/features and by not including the unwanted features/activities (waste). As LSD is a new and evolving area, there is less resource/data available regarding practices and principles for identification and elimination of waste. Also, there is no standard method available for handling waste. Hence, there is a need to explore the literature to find an explicit list of methods, principles and practices used to identify, classify and remove waste in LSD. This is helpful to bring everyone on one page and have a common understanding of how to fully incorporate lean concept in developing a software. The primary goal of this paper is to design a comprehensive Multivocal Literature Review (MLR) protocol to explore and identify the methods, principles and practices of LSD in handling waste. A MLR is conducted to find the methods, practices and principles used in order to identify, measure and remove waste in LSD. The protocol is currently in implementation stage. The study is looking forward to getting a list of methods, principles and practices used in handling waste in LSD after the conduction of this MLR.

Keywords: Lean Software Development (LSD); Waste; Principles; Practices and Methods; Multivocal Literature Review (MLR).

Introduction

The Lean Software Development (LSD) comes under the umbrella of agile frameworks. The focus of LSD is to increase customer value by reducing the waste i.e., developing and delivering only what the product needs (Gong & Janssen, 2014). Thus hence optimizes the development time and resources (Poppendieck et.al, 2003). Lean makes it fruitful for the organizations by not including the unwanted features/activities (waste) and add value to the product by only delivering value-adding activities/features (Womack & Jones, 1997).

The concept of “lean” has been originated from Toyota’s production system invented in the middle of twentieth century (Misaghi &

*Department of CS & IT, University of Malakand, Chakdara 18800, Pakistan, milyasmkd@uom.edu.pk

†Department of CS & IT, University of Malakand, Chakdara 18800, Pakistan, manzoorizzah@gmail.com

‡Corresponding Author: Department of CS & IT, University of Malakand, Chakdara 18800, Pakistan, rafiullah.cs@uom.edu.pk

Bosnic, 2014). Lean is now being adopted in other domains like software development, health care etc. because of its two main features i.e., reduction of waste and increasing customer value (Gong & Janssen, 2014).

The fundamental objective and principle of lean manufacturing is to first identify waste and then take steps to eliminate/reduce it. Manufacturing industry had its own principles and practices identification and elimination of waste. To fully implement and adopt these practices and principles in the software industry, it is needed to understand them. Lean manufacturing has seven wastes which are interpreted and converted to LSD waste by Mary and Poppendieck (2007). The seven wastes of LSD defined by (Poppendieck & Tom., 2007) as shown in the Figure 1, are partially done work, relearning, defects, handoffs, extra features, delays and task switching.

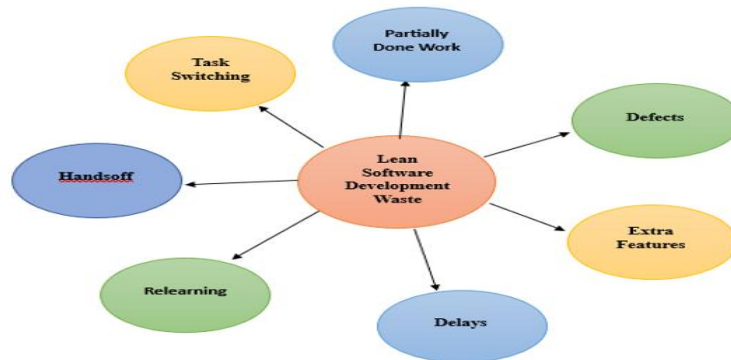


Figure 1: Seven wastes of LSD

It needs to be familiar and fully understand the lean practices which are used in software development to capture and identify different types of wastes. Moreover, the practices used for the removal or reduction of different waste types need to be known.

A Systematic Literature Review (SLR) and an industrial survey was conducted by (Nanduri, 2014) to find which lean development practices are used effectively in the literature and in the industry for the removal of waste. It is concluded by (Nanduri, 2014) that there is the divergence in the perception of practitioners and researchers regarding the lean practices in removal of wastes. Also, there is less focus on lean practices for the removal of seven types of wastes by the researchers.

Similarly, a SLR is conducted by Medidi (2015) in order to find the root causes of waste in LSD and the relation between the wastes. The study proposed that there is need for study to conduct quality assessment on the factors identified in order to develop strategies for the

mitigation of these factors. The study conducted focusing on the root causes of waste in whole software development life cycle (SDLC). So, in order to find in depth, a study can be conducted focusing on individual phases of SDLC.

A case study in 14 organizations is conducted by (Alahyari et al., 2019) in order to study waste using lean and agile approaches. The author concluded that, a joint and comprehensive concept is required to have a common understanding of how to identify, acknowledge and remove waste. Moreover, to attain viable and efficient waste management, both the interim and the prolonged perspectives are needed to be contemplated.

Extra effort is required to enhance the value provided to the customer. This extra effort is called overhead. Overhead is the indirect way of providing optimized value to the customer. According to LSD, this effort can be considered as waste, if it is implemented, it adds value to the customer. Enhancement in the implementation technique reduces the waste caused due to extra effort (Alahyari et al., 2019). To study the overheads in terms of LSD and agile software development, an exploratory study conducted by (Mummadi, 2021). The results of this study suggests that there is still more need to explore and study overheads in LSD. It is required to make a proper differentiation between wastes and overheads. There is also a need to identify if the overheads vary from organization to organization as the case study is conducted only in one organization. Moreover, there is also a need to prioritize and rank the overheads.

In order to promote lean concepts a literature review of research articles from different areas is conducted by (Deshmukh & Srivastava, 2021) to find gaps in the research and the enhancement and improvement area during implementation of lean concept in software development. It is concluded that no explicit list of lean principles is available. Furthermore, certain lean factors that are relevant for the adaption of Lean approach like waste types and lean principles, are needed to be figured out by the software industry in order to achieve “Lean” concept.

From the literature reviewed above, it can be concluded that there is no standard method for handling waste. There is no common definition of waste and practices for identification and elimination of waste. There is lack of research about overheads in LSD. Furthermore, there is a need to identify what type of waste is common to all and what waste can be specified to one domain. Domain specific waste can be according to the type of organization, the type of project e.g., customized project, general purpose project etc. or according to the size of project.

LSD is an emerging area and there is less work done specifically on the areas mentioned as gap above. Moreover, studies conducted so far on the area of interest have only included the published research findings

i.e. they used methodologies like SLR, empirical studies etc. So, instead of a SLR, a MLR is conducted as it helps us in covering both the published literature and grey literature (GL) in order to answer our research question. Moreover, the MLR bridges the gap between the perception of practitioners and researchers regarding the lean practices in identification and removal of wastes.

Research Methodology

This research study aims to develop standardized guidelines for waste removal in LSD. Thus, this research explores and identifies the methods, principles and practices of LSD to bring everyone on one page. MLR is conducted to accomplish the objectives and aim of research. This paper is intended to design a protocol i.e. a plan to conduct the MLR to explore both the published literature i.e. academic papers and unpublished resources i.e. GL for finding the methods, practices and principles of handling waste in LSD. For the development of this protocol, others researcher's studies are used for guidance (Matouq et.al, 2020; Kitchenham et.al, 2007; Garousi et.al, 2019; Calderón et.al, 2008; Abrar et.al, 2023; Azeem, et.al, 2022; Lohrasbinasab and Palacios 2020; Wang et al., 2022; Lambiase et.al, 2024; Azevedo et.al, 2024; Ullah & Ilyas, 2024).

Multivocal Literature Review

A MLR is another form of SLR. The MLR includes both formal/published literature e.g., conference papers and journal and unpublished literature i.e. GL e.g., blog posts, videos and white papers (Matouq et.al, 2020). According to Luxembourg GL is, "information which is produced in both print and electronic form at all levels of government, academia, business and industry, but it is not under the control of commercial publishers, i.e., where producing body's main activity is not publishing" (Schöpfel et.al, 2010). MLR is valuable for both practitioners and researchers as they provide a comprehensive overview of latest research findings and practical applications in a specified domain (Matouq et al., 2020).

The three stages for conducting a literature review according to (Kitchenham et.al, 2007) are: 1) planning, 2) conducting, and 3) reporting. In planning stage i.e. protocol, research questions are formulated, the data sources are defined, design search strings, and define inclusion-exclusion criteria and quality assessment criteria is defined. In the second stage, a review is conducted and executed according to the protocol designed. Here data are extracted from primary studies and data are synthesized. In the

last stage, the findings and results of MLR study are documented and reported. The overall process of MLR is shown in the Figure 2.

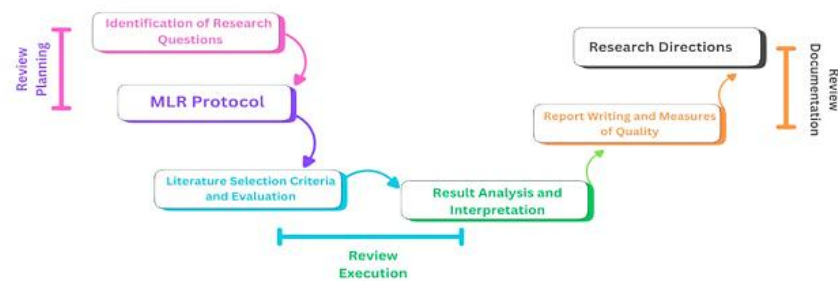


Figure 2: A Multivocal Literature Review Phases

Following the guidelines of Kitchenham and Charters (2007), various phase of the MLR protocol and their sub process are mentioned in Figure 3. Details of protocol development for the proposed study are mentioned in the following section.

Research Questions

The research questions of study formulated from the aim and objectives are as follow.

RQ1: What are the methods, practices and principles for the identification and measurement of waste in Lean Software Development?

This question is intended to find all the available practices, principles and methods which are used for identifying and measuring waste in LSD.

RQ2: What are the methods, practices and principles for the removal of waste in Lean Software Development?

This question is intended to find all the available practices, principles and methods which are used for removal waste in LSD.

Constructing Search Terms

The research questions are analyzed in the context of four perspectives.

- i. *Population:* Lean Software Development
 - ii. *Interventions:* Existing methods, principles and practices for identification, measurement and removal of lean waste
 - iii. *Outcome of relevance:* practices, principles, or methods for handling lean waste
- Experimental design:* Empirical studies, evaluative studies, case studies, expert opinions, blog posts.

Search Strategy

In this stage a search string is designed given below, to find the methods, practices and principles for the identification, measurement and removal of waste in LSD

The following search string is applied to various digital databases i.e. Google Scholar, ACM, Springer Link and IEEE Xplore libraries.

Search String

Finally, the following search string is design and applied to various online libraries.

Search string = ("Lean software" OR "Lean waste" OR "Software waste") AND (Practice OR Solution OR "implementation initiative" OR Advice OR Principle OR Direction OR Pattern OR Procedure OR Method OR Technique OR Approach OR Methodology OR Mechanism).

Search String Breakup

It is important to note that certain digital libraries such as “Science direct” does not allow a lengthy search string. For this reason, the search string was split into 3 sub strings.

String 1: ("Lean software" OR "Lean waste" OR "Software waste") AND (Practice OR Solution OR "implementation initiative" OR Advice).

String 2: ("Lean software" OR "Lean waste" OR "Software waste") AND (Principle OR Direction OR Pattern OR Procedure).

String 3: ("Lean software" OR "Lean waste" OR "Software waste") AND (Method OR Technique OR Approach OR Methodology OR Mechanism).

Literature Resources to be Searched

As per Kitchenham (Kitchenham et.al, 2007) , this study, carried out the search process of literature separately for both formal (white) literature and GL.

Published Resources

To find out pertinent literature to the research questions, the following digital libraries are searched.

- IEEE Xplore (<https://ieeexplore.ieee.org/Xplore/home.jsp>)
- ACM Portal (<https://dl.acm.org/>)
- Science Direct (<https://www.sciencedirect.com/>)
- Springer Link (<https://link.springer.com/>)
- Google Scholar (Search Engine)

Unpublished Resources (Grey Literature)

To locate the GL sources, Google search engine is used for articles or blogs.

Managing Search Results

The primary source references are stored electronically in a separate directory.

Publication Inclusion Exclusion Criteria

A separate criterion is adopted for the selection of published and GL.

Inclusion Criteria for Published Sources

Among all the studies which are retrieved by the search string, only those studies are selected for data extraction which are related to our research questions. Studies which discuss the LSD principles, methods and practices for identification and removal of waste are included. The criteria for including the studies are listed below.

- Articles that specifically focus on LSD are included
- The studies, relevant to research questions
- Studies whose full text is accessible online
- Studies that are written in English language
- The studies published from the year 2003 to 2024 are included only, as LSD methodology is first introduced in 2003 in creation of software

Exclusion Criteria for Published Sources

The exclusion criteria, which are explained below, are used to decide to exclude sources.

- Studies that are not written in English
- Studies that are not in the domain of Software Engineering
- Studies that irrelevant to the research questions
- Studies whose full text cannot be accessed
- In case of duplicate studies only the most recent one are selected
- Secondary studies i.e. SLR and tertiary as these studies are reflect duplicate findings in the primary studies

Inclusion Criteria for Unpublished Sources

Unpublished sources of GL according to the guidelines of (Garousi et.al 2019) are located by using widely used search engine called Google for appropriate sources. The inclusion criteria for GL is as follow.

- The website provides Ph.D./Master thesis case studies, blog posts, etc.
- Sources that are relevant and focus on LSD
- This study pertains to the search terms
- The study is written in English language

Exclusion Criteria for Unpublished Sources

Following exclusion criteria are used for GL

- The study fails to meet the research objective
- Studies that do not focus on LSD
- A study that doesn't clearly state its research approach
- Study that is not written in English
- Sources that are not fully accessible
- Textbooks and duplicate studies



Figure 3: MLR process of the proposed study.

Selection of Primary Sources

Initially, the papers are selected on the basis of the relevance of their title and abstract with the research questions. Then the entire article of the initially selected papers are reviewed, and final selection is made considering the above inclusion-exclusion criteria.

Quality Assessment (QA)

The papers selected in the final selection is assessed on their quality according to the quality assessment criteria established by us as outlined in Table 1. The quality of publications is assessed in parallel at the time of data extraction. This study using a three-point Likert scale ("yes," "partially," "no") to rate each component of the reviewed articles. To ensure meaningful results, values of 2 is assigned to "yes," 1 to "partially," and 0 to "no" for each component. Any paper that received an

average score of ≥ 0.5 is deemed to be acceptable quality and is included in the MLR.

Table 1: Criteria for Quality Assessment

S.No.	Criteria for Quality Assessment	Likert Scale
Q.1	Does the study have clear research objectives?	Yes = 2 Partially = 1 No = 0
Q.2	Are the practices, methods and principles for identifying and measuring waste in LSD clearly stated?	Yes = 2 Partially = 1 No = 0
Q.3	Are the practices, methods and principles for removing waste in LSD clearly stated?	Yes = 2 Partially = 1 No = 0

Data Extraction Strategy

The following strategy is adopted to extract the relevant data about the research questions.

Primary Study Data Extraction

The selected studies undergo a data extraction process to extract the essential data and answer the research questions listed below.

RQ1: Identification of the principles, methods and practices for identifying and measuring waste in Lean Software Development.

RQ2: Identification of the principles, methods and practices for removing waste in Lean Software Development.

The data extraction process involves reading the entire article and identifying the relevant data. The following data are extracted during the process and is stored in table form. The table consists of the fields of S. No, Paper _ID, date of review, title, authors, source reference, database, methodology, sample population, country/location of analysis, publication year, principles, methods and practices for identifying and removing measuring waste in lean software development, and medium type: blogs, television, YouTube and social platforms/networks.

Data Extraction Process

Data are extracted by a primary reviewer after reviewing the papers. Primary reviewer can take the guidance from secondary reviewer for addressing an issue occurred during data extraction. Secondary reviewers performed an inter-rater reliability test after the primary reviewer has completed the data extraction process. Few publications are randomly selected by secondary reviewer from the publication list selected by primary reviewer. Then data are extracted by the secondary reviewers from the randomly selected list. After the data extraction by the primary reviewer, the results are compared with those produced by the secondary reviewer. However, no variances are found.

Data Storage

The practices, methods and principles recognized for each publication are stored electronically in Microsoft Word document and SPSS are used for further analysis.

Data Synthesis

As there are two research questions, the synthesis are also categorized into two parts. For the Research Question1, the data are synthesized by creating one summary table having the columns (S. No, Practices/Methods/Principles, Frequency, Percentage) showing the list of all the practices/methods/principles along with their frequencies and percentages. The complete detail of every practice/method/principle mentioned in the Summary table is recorded in a separate table. The separate table contains the following columns i.e. Practices/Methods/Principles Group name, Reference S.No., Paper reference/Paper title. For the Research Question 2, the same process is performed as for the Research Question 1 mentioned above.

Preliminary Results

Currently, in this study MLR protocol is developed and initial results for certain protocol sections mentioned above are now available. After the implementation of search strategy outlined above on the specified digital libraries, a total 9361 results of formal (white) literature are obtained from the five digital libraries and 240 results for GL. Altogether the results for formal primary studies and GL are 9601.

Total 445 sources are initially selected on the basis of title and abstract. Afterwards, full text of initially selected sources was read for finding relevant studies. After that, the inclusion/exclusion criteria are applied on the relevant studies found and finally total 60 sources are selected for both formal literature and GL. Duplicate sources are also removed from the final list. Detail information about the initial and final selection for each digital library is stated in the Table 2.

Table 2: Data Sources with Publication Results

Resources	Total Results Found	Initial Selection	Final Selection
Science Direct	442	43	8
IEEE Xplore	84	66	10
Google Scholar	8230	214	20
Springer Link	403	55	6
ACM Digital Library	202	53	8
Google search engine	240	14	8
Total Papers	9601	445	60

Discussion

During the conduction of a MLR, one comes across few limitations and constraints and should try to reduce and mitigate such constraints. First constraint come across is to get results belonging to domains other than software development due to the fact that Lean methodology is now being adopted in different areas like health sector, construction etc. This risk is mitigated while formulating the search string by adding add key terms like software to get results specific to software development field. Moreover, to avoid the selection of irrelevant and poor-quality studies, the inclusion/exclusion criteria is formulated and define the quality assessment criteria. Furthermore, due to the limitation of time, only studies between the time period 2003-2024 are considered. Also, the constraint of biasness in extracting data by primary reviewer are reduce by cross-review done by secondary reviewer.

Conclusions

LSD is an emerging area and very few resources are present regarding its principles and practices for identification and removal of waste. Hence, there is a need to explore the literature to find an explicit list of methods, principles and practices used to identify, classify and remove waste in LSD. This paper outlines a plan to conduct the MLR protocol for the study. In this study, MLR protocol is developed and currently the study is in the execution of the protocol. Some preliminary results are obtained for few sections of the protocol. The expected outcomes from MLR are the identification of principles, practices and method for identification, measurement and removal of waste in LSD. The results of this study help in clearly distinguishing between value adding activities and waste in the process. The results provide us with a detail list of practices, methods and principles of LSD that are carried out by researchers and practitioners to handle lean waste. Motivation of this research is to provide a common understanding regarding lean waste handling techniques and to bridge the gap between different perspectives related to removal and identification of waste in LSD.

References

- Al-Matouq, H., Mahmood, S., Alshayeb, M., & Niazi, M. J. I. A. (2020). A maturity model for secure software design: a multivocal study. 8, 215758-215776.
- Alahyari, H., Gorschek, T., & Svensson, R. B. (2019). An exploratory study of waste in software development organizations using agile or lean approaches: A multiple case study at 14 organizations. *Information and software technology*, 105, 78-94.

- Azevedo, K., Quaranta, L., Calefato, F., & Kalinowski, M. (2024). A multivocal literature review on the benefits and limitations of automated machine learning tools. arXiv preprint arXiv:2401.11366.
- Abrar, M. F., Khan, M. S., Khan, I., Ali, G., & Shah, S. (2023). Digital information credibility: Towards a set of guidelines for quality assessment of grey literature in multivocal literature review. *Applied Sciences*, 13(7), 4483.
- Akbar, M. A., Mahmood, S., Meshram, C., Alsanad, A., Gumaei, A., & AlQahtani, S. A. (2022). Barriers of managing cloud outsource software development projects: a multivocal study. *Multimedia Tools and Applications*, 1-24.
- Barbara Kitchenham, S. M. C. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Retrieved from University of Durham Durham, UK.
- Calderón, A., Ruiz, M., & O'Connor, R. V. (2018). A multivocal literature review on serious games for software process standards education. *Computer Standards & Interfaces*, 57, 36-48.
- Deshmukh, M., & Srivastava, P. (2021). Literature Review of Lean Methodology and Research Issues for Identifying and Eliminating Waste in Software Development. *Intelligent Manufacturing and Energy Sustainability*, 375-388.
- Garousi, V., Felderer, M., Mäntylä, M. V. J. I., & technology, s. (2019). Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *106*, 101-121.
- Gong, Y., & Janssen, M. (2014). The use of Lean principles in IT service innovation: insights from an explorative case study. Paper presented at the Conference on e-Business, e-Services and e-Society.
- Kitchenham, B., Charters, S. J. K. U., & Report, D. U. J. (2007). Technical report title: guidelines for performing systematic literature reviews in software engineering, EBSE 2007-001.
- Lohrasbinasab, I., Acharya, P. B., & Colomo-Palacios, R. (2020). BizDevOps: a multivocal literature review. In *Computational Science and Its Applications–ICCSA 2020: 20th International Conference, Cagliari, Italy, July 1–4, 2020, Proceedings, Part VI* 20 (pp. 698-713). Springer International Publishing.
- Lambiase, S., Catolino, G., Palomba, F., & Ferrucci, F. (2024). Motivations, Challenges, Best Practices, and Benefits for Bots and Conversational Agents in Software Engineering: A Multivocal Literature Review. *ACM Computing Surveys*, 57(4), 1-37.

- Medidi, P. (2015). Waste in Lean Software Development: A Root Cause Analysis. In.
- Misaghi, M., & Bosnic, I. (2014). *Lean mindset in software engineering: A case study in a software house in brazilian state of santa catarina*. Paper presented at the Joint Conference on Knowledge-Based Software Engineering.
- Mummadi, S. (2021). Exploratory study of Overhead in Lean/Agile Software Development. In.
- Nanduri, V. R. (2014). Lean+ Agile vs Seven Wastes in Software Development. In.
- Poppendieck, M., & Poppendieck, T. (2003). *Lean software development: an agile toolkit*: Addison-Wesley.
- Poppendieck, M., Poppendieck, T. D., & Poppendieck, T. (2007). *Implementing lean software development: from concept to cash*: Pearson Education.
- Schöpfel, J., & Farace, D. J. (3rd edition ,2010). *Grey Literature," in Encyclopedia of Library and Information Sciences*. Paper presented at the CRC Press.
- Ullah, R., & Ilyas, M. (2024). Mob Programming Challenges, Success Factors, and Practices: A Multivocal Literature Review Protocol with Preliminary Results. *The Sciencetech*, 05(3), 2-27.
- Womack, J. P., & Jones, D. T. (1997). Lean thinking—banish waste and create wealth in your corporation. *Journal of the Operational Research Society*, 48(11), 1148-1148.
- Wang, Y., Mäntylä, M. V., Liu, Z., Markkula, J., & Raulamo-jurvanen, P. (2022). Improving test automation maturity: A multivocal literature review. *Software Testing, Verification and Reliability*, 32(3), e1804.