

A Grey Literature Review Protocol – Measuring Digital Products

1. Introduction

Digital disruption is happening worldwide, bringing a variety of products to the market that substitutes physical products. Digital product owners are working hard to achieve the company's goals and derive all possible advantages from the products. They use product metrics to evaluate and monitor the success of the product. However, there are various techniques and methods to use metrics, tools to capture and visualise them, and factors used to choose the right metric for suitable needs.

This study aims to extract data about digital product metrics usage in the domain of digital product management to accumulate the best practices and insights not readily available for the public. The systematic grey literature review was chosen to conduct the study. This review examines data about digital product metrics from relevant grey literature sources by following a systematic approach presented in Review Protocol. Review Protocol describes the research methodology, review planning, purpose of the review, search strategy, inclusion and exclusion criteria, data extraction strategy, data synthesis, and reporting approach.

2. Research Methodology

To conduct the study systematically and repeatably, it follows the guidelines according to Garousi [1] that was based on Kitchenham's systematic literature review guidelines [2]. The review considers three main phases: 1) planning the review, 2) conducting the review, and 3) reporting the review (see Table 1).

Table 1: Design of grey literature review.

Planning the review	- Identification of the need for GLR.
	- Formulation of the research questions and scope.
	- Search strategy: Definition and refinement of the search string.
	- Determination of the inclusion and exclusion criteria.
Conducting the review	- Usage of the search string
	- Performance of the study selection process.
	- Conduction of a quality assessment.
	- Data extraction
Reporting the review	- Writing down the findings of the report.

2.1. Planning the Review

This section of the review protocol will first motivate the research method chosen for the study and the research questions and then present the search method.

2.1. Motivation for the Grey Literature Review

Grey literature intends to investigate a broader scope of literature and provide more eligible and available examples of digital product metrics and their usage in the domain of digital product management. The information derived from this review will cover a scope that is not covered by the regular publications and will provide information from the practitioners in digital products. Furthermore, the necessity of GLR depicts its purpose in finding answers to which metrics product companies define. This information is never formally published as academic research [3]. The most popular definition of the grey literature review is the Luxembourg definition – “produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers, i.e., where publishing is not the primary activity of the processing body” [3], [4].

According to [5], [6] grey literature can be categorised according to its credibility (Figure 1). Sources associated with low credibility such as blogs, emails and tweets are classified as “Grey 1st tier”. While moderate credibility sources as reports, news articles, presentations, presentations Q/A articles from sources such as Product Talk, Stack Overflow form “Grey 2nd tier”. High credibility sources as books, magazines, LinkedIn articles [7], [1] present “Grey 3d tier”. In this review, sources from all three tiers will be included apart from the tweets and emails which can’t be found in Google Search and accessed by the reviewer.

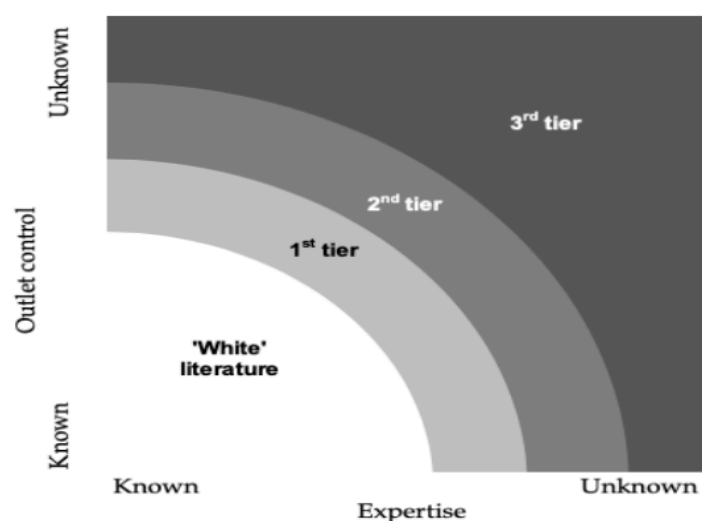


Figure1. Grey literature classification [1]

To ensure that the grey literature review (GLR) is suitable for the current study a Garousi checklist was applied, shown in Table 2 [1]. Garousi suggests that if one or two questions can be answered positively, the conduction of GLR is recommended, in other case, a Systematic Literature Review (SLR) should be performed.

First of all, the subject of digital product metrics in the domain of digital product management is “complex” enough because the industry of digital products is very dynamic and technological advances may lead to disruptive and ongoing changes. To succeed and survive, digital companies must adapt their business models and strategies to leverage the opportunities made possible by technology [8]. Secondly, contextual information (digital product management) is relevant to the subject of the study (digital product metrics). Thirdly, the synthesis of insights about digital product management will be relevant to the practitioners who apply metrics to

their day-to-day routine and use them for the success measurement of the products. In addition, for most Product Owners/Managers, their personal KPIs are defined as product KPIs and are measured using metrics [8].

Moreover, Google gives 346,000,000 outputs when searching for the “Digital Product metrics”. Thus, a large volume of practitioners’ sources indicates high interest in a topic. Consequently, by positively answering the four questions from Table 2, we are sure that the grey literature review is suitable for the current study.

Table 2: Checklist according to Garousi to decide whether a grey literature review should be performed [1].

ID	Question	Answer
1	Is the subject “complex” and not solvable by considering only formal literature?	Yes
2	Is there a lack of volume or quality of evidence, or a lack of consensus of outcome measurement in the formal literature?	No
3	Is the contextual information relevant to the subject under study?	Yes
4	Is it the goal to validate or corroborate scientific outcomes with practical experiences?	No
5	Is it the goal to challenge assumptions or falsify results from practice using academic research or vice versa?	No
6	Would a synthesis of insights and evidence from the industrial and academic community be helpful to one or even both communities?	Yes
7	Is there a large volume of practitioner sources indicating high practitioner interest in a topic?	Yes

An initial review of the grey literature and analysis of interest in learning more about digital product metrics indicate the high interest in insights about the topic of “digital product metrics”. A grey literature review can therefore contribute to the transfer of practical knowledge.

2.2. Research Questions and their motivation

This GLR focuses on identifying publications and web-based resources about digital product metrics and their handling in digital product management. A digital product is a software-enabled product or service that offers some form of utility to the people. It is one of the most sophisticated product types and refers to knowledge artefacts [9]. Sometimes the term “software product” is used to denote digital product, but since the software is also used for software development, which is not within my scope, I will henceforth use the term digital.

The research questions and their motivation are as follows:

- *RQ1: What aspects of digital products are measured?*

Digital products have various aspects that product owners need to track. Therefore, RQ1 identifies aspects of digital products that different product metrics measure and define those metrics. For example, some metrics reflect on the user experience[10], some metrics concern profitability and growth [11], some refer to product marketing success [12].

- *RQ2: What information do metrics provide to digital product owners?*

Various metrics can bring additional information for product owners. For instance, conversion can point out low user engagement on the platform [13]; this information identifies the purpose of the metric and helps to take further actions to contribute to the product development. RQ2 reveals what information metrics provide to product owners.

- *RQ3: What are the factors that impact the performance of digital products?*

Different factors impact the performance of the digital product as measured by the metrics. If product owners know these factors, they know what to do to improve the metrics. Hence RQ3 shed light on the factors impacting the performance of digital products.

- *RQ4: What are the tools used to measure metrics for digital products?*

Different product owners work with large sets of data, and various tools aid them, but not all of them are helpful for the measurement of all the metrics. Therefore, RQ4 investigates measuring tools to simplify the product owner routine.

- *RQ5: How are metrics for digital products visualised?*

Product owners collect and track metrics in dashboards. To depict metrics in the dashboard, the product owner has to visual theme it. Thus, RQ5 described visualisation practises of product metrics.

2.3. Search Strategy

The search strategy defines the unbiased approach for finding relevant grey literature sources on the topic of digital product metrics. It will specify which starts to be searched and how to search for them [14].

2.3.1. Search String Terms

To develop the search string, I followed the guidelines suggested by Kitchenham [2]. I started designing the search string in Google by assigning “all these words” criteria to “Digital product metrics” since this is the main topic of the study, which gave me 292,000,000 results and output included different language sources as well. I narrowed the language to only English. To eliminate output from software development bias. I identified that “Development KPIs” was a popular output representing the digital product scope but biases the output towards software development. Thus, I have decided to eliminate “development” from the search. Important to mention that Google provides an output of 292,000,000 results. It cuts them and has only 29 pages per 10 results on each.

The other name of a digital product is a software product. I had tried “software product metric” in the search and got results when the author talked about cycle time or team velocity and other software development team KPIs measurement. Thus, I decided to exclude the “process”, “team”, “quality” from the search. Since product managers use metrics to track success, I have included them in SS2 together with software and product.

SaaS (software as a service) is a type of digital product metric that should be included. After trying all previous combination actions in the search, I have noticed that SaaS is not an outcome. Hence, it is essential to include SaaS in the investigation as well.

Moreover, a Bing search can be used for the grey literature sources search. I have tried Bing and noticed that Bing does not cut the output providing more results than Google. However, it duplicates most of the articles, and I decided not to move forwards with it. Thus, the search string for the study is following:

Search string 1: “digital product “metrics” -development”. Google gave me 272 results. Below you can see how Advanced search in Google was designed (figure 2).

The screenshot shows the Google Advanced Search interface. Under "Find pages with...", the following criteria are entered: "all these words:" contains "digital product", "this exact word or phrase:" contains "metrics", and "none of these words:" contains "development". The "Then narrow your results by..." section includes dropdown menus for "language:" (English), "region:" (any region), "last update:" (anytime), "terms appearing:" (anywhere in the page), "SafeSearch:" (Show explicit results), "file type:" (any format), and "usage rights:" (all). A small "T" icon is visible in the top right corner.

Figure 2. Advance search for string 1

Below you can see how Advanced search in Google was designed (figure 3). Search string 2: “product metric software, OR product, OR success -team, -process, -quality”. Google gave me 253 results.

The screenshot shows the Google Advanced Search interface for a second search string. Under "Find pages with...", the criteria are: "all these words:" contains "product metric", "any of these words:" contains "software, product, success", and "none of these words:" contains "team, process, quality". The "Then narrow your results by..." section includes dropdown menus for "language:" (English), "region:" (any region), "last update:" (anytime), "terms appearing:" (anywhere in the page), "SafeSearch:" (Show explicit results), "file type:" (any format), and "usage rights:" (not filtered by licence). A small "1" icon is visible in the top right corner. At the bottom right, there is a blue button labeled "Advanced Search".

Figure 3. Advance search for string 2.

Search string 3: “SaaS product “metric”. Google gave me 262 results. Below you can see how Advanced search in Google was designed (figure 4).

The image shows the Google Advanced Search interface. Under the heading "Find pages with...", there are five input fields: "all these words:" containing "SaaS product", "this exact word or phrase:" containing "metric", "any of these words:", "none of these words:", and "numbers ranging from:". Below this, under "Then narrow your results by...", there are several dropdown menus: "language:" set to "English", "region:" set to "any region", "last update:" set to "anytime", "terms appearing:" set to "anywhere in the page", "SafeSearch:" set to "Show explicit results", "file type:" set to "any format", and "usage rights:" set to "all".

Figure 4. Advance search for string 3

2.3.2. Search Sources

I will use Advanced Google search for the search in the incognito mode. Incognito mode eliminates cookies' influence on the search.

2.4. Selection Criteria

Applying the search described above will return a significant number of results (36,900) which is far more extensive than manageable. The relevant articles should be now selected. Inclusion and exclusion criteria are dedicated to identifying relevant studies that provide sufficient information to address research questions. Given the aim of the study, the following inclusion criteria (IC) and exclusion criteria (EC) were identified (see table 3).

Since the focus area of the study is digital product management which leads to the requirement that the articles regarding product metrics should be in the domain of digital product management hence, IC1 validates that point by asking, “Is the study within the domain of digital product management?”. Moreover, articles must contain enough information about the product metrics because only some useful data can be extracted from the source. Thus, IC2 asks if the report contains sufficient information about measuring digital products, i.e., explain and discuss at least one metric?

Exclusion criteria are dedicated to examining if the information in the text is understandable, accessible, and complete, sophisticated and sufficient. Thus, sources that are not in English are out of scope because they could not be understood, which leads to the EC1 that asks if the source is not in English. Moreover, grey literature sources may not be complete, or to access them, the content should be purchased. That's why EC2 asks, “Is the full version of the page accessible?”. If the source contains the duplicate content of another one, it should not be included in the grey literature review, which is covered by EC 3.

Finally, sources should contain enough information. Hence, the EC3 (“Is the page less than 500 words?”) is applied. To check how many characters, the page has, I am going to use the “Word Counter Plus” plugin.

Table 3: Inclusion and Exclusion Criteria.

Criteria Identifier	Criteria	Criteria reasoning
IC 1	Is the article within the domain of digital product management?	The article must address the usage of metrics within the context of managing a digital product.
IC 2	Does the article contain sufficient information about measuring digital products, i.e., explain and discuss at least one metric?	The article must include enough content for answering the research questions.
EC 1	Is the article not an English source?	The source should be understandable.
EC 2	Is the full version of the article not accessible?	Sometimes some sources can be accessed by paying, login, or any other condition. If this is the case, the page can't be fully understood.
EC 3	Is the article a duplicate?	There is no need to analyse the two same studies.
EC 4	Is the article a grey literature source?	If official publishers create the article, it is out of the scope of the current research.
EC 5	Is the article too short (less than 500 words)?	The usual one space page contains 500 words. If the literature resource has fewer words, then it is not sophisticated enough for the analysis. Online Google Plugin will be used for measuring the number of words "Word Counter Plus."

These criteria should be applied in a three-stage process:

1. The header inclusion criteria (IC1) screening
2. Full-text inclusion criteria screening
3. Full-text exclusion criteria screening [7]

First, I need to screen the header (title) of the page, which usually gives the main topic of the source. If the header suits the inclusion criteria (IC1), then I can proceed with the further evaluation of the text. Secondly, I click and open the article and start screening by ECs (EC1, EC2, EC3, EC4). In case the article does not suit the EC, then I proceed with it further. The last step is to check if the paper fits IC2. In case of a positive answer, I save the report down by pressing print and save it as pdf. The saving step is needed because the link might become broken, disappear.

During the screening process, I will mark the number of the total papers reviewed, the number of documents that failed after each exclusion and inclusion criteria and provide this information in the final report of the study. The process of criteria validation is logged to the Excel table that contains the link to the page, its traceability number, and comment regarding each of the criteria. Based on the table received, the final batch of the grey literature sources to analyse should be identified by applying the filter.

3. Data Extraction Strategy

From the final list of sources, I will extract relevant data for the study. To ensure an unbiased and sophisticated extraction strategy, the data extraction form was designed [14], [15], [16]. The form includes two parts: data about the source (metadata), data supporting the research questions. The first part provides information about the web-based source as the identifier, title of the article, authors, link when the page was accessed. The second part consists of the fields that will help to gather information for the research questions of the study, check-in table 4.

To answer RQ1, I will gather metric name and definition, data required, frequency of data, source of data, formula. To elaborate on the RQ2, I will get information on the aspect of digital metrics to measure. By getting the data about information that metric provides, I will approach answering RQ3. Data about factors that influence the performance of the subject measured by the metric will help me to answer RQ4. Columns “Strengths of the metric” and “Weaknesses of the metric” will highlight data for the RQ5. The “Tools for the measurement” field will gather input for RQ6, while the “How to visualise” will gather information for RQ7.

Table 4: Data Extraction Form.

Data about the source (metadata)	
Data about the page	Description
Identifier	Unique id from the table (should be the same as during the process of criteria validation)
Title	Title of the page (header)
Authors	By whom the page was created
Link	Link to the page so that it is possible to find it later

Date accessed	Date when the article was accessed.
Data supporting research questions	
Metric name and definition	Definition of the metric to support the answer to the RQ1 and provide more context and knowledge about the metric
Data required	Data needed for the metric calculation. Information is required to support RQ1 and provide more knowledge about the metric data measurement.
Frequency of data	How often the metric should be applied. Information is required to provide more knowledge of RQ1 and how to use the metric and measure needed data.
Formula	The formula of metric calculation. Information is required to provide more knowledge on the RQ1.
The aspect of a digital project to measure	Aspects of the digital product that the metric measures. Information Is required to provide answers on RQ1 and help to group metrics by the elements they measure.
Information that the metric provides	What information metric can provide to the product owner. Data includes knowledge on the RQ2.
Factors that influence the performance of the subject measured by the metric	Describe factors that influence the performance of digital products measured by metrics and reflect on RQ3.
Tools for measurement	Tools that are used for the metrics measurement. Information to support RQ4.
How to visualise	Techniques of metrics visualisations. Information to support RQ5.

The data will be extracted to the Excel table, where each of the columns will be the form field, and each of the sources will be a reviewed paper. If the article covers multiple metrics, each metric will have its row in the table. Essential to mention if the paper will include the cross-

references and after their validation with inclusion and exclusion criteria will be suitable for the study, they are going to be included in the extraction table as well.

4. Conducting the Review

In this part of the review protocol, it will be discussed how the Strategy of finding grey literature sources presented in the “Planning the review” part of the document was applied and its outcomes.

4.1. Screening Execution

I have conducted a search procedure in 3 milestones.

- Milestone 1: screening of Search String 3. Date: 24.11.2020
- Milestone 2: screening of Search String 1. Date: 19.12.2020
- Milestone 3: screening of Search String 2. Date: 23.01.2021

I have used incognito mode in Google.

The process in each milestone was done as follows.

1. Open the Google Advanced Search in the Incognito Mode of Google
2. Enter the Search String parameters as in the screenshots of the “Search string terms” of the report
3. Start a screening process:
 - 3.1. Check the title of the article without opening it. If it does not suit IC1, mark that down and do not proceed further.
 - 3.2. Open the article in the separate tab
 - 3.3. Validate EC4. If the answer to the EC4 is “yes”, then do not proceed further and open a new article
 - 3.4. Validate EC1. If the answer to the EC4 is “yes”, then do not proceed further and open a new article
 - 3.5. Validate EC2. If the answer to the EC2 is “yes”, then do not proceed further and open a new article and open a recent article
 - 3.6. Validate EC3. If the answer to the EC3 is “yes”, then do not proceed further and open a new article
 - 3.7. Check if the article meets IC2. If not, then do not proceed further
4. Save a PDF version of the article in a folder with the name *<SS + number of the string>*.
5. Execute data extraction
 - 5.1. Remove articles that do not fit selection criteria during the data extraction phase
 - 5.2. Add articles that match criteria and are backward links

4.1.1. Search String 1 Screening Results

After conducting a screening procedure based on Search string 1, I have examined 27 pages in Google with total hits of 272. Some pages had 11 hits, and some had ten hits. Seventy-seven articles were excluded from the final batch violating IC1. 15 articles were not in English, which led to the exclusion based on the EC1. 18 articles did not have a full version that led to the exclusion based on the EC2. Thirty-four articles were duplicated and were eliminated based on

the EC3. Seventeen articles did not have sufficient information found and violated IC1. That led to the fact that 95 articles formed a final batch of the articles. More detailed statistics are shown in figure 4.

4.1.2. Search String 2 Screening Results

After conducting a screening procedure based on Search string 2, I have examined 26 pages in Google with total hits of 253. 134 (figure 2) articles were excluded from the final batch violating IC1. 7 articles appeared to be a not grey literature source which led to the exclusion based on the EC4. 26 articles did not have a full version that led to the exclusion based on the EC2. 9 papers were duplicated and were eliminated based on the EC3. That led to the fact that 67 articles formed a final batch of the articles.

4.1.3. Search String 3 Screening Results

After conducting a screening procedure based on Search string 3, I have examined 24 pages in Google with 262 total hits. 52 (figure 3) articles were excluded from the final batch violating IC1. 4 articles appeared to be a not grey literature source which led to the exclusion based on the EC4. 4 were eliminated because of EC4. 11 articles did not have a full version that led to the exclusion based on the EC2. 5 articles were duplicated and were eliminated based on the EC3. 13 articles did not have sufficient information found and violates. The final number of articles is 140.

4.1.4. Screening During Data Extraction

During the data extraction phase, 28 articles were excluded as non-meeting selection criteria and 1 were added as backward links that fit selected criteria. As a result, 281 articles formed the base for data extraction (figure 5).

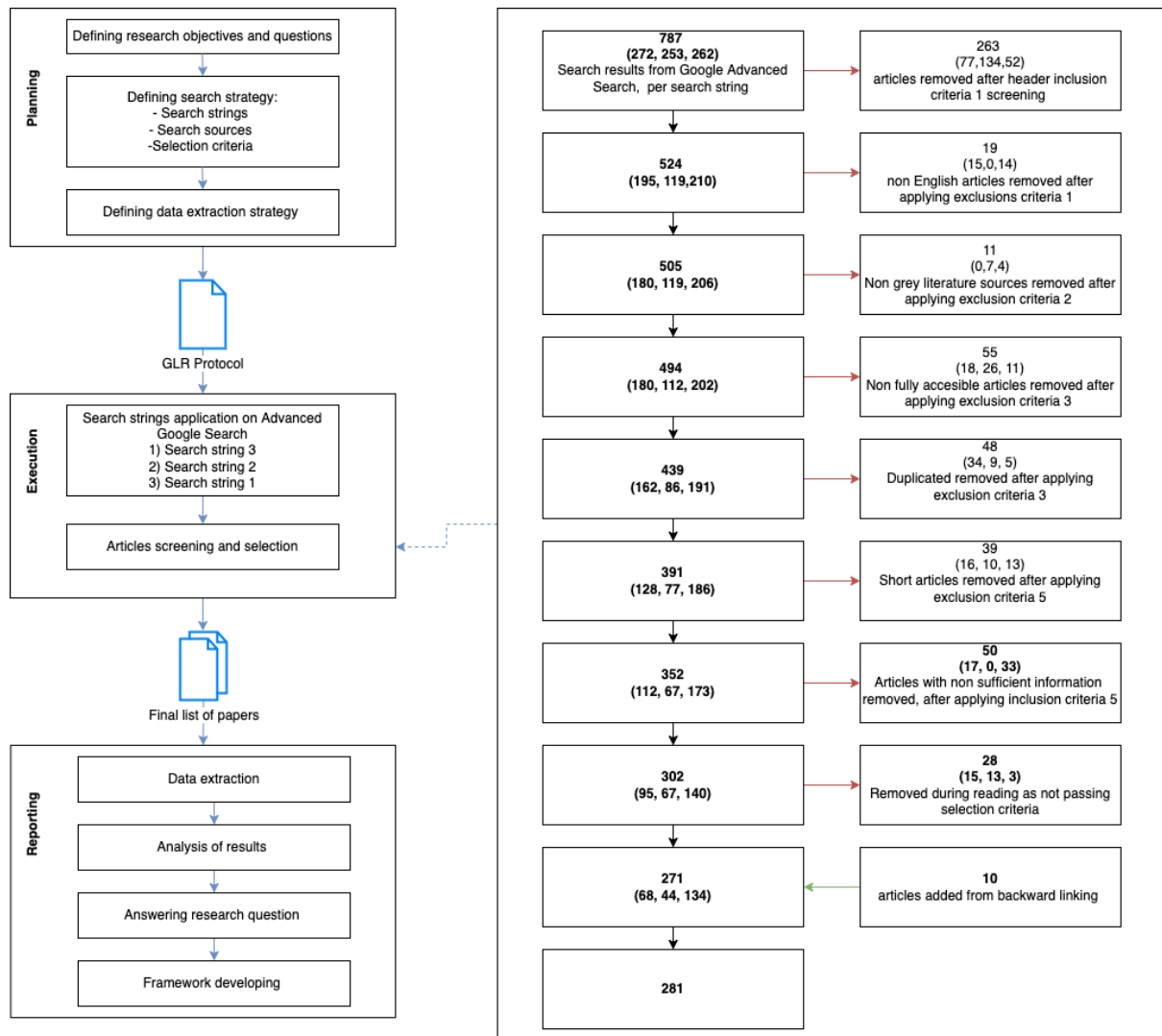


Figure 5. Screening Process Results.

Reference list

- [1] V. Garousi, M. Felderer, and M. V. Mäntylä, "Guidelines for including grey literature and conducting multivocal literature reviews in software engineering," *Inf. Softw. Technol.*, vol. 106, pp. 101–121, Feb. 2019, doi: 10.1016/j.infsof.2018.09.006.
- [2] Kitchenham B., "Guidelines for performing Systematic Literature Reviews in Software Engineering | BibSonomy." <https://www.bibsonomy.org/bibtex/aed0229656ada843d3e3f24e5e5c9eb9> (accessed Nov. 04, 2021).
- [3] K. Godin, J. Stapleton, S. I. Kirkpatrick, R. M. Hanning, and S. T. Leatherdale, "Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada," *Syst. Rev.*, vol. 4, no. 1, p. 138, Oct. 2015, doi: 10.1186/s13643-015-0125-0.
- [4] J. Schöpfel D. J. Farace, "Grey Literature" in *Encyclopedia of Library and Information Sciences*. CRC Press, 2010.

- [5] R. J. Adams, P. Smart, and A. S. Huff, “Shades of grey: Guidelines for working with the grey literature in systematic reviews for management and organizational studies,” *Int. J. Manag. Rev.*, vol. 19, no. 4, pp. 432–454, 2017, doi: 10.1111/ijmr.12102.
- [6] V. Garousi, M. Felderer, M. V. Mäntylä, and A. Rainer, “Benefitting from the Grey Literature in Software Engineering Research,” *ArXiv191112038 Cs*, Nov. 2019, Accessed: Dec. 28, 2021. [Online]. Available: <http://arxiv.org/abs/1911.12038>
- [7] J. Munch, S. Trieflinger, and B. Heisler, “Product Discovery – Building the Right Things: Insights from a Grey Literature Review,” in *2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, Cardiff, United Kingdom, Jun. 2020, pp. 1–8. doi: 10.1109/ICE/ITMC49519.2020.9198328.
- [8] S. Bhattacharjee, R. D. Gopal, J. R. Marsden, and R. Sankaranarayanan, “Digital goods and markets: Emerging issues and challenges,” *ACM Trans. Manag. Inf. Syst.*, vol. 2, no. 2, pp. 1–14, Jun. 2011, doi: 10.1145/1985347.1985349.
- [9] R. Boyett, “What is Digital Product in 2019?,” *Medium*, Jun. 13, 2019. <https://medium.com/ideas-by-idean/what-is-digital-product-in-2019-62240175bd33> (accessed Dec. 06, 2020).
- [10] W. Albert and T. Tullis, *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Newnes, 2013.
- [11] Y. Anemunmi, “3 Key Elements of Product Adoption to Promote Retention and Growth”, [Online]. Available: <https://betterproduct.community/resource/product-adoption/>
- [12] D. Daniels, “7 Metrics Every Product Marketing Manager Should Know”, [Online]. Available: <https://www.pragmaticinstitute.com/resources/articles/product/7-metrics-every-product-marketing-manager-should-know/>
- [13] “How to Increase Engagement & Conversions with Interactivity,” *The Good*. <https://thegood.com/insights/increase-engagement-interactivity/> (accessed Dec. 27, 2021).
- [14] A. Kofod-Petersen, *How to do a Structured Literature Review in computer science*. 2015.
- [15] “(PDF) Lessons from applying the systematic literature review process within the software engineering domain | Pearl Brereton - Academia.edu.” https://www.academia.edu/4132941/Lessons_from_applying_the_systematic_literature_review_process_within_the_software_engineering_domain (accessed Nov. 04, 2021).
- [16] J. Randolph, “A Guide to Writing the Dissertation Literature Review”, doi: 10.7275/B0AZ-8T74.