How Has the Health of Software Ecosystems Been Evaluated? A Systematic Review

Simone da Silva Amorim
Federal Institute of Bahia
Federal University of Bahia
Salvador, Brazil
simone.amorim@ifba.edu.br

Félix Simas S. Neto
Federal Institute of Bahia
Salvador, Brazil
felixneto@ifba.edu.br

John D. McGregor
Clemson University
Clemson, USA
johnmc@clemson.edu

Eduardo Santana de Almeida
Federal University of Bahia
Salvador, Brazil
esa@dcc.ufba.br

Christina von Flach G. Chavez
Federal University of Bahia
Salvador, Brazil
flach@ufba.br

ABSTRACT

The health of the software ecosystems concerns to the growing and continuity to exist remaining variable and productive over time. Research on this area is becoming more important. Even today, no studies have been available summarizing the research on evaluation approaches for the health of software ecosystems. The objective of this study is to structure and analyze the available literature on this field identifying the state-of-the-art of the research. We conducted a systematic literature review to obtain an overview of the existing studies in this area. 23 studies were selected as primary studies by applying inclusion, exclusion and quality criteria. The findings show that the research area is quite immature. There are few approaches and tools to support the evaluation work. In these studies, only 3 reported a complete evaluation of the health of ecosystems, 5 studies were considered as initial proposals, and the others evaluated the health partially.

CCS CONCEPTS

• **Software and its engineering → Operational analysis**; **Empirical software validation**; *Software evolution*; Open source model;

KEYWORDS

Software Ecosystems Health, Systematic Literature Review, Software Evaluation

ACM Reference format:
Simone da Silva Amorim, Félix Simas S. Neto, John D. McGregor, Eduardo Santana de Almeida, and Christina von Flach G. Chavez. 2017. How Has the Health of Software Ecosystems Been Evaluated? A Systematic Review. In *Proceedings of SBES'17, Fortaleza, CE, Brazil, September 20–22, 2017,* 10 pages.
https://doi.org/10.1145/3131151.3131174

SBES'17, September 20–22, 2017, Fortaleza, CE, Brazil

ACM ISBN 978-1-4503-5326-7/17/09…$15.00
https://doi.org/10.1145/3131151.3131174

1 INTRODUCTION

Software ecosystems have gained significant attention over recent years. There are several concepts for software ecosystems, for example Messerschmitt et al.[16] states that *"Traditionally, a software ecosystem refers to a collection of software products that have some given degree of symbiotic relationships"*. It is claimed that software ecosystems provide several benefits to organizations that adopt this approach. They increase the attractiveness for new users, accelerate innovation, share cost of the innovation, and decrease the cost of software maintenance by sharing this activity with third-party[3]. In this scenario many organizations are migrating to this approach. We can see the success of many open source and commercial organizations such as Hadoop, Eclipse, Apple, and so on. Hadoop market *"is growing at almost 55 per cent a year and is expected to be worth $20.9 billion in 2018"*[19]. As well as, we see an icon of the closed code, the giant Microsoft, opens their boundaries in recent years adopting the open source ecosystems approach[17]. Nowadays, Microsoft is on the top of the list of organizations with the most contributors in the open source approach[8].

In this context, where several organizations are adopting the software ecosystem approach, the health of the software ecosystems is a factor of extreme importance. Iansiti and Levien[10] introduced the idea of what a healthy ecosystem provides *"durably growing opportunities for its members and for those who depend on it"*. Evaluating the health is a fundamental activity considered as one key feature for organizations, partners, or even common developers. This allows to know the real health state for all stakeholders that desire to engage into the ecosystem. However, evaluating this health involves complex and challenging tasks. These tasks should also be supported by appropriate approaches, techniques, and tools.

Furthermore, there are no efforts to systematically analyze how the health of software ecosystems have been evaluated. There is not a systematic proposal to collect and synthesize existing health evaluation approaches. These approaches should investigate what benefits and limitations can contribute to improve the health of ecosystems. Besides, the research outcomes should be published to software ecosystem community. Several organizations can use a consolidated evaluation process to consider participating, and investing resources in the ecosystem. Moreover, the results of an effective health evaluation can help the organization to evaluate and compare platforms before basing new products on the assets available in the ecosystem. A clear and accurate health evaluation is an important input for investment decisions.

In face of the importance to get an effective evaluation process for the health of ecosystems, we decided to conduct a systematic literature review to know the scenario of existing approaches and to provide a picture of the state-of-the-art of the existing approaches. The outcomes reported in this work can be reference to support new evaluating approaches.

The remainder of the work is organized as follows: Section 2 presents some background and related work on the health of ecosystems. The objectives and methodology are described in Section 3. Section 4 describes the outcomes gathered from the systematic review. Section 5 presents the summary of study findings and the implications from research and practice. Section 6 describes the limitations of the methodology used. Finally, Section 7 offers a brief description of the outcomes and some concluding remarks.

2 BACKGROUND AND RELATED WORK

In this section, we describe a short background and some related work for the health of software ecosystems. We start with the systematic review of Hyrynsalmi et al.[9]. Considering this work and our own observation, there is not an agreed upon definition for "health" in an ecosystem context. Different studies use different concepts. We added a research question to gather different definitions in the studies. The concept of Iansiti and Levien[10], described in previous section, serves as basis for the majority of the primary studies. They introduced an evaluation framework for ecosystem health and defined indicators such as: robustness, productivity, and niche creation. KDE ecosystem[1] is an example of healthy ecosystem. They have a non-profit organization that supports business and financial issues to improving their robustness. KDE also have some practices to increase the productivity providing a set of tools for development and tests. Last, they have specific policies to attract new developers and receive new projects, increasing their niche creation capacity.

In addition, Manikas and Hansen[15] have taken a different perspective on the health of software ecosystems. They state that software ecosystems are different from business ecosystems because the actors are different from the products. The software component and the actor have different influence on the ecosystem. A software component can have a positive influence, but the actor responsible for that component may not have the same positive influence on the ecosystem health. They proposed a health framework composed by three main elements that affects the health of software ecosystems, they are: the actors, the software and the orchestration. Other concepts can be seen in this study later.

Equally important, in 2014, Fotrousi et al.[6] conducted a mapping study. They found 34 studies published from 2004 on the use of KPI for software-based ecosystems. The goal of this study was to gather existing research on KPI-based software ecosystem assessment. Key performance indicators (KPI) are initial indicators that can help to foreseen the sustainability and the health state of software ecosystems. This mapping study provides an overview of the literature on KPI for software ecosystems, explaining how KPI are used in the management of software ecosystems. Despite introducing several health indicators, this study did not mention any approach used to evaluate the health of software ecosystems, as well as key areas, practices and tools to support the evaluation process.

Similarly, a systematic review was conducted by Franco-Bedoya et al. in 2014[7]. They introduced a quality model for the quality assessment of OSS ecosystems (QuESo). Building this model, they gathered several quality measures from the systematic review, classified and organized the set of measures to create the QuESo. The main question was: *"What measures or attributes are defined to assess or evaluate open source software ecosystems?"*. They reviewed 53 studies of which 17 provided relevant measures to evaluate the quality of open source ecosystems. After using two criteria to select measures, they chose 68 different measures to compose the QuESo. This systematic review focused only on obtaining health indicators, without provide other information about ecosystem health evaluation.

Finally, another systematic review was conducted by Hyrynsalmi et al. in 2015[9]. This review aimed to support a common sense for the health of ecosystems definition. They tried to characterize this concept through empirical evidences. Their focus was to understanding how the health of ecosystems has been defined. Our work is similar to this work when we also raised a common understanding for the concept of health of ecosystems. However, the similarity is only this. Our focus is on research approaches to evaluate the health of ecosystems. We investigated deeply the existing approaches to discover key areas, practices, metrics, and tools that compose and support these approaches. Despite these three literature reviews present valuable information about the health of ecosystems, they did not investigate existing evaluation approaches and their details. All of studies cited here supported our work providing references for the definition of health of ecosystems and their indicators.

3 RESEARCH METHOD

This study aims to obtain an overview of the research literature on evaluation of the health of software ecosystems. Performing the systematic review, we followed three phases: planning, conducting, and reporting based on a systematic review protocol. We conducted our study based on the guidelines of Kitchenham et al.[13]. The remainder of this section describes our protocol of the review.

3.1 Research Questions

Identifying evaluating approaches for the health of software ecosystems, we defined a high-level research question: *How have the health of software ecosystems been evaluated?* This research question was divided into 6 specific research questions. Table 1 shows all research questions of this study with the motivation for each one.

3.2 Scope of Study

The goal of this systematic review is to clarify the research area around evaluation of the health of software ecosystems. There are several studies in this area with different means for health concept. The scope of our study is focused on approaches used to evaluate

[1]http://www.kde.org

Table 1: Research Questions

Nr.	Research question	Motivation
RQ1	What definitions of the term "health of software ecosystem" exist?	There are several definitions for the health of software ecosystems. We aim to find out which concepts each study is using in its design.
RQ2	What are the approaches to evaluating the health of software ecosystems?	With this question, we intend to identify the approaches, methods, models or frameworks which are used for evaluating the health of software ecosystems.
RQ3	What key ares are emphasized by existing approaches as having an influence on evaluating the health of software ecosystems?	By a key area, we mean a characteristic that is common across a set of studies or processes that has an influence on the health of software ecosystems.
RQ4	Which business, technical and social practices are used by existing approaches to evaluate the health of software ecosystems?	Through this question, we hope to understand which practices have a substantial role in the software ecosystem health that is not addressed by the existing approaches of evaluating the health nowadays.
RQ5	What are the metrics used to measure the health of software ecosystems?	In this question, we intend to identify metrics commonly used in evaluation of the health of software ecosystems.
RQ6	Are there tools that support the evaluation process of the health of software ecosystems?	There are several tools that support different models or approaches of evaluation. In this sense, we would like to investigate which tools are used to support the evaluation processes of the health of software ecosystems.

the health of software ecosystems. The concept of software ecosystems health must be included in the study. We did not consider approaches describing and assessing process structures of other nature for software ecosystems. There are some papers that do not present evaluation approaches explicitly defined. However, we also examined studies that addresses some metrics to evaluate some aspect of the health of software ecosystems.

3.3 Search Strategy

Following the review protocol, we built our research string. We included the term "open source systems" to get studies that evaluate the health of open sources ecosystems. Our intention was also to capture studies that do not use explicitly the term "software ecosystem" in their description. Open source systems can also be classified as software ecosystems[12]. The goal was to encompass all types of software ecosystems. We adapted the string to work with each research database mechanism. The string is: *(evaluation OR evaluate OR measurement OR evaluating OR measure OR assessment OR assess) AND (approach OR method OR framework OR model OR practices) AND (health OR healthy) AND (software ecosystem OR open source ecosystem OR open source systems)*

The following electronic databases were searched:

- ACM Digital Library
- IEEEXplore
- ISI Web of Science
- ScienceDirect
- SpringerLink
- Scopus

In addition, we used the snowballing technique [22] and hand-searched relevant papers for two events: international conference proceedings on software business (ICSOB) and international workshop on software ecosystems (IWSECO). The goal was to get papers not captured by the electronic research. Following the guidelines of Kitchenham et al.[13], two authors analyzed the papers separately. After, they cross checked their results to reduce the likelihood of bias. Also, during data extraction, another author performed a data extraction. He got a random sample of a primary studies to analyze. Following, their results cross-checked. The search process is shown in Fig. 1.

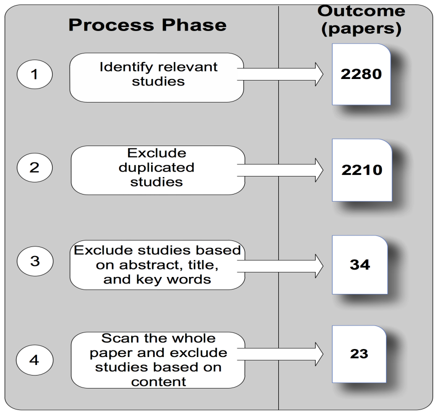

Figure 1: Search process

Initially, to conduct the search process, we used a tool called StArt (State of the Art through Systematic Review)[2]. This tool imported the files generated from databases, and also allowed the inclusion of papers manually. In the first phase, after getting all research papers from databases and conferences, we collected 2280 papers. In the phase 2, 70 papers were removed as duplicates. In the phase 3, after applying the inclusion/exclusion criteria, we excluded studies based on abstract, title, and key words, resulting in 34 papers to analyze. In the phase 4, we scanned the whole papers and excluded studies do not related to evaluation of the health of software ecosystems, based on the exclusion criteria. We also excluded one paper that we consider duplicated, because it described the same approach with similar data to another paper. The final result was 23 papers relevant for the detailed quality assessment, see A. In spite of the concept of "Health" for ecosystems having been introduced in 2002, only lately this topic has attracted attention of the research community. Figure 2 illustrates the number of papers by year of publication. Furthermore the Table 2 lists the paper distribution by source.

3.4 Exclusion and Inclusion Criteria

Exclusion and inclusion criteria were defined to ensure that relevant papers will be analyzed. They were defined based on the research topic, singleness, language, period of time, and work format. We use the following inclusion criteria: (i) The study must explore practices, theory, approaches or issues related to evaluation of the health on software ecosystems; (ii) The study must be unique, i.e.

[2] http://lapes.dc.ufscar.br/tools/start_tool

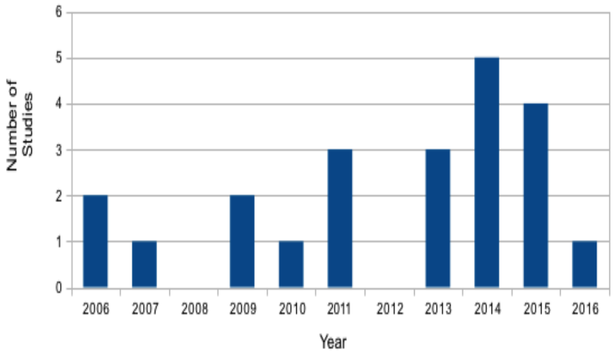

Figure 2: Timeline of publications in evaluation approaches(number of studies).

Table 2: Paper distribution

Source	Count	Percent(%)
International Conference on Software Business (ICSOB)	3	13.00
International Conference on Management of Emergent Digital EcoSystems (MEDES)	1	4.35
European Conference on Software Architecture Workshops (ECSAW)	2	8.70
International Conference on Research Challenges in Information Science (RCIS)	1	4.35
International Workshop on Building Sustainable Open Source Communities (OSCOMM)	1	4.35
International Conference on Malicious and Unwanted Software (MALWARE)	1	4.35
International Workshop on Software Ecosystems (IWSECO)	1	4.35
International Conference on Open Source Systems (OSS)	2	8.70
ACM international workshop on Software-defined ecosystems (BigSystem)	1	4.35
International Conference on Software Technologies (ICSOFT)	1	4.35
International Symposium on Empirical Software Engineering and Measurement (ESEM)	2	8.70
European Network on Chaos and Complexity Research and Management Practice Meeting	1	4.35
Journal of Theoretical and Applied Electronic Commerce Research	1	4.35
International Journal of Open Source Software and Processes	1	4.35
Information and Software Technology	1	4.35
International Journal of Web Information Systems	1	4.35
Harvard Business School	1	4.35
Journal Computer	1	4.35
Total	23	100

when a study has been published in more than one venue, the most complete version will be used; (iii) Papers must be written in English.

In the other hand, we defined the following exclusion criteria: (i) Studies that do not address health and evaluation in software ecosystems are going to be excluded; (ii) Studies that mention health or evaluation in software ecosystems, but do not discuss any type of method, activity, experience, or approach concerning evaluation of software ecosystems health; (iii) Studies that were only available as abstracts or Power Point presentations; (iv) Short papers (less than three pages long); (v) Studies that were not published in the period between 2001 and 2016 are going to be excluded; (vi) Studies presented in languages other than English.

3.5 Quality Assessment

One of the steps of Kitchenham's guidelines is the quality assessment of the primary studies[13]. We developed a quality checklist to ensure that minimal parts of a research study are present in the primary study. The 10 quality criteria that we used are: purpose, methodology, motivation, technology definition, limitations, data collection, data analysis, sampling, sampling justification, and completeness. The quality checklist questions are available at *http://www.professores.ifba.edu.br/simoneamorim/review/*.

First of all, our quality instrument was used to support the selection process providing a detailed inclusion/exclusion criteria. The quality data were collected during data extraction using our quality checklist. The answers of the checklist could be "yes" or "no". Three criteria (Q2, Q3, Q4) were used as the basis to support the decision of including or excluding a primary study. We assessed 23 studies for quality and did not exclude any paper. However, we observed that 4 papers had less than 50% of points in the quality assessment. The reason was that some papers only introduced a proposal, but they did not validate the evaluating approach. We decided to keep these studies in the systematic review because we consider the idea proposed relevant for our study goals. Table 3 shows the results of the quality assessment for each paper.

Table 3: Results of the Quality Assessment

Study	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total%
[S1]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	90
[S2]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S3]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S4]	Y	Y	Y	Y	N	N	N	N	N	N	40
[S5]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S6]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	90
[S7]	N	Y	Y	Y	N	N	N	N	N	Y	40
[S8]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S9]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S10]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S11]	N	Y	Y	Y	N	N	N	N	Y	Y	50
[S12]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S13]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S14]	Y	Y	Y	Y	N	Y	Y	Y	N	Y	80
[S15]	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	90
[S16]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S17]	N	Y	Y	Y	N	N	N	N	N	Y	40
[S18]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	90
[S19]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S20]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S21]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
[S22]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	90
[S23]	N	Y	Y	Y	N	N	N	Y	Y	N	40

3.6 Data Extraction

In this phase, we extracted data from 23 primary studies and included data in a predefined extraction form. The data included were related to bibliography, objectives, problems, results, and the answers of the 6 research questions. The data extraction available at *http://www.professores.ifba.edu.br/simoneamorim/review/*.

3.7 Synthesis of Findings

According to Cruzes et al. the research synthesis is the process of summarize, integrate, combine, and compare the findings of different studies on a specific topic or research question. This process is the heart of systematic reviews and much attention must be given to choose the appropriate method of synthesis[4]. Based on the nature of our research questions, we chose the method of narrative synthesis[20]. Narrative synthesis is a form of story telling based on evidence as a way of assuring through trustworthy arguments the obtainment of plausible outcomes. In this study, we adopted the narrative synthesis based on guidelines provided by Popay et al.[18]. The results of this narrative synthesis is described in the next section.

4 FINDINGS

This section presents the results to answer each research question listed in Table 1.

4.1 RQ1: What definitions of the term "health of software ecosystem" exist?

From the studies found that reported some kind of evaluation of the health of software ecosystems, we gathered several different concepts of health adopted by them. In 2015, Hyrynsalmi conducted a systematic review that reported different concepts for ecosystem health[9]. Additionally, Manikas and Hansen discussed the types of components for health of ecosystems[15]. These work cover a large range of possible definitions. However, to contextualize and improve the understanding of our results, we also provide the health concepts found in our primary studies.

The first health concept among primary studies comes from Iansiti and Levien. They consider health as the ability to grow the ecosystem and keep it attractive for all members of the community. They also provide three health indicators: robustness, productivity, and niche creation[10]. Moreover, we captured the concept of Manikas and Hansen. They state that software ecosystem health is composed of the health of their components. The interaction among these components affects the entire software ecosystem. Finally, the general concept adopted to open source software ecosystems is the survivability over time. The majority of the authors consider that longevity is synonymous of good health for a software ecosystem.

In summary, several papers provide varying views for the health concept. They performed some kind of health evaluation regarding these concepts. However, we found three papers that evaluate the health, but we cannot identify exactly what the term "health" means to them. This diversity of concepts influences the arrangement of the elements that compose the software ecosystem evaluating approaches. Table 4 shows the concepts and the papers related to them.

Based on the concepts presented by studies, we can identify the main features of the health of software ecosystems. They are:

Table 4: Concepts of Health of Software Ecosystems

Identifier	Health Concept	Studies
C1	*A healthy software ecosystem provides durably growing opportunities for its members and for those who depend on it. They have three indicators for ecosystem health: robustness, productivity, and niche creation.*	[S9] [S22] [S18] [S4] [S2] [S20] [S12] [S15]
C2	*The software ecosystem health is the ability of the ecosystem to endure and remain variable and productive over time. This health concept is composed of elements interacting to build the health for the whole ecosystem. They are: each individual actor, network of actors, each individual software component, platform, software network, and orchestrator.*	[S5] [S11]
C3	*The health for OSS software ecosystems is composed of the ability of the project to survive throughout time. This means the survivability of the ecosystem for the next several time periods.*	[S1] [S6] [S21] [S19] [S21] [S10] [S16] [S17]
C4	*Ecosystem health is defined as "long-term financial well-being of the business ecosystem and the long-term strength of the network".*	[S3] [S13]
C5	*Ecosystem health refers to the global condition of an ecosystem, provides a powerful theoretical and practical framework for monitoring system activity, identifying and predicting areas for improvement, and evaluating changes in ecosystems.*	[S8]
C6	*No concept presented.*	[S7] [S14] [S23]

productivity, attractiveness, survivability, longevity, successful, improvement, connectedness, interactions, and dependencies. Combining these features, we can generate a synthesized concept for the health of software ecosystems: *"A healthy software ecosystem has the capacity of keeping their productivity and attractiveness, facing problems, disruptions and junctions. At the same time, they also monitor and implement advances of their strategies to achieve the success over time"*. This success should include all their internal elements considering their interactions and dependencies.

4.2 RQ2: What are the approaches to evaluating the health of software ecosystems?

This research has identified 23 studies reported to deal with health evaluation for software ecosystems. The goal of this study is to gather the main characteristics of the primary studies considering different perspectives such as issues addressed by them, empirical research methods used, evaluation scenarios, studied objects, and evidence validated in some real-world software ecosystem. The scope of this study does not include a detailed analysis and discussion on the different aspects of the published evaluation. Table 5 shows the characteristics of the approaches. We observed that only 5 approaches are formally defined as an entire evaluation proposal for an ecosystem. Other 18 approaches evaluate only some kind of feature or health metric for software ecosystems.

Based on data from primary studies, we observed that the evaluation scenarios are composed mainly by open source ecosystems. Only 7 studies are conducted in scenarios with commercial software ecosystems. Only an unique scenario is applied to commercial and open source ecosystems at the same time. Open source ecosystems

Table 5: Overview of the approaches

Study	Software Ecosystem	Type	Domain	Issues Addressed	Name	Validated
S1	Vuforia	Commercial	Augmented Reality	The developers network and applications		yes
S2	Axis	Commercial	Embedded systems	Governance activities		yes
S3	Azure, Cloud Foundry, dotCloud, Engine Yard, Google App Engine, Heroku, Nodejitsu and OpenShift	OSS	PaaS providers	Developer and Project activities		yes
S4				Creation, monitoring and evaluation for digital business ecosystems		no
S5	Apache Cordova	OSS	Mobile framework	Software network; Keystone and dominator activities		yes
S6	Markmail, Jira and GIT	OSS	Service Oriented Computing and Quality of Service	Performance indicators	SALMonOSS	partially
S7		OSS		Interactions among stakeholders		no
S8	Antivirus	Commercial	Antivirus	Diversity, stability, and activity		yes
S9	WordPress, Joomla and Drupal	OSS	Content Management System	Health metrics		yes
S10	Topcased and Papyrus	OSS	Model Driven Development	Activity, company influence in ecosystems, and interaction between ecosystems		yes
S11	Cytoscape Consortium	OSS	Genetics and biology	Health metrics		yes
S12	GNOME	OSS	Linux desktop environment	Quality of the platform, community, and network	QuESo	partially
S13	Magento, PrestaShop, and WooCommerce	Commercial	E-commerce systems	Health metrics		yes
S14	90 OSS projects	OSS	OSS	Social network analysis		yes
S15	Dutch IT companies	Commercial	IT systems	Health metrics		yes
S16	Nagios	OSS	IT infrastructure	Social network activity		yes
S17		OSS	OSS systems	Sustainability, process maturity, and maintenance capacity	QualOSS	no
S18	Apple, Google, BlackBerry and Nokia	Commercial and OSS	Mobile systems	Business process and Health metrics		yes
S19	Apache Lenya, Apache log4J, Apahce Excalibur, and Apache OJB	OSS	OSS systems	Health metrics	FOSSDA	yes
S20	Python, Gnome, Wordpress, Joomla, Drupal	OSS	OSS systems	Health metrics	OSEHO	yes
S21	Apache Tomcat, Apache HTTP Server, Apache Xindice, and Apache Slide	OSS	OSS systems	Health metrics		yes
S22		Commercial	IT systems	Health metrics		partially
S23		Commercial	OSS systems	Relationship among stakeholders		no

have free data available increasing the number of software ecosystems evaluated. Issues addressed by evaluations approaches have different focus, but the majority are concentrated in evaluating health metrics. These metrics vary in accordance with the author. Often they mention the health indicators proposed by Iansiti and Levien[10] and aggregate other metrics to complement their study. Another focus addressed by studies are social activities and network influencing the health.

Moreover, we observed that almost all approaches are validated by data collected in real-world ecosystems. Only 4 approaches do not have any kind of validation. They are studies that propose a model to be built in the future. For our understanding, a validated approach present some data to illustrate or prove the metric proposed by the study. The majority of the studies evaluate a set of health metrics considering some aspects of the health concept. No approach evaluates the health for the entire ecosystem. There are only two approaches that introduce this proposal, but the studies are just proposals without data and research methods. Other studies present a wealth of data such as the OSEHO [S20], FOSSDA [S19], and Wahyudin's work [S21]. However, they also explain the challenges that prevent performing a complete evaluation of the health of ecosystems.

Regarding empirical methods, we found 5 types of empirical approaches used in primary studies. We gathered the types of empirical studies from the work of Easterbrook et al.[5] and Zannier et al.[23]. Our results pointed out that *Experience* and *Example application* are the most frequently used means of evaluation followed by *Case studies*. Three papers indicated that this research topic gained the attention of the researchers. Lately, there are different evaluation approaches been developed for the health of ecosystems. Table 6 shows different empirical research methods used in the primary studies.

4.3 RQ3: What key areas are emphasized by existing approaches as having an influence on evaluating the health of software ecosystems?

Key areas can be defined as sets of fields investigated and included in the evaluation of the health of ecosystems. It is important to know

Table 6: Empirical Research Methods

Empirical Methods	Concepts	Studies
Case Study	*An empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident*[5].	[S2] [S10] [S18]
Discussion	*Provides some qualitative, textual, opinion-oriented evaluation. E.g. compare and contrast, oral discussion of advantages and disadvantages*[23].	[S23]
Example application	*Describes an application and provide an example to assist in the description, but the example is "used to validate" or "evaluate" as far as the authors suggest*[23].	[S5] [S6] [S12] [S15] [S19] [S20] [S22]
Experience	*Relates the use of a previously-reported tool or technique in a practical software project. The conclusions are supported by comparative data and/or statistics*[21].	[S1] [S3] [S8] [S9] [S11] [S13] [S14] [S21]
Proposal	*Presents and justifies the need to study a research problem and to present the practical ways in which this research should be conducted*[14].	[S4] [S7] [S17]

the directions and subjects that researchers are considering to build the evaluation approaches. An analysis of the studies showed 4 key areas addressed by research: productivity, community, business, and quality. Among the main topics studied by authors, productivity in the community was the most explored. Productivity is one of the health indicators introduced by Iansiti and Levien[10]. Other authors presented different methods to measure the productivity of software ecosystems. They used several metrics to estimate a value for productivity [S1] [S2] [S3] [S5] [S7] [S9] [S10] [S11] [S19] [S20] [S21].

Another key area was raised frequently in the primary studies - issues of the community such as satisfaction, relationship, diversity, and roles in the community. These issues addressed many metrics that represented part of the health of software ecosystems [S1] [S3] [S7] [S19] [S10] [S14] [S16] [S20] [S23]. Some studies addressed issues related to business side such as market share, governance, financial management, and others[S2] [S4] [S9] [S20] [S22]. Some studies also addressed performance indicators, process maturity and sustainability as well as quality issues[S6] [S12] [S17]. Finally, we cannot identify specific key areas in some studies [S8] [S13] [S15] [S18].

4.4 RQ4: Which business, technical and social practices are used by existing approaches to evaluate the health of software ecosystems?

Jacobson states that a practice *provides a way to systematically and verifiable address a particular aspect of a problem.* They address a particular aspect of a problem, instead of addressing the entire problem[11]. We consider that practices have directly influenced the health of software ecosystems. A practice produces a result, good or bad, that can be expressed by a metric. The way to guide a software ecosystem to achieve a good health probably will be through practices. However, practically all approaches found in primary studies directed their efforts to evaluate the health by metrics. Only two studies included practices directly in their evaluation [S2] [S17]. In [S2], Wnuk et al. investigated what *governance activities* can improve the health of ecosystems. They connected governance and health, considering *governance* as all activities of management of the ecosystem[12]. Aside from that, *health* is the capacity of growing and prosperity[10]. The *governance* is used to achieve *health*. They classified practices in accordance with health indicators defined by Iansiti and Levien[10]. We describe these 19 practices below:

(1) **Robustness -** Create partnership model, Do marketing, Grow profits, Establish partner development programs, Form alliances, Stabilize APIs, Raise entry barriers, Make partners explicit, and Propagate operation knowledge.
(2) **Productivity -** Organize developer days, Collaborative marketing, Create sales partner program, and Create new sales channels.
(3) **Niche Creation -** Expand applicability, Make strategy explicit, Create API, Co-development, Develop complementary platforms, and Develop new business models.

An another study, [S17] proposed to include the health evaluation into the process maturity evaluation for ecosystems. They described some processes that will be covered by the proposed model. Some example of these processes are: change submission and review, peer review of changes, propose significant enhancements, test the programs produced by the project, and plan releases. The authors did not describe the detailed practices in these processes. Furthermore, they did not explain how the processes would be used to evaluate the maturity of a software ecosystem. In summary, practices have an influential role in the health of ecosystems. In addition, this role is not addressed by existing approaches for evaluating the health nowadays.

4.5 RQ5: What are the metrics used to measure the health of software ecosystems?

The majority of primary studies evaluate the health of ecosystems using metrics. There are tens of metrics defined by different authors, including a systematic review conducted by Franco-Bedoya et al.[S12]. This study aimed to know what measures or attributes are defined to assess or evaluate open source software ecosystems. They found 68 different measures used to build their evaluating model. Practically, all primary studies presented some kind of metric in their evaluation approach. We observed that some metrics are repeated in other studies, or exist similar metrics with different names. Some metrics have the same name, but they are calculated per different period of time (day, week, month, year). Other metrics were defined specifically for the ecosystem context or approach. These metrics cannot be easily operationalized for other ecosystems.

Answering this RQ5, we raised 211 metrics from all primary studies. We started working from [S12] study and included metrics captured in other primary studies. We did not distinguish the type of ecosystem that the metric can be applied. As well as, we did not consider if these metrics can be measured in the practice of real-world ecosystems. Metric's names were kept in accordance with the original study to avoid misunderstanding. Many studies introduced metrics and focused on measuring their values in the ecosystem. However, we observed that there is a lack of critical work to investigate issues such as: which metrics are better to be applied in the health evaluation, which metrics can be really measured in real-world ecosystems, which implications of the interactions among these metrics, how a metric can influence another metric, and so on. There are many open research issues regarding the work with health metrics. Table 7 shows an example for 5 health metrics. All metrics collected are available at *http://www.professores.ifba.edu.br/simoneamorim/review/*.

Table 7: Health Metrics for Evaluation Approaches

Name	Description	Studies
Files per Version	Number of files per version.	[S12]
Files Changed	Number of files that has been changed.	[S12]
Member Effort	The effort of member m in community c.	[S12]
Value Creation	The overall value of new options created.	[S20]
Zeta Model	Bankruptcy classification score model.	[S12]

4.6 RQ6: Are there tools that support the evaluation process of the health of software ecosystems?

There are many ways of performing an evaluation of software ecosystems. Existing tools can help and make the process of evaluation faster. We gathered from the primary studies tools used to support the evaluation process. There is not a tool that support the whole evaluation process. We observed that 9 primary studies described explicitly some kind of tool used in some point of the evaluation process. The majority of tools perform data extraction from repositories. Many tools were developed by the authors based on GIThub[3] repositories from which they extracted data. One study, [S23], extracted data from a website specialized to analyze open source projects, the FLOSSmole[4]. It also used the CVSAnalY[5] to extract statistical information out of CVS repositories. Furthermore, [S19] study used the ProMonCo, a project monitoring cockpit that provides integrated indicators to analyze the project status[2]. In summary, the results for Question RQ6 show that only 4 tools developed by third parties were used, and 5 authors developed their own tool. None of them informed any functionality of these tools. Other work did not describe anything about tools. There are not tools to support important evaluation issues such as: data analysis and data visualization, conducting and maintenance of the evaluation, and even provide post-evaluation activities. Table 8 shows the studies and tools used in the evaluation process.

5 DISCUSSION

This section discusses and summarizes contributions of this systematic review to the software ecosystems research community.

5.1 Summary of Study Findings

The evaluation of the health of software ecosystems has produced few approaches. The majority of the approaches expect some kind of rigorous and formal evaluation. Quite a few studies provide some data validation, other focus only in parts of the ecosystem, and many of them do not present a "formal" evaluation approach. A large majority of the studies can be categorized as "example application" or "experience". They present a lack of certain pieces of information to be considered a complete study. They propose several health metrics and show data extracted for these metrics. However, they do not show how these metrics can be part of a complete evaluation approach. They do not also exhibit the health state of the entire ecosystem.

[3]https://github.com/
[4]http://ossmole.sourceforge.net
[5]http://cvsanaly-web.stage.tigris.org/

Table 8: Tools used in the evaluation process

Study	Tools
[S1]	Import.io(https://www.import. io/)
[S3]	Authors developed their own tool based on Ruby
[S6]	Authors developed their own tool
[S8]	Microsoft Software Removal Tool
[S9]	Authors developed their own tool based on Java and PH
[S13]	Authors developed their own tool based on Python and Scrapy(https://scrapy.org/)
[S19]	Project Monitoring Cockpit (ProMonCo)[2]
[S21]	Authors developed their own tool for mining the web-based developers mailing list
[S23]	FLOSSmole and CVSAnalY

The mass of the studies are in early stages of development, explaining the lack of robust assessments. We have identified only three complete approaches: FOSSDA [S19], OSEHO [S20], and Wahyudin's approach [S22]. Other approaches are focused on measuring specific health aspects, or define only a proposal. However, they did not present data evaluation. Furthermore, the results also exhibit a general lack of studies replication. Many studies use concepts proposed by other authors, but they did not perform a replication of the original study. Two studies, [S11] and [S13], present a partial application of the original study. However, they also introduced extensions and discussed issues with the primary approach.

Regarding the studies quality, some topics were not well described. Aspects such as issues of bias, limitation, validity, and reliability were not always addressed. None of them presented and explained data collection. We consider that these topics should be treated by the studies, mainly to avoid the risk of bias, subjectivity and/or confounding. Despite the quality assessment of the studies 78% present more than 80% of quality (see table 3). In spite of the problems and absence of information, the outcomes related here have several implications for researchers and practitioners. They can help to identify important research areas. So, the empirical software engineering community can be encouraged to improve the evaluation of the health of software ecosystems.

5.2 Implications for Research and Practice

This work shows a set of research gaps and the necessity of more empirical studies for evaluating the health of ecosystems. In our opinion, issues raised by this work could guide new studies in the near future. We found some important research gaps such as the lack of practices, metrics and tools adequate to conduct the process of health evaluation.

Starting with the proper concept of the health of software ecosystems, we found some discrepancies. There is not a common sense about health definition. The RQ1 captured various concepts adopted into primary studies. It would be useful to have an agreed upon concept on which to build solid evaluation approaches. Practically all studies have concentrated their efforts on getting metrics for health evaluation. They discovered several metrics, but there is not a consensus about which metrics are better to evaluate. Working with metrics has challenges described by Jansen's work, in [S20]. There are some problems such as: difficulty in defining which metrics can be measured, how to work with highly abstract metrics, which is the better way get value for a metric considering different ways of measuring, and so on.

Other research areas on evaluating of the health of ecosystems are in the initial state, or even nonexistent. For example, the study of appropriate practices. One unique approach addressed explicitly practices on the health evaluation. Wnuk et al., in [S2], connected the health of ecosystems with some governance activities. They evaluated the governance model of an ecosystem. There are other studies[1, 12] that provide ecosystems governance models. These studies consider processes and practices to preserver and improve the ecosystem health. However, these models do not address health evaluations. They present models for ecosystems management that aim to achieve the health. The influence of the practices on the health of ecosystems should be more deeply explored. They interfere directly on the actions that guide the ecosystem. Metrics will only describe results of the direct application of practices.

In addition, this study shows the absence of tools to support the evaluation process. Some approaches developed their own tool to extracted data. Few tools existing in the market were also used. The majority of the studies did not reveal what tools their used in the evaluation process. There is a necessity of robust assessment to gather adequate tools to support all parts of he evaluation. Data analysis tools, visualization tools, and a registry of evaluation tool are required. An evaluation in a ecosystem should not be done manually. This can bring many mistakes and bias to the evaluation. The construction of a suitable environment to support all stages of the process is essential to obtain a good quality assessment.

For practitioners, this systematic review shows that many encouraging studies of evaluating the health of ecosystems have been reported. Although the evaluations are immature, the review results suggest that the metrics' analysis can identify problems and improve their health. In order to increase the usefulness of the research in the practice, it is necessary to have a sufficient number of high quality studies. These studies will provide a solid basis of knowledge for problems investigation and improvement proposals affording the evaluation. Additionally, we emphasize that to perform research in a large environments such as an ecosystem is a big challenge. Many people in different roles should be mobilized and with a large range of variables. Researchers and practitioners need to work together in collaboration to overcome challenges through the mutual support. This will contribute to achieve a good level of maturity in the future.

6 LIMITATIONS

Although we have followed Kitchenham's guidelines[13] to conduct this systematic review, the process can suffer by some limitations. Bias may have been introduced during the selection process of primary studies causing problems such as: errors in data extraction, omission on the approaches classification, and flaws in the quality assessment process. To avoid bias, the papers selection was perform by two researchers in parallel. After selection, the results were compared to resolve differences. In other phases, variations were discussed to find a consensus.

In addition, there are some situations beyond our control that can interfere in the research quality. For example, research string terms can be confused with terms with similar meanings but that belong to another research topic. Also, search engines quality could have influenced the completeness of the identified primary studies. In addition, there are papers with different problems such as: lack of sufficient details about the design and execution of the reported studies; some approaches that are not described adequately; issues of validity are not addressed; and data collection and analysis methods do not reveal important details. These problems interfered in our process of analysis, so some subjective decisions were taken by the researchers to perform the data synthesis. To minimize these limitations, all the doubts were discussed by researchers to resolve all ambiguities.

7 CONCLUSION

Software ecosystems have attracted increased attention in recent years. However, research in this area is awakening and many research gaps should be filled. Mainly concerning the health of software ecosystems and its implications. There has been no comprehensive attempt to systematically investigate *how* the health of ecosystems has been evaluated and *what* approaches are available. In order to fill this gap, we conducted a systematic review to investigate evaluating approaches for the health of ecosystems.

In the beginning, we identified 2280 studies from searching the literature. 23 primary studies were qualified for the data extraction phase after being filtered by inclusion/exclusion criteria. Only 3 studies reported a complete health evaluation for software ecosystems. 5 studies were considered proposals to be developed in the future. The majority of the studies addressed partially some aspect of the health, exploring metrics into evaluation. In summary, we identified 6 different concepts for the health of software ecosystems, 211 health metrics, 5 approaches with a name formally defined, 4 key areas addressed by research, 19 governance practices, and 9 support tools were used by studies.

Our systematic review revealed an absence of formal studies with scientific rigor as case studies and experiments. Some published studies did not show a robust explication of limitations or validation or a detailed data collection and analysis process. So, we perceived the immature state of the evaluation approaches. Our findings provide important contributions to academic researchers and practitioners. We described many research gaps that should be explored by researchers, as well we provided useful information about different aspects of the evaluation approaches. For practitioners, we emphasize that an analysis of metrics can help to identify problems and improve the ecosystem health. However, there is the necessity of more high quality studies to consolidate the evaluation approaches. We believe that research in this area have high potential for industrial adoption and will bring plenty of benefits for all.

A STUDIES INCLUDED IN THE REVIEW

[S1] L. Soussi, Z. Spijkerman, S. Jansen, A case study of the health of an augmented reality software ecosystem: Vuforia, in: proceedings of the 7th International Conference on Software Business (ICSOB), 2016, pp. 145-152.

[S2] K. Wnuk, K. Manikas, P. Runeson, M. Lantz, O. Weijden, H. Munir, Evaluating the governance model of hardware-dependent software ecosystems a case study of the axis ecosystem, in: proceedings of the 4th International Conference on Software Business (ICSOB), 2014, pp. 212-226.

[S3] G. Lucassen, K. van Rooij, S. Jansen, Ecosystem health of cloud paas providers, in: proceedings of the 4th International Conference on Software Business (ICSOB), 2013, pp. 183-194.

[S4] A. D'Andrea, F. Ferri, P. Grifoni, T. Guzzo, Digital eco-systems: the next generation of business services, in: Proceedings of the Fifth International Conference on Management of Emergent Digital EcoSystems (MEDES), 2013, pp. 40-44.

[S5] K. Manikas, D. Kontogiorgos, Characterizing software activity: The influence of software to ecosystem health, in: Proceedings of the 2015 European Conference on Software Architecture Workshops (ECSAW), 2015.

[S6] M. Oriol, O. Franco-Bedoya, X. Franch, J. Marco, Assessing open source communitiesÃŢ health using service oriented computing concepts, in: Proceedings of the 2014 IEEE Eighth International Conference on Research Challenges in Information Science (RCIS), IEEE, 2014, pp. 1-6.

[S7] B. Lundell, B. Forssten, J. Gamalielsson, H. Gustavsson, R. Karlsson, C. Lennerholt, B. Lings, A. Mattsson, E. Olsson, Exploring health within oss ecosystems, in: Proceedings of the 1stInternational Workshop on Building Sustainable Open Source Communities (OSCOMM), 2009, pp. 1-5.

[S8] F. L. Levesque, A. Somayaji, D. Batchelder, J. M. Fernandez, Measuring the health of antivirus ecosystems, in: Proceedings of the 10th International Conference on Malicious and Unwanted Software (MALWARE), IEEE, 2015, pp. 101-109.

[S9] G. L. Sonny Van Lingen, Adrien Palomba, On the software ecosystem health of open source content management systems, in: proceedings of the 5th International Workshop on Software Ecosystems (IWSECO), 2013, pp. 45-56.

[S10] A. M. Jonas Gamalielsson, Bjrn Lundell, Open source software for model driven development: A case study, in: Proceedings of the 7th IFIP WG 2.13 International Conference (OSS), 2011, pp. 348-367.

[S11] J. Y. Monteith, J. D. McGregor, J. E. Ingram, Proposed metrics on ecosystem health, in: Proceedings of the 2014 ACM international work- shop on Software-defined ecosystems (BigSystem), 2014, pp. 33-36.

[S12] O. Franco-Bedoya, D. Ameller, D. Costal, X. Franch, Measuring the quality of open source software ecosystems using queso, in: Proceedings of the 10th International Conference on Software Technologies (ICSOFT), 2015, pp. 39-62.

[S13] D. Alami, M. Rodrguez, S. Jansen, Relating health to platform success: Exploring three e-commerce ecosystems, in: Proceedings of the 2015 European Conference on Software Architecture Workshops (ECSAW), 2015.

[S14] S. Onoue, H. Hata, K. Matsumoto, Software population pyramids: the current and the future of oss development communities, in: Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), 2014.

[S15] E. den Hartigh, M. Tol, W. Visscher, The health measurement of a business ecosystem, in: Proceedings of the European Network on Chaos and Complexity Research and Management Practice Meeting (ECCON), 2006, p. 139.

[S16] J. Gamalielsson, B. Lundell, B. Lings, The nagios community: An extended quantitative analysis, in: proceedings of the 6th International IFIP WG 2.13 Conference on Open Source Systems (OSS), 2010, pp. 85-96.

[S17] M. Soto, M. Ciolkowski, The qualoss open source assessment model measuring the performance of open source communities, in: proceedings of the 3rd International Symposium on Empirical Software Engineering and Measurement (ESEM), 2009, pp. 498-501.

[S18] P. R. J. Campbell, F. Ahmed, An assessment of mobile os-centric ecosystems, Journal of Theoretical and Applied Electronic Commerce Research 6 (2011) 50-62.

[S19] T. Moser, S. Biffl, D. Winkler, W. D. Sunindyo, Analyzing oss project health with heterogeneous data sources, International Journal of Open Source Software and Processes 3 (2011) 1-23.

[S20] S. Jansen, Measuring the health of open source software ecosystems: Beyond the scope of project health, Information and Software Technology 56 (2014) 1508-1519.

[S21] D. Wahyudin, K. Mustofa, A. Schatten, S. Biffl, A. M. Tjoa, Monitoring the health status of open source web engineering projects, International Journal of Web Information Systems 3 (2007) 116-139.

[S22] M. Iansiti and R. Levien, Keystones and Dominators: Framing Operating and Technology Strategy in a Business Ecosystem.Harvard Business School 03-061 (November 2002).

[S23] K. Crowston, J. Howison, Assessing the health of open source communities, Journal Computer 39 (2006) 89-91.

REFERENCES

[1] Alfred Baars and Slinger Jansen. 2012. A Framework for Software Ecosystem Governance. (June 2012), 168-180 pages.

[2] Stefan Biffl, Wikan Danar Sunindyo, and Thomas Moser. 2010. A Project Monitoring Cockpit Based On Integrating Data Sources in Open Source Software Development. (July 2010), 620-627 pages.

[3] Jan Bosch. 2009. From Software Product Lines to Software Ecosystems. In *Proceedings of the 13th International Software Product Line Conference (SPLC) (SPLC '09)*. 111–119.

[4] Daniela S. Cruzes and Tore Dybå. 2011. Research synthesis in software engineering: A tertiary study. *Information and Software Technology* 53 (2011), 440–455. https://doi.org/10.1016/j.infsof.2011.01.004

[5] Steve Easterbrook, Janice Singer, Margaret-Anne Storey, and Daniela Damian. 2008. Selecting Empirical Methods for Software Engineering Research. In *Guide to Advanced Empirical Software Engineering*, Forrest Shull, Janice Singer, and Dag I.K. Sjøberg (Eds.). Springer, Chapter 11, 285–311. https://doi.org/10.1007/978-1-84800-044-5_11

[6] Farnaz Fotrousi, Samuel A. Fricker, Markus Fiedler, and Franck Le-Gall. 2014. KPIs for Software Ecosystems: A Systematic Mapping Study. (June 2014), 194-211 pages. https://doi.org/10.1007/978-3-319-08738-2_14

[7] Oscar Franco-Bedoya, David Ameller, Dolors Costal, and Xavier Franch. 2014. Measuring the Quality of Open Source Software Ecosystems Using QuESo. (August 2014), 39-62 pages. https://doi.org/10.1007/978-3-319-25579-8_3

[8] Abhimanyu Ghoshal. 2016. Microsoft has more open source contributors on GitHub than Facebook and Google. http://thenextweb.com/microsoft/2016/09/15/in-your-face-google/. (2016). Online; accessed 9 October 2016.

[9] Sami Hyrynsalmi, Marko Seppänen, Tiina Nokkala, Arho Suominen, and Antero Järvi. 2015. Wealthy, Healthy and/or Happy - What does 'Ecosystem Health' Stand for? (June 2015), 272-287 pages. https://doi.org/10.1007/978-3-319-19593-3_24

[10] M. Iansiti and R. Levien. 2002. Keystones and Dominators: Framing Operating and Technology Strategy in a Business Ecosystem. *Harvard Business School* 03-061 (November 2002).

[11] Ivar Jacobson, Pan Wei Ng, and Ian Spence. 2007. Enough of Processes - Lets do Practices. *Journal of Object Technology* 6, 6 (July-August 2007), 41–66.

[12] Slinger Jansen and Michael A. Cusumano. 2013. Defning Software Ecosystems: A Survey of Software Platforms and Business Network Governance. In *Software Ecosystems: Analyzing and Managing Business Networks in the Software Industry*, Slinger Jansen, Sjaak Brinkkemper, and Michael Cusumano (Eds.). Edward Elgar Publishing, Chapter 1, 13–28. https://doi.org/10.4337/9781781955635.00008

[13] Barbara Kitchenham and Stuart Charters. 2007. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. Technical Report EBSE 2007-001. Keele University and Durham University Joint Report.

[14] USC Libraries. 2005. Research Guide - Organizing Your Social Sciences Research Paper: Writing a Research Proposal. http://libguides.usc.edu/writingguide/researchproposal. (2005). Online; accessed 7 October 2016.

[15] Konstantinos Manikas and Klaus Marius Hansen. 2013. Reviewing the Health of Software Ecosystems - A Conceptual Framework Proposal. (June 2013), 33-44 pages. https://doi.org/10.1.1.415.7578

[16] D. Messerschmitt and C. Szyperki. 2003. *Software Ecosystem - Understanding an Indispensable Technology and Industry*. MIT Press Cambridge.

[17] Microsoft. 2016. Openness at Microsoft. https://www.microsoft.com/en-us/Openness/OpennessAtMicrosoft. (2016). Online; accessed 9 October 2016.

[18] Jennie Popay, Helen Roberts, Amanda Sowden, Mark Petticrew, Lisa Arai, Mark Rodgers, Nicky Britten, Katrina Roen, and Steven Duffy. 2006. *Guidance on the conduct of narrative synthesis in systematic Reviews. A Product from the ESRC Methods Programme*. Technical Report. Lancaster University. https://doi.org/10.13140/2.1.1018.4643

[19] Transparency Market Research. 2013. Global Hadoop Market – Industry Analysis, Size, Share, Growth, Trends and Forecast 2012 – 2018. White Paper. (August 2013). http://www.transparencymarketresearch.com/hadoop-market.html

[20] M. Rodgers, A. Sowden, M. Petticrew, L. Arai, H. Roberts, N. Britten, and J. Popay. 2009. Testing methodological guidance on the conduct of narrative synthesis in systematic reviews. *Evaluation* 15 (2009), 047–071. https://doi.org/10.1177/1356389008097871

[21] Mary Shaw. 2003. Writing Good Software Engineering Research Papers. (June 2003), 726-736 pages. https://doi.org/10.1007/s10009-002-0083-4

[22] Claes Wohlin. 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. (May 2014), 321-330 pages. https://doi.org/10.1145/2601248.2601268

[23] Carmen Zannier, Grigori Melnik, and Frank Maurer. 2006. On the Success of Empirical Studies in the International Conference on Software Engineering. (may 2006), 20-28 pages. https://doi.org/10.1145/1134285.1134333