INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2026, VOL. 72, NO. 3, PP. 1-11
Manuscript received May 14, 2026; revised June, 2026. doi: 10.24425/ijet.2026.157954

Meta-analysis of publications on software logging

Janusz Sosnowski, Jan Lewandowski, and Łukasz Reszka

***Abstract*—Logs are widely used to monitor software systems. They provide rich information to assess system runtime behavior and detect anomalies. Log generation and significance have been studied in many publications. The multiplicity of appearing publications hampers tracing research advances. Review papers on software logging address these problems. Nevertheless, the problem of publication selection tailored to the readers' interest still arises. To tackle this, we have proposed combining a publication taxonomy with a meta-analysis of derived characteristic features. This approach has been enhanced by the development of specialized analysis tools based on natural language processing. These tools can also be used for analyzing publications in other domains.**

***Keywords*—logging practices; log mining; anomaly detection; literature surveys**

I. Introduction

WHILE developing software systems divers logging statements are inserted in the source code. They produce execution and event logs which are useful in system diagnostics (anomaly detection, failure prediction), operational or user's behavior monitoring, etc. An important issue is providing proper logging statements to assure high observation precision at acceptable performance overhead. Log studies address problems related to questions of what, where, how and whether to log [1]. In the literature, various methods, algorithms and tools are presented, they are illustrated with practical assessment of their effectiveness (in academic and industrial environments).

Many literature reviews (systematic and mapping studies) on software logging have been presented in recent 20 years. In most cases, they were prepared according to well-accepted guidelines, e.g. [2, 3]. Moreover, they usually are constructed to answer the posed research questions, such as categorizing logging research into considered topics, comparing proposed approaches, identifying open problems, publication trends and research gaps, and planning future works. They differ in considered publications features during search strategies (time ranges, venues, author experience and affiliations). Usually, primary publications are selected from available databases according to search strings applied to publication titles, abstracts and keywords. Initial selection returns high number of publications. Hence, additional inclusion/exclusion criteria are used (e.g. high ranked publications). Final subjective screening leads to a limited number (typically 50 – 200) of publications considered for deeper text analysis.

Authors are with Warsaw University of Technology, Poland (e-mail: Janusz.sosnowski@pw.edu.pl, jan.lewandowski5.stud@pw.edu.pl, lukasz.j.reszka@gmail.com).

Having analyzed literature surveys on software logging we noticed the need of aggregating the knowledge from these surveys and upgrading it with newly appearing primary papers, etc. For this purpose, we have introduced a hierarchical taxonomy of publications (considering diverse perspectives and a broad range of topics). It is supported with a meta-analysis involving indicative features related to publication appearance, venues, authorship and content. Special attention is given to tracing and assessing considered references. The proposed methodology and the developed software tools offer new capabilities for conducting literature surveys.

The paper is organized as follows. Section II describes the background and related work combined with multidimensional publication taxonomy. Section III presents the proposed methodology and developed tools. It is followed with advanced reference analysis in Section IV. Section V illustrates results related to meta-analysis of software logging publications. Summary and future works are discussed in Section VI.

II. Background and Related Works

Considering the goal of publications, we distinguish between primary and review papers. In primary publications, the authors present results of their studies related to specified problems (research question). Review publications (secondary studies) are targeted at interpreting, evaluating and synthesizing primary ones sometimes they refer also to previous reviews (tertiary studies) [3]. We have studied systematic literature reviews (SLR), literature mapping studies (LMS) and meta-analysis (MA). SLR follows a highly structured and exhaustive paper search based on predefined inclusion/exclusion criteria. It focuses on specified research questions, e.g. overview of the current state of knowledge (qualitative aspects of considered problems), identifying challenges and gaps, and generating insights. LMS provides an overview of a subject as it is reported in primary studies. The meta-analysis of a literature review synthesizes data from multiple studies to assess publication features. Literature reviewing is formalized in several guidelines (e.g. [2, 3]). In practice, we encounter reviews combining these approaches (SLR and MA).

In the case of domains involving high publication activities, we can encounter general and problem oriented literature surveys. General surveys cover a wide range of considered topics as opposed to problem oriented surveys focusing on a selected problem or correlated topics. This is the consequence of topic taxonomy within the considered domain. Such taxonomy can be coarse or fine grained depending upon the

problem aggregation levels. Coarse grained taxonomy can be derived from general surveys. Further decomposition follows problem oriented and primary publications. This leads to some hierarchical classification:

Log statement issues: statement insertion policy (what, where, why and whether to log), log position predicting, identification of log deficiencies, log updating (adding, deletion, modification), log cost (code and time overhead), debugging and diagnostic accuracy, log interpretation and understanding;

Mining log files: log parsing (based on regular expressions, frequent pattern mining, or NLP), anomaly detection, log analysis (automatic, semiautomatic), reporting and prediction precision (verbosity, duplications, redundancy), log deficiencies, log processing methods (AI, ML, NLP), log benchmarks (academic, industrial);

Logging applications: security issues, data base, big data, real time, embedded, mobile systems, ML and AI systems, workflow analysis, system performance analysis, monitoring user or system behavior, troubleshooting, debugging;

Log files management: creating log collections, rolling, removing, archiving, compression, log maintenance;

Emerging problems and challenges: divers author suggestions and observations.

In the literature, we can also meet other names of these problem categories, or their aggregations, e.g. logging practices, log instrumentation [4], log analysis [5] (placement of logging statements, overhead costs, log compression, log mining, anomaly detection). Deeper analysis and filtering can lead to cross-sectional (multilayer) decomposition, e.g. anomaly detection based on machine learning and artificial intelligence approaches, academic or industrial applications, considered topics vs publication or author rankings. This exploration can be reinforced by predefined keywords or n-grams derived from the publication text (compare Section III).

TABLE I
CLASSIFICATION OF SURVEY PUBLICATIONS

Class	Reference list
General surveys	[6] - 2016; [7] -2017; [8], [9] – 2021; [5] – 2022; [1] – 2023; [4] - 2024
Problem Oriented surveys	Parsing: [10] – 2019; [11] – 2020; [12] – 2021; [13] – 2022; [14, 15, 16] – 2023; [17, 18, 19] – 2024; Analysis: [20] – 2017; [21, 22] – 2020; [23, 24] – 2021; [25, 26, 27, 28, 29] – 2022, [30, 31] – 2024 Security [32] - 2016, [33, 34, 35] - 2020, Anomaly det.: [36, 37], - 2020; [38] – 2021; [39, 40, 41], - 2022; [42, 43, 44, 45]-2024,

Within log analysis divers methods are presented, they are supported with developed tools and assessed using specified log datasets. Table I illustrates classification dimensions (taxonomy profiles) of survey papers. The cited references are grouped according to the topic scope and specified publication years. In [46], we extended this table to cover over 100 recent primary papers. Moreover, we provided publication mapping to considered 15 log analysis approaches, 19 developed/used tools and 28 log benchmarks. It is supplemented with descriptions summarizing main points of topics/issues studied in cited references.

Problem oriented surveys in Table I cover quite wide scope of relevant issues. Progress of new studies and applications triggered surveys on appearing or specific logging aspects, e.g. log smell detection related to identified bad practices [47], log severity taxonomy and adjustments during development [48], log anomaly detection using LLM models and Retrieval Augmented Generation [49], log-correlation tools for cyberattack detection and prediction in large networks [50], AI techniques for microservice log analysis [51]. The introduced taxonomy deals with diverse categorization dimensions (perspectives): publication category (primary, survey), considered topics, presented solutions, publication time, etc. It facilitates grasping the range of relevant research and gaining comprehension (understanding) of the considered domain, as well as searching for and exploring relevant primary papers or triggering new studies.

Survey publications differ in search queries, selected references (publication time ranges, venues), analysis methods, research questions/goals. The topics considered (fine-grained view) are usually combined in predefined categories (higher level corresponding to coarse grained problem decomposition). Sometimes, the same topics can be included in different classification groups of the surveys. Appearing taxonomy differences result from author subjective views, research center tradition and experience, etc. Such surveys can provide creative ideas and enable exploring topic progress in time.

III. METHODOLOGY AND TOOLS

Log mining and logging techniques is a tough and important topic widely discussed in comprehensive literature, e.g. Scopus, IEEE, Springer repositories. So, in practice some filtering is needed. Usually, publications are filtered in three phases: elimination of duplicated publications, filtering according to inclusion/exclusion criteria, and finally review of selected papers, e.g. based on specified additional features (screening titles, abstracts and keywords). The search criteria can be quite general, e.g. specifying overall domain (software engineering, computer science), publication time span, publication language, only peer reviewed or regular papers, primary publications. Additionally, we can select publications from specified venues or authors. Here, classical guidelines, e.g. [2, 3]) are useful.

For the selected papers that are deemed relevant and reporting profound studies, we can apply the process of snowballing for further searches. Backward snowballing checks the reference lists of included articles to identify additional studies for verification. Forward snowballing identifies subsequent studies that have cited included articles. Author snowballing involves checking other papers of authors included in the collected set. We can extend snowballing to additional searching other papers in publication locations or author affiliations reported (dominated) in the collected set. Here, arises the problem of aggregating sets of publications on the considered domain as well as selecting the most interesting /innovative ones, etc. This process needs extracting appropriate metadata from publications. The derived metadata characterize various features of the publication. In particular, they provide insights into research topics, the scope of considered problems, research advances, practical usage and experience, as well as aggregated and cross-sectional statistics. Metadata facilitate and support literature searches tailored to research needs [46]).

Tracing literature related to realized projects and research, we found the need of supporting tools targeted at diverse tasks: publication search and data extraction, classification of considered problems, deriving and analyzing publication features. For this purpose we developed two complimentary tools: PubAnalyzer – deals with these tasks by analyzing publication repositories, PdfAnalyzer – explores cited references and other features in publication pdf documents.

PubAnalyzer integrates the created SQlite database (using sql model) and a set of python scripts which assure: 1) accessing specified publication repositories (e.g. IEEE, Scopus, WoS) and selecting appropriate set defined by search strings, deriving specified publication features and storing them in the database, 2) multidimensional analysis. The extracted data includes: publication title and type, author names, author affiliations, publication venue, date, number of references, number of citations (optionally other defined features can also be included). The data analysis framework bases on developed python scripts integrated with available tools for text exploration – packets: *gensim*, *nltk*, *sklearn*, *sentence_transformers* and *UMAP* (*umap-learn.readthedocs.io/en/latest/*), statistical packet pandas (*pandas.pydata.org/*), visualization of clustering effects - packets *seaborn* and *matplotlib*, etc. The developed *Analysis* module performs thematic modeling (LDA model of *genism*), derives n-grams, etc. Results are visualized using appropriately configured Excel and VOSviewer programs.

Specified search strings are usually not sufficiently accurate, so in practice we get publication sets comprising some positions beyond the analyzed domain. To remove irrelevant papers, we use grouping algorithm HDBSCAN from *scikit-learn* packet (model all-distilroberta-v1 with 768 dimensions). This algorithm was sufficiently accurate while applied simultaneously to title, abstracts and keywords (details of tuning this algorithm we give in [46]). It provides clusters of publications which need verification. Manual verification of these clusters including analysis of bi-grams with the highest NPMI (normalized point mutual information from *nltk* packet) in the noise cluster confirmed good accuracy. Using search string "log analysis" with some simple inclusion exclusion criteria we extracted 2047 and 1249 publications from Scopus and WoS repositories, respectively (in total 3296). They were related to the period 1978-2024. The collected publications comprised also non relevant publications. The clustering algorithm provided 16 clusters, the main cluster comprised 1016 points (clearly separated from other clusters), The remaining 15 clusters represented irrelevant domains of publications, e.g. communication channels, geophysical research, signal filtering, biometrical recognitions, etc. The set of the main cluster publications we denote S1016 – related to software logging.

An important issue in publication analysis is their topic classification at fined-grained level. We tried to do this using basic publication features (titles, abstract, keywords) and clustering mechanism. Unfortunately, the results obtained were not satisfactory: large noise cluster, many clusters partially overlapping topics (identified by manual checking). Finally, we decided to identify paper topics analyzing n-grams and text mining metrics (e.g. tf-idf term frequency-inverse document frequency, or NPMI) derived from the title, abstract and keywords. Combining these features with keywords facilitated creating search terms for papers related to specified topics [46].

The developed platform **PdfAnalyzer** is composed of two frameworks: RefExt – extracts references comprised in publication pdf files, RefAnalyzer – performs publication classification and derives specified statistics. RefAnalyzer processes pdf files using services of the parsing server available in GROBID library. This library utilizes a cascade of sequence labeling machine learning models trained on small, manually-labeled datasets derived directly from pdf layouts. These datasets use TEI XML format allowing models to map pdfs into structured metadata. The XML format simplifies searching/ filtering processes and error elimination. The repository created with XML files is independent of the publisher's format and can be converted in other formats (e.g. JSON). GROBID comprises models responsible for divers metadata: header (title, author names, publication date, etc.), references, full text (including tables and figures). Reference extraction process was satisfactory (verified on many documents), however, scarcely some deficiencies appeared, e.g. erroneous inclusion data from footnotes of table contents, false positions comprising only author names. We have eliminated this by including validation heuristics in the basic algorithm.

Within the hierarchical tree TEI generated by GROBID, the system utilizes specific labeling models to classify publications: journal, conference, book, arXive, other. This classification is primarily handled through the level attribute found within the <monogr> element. To improve the accuracy of the primary algorithm we included regular expressions specifying conference papers (tags related to conference, symposium, workshop, proceedings, etc.). The improved algorithm assured 95% accuracy in type classification. It was validated with labels generated by the gemini-2.5-pro Google model.

An important problem is parsing specifications of references to extract required data, such as authors, titles of publications, venues, publication dates, etc. Another issue arises with sporadically appearing duplications in derived references, e.g. due to additional reference lists in included appendices in the paper. This is especially important when combining (aggregating) or comparing reference lists of publications. We analyzed various string similarity metrics (e.g. Levenstein, partial_ratio, token_sort_ratio, token_set_ratio) to detect duplications. The experiments led us to select *Wratio* (averaged weighted of four metrics). To ensure high accuracy, we experimentally tuned the compatibility threshold to 86%. With GROBIT parsing, we achieved high accuracy in identifying the required data, as verified through a manual audit. Additionally, we explored the potential of LLMs through an experimental fine-tuning of the Phi-3 model (using Qlora technique) achieving 80-86% accuracy.

Moreover, we supported GROBIT with additional programs to interpret identified metadata, correct errors, filtering and unifying data. Special attention was paid to preprocessing conference and journal names, which are often specified differently (e.g. full names, shortened names, acronyms). Therefore, we deleted some words in publication specifications, such as "international", "proceedings", the conference serial number.

Sometimes the extracted original data on references are erroneous or not complete, e.g. lacking publication date, lacking authors, lacking information qualifying publication class (denoted as others). Fortunately, they constitute small percentage and can be analyzed manually if required. The final

lists of derived and labeled publications have been verified manually to eliminate some deficiencies, e.g. skipped duplications (0.2-1%). This verification is relatively simple and can be automatized in the future.

An important issue in publication evaluation is analysis of cited references. Here, arises the problem of extracting the cited references and deriving their characteristic features. Typically, references are included in a separate publication section However, some papers provide also additional reference list included in appendices (e.g. [25]). or in dedicated repositories [7] available via specified links. Moreover, they can appear in footnotes or in the main paper text. Hence, extracting cited references from the pdf file of the publication creates some problems needing more complex extraction algorithms. Creating complete list of cited references we have to eliminate appearing duplications.

It is worth noting that diverse editors use various formats of cited references, e.g. authors can be specified in the form: the first, middle and last name (or in other order), the first names quite often are restricted to the initial letters. The author list is followed by the publication title (sometimes in apostrophes), venue (e.g. the name of the journal or conference), and other features (e.g. volume, doi number). The publication venue can be specified using the original name (e.g. journal, conference, arXive number) or using acronyms. The publication date can appear just before the title or at the end of the whole specification (sometimes in brackets). All the included specifications can be distinguished with diverse separators (e.g. commas, semicolons, etc.). Subsequent references in the list can be explicitly numbered (usually in square brackets) or simply specified in separate paragraphs. Some editors (e.g. IEEE) use well defined formats, however in general we observed big diversity and even some anomalies, or errors. In our approach we consider all these diversities and create the list of references in a unified form stored in the excel file with columns attributed to the extracted elements (authors, title, venue, etc.) or JASON files. They are useful for further processing and filtering.

Publication references can be analyzed in three dimensions: individual, aggregated and comparative view. The individual view presents reference features for the specified publication; the aggregated view relates to specified set of considered publications (e.g. combined list of references included in these publications). The comparison view shows differences in cited references between two specified publications or sets of publications. Most advanced processing is needed in aggregated and comparative views. The derived basic reference features are illustrated in Section IV.

IV. Reference Analysis

Reference analysis is illustrated for three classes of software logging publications: general reviews, problem oriented surveys and primary publications. In [46], we have studied seven general surveys published in high ranked journals (5) and conferences (2). Rows of Table II present basic features of cited references in the specified publications. The i-th column corresponds to the considered publication or a set of publications (Pub_i) and provides derived features: the number of cited references (N), the range of publication dates (D) of included references in Pub_i, the number of common references (C) in publication Pub_{i+1} and Pub_i, the number of unique references (U) cited in publication Pub_{i+1} as compared with Pub_i. Sometimes, redundant references appear in publications (N > C + U), e.g. papers [6], [7] and [8] comprised 1, 5 and 2 repeated references, respectively. The extended version of the comparison table includes lists of common and unique references (Excell form).

Subsequent columns in Table II represent general surveys correlated with highly influential authors (high rank publication locations). Subsequently compared papers (ordered chronologically) showed relatively low number of common publications (C/N in the range 5-28%). This results from the assumed paper selection schemes applied in these surveys (Section III), they were specified by diverse search strings, inclusion/exclusion criteria, publication repositories. Typically, in surveys the initial selection produced rich sets of publications (e.g. several thousand) which were further screened by the authors (subjective analysis based on title and abstract analysis). Final full text verification limited this set typically to 40-200 references considered as significant and representative. Hence, this reduction resulted in high dispersion of analyzed papers.

TABLE II
REFERENCE FEATURES FOR GENERAL SURVEY PUBLICATIONS

	[6]	**[7]**	**[8]**	**[9]**	**[5]**	**[1]**	**[4]**
N	41	58	137	125	205	57	227
D	1987 2015	1989 2016	1995 2021	1983 2020	1972 2020	1987 2021	2004 2023
U	40	48	120	100	157	47	218
C	0	5	15	25	46	10	9

The low percentage of common references suggests creating aggregated sets od references covering the specified set of publications which can provide a wider view on the analyzed research domain/problem. Combining references from the seven general surveys (listed in Table I) we have created set Sum1 which comprised 614 unique references (dated 1972-2023). The original survey developed in our institute [46] (covering 146 survey and primary publications related to 2003-2025) compared with Sum1 references showed 107 unique new references. Aggregating references of our survey [46] with Sum1 set we obtained set Sum2 comprising 722 references (published in 1972-2025) which can be considered as well screened and significant. This set can be submitted to a deeper analysis. The aggregation process (Section III) resulted in elimination of about 30% redundant papers as compared with simple reference concatenation of the considered surveys.

Most survey papers ignore preprint references (e.g. published in arXives), which may be interesting. We have found an interesting survey [52], published in 2022, and not referenced by other papers. In fact, it is relatively long (34 pages) and well organized survey with an original taxonomy. Comparing this survey with highly ranked survey [4] we identified 180 new unique references within 257 cited ones (1972-2021). Comparing it with our survey [46] we identified 223 unique papers, comparing with set Sum2 we identified 120 unique ones. So, it is reasonable to include it into an extended repository of publications on software logging for further considerations.

TABLE III
COMPARISON OF REFERENCE FEATURES OF AGGREGATED PROBLEM ORIENTED SURVEYS WITH SUM2

	Sum2	**PAR**	**AD**	**SEC**	**AL**
N	**722**	396	552	174	661
D	**1972** **2025**	1966 2024	1988 2024	1987 2020	1946 2024
U		219	442	123	493
C		177	110	51	168
u		0.30	0.62	0.17	0.69
c		0.25	0.15	0.07	0.24

We have performed similar analysis for problem oriented surveys listed in Table I. The derived aggregated sets of references corresponding to these surveys are denoted as follows: parsing PAR = {[10, 11. 12, 13, 14, 15, 16, 17, 18, 19]}, log analysis LA = {[20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31]}, security SEC = {[32, 33, 34, 35]}, and anomaly detection AD = {[36, 37, 38, 39, 40, 41, 42, 43, 44, 45]}. An interesting issue was to compare references between different types of surveys. Some comparison results are shown in Table III. It presents comparison of aggregated references corresponding to specified publication sets compared with Sum2 reference list (aggregated Sum1 and [46]) – the first highlighted column. Here, we have included two additional parameters: ratio of new (u) and ratio of common (c) references in the ith column as compared with the set of the first column (Sum2); u=Ui/N0. c = Ci/N0, where N0 is the number of references in set Sum2, Ui, Ci parameters related to compared i-th column reference sets .

Table III shows significant percentage of unique (not covered by Sum2 list) references comprised in problem oriented (aggregated) lists (U/N in the range 55-83%). This confirms high reference orthogonality between problem oriented and general surveys. Similar comparisons between problem oriented surveys showed even higher difference (74.5-92.6 %). Hence, problem oriented surveys provide really deeper insight into the considered problems. Table IV presents comparison of aggregated sets of references between problem oriented surveys. The results are presented in three comparison groups. Aggregated sets in subsequent columns of the group are compared with the aggregated reference sets attributed to the group (bolded). Here, parameters u and c are referred to the number of references (N0) included in the comparison group: PAR – 396, AD – 552, SEC – 174.

TABLE IV
COMPARISON OF REFERENCE FEATURES
BETWEEN PROBLEM ORIENTED SURVEYS

	PAR			**AD**		**SEC**
	AD	SEC	LA	LA	SEC	LA
N	552	174	661	661	174	661
D	1988 2024	1987 2020	1946 2024	1946 2024	1987 2020	1946 2024
C	464	120	571	600	144	612
U	88	54	90	61	30	49
u	1.17	0.3	1.44	1.09	0.26	3.52
c	0.22	0.14	0.23	0.11	0.05	0.28

Comparing references between primary papers, related to the same topic, facilitates identifying similarities and differences in referenced publications which is useful in publication evaluation (e.g. originality, considered research scope, topicality). For example comparing two primary papers on parsing [53] and [54] published in 2025 differed in 81% references, in fact they present diverse parsing methods (based on log zip vs token conversion). Papers [55, 56, 57], published in 2023, 2024, 2025, respectively, related to anomaly detection comprised 81-87% different references.

V. PUBLICATION FEATURES

The developed publication analysis frameworks provide the capability of automatic data exploration in various perspectives: e.g. appearance, venues, author features, research topics.

A. Publication appearance

An interesting issue is deriving distribution of paper appearance in time perspective, for example their appearance for subsequent years. Such plots are provided in many SRLs. We derived them with developed tools (PubAnlyzer and PdfAnalyzer). Fig. 1 presents distribution of four publication types (journals, conferences, etc.). The derived publication distribution for Sum2 publications showed that the number of published papers increased over time, however some slowdown was observed in recent five years. This slow down results from the fact that references in Sum2 were extracted from the survey papers published mostly before 2023. Such analysis presented by other authors of selected publications surveys also showed some decrease one year before the publication date (e.g. [52]). Moreover, recent primary publication are not always registered in publication repositories (delay effect). This effect was also observed in similar plots related to aggregated sets of publications on problem oriented surveys. Recent surveys within the set Sum2 are in minority so in total they comprise lower percentage of newer publications and the older references appear in several older surveys.

In Fig. 2 we give distribution of published papers in journals and conferences for S1016 set comprising selected papers on log analysis (Section III). Publications in conferences dominated (72%). Lower percentage of conference papers (65%) was identified within 110 screened publications in [52], in the case of all aggregated sets considered in Section III it was 58%. This results from the fact that usually authors of survey papers pay higher attention to journal papers in the screening process. This was also confirmed in Fig. 1 related to the set of publications in Sum2 (aggregated references of general surveys).

Fig. 3 gives distribution of citations for publications from S1016 in relation to their publication years. Citations not exceeding 9 relate to a significant percentage of papers: 72% (journals) and 80% (conferences). The presented plots show the average citations of journal (av-jour) and conference (av-conf) papers, as well as summarized citations for all papers published in the specified year within journals (sum-jour) or conferences (sum-conf). We observe some fluctuations, they result from sporadic appearance of highly cited paper in the relevant year, For the last 4 years, the citations decrease is natural. A deeper study we presented in [46].

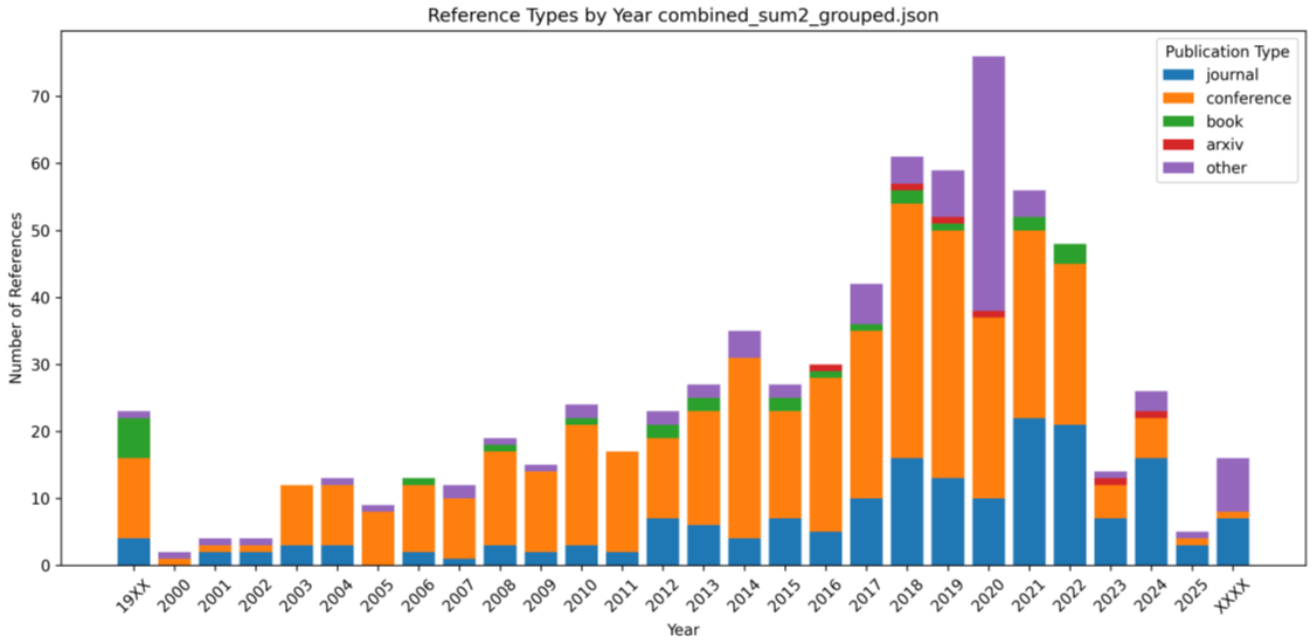

Fig. 1. Distribution of publications within general surveys (set Sum2)

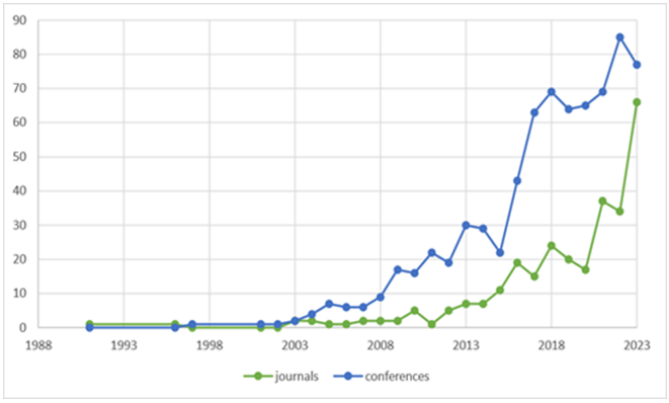

Fig. 2. Distribution of publications within set S1016

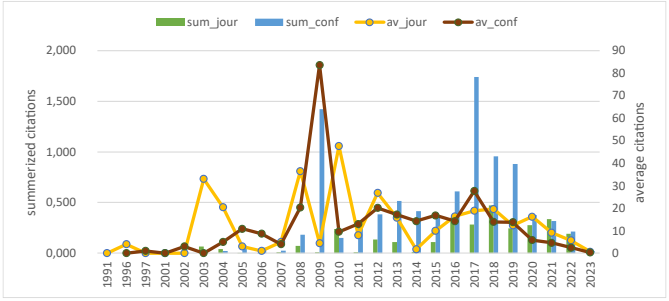

Fig. 3. Distribution of citations for publications in S1016

We analyzed the ranking of the most cited publications and authors for S1016. To outline the dynamics of citations (increasing over time), we provide the numbers of citations for 2024, 2025 and 2026. Moreover, another metric of paper interest is the number of text views. For 10 highly cited papers (published in 2009-2017) on logging, we found citations in the range 146-996 (at the end of 2024). Statistics for the three most cited papers were as follows:

[58] – citations: 876, 1055, 1425; views 24953, 32867,
[59] – citations: 753, 894, 996; views 4263; 5791;
[20] – citations 370, 467, 725; views: 5800, 7404

Citation numbers correspond to years 2024, 2025 and March 2026, respectively. The number of reported views relate to the end of 2024 and March 2026.

Within the considered set of publications S1016, we identified 77 editors. Top eight editors published 189 (66%) journal papers and over 600 (82%) conference papers. Only nine editors published more than six papers, for 45 editors, only a single paper was attributed. Hence, we can focus our interest on those most productive. The top five editors with the highest sum of citations (aggregated over all papers on logging published by the specified editor) were: IEEE (429), ACM (104), Elsevier (59), Springer (135), and MDPI (16). We give more details in [46]. Another interesting aspect is distribution of publications on their locations (venues) correlated with their rankings. We discuss this in subsequent subsections.

B. Publication venues and positions

We can identify publication locations (journal, conference names and others) using developed tools (Section III). In the above mentioned set S1016 (roughly screened) we have identified 617 sources, however only 80 (13%) contributed at least 3 publications, 45 (7.2%) related to 4-17 publications, most (73%) provided 1 paper. Locations with the highest number of relevant publication are worth systematic observation.

More representative distribution of publication venues can be derived from the set of well screened paper selection related to derived aggregated sets of publications from representative surveys (Section IV). We provide the distribution of publications in 15 most representative locations derived from the aggregated set ALL (combined sets: Sum2, PAR, AL, AD, SEC and [52]) which covers a wide spectrum of publications analyzed in recognized surveys (Section IV):

Journals

1. IEEE Access - 34
2. IEEE Trans. on Software Engineering (TSE) - 28
3. Information and Software Technology 18
4. Applied Sciences (Switzerland) - 11
5. Journal of Systems and Software - 10
6. IEEE Trans. on Net. and Service Manag. (TNSM) - 9
7. ACM Computing Surveys - 8
8. IEEE Trans. on Dependable and Secure Computing - 6
9. Computers & Security - 6
10. Lecture Notes in Computer Science (LNCS) - 6
11. Automated Software Engineering - 5
12. Communications of the ACM- 5
13. Applied Intelligence - 4
14. IEEE Trans. on Knowledge and Data Engineering - 4
15. Sensors - 4

Conferences

1. Int. Conf. on Automated Software Eng. (ASE) – 45
2. ACM/IEEE Int. Symp. on Empirical Software Engineering and Measurement (ESEM) - 35
3. Int. Conf. on Software Engineering (ICSE) - 28
4. Joint European Soft. Eng. Conf. and Symp. on the Foundations of Software Engineering (FSE) - 28
5. Int. Symp. on Soft. Rel. Engineering (ISSRE) - 21
6. Int. Conf. on Big Data (Big Data) - 18
7. Int. Conf. on Know. Discovery and Data Mining - 16
8. Int. Conf. on Soft. Maint. Evolution (ICSME) - 14
9. Int. Conference on DataMining (ICDM) - 13
10. IFIP Int. Conf. on Depend. Syst. and Net. (DSN) - 13
11. Int. Conf. on Cluster Computing (CLUSTER)- 11
12. Int. Conference onWeb Services (ICWS) - 11
13. Int. Parallel and Distributed Proc. Symp. (IPDPS) - 10
14. SIGSAC Conf. on Computer and Communications Security - 7
15. USENIX Symp. on Operating Systems Design and Implementation (OSDI) - 7

Similar distributions are also encountered in individual survey papers. They may differ due to variations on selected papers for the analysis (individual author preferences), survey problem focus (higher differences were observed for security related publications), etc. Bigger differences relate to locations attributed to lower number (1-5) of published articles. In the considered set of publications ALL, the total number of identified journals and conferences is quite big: 477 unique conferences (in 214 only single papers were published), 360 unique journals (205 single papers). It is much more than reported in individual surveys, e.g. in [52] authors analyzed 112 papers from 48 venues (18 covered 2-13 publications).

Due to the large number of publication locations (venues) the search process is difficult and should be supported with other paper attributes specified in the search process. Restricting searches to locations with most frequently appearing publications can skip some interesting publications (we reported some examples in [46]). Moreover, we should be conscious of journal/conference changes in the scope of considered topics in time. Sporadically, new publication locations appear due to splitting the research areas or including new ones (e.g. AI, LLM) as well as extending the scope of conferences.

There are also other prestigious journals and conferences that provide good papers in the range 1-6. This statistic confirms a large dispersion of papers across many sources, which may have diverse ranks. It is reasonable to assess these ranks. The significance of the publication source (venue) can be evaluated using various metrics, including impact factor (IF) or cite score for journals and core rank for conferences. Of the 730 conference papers in S1016 only 220 (30%) were classified in CORE with the distribution: A* - 11%, A – 24%, B – 24%, C – 27%, and other (not ranked) 14%. Most journals do not have IF, but among those with higher IF are ACM Computing Surveys (16.6, 3), IEEE Trans on Knowledge and Data Eng. (8.9, 3), IEEE Trans. on Software Eng. (7.4, 10), IEEE Trans. on Dependable and Secure Computing (7.3, 4), Computers and Security (5.6, 6), IEEE Trans on Network and Serv, Manag. (5.3, 11), Empirical Software Engineering (4.1, 9), IEEE Access (3.9, 17), and Jour. of Systems and Software (3.5, 6). In the brackets we provide IF and the number of identified papers.

Another metric can be the distribution of paper citations across sources. Considering sources with at least 100 summarized citations of relevant papers, we identified 30 sources (10 journals) with scores 100-1002 [46]. The number of papers attributed to each source ranged from 1 to17, most papers falling between 2 and 5. Statistics of citations per single paper in this set were as follows: median 1-753, Q3 quartile 5 - 753. High values were related to some venues providing single papers with high citation counts (e.g. ACM SIGOPS – 753, ACM SGMOD 247). This should be correlated with other statistics while searching papers.

C. Author features

Analyzing author features, we considered several aspects: the number of coauthors, affiliation country and institution, author citations, and author activity. Using PubAnalyzer and publication set S1016 we found that most papers were coauthored (median number of coauthors 4). Single-author papers constituted a small percentage 4.4% (4 journals, 40 conferences). For 83% of journal papers, the number of coauthors was between 2 and 5, 85% of conference papers had from 2 to 6 coauthors. The maximum number of coauthors was 11 and 17 for journal and conference papers, respectively. Papers with coauthors from the same affiliation constituted 32.2% and 46.7%, while the number of affiliations within the range 2-5 constituted 65.4% and 50.6%. The maximum number of affiliations ranged up to 6 and 8 in journal and conference papers, respectively. Therefore, a significant percentage of papers were contributed by collaboration of two or more institutions.

Another view is the statistics of countries of affiliations. Authors of the publications in S1016 were associated with 69 countries (over 27 countries contributing 10 or more papers). Dominated China followed by United States , India, Canada, Japan, South Corea, and 11 European countries. Most papers were attributed to larger countries; however, the citation indexes of China and India were relatively low. Similar statistics we derived at the level of institutional affiliations [46].

In S1016 we identified 2394 authors. Top 200 authors collected 50-989 citations of all their included papers on logging, 1084 authors collected 1-10 citations and 585 published not cited papers on logging. Citations give some view on publication popularity, those with high values can be used in author snowballing. The author's contribution can be also assessed in the number of coauthored papers. 32 authors

produced 10-27 papers, 93 authors 5-9, 1797 produced single papers. Similar analysis performed with RefAnalyzer applied to the set of aggregated references in considered surveys ALL (compare Section B) can be considered as most influential, and they indicated 1114 authors who published 2 or more papers (3259 authors published single papers). The 20 most active authors published from 22 to 59 papers (H. Zhang 59, M. Lyu 34). Comparing this list with similar one (most cited papers) derived from S1016 papers we found only 5% of overlapping authors. It was also similar percentage with the list of most productive authors in S1016. This confirms that high citation numbers are not directly correlated with the authors with the highest number of papers, some authors produced only a single highly cited paper, Huang L, Jordan M., Patterson K. authored the same highly cited (753) paper published in 2009 on console logs [59]. Zhang H. ranked on 9-th position (842 citations) authored 27 papers (in 2024 [46]). The list of authors who published more than 3 papers cited in more than 40 times comprised only 8 authors.

We can also analyse correlations between most cited authors and affiliations within publications on logs. We have identified 19 affiliations with 5-10 authors and aggregated citation contributions 557-989, e.g. Microsoft Research (10, 557-989), University of Hong Kong Shenzen (9, 568-989), University of Newcastle (6, 637-989), Univ. of Illinois (5, 568-949). Correlating authors with the highest number of contributed papers (on logging) with affiliations we identified 20 affiliations (among 200) with at least 3 authors, e.g. Beijing Univ. Of Posts and Telec. (8, 115-464), Microsoft Research (5, 254-842), Concordia Univ. (4, 115-842), Tsingua Univ. (4, 154- 842). Such statistics can facilitate the selection authors or affiliations for the paper snowballing or adjusting the monitoring of future publications. The most productive author was H. Zhang (27 publications, 842 citations - 31.2 average per paper). The most cited authors were M. Lyu and J. Lou: 989 and 949, respectively (cited publications on logging).

Additionally, we can derive author activity period - the time between the first and the last publication. Here, we should consider authors of 2 or more papers (25%). Short activity (1-4 years) corresponded to 53.1% of authors, medium activity (5-8 years) – 25.5%., remaining 9-21 years. Another aspect is the average time span between subsequent papers: for 38.9% of authors, this was no more than 2 years; 2-4 years for 37.2%; and 4-6 years for 20.3% authors. This confirms that there is a group of authors active in logging problems. This can be further analysed to identify problem classes associated with these authors.

The presented statistics are useful to restrict the number of traced authors to those with the highest publishing or citation indices. Here, we can also use cross sectional statistics, e.g. correlating results with specified publication time span, high rank affiliations, high rank publication locations (venues, specified topics).

D. Paper topic profiles and attributes

The presented statistics on publications and authors are useful for paper searches (domain, publication venue, specified characteristic features/attributes), identification of problem trends, activities of authors, relevant institutions, and research trends. Publication assessment needs extracting content related features, e.g. keywords.

Keywords provided by the authors are helpful for topic classification. Having analysed keywords related to S1016 papers considered in [46], we have identified 2547 different keyword phrases, 536 appeared at least in two publications. Many keywords included several words e.g. phrase “log analysis” appeared in 44 papers, however in combination with other words appeared in 317 papers. These longer keywords were more descriptive, e.g. Event log analysis (7), System log analysis (3), Automated log analysis (3), Security log analysis (3), Network log analysis (3), Web log analysis (3), Semantic log analysis (2) - with specified frequencies in brackets. Log parsing phrase appeared in 10 papers, however 63 keywords comprised log parsing within all keyword sequences Nevertheless, in many cases keywords are coarse grained. In some publications, we can find additional editor keywords, in fact they are more general (coarse grained: e.g. software engineering, computer methodology).

To find search strings characterizing publications on specific topic class we derived n-grams based on tf-idf (using PubAnalyzer). Having classified the set S1016 papers in 8 topic classes (anomaly detection, log parsing, security, anomaly prediction, anomaly diagnosis, log compression, survey paper, machine learning), we searched most representative n-grams. For each topic class we searched n-grams generated within paper titles, abstracts and keywords. We found that topic correlated phrases appeared in titles relatively scarcely, For anomaly detection, log parsing and security topics we obtained the following phrase coverage (x/y where x- the number of identified phrase appearance within y publications related to the considered topic): anomaly 161/248; parsing 38/116, parser 18/116; security 72/389, here we identified many other correlated phrases: threat 13, attack 25, malicious 9, forensic 42, intrusion 13. Better coverage assured checking abstract: anomaly 219/248; parsing 89/116; security 255/359. Checking keywords, we identified dispersed several phrases, the single ones directly specifying topic were relatively high but lower than in abstracts. However, some basic set of keywords (3-4) assured quite big coverage, e.g. for security 3 keyword phrases assured 269/389 coverage; log parsing 79/116. For anomaly diagnosis 4 keywords assured 23/81 coverage, as opposed to 2 abstract phrases 55/91 and 3 phrases (root cause, cause analysis, diagnosis) within titles 34/81.

We performed deeper analysis of author keywords for the set S92 (selected from the set S1016) comprising high ranked 92 papers on logging cited 25 and more times, which we analysed in detail considering full texts. The analysed data comprised basic features as outlined in Section III (title, abstract, keywords, numbers of citations, references, and pages) along with additional features extracted manually from the publication text: paper section and subsection titles, captions of figures, tables, code snippets, distinctive text snippets, etc. The developed analysis scripts allowed us to evaluate the impact of these extended paper features on paper assessment. Having analysed attributed keywords by authors, we identified 292 diverse keyword phrases mostly 2 and 3 word sequences, e.g. log analysis (38), anomaly detection (20), log parsing (13), deep learning (8), data mining (5), log (5), and log management (4). Another problem was checking the distribution of keywords correlated with publication classes. For three classes of papers we have got the following statistic of characteristic keywords: anomaly detection (24, 70/133), log parsing (15, 39/56), surveys

(14, 52/67), where the first number in the brackets denotes the number of considered papers, x/y specifies the number of unique keywords x (they were informative - typically 1-3 word sequences) versus the number of their appearance y. On average we identified 4-5 keywords per article. In the case of survey papers high topic dispersion was observed in keywords (some general phrases). For illustration, we present an excerpt of keyword distribution for anomaly detection publications (with specified frequencies in brackets): anomaly detection (17), log analysis (7), log parsing (3), lstm (3), error detection (2), cnn (2), cloud operation (2), abnormal task (1), anomaly identification (1), anomaly classification (1), event log forensic (1), execution anomaly detection (1), log anomaly (1), log instability (1), log sequence anomaly detection (1), trace anomaly index (1).

Depending upon the publication, the provided author keywords were more or less informative, however usually too general. e.g. log analysis, data mining, data quality, and software log. Some keywords related to the specificity of methods, e.g.: graph neural network, hierarchical transformers, log structured merge tree, experimentation, and clustering. Using PubAnalyzer, we tried to find more descriptive phrases by deriving n-grams, (n=1-3), ordered according to tf-idf and NPMI parameters. Derived n-gram sets were more extensive, and the calculated NPMI and tf-idf parameters allowed us to reduce these sets by selecting those with the higher values. In this way, we analyzed up to several hundred n-grams for each paper category and identified the most interesting ones that could be used in deeper searches, e.g., parse tree, parse accuracy, case study, feature extraction, parsing error, and route cause analysis. Checking those with higher parameters, we found that they provided more detailed/technical information compared to the author's keywords. Unfortunately, at first glance many of them were irrelevant, so manual selection was needed to extract interesting ones. In general, they related to the efficiency of the proposed solutions (e.g. time, performance, effectiveness, memory utilization, accuracy, precision, evaluation, and parameter tuning effort), practical aspects of the research (running example, comparison, industrial logs), types of resolved problems and used methods (classification, parsing, anomaly detection, false alarm rate, support vector machine, hidden Markov model, recurrent neural network).

Paper significance can correlate with the number of citations and text views. However, it is also important to trace dynamics of these parameters in time. Recently published papers usually have low values of these parameters but can be interesting. Some additional metric of paper potential importance is the rank of publication location, author affiliation, author reputation, and activity features (compare Section C).

Beyond citation we can also monitor other paper content features. The basic ones are the number of included references, (4, 205), text pages (4-49), tables (0-19), figures (0-30). In the brackets we give the minimal and maximal numbers derived from the 92 highly cited papers (25-876). Papers with a higher number of illustrative elements seem to be more attractive, and worth further review; however, we did not find significant correlations with the number of citations. For example a very good survey published in Arxive [52] based on 230 reference provides a deep insight in multivariate taxonomy comprised informative section titles, and multi sentence captions of tables and figures, which characterize valuable contribution (other examples we give in [46]). These characteristic features have been extracted from the contents of pdf documents. Collecting such excerpt from the paper provides well-thought out phrases/sentences characterizing substantive contributions. They can be submitted to NLP analysis to improve assessment of paper content, e.g., linear and circular logging, classification of cloud log forensics, logging libraries and utilities, log verbosity level and description prediction, log statement automation, logging code progression, mining log files for different applications, logging habits, application-specific logging, representation of log files, log summarization and visualization, quality of logging statements, metrics for log prediction, user statistics and behavior, code coverage with logging statements, business decision-making, log feature vectors, system runtime behavior.

In view of the presented features and relevant statistics, we should be aware that beyond highly ranked publications attributed to recognized venues (papers with high IF, highly cited or observed authors) there are many papers worth tracing, especially recent publications or those by new authors. These include papers proposing innovative solutions, applied to emerging systems, proposing incremental contributions, etc. By analyzing reference sets, we can examine the distribution of publication dates, relevant topics, venues, and so on. Recent references provide more up-to-date view of the logging issues considered. In the derived publication sets on logging (Section IV) references reported in 2020 or later constituted a significant percentage: Sum1 (25.6%), [46] (60.7%), [4] (49.3%), Sum2 (31.6%), ALL (35.6%). They can gain higher priority in deeper exploration.

Sometimes, interesting papers appear in unexpected venues. Hence, deeper analysis of publication texts are needed correlated with the gained knowledge from earlier searches and problem taxonomy in surveys. The proposed tools and provided statistics are helpful in this process. Deeper insight into considered problem or presented solutions can be supported by tracing additional features, e.g. derived from figure, table and algorithm captions.

VI. Conclusion

Creating, maintaining and using logs is widely discussed in numerous surveys and primary papers. Here, the problem arises of understanding the current state of software logging in research and practice, as well as finding solutions or suggestions targeted at encountered problems. To facilitate efficient studies of relevant literature, we propose multidimensional searches based on meta-analysis covering a broad scope of publication features. This process is supported by careful publication selection (screening, filtering) based on factors indicating meaningful and substantive contents. To a large extent, they result from studies presented by influential and experienced authors with recognized achievements (contributions).

The developed methodology facilitates adapting literature search to review intent, such as grasping the knowledge on logging (introductory or advanced level), tracing specified topics, identifying research progress and challenges, finding correlations with developed projects, choosing optimal venue for submitting publications, supporting reviewing process (literature relevance), and establishing contacts for collaboration. This is supported by the introduced publication

taxonomy (coarse and fined grained), developed tools, and optimized search strings derived from advanced keyword and n-gram exploration. These concepts are illustrated with many results and referenced in our comprehensive SLR [46]. This approach can also serve as a foundation for integrating with LLM models and AI methods. Additionally, our approach can be applied to surveys in other technical domains.

REFERENCES

[1] S. Gu, G. Rong, H. Zhang, H. Shen, "Logging practices in software engineering: A systematic mapping study", IEEE Trans. on Softw. Eng., vol. 49, no. 2, pp. 902-923, 2023.

[2] B. Kitchenham, L. Madeyski and D. Budgen, "SEGRESS: Software Engineering Guidelines for REporting Secondary Studies," in IEEE Transactions on Software Engineering, vol. 49, no. 3, pp. 1273-1298, 1 March 2023, doi: 10.1109/TSE.2022.3174092.

[3] M. J. Page et al., "PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews," Brit. Med. J., vol. 372, art. no. n160., 2021, [Online]. Available: https://www.bmj.com/content/372/bmj.n160

[4] M. A. Batou, M. et al., "A literature review and existing challenges on software logging practices", Empirical Software Engineering, 29: 103, pp. 1-63, 2024, doi: 10.1007/s10664-024-10452-w

[5] S. He, P. He, Z. Chen, T. Yang, Y Su, M. R. Lyu, "A survey on automated log analysis for reliability engineering", ACM Computing Surveys, vol. 54, no. 6, pp. 1–37, July 2022.

[6] P. He, J. Zhu, S. He, J. Li; M. R. Lyu, "An Evaluation Study on Log Parsing and Its Use in Log Mining", In Proc. 46th Annual IEEE/IFIP Int. Conf. on Dependable Systems and Networks (DSN), pp. 654–661, 2016. doi: 10.1109/DSN.2016.66.

[7] G. Rong, Q. Zhang, X. Liu, S. Gu, "A systematic review of logging practice in software engineering", In Proc. 24th Asia-Pacific Softw. Eng. Conf., (APSEC), 2017, doi: 10.1109/APSEC.2017.61.

[8] B. Chen, Z. M. Jiang, "A survey of software log instrumentation", ACM Computing Surveys, vol. 54, no. 4, July 2021, ISSN: 0360-0300. doi: 10.1145/3448976.

[9] J. Candido, M. van Deursen, A. Aniche, "Log-based software monitoring: A systematic mapping study", PeerJ Comput. Sci., vol. 7, 2021, Art. no. e489, doi: 10.7717/peerj-cs.489.

[10] J. Zhu, S. He S.; J. Liu, P. He, O. Xie , Z. Zheng, M. R. Lyu, "Tools and benchmarks for automated log parsing", In Proc. IEEE/ACM 41st Int. Conf. on Softw. Eng.: Softw. Eng. in Practice (ICSESEIP), pp. 121-130, 2019. doi: 10.1109/ICSE-SEIP.2019.00021.

[11] D. El-Masri, F. Petrillo, A. Hamou-Lhadj, A. Bouziane, "A. systematic literature review on automated log abstraction techniques", Inf. Softw. Technol., vol. 122, Mar. 2020, Art. no. 106276.

[12] S. Lupton, H. Washizaki, N. Yoshioka, Y. Fukazawa, "Online log parsing: Preliminary literature review", In Proc. IEEE Inter. Symp. Softw. Rel. Eng. Workshops (ISSREW), pp. 304–305, 2021.

[13] Y. Fu, M. Yan; J. Xu, J. Li, Z.; Liu, X. Zhang, D. Yang, "Investigating and improving log parsing in practice", In*:* Proc. ACM Joint European Softw. Eng. Conf. and Symp. on the Foundations of Softw. Eng., pp. 1566–1577, 2022. doi: 10.1145/3540250.3558947

[14] Y. Fu, M. Yan, Z. Xu, X. Xia, X. Zhang, D. Yang, "An empirical study of the impact of log parsers on the performance of log-based anomaly detection", Empirical Software Engineering, vol. 28, no. 1, 2023, doi: 10.1007/s10664-022-10214-6

[15] T. Zhang, H. Qiu, G. Castellano, M. Rifai, C. S. Chen, F. Pianese, "System log parsing: A survey". IEEE Trans. on Knowledge and Data Eng. vol. 35, no. 8, pp. 1–20, 2023. doi: 10.1109/tkde.2022.3222417.

[16] V-H. Le, H. Zhang, "An evaluation of log parsing with ChatGPT", June 2023, doi: 10.48550/arXiv.2306.01590

[17] P. Mudgal, R. H. Wouhaybi, " An assessment of ChatGPT on log data", In AI-generated Content. Int. Conf. AIGC 2023. Communications in Computer and Information Science, vol 1946. Springer, 2024, doi: 10.1007/978-981-99-7587-7_13

[18] Y. Zhou, Y. Chen, X. Rao, Y. Zhou, Y. Li, Ch. Hu, "Leveraging large language models and BERT for log parsing and anomaly detection", Mathematics, *12*(17), 2758; 2024.

[19] Z. Jiang, et al., "A large scale evaluation for log parsing techniques, how far are we?", In Proc. 33rd ACM SIGSOFT Int. Symp. on Softw. Testing and Analysis (ISSTA Technical Papers); pp. 223 – 234, 2024.

[20] P. He, J. Zhu, Z. Zheng, M. R. Lyu, "Drain: An online log parsing approach with fixed depth tree", Proc. IEEE Int. Conf. on Web Services (ICWS), pp. 33–40, 2017. doi: 10.1109/icws.2017.13.

[21] V. Bushong, et al., "On matching log analysis to source code: A systematic mapping study", In: Proc. Int. Conf. on Research in Adaptive and Convergent Systems, pp 181–187, 2020. doi: 10.1145/3400286.3418262

[22] D. Das, et al., "Failure prediction by utilizing log analysis: A systematic mapping study", In Proc. Int. Conf. Res. Adapt. Convergent Syst., pp. 188–195, Oct. 2020. doi: 10.1145/3400286.3418263

[23] E. Mendes, F. Petrillo, "Log severity levels matter: A multivocal mapping", In Proc. IEEE Int. Conf. on Softw. Quality, Reliability and Security (QRS), pp. 1002–1013, 2021.

[24] J. Raffety, B. Stone, J. Svacina, C. Woodahl, T. Cerny, P. Tisnovsky, "Multi-source log clustering in distributed systems", Lecture Notes in Electr. Eng. Information Science and Applications, pp. 31-41, 2021. doi: 10.1007/978-981-33-6385-4_4

[25] Ł. Korzeniowski, K. Goczyła," Landscape of automated log analysis: a systematic literature review and mapping study". IEEE Access, vol. 10, pp. 21892-21913, 2022. doi: 10.1109/access.2022.3152549

[26] S. He, S.; et al., "An empirical study of log analysis at Microsoft", In Proc. 30th ACM Joint European Softw. Eng. Conf. and Symp. on the Foundations of Softw. Eng., pp. 1465–147, 2022.

[27] E.Chuah, A. Jhumka, M. Malek, N. Suri, "A survey of log-correlation tools for failure diagnosis and prediction in cluster systems", IEEE Access, vol. 10, pp. 133487 – 133503, 2022. doi: 10.1109/ACCESS.2022.3231454

[28] S. Silalahi, R. M. Ijtihadie, T. Ahmad; H. Studiawan, "A Survey on Logging in Distributed System", 1st International Conference on Information System & Information Technology (ICISIT*)*, Yogyakarta, Indonesia, pp. 7-12, 2022. doi: 10.1109/ICISIT54091.2022.9873095

[29] R. R. Abdalla, A. K. Jumaa, "Log file analysis based on machine learning: A survey", UHD Journal of Science and Technology, 6(2), pp.77–84, 2022. doi: 10.21928/uhdjst.v6n2y2022.

[30] P. Foalem, F. Khomh, A. Nguimbous, H.. Li,. "An empirical study on logging vvolution on Stack Overflow: trends, topics, and challenges", Nov. 2024, Available at SSRN: https://ssrn.com/abstract=5008961 or http://dx.doi.org/10.2139/ssrn.5008961.

[31] Y. Li et al., "Exploring the Effectiveness of LLMs in Automated Logging Statement Generation", An Empirical Study. IEEE Trans. Softw. Eng. vol. 50, no. 12, pp. 3188-3207, 2024.

[32] A. Tall, J. Wang, D. Han, "Survey of data intensive computing technologies application to security log data management", . In Proc. 3rd IEEE/ACM Int. Conf. on Big Data Computing, Applications and Technologies, pp. 268–273, 2016. doi: 10.1145/3006299.3006336

[33] J. Svacina, et al., "On vulnerability and security log analysis: A systematic literature review on recent trends", In Proc. Int. Conf. on Research in Adaptive and Convergent Systems, pp. 175–180, 2020.

[34] M. Landauer, F. Skopik, M. Wurzenberger, A. Rauber, "A. System log clustering approaches for cyber security applications: A survey", Computers and Security, vol. 92, 2020, ISSN: 01674048.

[35] M. Raut, S. D. Havale, A. Singh, A.; Mehra, "Insider threat detection using deep learning: A review", In Proc. 3rd Int. Conf. on Intelligent Sustainable Systems (ICISS), pp. 856-863, 2020.

[36] R. B. Yadav, P. S. Kumar, S. V. Dhavale, "A survey on log anomaly detection using deep learning", In *Proc.* IEEE 8th Int. Conf. on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), pp. 1215–1220, 2020. doi: 10.1109/ICRITO48877.2020.9197818

[37] A. Meenakshi, C. Ramachandra, S. Bhattacharya, "Literature survey on log-based anomaly detection framework in cloud", In Computational Intelligence in Pattern Recognition, Proc. CIPR 2020, Advances in Intelligent Systems and Computing, vol. 1120, Springer, pp. 143–153, 2020. doi: 10.1007/978-981-15-2449-3_12.

[38] S. Ranga, M. Nageswara Guptha, M. S. Hema, "A Survey on automatic abnormalities monitoring system for log files using machine learning", Turkish Online Journal of Qualitative Inquiry (TOJQI), Vol., Issue 6, pp. 8386-8393, July 2021.

[39] P. Raut, A. Mishra, S. Rao, S. Kawoor, S. Shelke, S.; M. Deore, V. Kumar, " Review on log-based anomaly detection techniques", In: Proc. of 2nd

Int. Conf. on Sustainable Expert Systems . Lecture Notes in Networks and Systems, vol. 351, Springer, pp. 893-906, 2022.

[40] M. Majd, P. Najafi, S. A. Alhosseini, S.A.; F. Cheng F.; C. Meinel, "A comprehensive review of anomaly detection in web Logs", In Proc. IEEE/ACM Int. Conf. on Big Data Computing, Applications and Technologies, pp. 158-165, 2022. doi: 10.1109/BDCAT56447.2022.00027

[41] J. M. Jose, S. R. Reeja, "Anomaly detection on system generated logs—A survey study", In: Mobile Computing and Sustainable Informatics. Lecture Notes on Data Engineering and Communications Technologies, vol 68. Springer, pp. 779-793, 2022.

[42] E. Karlsen, X. Luo, N. Zincir-Heywood, M. Heywood, "Benchmarking large language models for log analysis, security, and interpretation", J Netw Syst Manage, vol. 32, issue 3, paper 59, 2024. doi: 10.1007/s10922-024-09831-x

[43] S. Partovian, A. Bucaioni, F. Flammini, J. Thornadtsson, "Analysis of log files to enable smart-troubleshooting in Industry 4.0: a systematic mapping study", IEEE Access, vol. 12, pp. 147640-147658, 2024.

[44] Z. A. Khan, et al., "Impact of log parsing on deep learning-based anomaly detection", Empirical Software Engineering, 29, 139, 2024. doi: 10.1007/s10664-024-10533-w

[45] S. Lupton, H. Washizaki, N. Yoshioka; Y. Fukazawa, "Landscape and taxonomy of online parser-supported log anomaly detection methods", IEEE Access, vol. 12, pp. 78193-78218, 2024.

[46] J. Sosnowski, Ł. Reszka, "Software Logging: a Multidimensional Literature Survey", internal report, Inst. of Comp. Science, Warsaw University of Technology, 2026

[47] N. Madi, M. Binkhonain, "A systematic literature review on logging smell detection", Information and Software Technology, Volume 190, 107961, February 2026, doi: 10.1016/j.infsof.2025.107961

[48] E. Mendes, M. Vasconcellos, F. Petrillo, S. Hallé, "From description to prescription: Unraveling log severity adjustments in open-source software", Journal of Systems and Software Volume 231, 112643, January 2026, doi: 10.1016/j.jss.2025.112643

[49] M. De la Cruz Cabello, T. P. Sales, M. R. Machado, "AIOps for log anomaly detection in the era of LLMs: A systematic literature review", Intelligent Systems with Applications, vol. 28, art. 200608, December 2025, doi: 10.1016/j.iswa.2025.200608

[50] E. Chuah. H. Kalutarage, C. Maple, "A systematic literature review of log-correlation tools for cyberattack detection and prediction in large networks", Journal of Information Security and Applications, Volume 92, 104096, July 2025, doi: 10.1016/j.jisa.2025.104096

[51] Md Arfan Uddin, et al., "Microservice logs analysis employing AI: A systematic literature review", Journal of Systems and Software, Volume 236, 112786, June 2026, doi: 10.1016/j.jss.2026.112786

[52] S. Gholamian, P. A. S. Ward, "A Comprehensive survey of logging in software: From logging statements automation to log mining and analysis", 1 Jan. 2022, arXiv:2110.12489, doi: 10.48550/arXiv.2110.12489

[53] D. Gao, Ch. Liu, N. Chen, X. Hu, "LogGzip: Towards log parsing with lossless compression", Journal of Systems and Software, Vol. 223, 112349, 2025. doi: 10.1016/j.jss.2025.112349

[54] M. Wei; et al., "TCMS: A Multi-Sequence Log Parsing Method Based on Token Conversion". IEEE Transactions on Dependable and Secure Computing, vol. 22, no. 3, pp. 3028-3045, May-June 2025. doi: 10.1109/TDSC.2024.3520628

[55] J. Ye, Ch. Liu, Y. Zhang, "LogOW: A semi-supervised log anomaly detection model in open-world setting". Journal of Systems and Software, vol. 222, 112305 April 2025, doi: 10.1016/j.jss.2024.112305

[56] L. Yang, J. Chen, Sh. Gao, Z. Gong, H. Zhang, Y. Kang, H. Li, "Try with Simpler - An evaluation of improved principal component analysis in log-based anomaly detection". ACM Trans. on Softw. Eng. and Methodology, vol. 33, Issue 5, Article 115, pp. 1 – 27. June 2024. doi: 10.1145/3644386

[57] Y. Sun, J. Wai, K. Hi, K. You, "SemiRALD: A semi-supervised hybrid language model for robust anomalous log detection". Information and Software Technology, vol. 183 (6), 107743, July 2025, doi: 10.1016/j.infsof.2025.107743

[58] M. Du, F.; Li, G. Zheng, V. Srikumar, V. "DeepLog: Anomaly detection and diagnosis from system logs through deep learning". In Proc. CCS'17: ACM SIGSAC Conf. on Computer and Communications Security, pp. 1285–1298, 2017. doi: 10.1145/3133956.3134015

[59] W. Xu, L. Huang, A. Fox, D. Patterson, M. I. Jordan,. "Detecting large-scale system problems by mining console logs". In Proc. of 22nd ACM SIGOPS Symp. on Operating Systems Principles, (SOSP'09), pp. 117–131, 2009. doi: 10.1145/1629575.1629587