Modernizing Legacy Systems with Microservices: A Roadmap

Daniele Wolfart
PPGComp - Western Paraná State University, Brazil

Wesley K. G. Assunção
PPGComp - Western Paraná State University, Brazil

Ivonei F. da Silva
PPGComp - Western Paraná State University, Brazil

Diogo C. P. Domingos
PPGComp - Western Paraná State University, Brazil

Ederson Schmeing
PPGComp - Western Paraná State University, Brazil

Guilherme L. Donin Villaca
PPGComp - Western Paraná State University, Brazil

Diogo do N. Paza
PPGComp - Western Paraná State University, Brazil

ABSTRACT

Legacy systems are long-lived applications, with obsolete technology and degraded architecture. These systems hamper digital transformation and innovation, and require a great amount of resources for maintenance. The modernization of monolithic legacy systems is a strategy to promote better evolution and maintenance, taking advantage of new technologies such as microservices. Microservice architectural style is a paradigm to develop systems as a suite of small and autonomous services, communicating through a lightweight protocol. However, the migration of legacy systems to microservices is complex. Although we can find several studies on this topic, they usually focus on specific activities, e.g., the identification of the microservice boundaries in the legacy code. Also, existing pieces of work do not cover real-world scenarios, since they do not take into account organizational, operational, and technical aspects. To overcome this limitation, in this paper we present a roadmap for modernizing monolithic legacy systems with microservices. The roadmap is distilled from the existing body of knowledge, describing common activities and input/output information. The proposed roadmap is composed of eight activities, grouped in four phases, namely initiation, planning, execution, and monitoring. The main contributions are: (i) serve as a basis for practitioners to plan, execute, and monitor the modernization process; (ii) be a reference for researchers to design new studies; and (iii) motivate tool builders to deal with existing needs.

KEYWORDS

Software Migration, Software Evolution, Cloud Computing

1 INTRODUCTION

The large majority of existing industrial systems are long-lived applications usually developed in a monolithic architecture. Due to extensive maintenance and obsolete technology, these architectures decay and degrade [45]. Legacy systems hamper digital transformation, makes innovation more difficult, and require a great amount of resources for maintenance, draining investment that otherwise could be used in system evolution, e.g., developing new features [21, 32]. To remain competitive, companies have been modernizing these monolithic legacy systems [27]. “Software modernization attempts to evolve a legacy system, or elements of the system, when conventionally evolutionary practices, such as maintenance and enhancement, can no longer achieve the desired system properties” [45]. During modernization, a legacy system is re-engineered [6] or migrated [25] to a modern architecture or platform. The modernization leads to benefits such as easing engineering activities, satisfying user needs, or achieving business goals [44].

Nowadays we can observe a trend on adopting microservice architectures as strategy for modernizing monolithic legacy systems [10, 27, 50]. Microservices became popular around 2014 [29], mainly due to their adoption by big players, as for example Netflix [51] and Uber [17]. Moreover, other factors have contributed to the popularity of microservices. Firstly, microservices are inspired by the service-oriented architecture (SOA) principle. According to this principle, complex applications are built as a suite of small, loosely coupled, and autonomous components that encapsulate business capabilities and communicate with each other using language-agnostic APIs [13]. Splitting large applications into microservices provides independence for agile teams and optimizes
the deployment process [36]. Secondly, some studies argue that microservices reduce maintenance and evolution effort, increase availability of services, ease the inclusion of innovation, ease the incorporation of DevOps in the development lifecycle, and facilitate scalability [48]. Finally, there are many open-source tools to aid the deploying, managing, and monitoring of the microservices, e.g., Docker, Kubernetes, Prometheus, and Grafana, to cite some [50].

Microservices emerged in industry but only recently have caught the attention of the software-engineering research community [11, 29]. We can find, in the literature, secondary studies on the topic of migrating monolithic system into microservices [11, 19, 41], classification of refactoring approaches for the modernization with microservices [20], and industrial reports/surveys with practitioners [4, 10, 19, 50]. Despite the growing interest in the modernization with microservices in both practice and research, there is still a lack of comprehensive studies on how to conduct such a modernization. Another limitation is that existing pieces of work usually do not take into account organizational and operational perspectives of the modernization, focusing mostly on technical aspects of microservices. Furthermore, existing studies mostly focus on specific activities of the process, e.g., identification of the microservices boundaries in the code. Based on this, companies lack a roadmap for real world scenarios, covering the entire modernization process and dealing with organizational, operational, and technical aspects. These limitations might hamper or make complex the proper use of microservices as a strategy to modernize legacy systems.

Based on aforementioned limitations, the goal of this paper is to present a roadmap, i.e., a process to conduct the migration, for guiding practitioners on modernizing legacy systems with microservices. For the definition of such a roadmap, we collected, analyzed, discussed, and organized the existing literature on this topic. Our study (presented in Section 2) relies on 62 papers found by conducting a systematic mapping. The resulting roadmap (described in Section 3) is composed of eight activities and covers the modernization phases of initiation, planning, execution, and monitoring. We also describe the input and output information for each activity.

The contribution of the roadmap defined in our study (highlighted in the related work, in Section 5) is: (i) for practitioners, our roadmap can serve as a basis to decide, plan, execute the modernization process, and monitor the microservices. Instead of a starting blindfolded, we provide a comprehensive set of activities encompassing existing specialized literature; (ii) for researchers, our results can serve as a starting point to design new studies, identify gaps, e.g., focusing on the activities less investigated in the literature; and (iii) for tool builders, serve as a reference to develop automated support for the migration/modernization process.

2 STUDY DESIGN AND EXECUTION
As our goal is to define a roadmap for modernizing legacy systems with microservices, from the perspective of the existing body of knowledge in the literature, we posed the following research question:

RQ: Why and how are monolithic legacy systems migrated to microservices? For answering this RQ, we considered the driving forces that motivates the modernization, common activities conducted when modernizing with microservices, and the input used and output generated in each activity.

<table>
<thead>
<tr>
<th>Database</th>
<th>URL</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACM Digital Library</td>
<td><a href="http://dl.acm.org">http://dl.acm.org</a></td>
<td>53</td>
</tr>
<tr>
<td>IEEE Xplore</td>
<td><a href="http://ieeexplore.ieee.org">http://ieeexplore.ieee.org</a></td>
<td>72</td>
</tr>
<tr>
<td>Web of Science</td>
<td><a href="http://www.isiknowledge.com">http://www.isiknowledge.com</a></td>
<td>35</td>
</tr>
<tr>
<td>Science Direct</td>
<td><a href="http://www.sciencedirect.com">http://www.sciencedirect.com</a></td>
<td>99</td>
</tr>
<tr>
<td>Scopus</td>
<td><a href="http://www.scopus.com">http://www.scopus.com</a></td>
<td>52</td>
</tr>
<tr>
<td>SpringerLink</td>
<td><a href="http://www.springerlink.com">http://www.springerlink.com</a></td>
<td>111</td>
</tr>
</tbody>
</table>

To achieve this goal and answer the posed RQ, the methodology of our study is composed of the following steps:

1. **Primary sources selection**: identification, screening, and validation of studies from digital libraries.
2. **Data extraction**: reading of papers to extract relevant data to answer the posed RQ.
3. **Data Classification and roadmap definition**: the data extracted were classified and analyzed in order to establish a comprehensive roadmap for modernizing legacy systems with microservices.

2.1 Primary Sources Selection
For the selection of primary sources, we followed a methodology based on a systematic mapping, according to the process proposed by Petersen et al. [39]. A systematic mapping study is designed to provide an overview of a research field [38, 39]. From the RQ of our study, we derived the following sets of keywords: (i) “monolith”, (ii) “migrate, transform, refactor, decompose, extract, partition”, and (iii) “microservices”. To define the search string, we composed these keywords with their lexical and syntactic alternatives (synonym, plural, gerund, etc.). For defining the search string, we also considered eight related pieces of work in the topic [11, 14, 19, 22, 33, 41, 46, 48]. The string was used for searching studies in six digital libraries, as presented in Table 1. The search on these libraries was conducted on March, 28th 2020. A total of 422 studies were retrieved from all databases. Other digital libraries could be used, however, we chose the ones frequently considered in software engineering reviews.

For the screening of the studies, we adopted four inclusion criteria (IC) and two exclusion criteria (EC): IC-1: papers written in English and peer reviewed; IC-2: the goal of the study is the migration from monolithic to microservices; IC-3: the study reports on strategies, guidelines, processes, etc., to migrate from monolithic to microservices; IC-4: full-text available; EC-1: secondary or tertiary studies; and EC-2: tutorials, keynotes, position papers.

From the 422 retrieved studies (see Table 1), after applying the ECs, 79 studies were removed (422-79 = 343). Then, by considering the ICs for the title, keywords, and abstract, 260 studies were removed (343-260 = 83). Finally, after entirely reading the studies, we removed another 21 papers (83-21 = 62). The final set of papers were composed of 62 primary sources, presented in Appendix 6.

To validate the final set of primary sources, we selected four secondary studies that completely [20, 41], or partially [11, 18],
focus on studies on the topic of modernization/migration from monolithic to microservices. We observed the overlap between studies identified in the libraries and the ones already mapped in the secondary studies. Table 2 presents the secondary studies, together with the number of studies they discuss and the studies we identified as relevant for our work (overlap). The selection of the primary sources from the secondary studies also take into account the ICs and ECs, resulting in 42 distinct studies in the overlap.

In the four secondary studies presented in Table 2, in which [11] is an extension [18]), a total of 4, 17, and 14 relevant studies were exclusively identified from one source, namely #1, #2, and #3, respectively. In addition, 2 studies were identified from both in #1 and #3, 1 study from #2 and #3, and 4 studies among the three secondary studies. From this selection of studies, we identified 42 relevant studies. Compared with the results of our systematic mapping study, all of the 42 papers identified as relevant are among the 62 primary sources. The 20 studies not identified in the secondary studies, are mainly from 2019 and 2020. This was expected, since the secondary studies published in 2019 have not been mapped by these more recent studies. These results strongly confirm that we composed a relevant set of primary sources for our study.

2.2 Data extraction

To answer the RQ of our study, we extracted three pieces of information from each of the 62 primary studies, as follows:

- **Driving forces**: the motivations to start or that originated the modernization of the monolith legacy system with microservices.
- **Process**: which were the activities/steps described in the paper to conduct the modernization with microservices.
- **Input/Output information**: the pieces of information used as input and generated as output for each activity/step along the modernization.

The extraction was done by six authors of this paper, and validated by two experts on microservices. The authors in charge of the extraction, received a training of 20 hours about the topic of modernizing with microservices, to level their background in the topic. This training was based on the discussion of three papers [2, 5, 48].

2.3 Data Classification and roadmap definition

To analyze the driving forces and distill a roadmap, we perform a classification and a qualitative analysis of the raw data to find significant concepts and explore their relationships [7]. Three authors focused on classifying the driving forces, two authors classified the input and output information, and two authors dealt with the classification of activities/steps. For cross-validation, at least two authors inspected each primary study to collect the evidence. Also, at least other two authors discussed the data extract for refining the classification. The final set of driving forces and the definition of the roadmap were obtained by an incremental process, in refinement cycles. All this process was aided by the use of spreadsheets.

3 RESULTS AND ANALYSIS

This section describes the results and analysis of our study. To answer our research question, firstly we describe the driving forces (why) to modernize their legacy. Secondly, we focus on the activities and information used (how) for conducting the modernization.

### 3.1 Driving Forces

We identified 11 driving forces for modernizing legacy systems with microservices. Table 3 presents these driving forces sorted by the most mentioned ones.

<table>
<thead>
<tr>
<th>Driving Force</th>
<th>Total</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent and automated deploy</td>
<td>23</td>
<td>S[52, 54, 59, 63–65, 68–71, 73, 74, 77, 81, 89–92, 97, 98, 102, 103, 105, 109]</td>
</tr>
<tr>
<td>Easier maintenance and evolution</td>
<td>17</td>
<td>S[64, 66, 67, 70, 73, 75, 78, 83, 85, 90, 94, 96, 104, 107, 109, 112]</td>
</tr>
<tr>
<td>Independence of teams</td>
<td>13</td>
<td>S[63, 69, 79, 81, 84, 89, 90, 97, 101, 103, 107, 111, 112]</td>
</tr>
<tr>
<td>Loosely coupled services</td>
<td>12</td>
<td>S[52, 56, 57, 69, 70, 72, 77, 80, 81, 86, 103, 106]</td>
</tr>
<tr>
<td>Cohesive services</td>
<td>10</td>
<td>S[56, 61, 69, 72, 77, 80, 82, 88, 90, 99]</td>
</tr>
<tr>
<td>Technology flexibility</td>
<td>9</td>
<td>S[53, 60, 68, 81, 90, 98, 103, 108, 113]</td>
</tr>
<tr>
<td>Infrastructure facilities</td>
<td>8</td>
<td>S[60, 62, 64, 93, 98, 102, 108, 113]</td>
</tr>
<tr>
<td>Agility enabler</td>
<td>7</td>
<td>S[62, 67, 69, 71, 82, 98, 101]</td>
</tr>
<tr>
<td>Easier reuse</td>
<td>3</td>
<td>S[59, 64, 67, 94]</td>
</tr>
<tr>
<td>Reduced time to market</td>
<td>3</td>
<td>S[68, 69, 71]</td>
</tr>
</tbody>
</table>

Considering operational aspects of the software development/delivery, scalability and deployment were mentioned in 43.55% and 37.10% studies, respectively, they are the two most common driving forces. For example, these two driving forces are related to faster S[81, 98, 103] and independent deploys S[52, 54, 65, 74, 77, 90, 92, 97, 103, 105, 109], continuous delivery S[63, 64, 73, 77, 103], software replication S[56], and increase the elasticity of large applications S[62, 78, 86, 108]. In the same category, independence of teams (20.97%) and infrastructure facilities (12.90%) are also mentioned. In the former, we can cite as example the team autonomy S[52, 69, 81, 90, 91, 97, 101, 103, 109]. For the latter, easy configuration of cloud services S[113], high reliability S[108] and availability of services S[60, 62, 64], and optimization of computational resources S[113].
From a technical perspective, easier maintenance and evolution (27.42%) are explicitly mentioned as motivation to adopt microservices. This driving force is a consequence of optimized modularity properties such as cohesive and loosely coupled services (together these forces are mentioned in 35.48% of studies). Interestingly, only in five primary sources both coupling and cohesion are mentioned as a concern [56, 69, 72, 77, 80]. Furthermore, technology flexibility (14.52%) and easier reuse (6.45%) are also pointed as driving forces that companies rely on to migrate to microservices.

Organizational driving forces are less often mentioned in the primary studies. For example, microservices are seen as agility enablers (11.29%) that are related to faster deployment cycles [69, 71, 98]. Finally, reduced time to market (≈5%) is related to the use of microservices for different purposes/systems [71]. Another example may be the incentive to the DevOps culture, with an emphasis on collaboration between developers and teams, ensuring shorter lead time and greater agility in software development [69].

In summary, we can observe that several driving forces motivate the modernization. These driving forces are based on different perspectives, namely operational, technical, and organizational. Interestingly, there is not a dominant driving force that is mentioned in all papers. Yet, differently from some studies advocating mostly for technical benefits of microservices, e.g., use of different technologies [4, 5, 14, 33, 41], we can see that supporting operational tasks are the most common driving forces to adopt microservices.

### 3.2 Activities and Information

From the 62 primary sources, eight studies do not clearly present a modernization process, namely [53, 64, 68, 75, 76, 82, 91, 105]. They only discuss problems with legacy systems, benefits of microservice adoption, or lessons learned with a modernization with microservice. Based on that, our analysis of activities and information of the modernization process relies on 54 primary sources. As described in Section 2, our study aims at describing how the modernization of legacy systems with microservices is done. In this sense, the information collected in the primary sources was the basis to define a roadmap for modernization.

In order to define the road map, we performed a comprehensive analysis of data extracted from the primary sources (see Section 2.3). Then, we defined a roadmap composed of eight activities, covering four different phases of the modernization with microservices. Figure 1 presents an overview of the phases and its activities for our proposed roadmap. The phases were proposed after we defined all the eight activities of the road map. When we grouped the activities, we observed a resemblance to the project management life cycle presented in the well-known Project Management Body of Knowledge (PMBOK Guide) [26, 40]. This was the basis for naming the phases of the proposed roadmap. To corroborate our analysis, Table 4 presents the studies that discuss each activity of the proposed roadmap. As expected, the first activities have received much more attention than the last ones. This indicates that the research on the topic of modernizing legacy systems with microservices is still in early stages, highlighting the need for more studies. In the following, we describe details of each activity, and discuss the observed input and output information.

![Figure 1: Proposed roadmap for modernizing legacy systems with microservices](image-url)
2. Understand the legacy system: the goal of this activity is to analyze and explore the legacy system to understand its implementation, its features, how these features interact with each other, and how the monolithic legacy architecture is organized. This activity is fundamental to the next ones.

According to a primary source S[62], the first task to understand the legacy system is to list all the features for which the system was designed for. The following task is the identification of the implementation artifacts of each feature to subsequently decouple/extract them as microservices, which is known as feature location [12, 43]. Practitioners can also use reverse engineering techniques to create high-level artifacts, e.g., UML diagrams, from low level artifacts, e.g., source code, to ease the understanding of the legacy system [6, 28]. For example, in S[73] engineers used code analysis tools and existing documentation to obtain a high-level view of system structure. Some studies describe the use of Domain-Driven Design (DDD), which recommends that the structure of the software code should match the business domain [15]. The primary sources S[69, 72, 113] use DDD to structure legacy system features in domains and subdomains with clear context and boundaries. These contexts help to understand and address complexities based on business intentions.

Regarding the information used as input to Understand the legacy system, studies do not make clear which artifacts they used S[56, 63, 74, 79, 87, 88]. Among the identified artifacts, source code is the most common. In addition to development artifacts, one study also mentions the use of slides, presentations, and contracts with customers, and unwritten knowledge among developers and engineers S[73]. Considering the output information, in addition to the artifacts presented below, stakeholders involved in the modernization must have a clear understanding of legacy system requirements S[79].

- **Input information**: source code S[67, 69–71, 73, 84, 85, 95, 101, 104, 109, 113]; requirements and textual documents S[65, 73, 77]; test cases S[84]; UML diagrams S[73, 95, 102]; use case specifications S[92, 112]; logs S[109, 113]; and database models/schema S[69, 73, 90].

- **Output information**: data flow graphs, execution traces, and call graphs S[56, 65, 88, 92]; usage scenario models, load profiles, and performance constraints S[84, 88]; visual representation of the system structure S[74, 88]; business capabilities and system operations S[56, 112]; feature model S[87]; and the most used functionalities, domains and subdomains, and bounded contexts S[69].

3. Decompose the legacy system: this activity aims to break down the monolithic architecture into small and cohesive units, which will be transformed into microservices. The decomposition/identification of microservices in monolithic legacy systems is acknowledged as one of the most complex activities of the modernization process S[109]. This is one of the most investigated activities, being the focus of 85% of primary sources (46 out of 54, see Table 4).

Although the topic of decomposing software system in modules has been discussed for a long time [37], the limits among modules in monolithic architectures are flexible, and the systems eventually evolve into a “big ball of mud” [16]. When these modules are microservices, the limits are physical, splitting their implementation, distribution, management, and monitoring S[61]. However, Decompose the legacy system is not only about partitioning the system to facilitate maintenance, but also defines how the system will be able to evolve and scale S[61]. Some authors recommend that the identification of microservices should be an incremental process S[78]. They argue that due to the complexity of this activity, practitioners will learn how to do it better for each microservice decoupled from the legacy S[78].

An important issue in the activity of identifying the microservices is the definition of the correct level of granularity, that is, the trade-off between size and number of microservices S[61]. Eliciting strong interface limits at the right level of granularity is a challenge inherited from SOA [13]. One study highlights that the proper decomposition of the legacy system into microservices, in the right granularity, is paramount to companies’ sustainability S[109]. Besides, when identifying what would be the correct level of granularity, practitioners should consider a checking task regarding the microservice smells [47], i.e., antipatterns, such as megaservice and microservice greedy.
The starting point to decompose the legacy system is to define the responsibility for each microservice, indirectly defining the size of the microservices [78]. Some authors recommend the microservices to be designed around business capabilities, and allow their independent development [76, 90]. A strategy to carry out the decomposition and define the granularity of microservices is by following the Principle of Single Responsibility (SRP) [70]. SRP states that each system unit must have no more than one reason to change [30]. This means that if we have different software components, each one must have an independent responsibility and need. In addition, to enable independent development, each microservice can be developed using technologies that best suit the characteristics of the business capability [76]. Another strategy is the identification of microservices based on system operations and their data [112]. The sources of information here are models that describe the system operations, i.e., the system API, and the variables that contain the information that the operations write and read. The decomposition is achieved by grouping operations so that each microservice can internally access and change only its own data. When a microservice needs to refer, i.e., read or write, to a data that belongs to another microservice, this is achieved only through the microservice API. This allows a good decomposition that guarantees low coupling, as well as high cohesion.

To aid the identification of microservices, some tools and approaches have emerged. An example is Service Cutter [98], a tool to identify microservices that are highly cohesive and loosely coupled. An approach with the same goal is proposed in [61], which is based on the semantic similarity of foreseen/available functionality described through OpenAPI specifications. In addition, some tools use clustering algorithms to “cut” the monolith [67, 77, 98]. Differently, the study [102] presents a method to identify microservices based on the analysis of UML class diagrams.

To corroborate our findings, a recent survey with practitioners indicates that defining microservice granularity is a multi-criteria task [50]. The decomposition of a legacy system into microservices should consider, for example, resource consumption, team structure, delivery cycles, business capabilities, and data access. As we discussed, this activity is related to technical, operational and organizational aspects of companies.

- **Input information**: the output of the previous activities [54, 55, 57, 94]; desired non-functional requirements [54, 66, 79, 80, 111]; cloud service provider requirements [93]; and API specifications [61].
- **Output information**: microservice candidates [54, 55, 66, 72, 77, 81, 97, 103] or identified microservices [57, 58, 100, 104].

4. **Define the microservice architecture**: this activity is responsible for defining a microservice-based architecture with the design decisions to meet functional and non-functional properties defined in previous activities. In a simple view, a microservice architecture describes the microservices, their APIs, and communications, which is usually done by using a lightweight protocol, e.g., HTTP/RESTful [76, 87, 113]. Here, practitioners also have to define a strategy of service discovery. In this architecture, the user interface, business logic, and databases (data, logic, and behavior) are packaged in a single application/microservice and deployed on a server [112]. In addition, despite the benefit of technology flexibility enabled by microservices (see Table 3), Wang et al. [50] identify that organizations should restrict the number of programming languages in order to facilitate maintenance and reuse.

Although the previous activity is designed to identify/decide on how to decompose the legacy system, here practitioners can refine the scope of the microservices, reasoning about the architectural role of each microservice, taking into account heterogeneity and decentralized governance [78]. Despite the fact that our proposed roadmap has not considered stakeholders for each activity, [50] recommends the definition of owners for each microservice, with well-defined service responsibilities and boundaries.

As part of the architecture definition, practitioners should decide about having a local infrastructure, or using a cloud service provider, such as Amazon Web Services, Microsoft Azure, or Google Cloud. The infrastructure must provide resilience and high availability in microservices, allowing for scalability management and continuous deployment [112].

- **Input information**: the output of the previous activities [109]; non-functional requirements such as availability, security and flexibility [87, 88]; microservices interactions, dependencies, and constraints [87]; Data access S[94]; Facade pattern S[90, 98].
- **Output information**: microservice architecture model S[60, 79] with design decisions regarding communication protocol S[60, 72]; and cloud service provider/infrastructure S[93].

5. **Execute the modernization**: the goal of this activity is to conduct the modernization by re-engineering the monolithic legacy system into microservices. According to S[111], microservices reduce the development complexity in comparison to monolithic architectures. However, for the modernization that relies on existing artifacts, this can lead to more effort in comparison to greenfield projects.

Some authors reinforce the adoption of continuous integration and deployment during the migration S[59, 60, 72], using tools such as GitLab and Jenkins. Another point is automation of the configuration management, for instance, using Docker containers S[93] with Spring Boot S[59], and Artifactory for package repository S[59]. It is important to highlight that microservice architecture enables DevOps [1, 49], in this sense, integration between development and operation practices can be leveraged during the modernization [3]. Similarly to monolithic architectures, refactoring should be performed regularly and documentation is as important as the source code S[58]. In addition, when implementing microservices, code management will be necessary. Challenges like managing common code shared by multiple microservices and preventing/dealing with breaking API changes will be faced by developers. In this case, API gateways and “sidecar technology” should be considered [50].

A primary source describes a model-based approach for the automatic modernization of monoliths to microservices S[63]. This approach relies on a domain-specific language, namely JetBrains MPS. A domain-specific language is a computer language created/specialized to a particular application domain, differently from a general-purpose language, which is broadly applicable across domains [31].
Modernizing Legacy Systems with Microservices: A Roadmap

6. Integrate the microservices and the legacy: this activity is responsible for the integration between microservices and the legacy system. As the modernization with microservices is an incremental process S[78], following a strangler pattern S[19], practitioners have to decide on how to put together the legacy and the modernized parts of the system. This is one of the least explored and briefly discussed activities in the primary sources (see Table 4). As we discussed in previous activities, architectural choices need to facilitate continuous delivery and deployment during the modernization S[23]. A strategy to aid the activity of Integrate the microservices and the legacy is the use of feature toggles S[24] and backward/forward compatibility S[42]. Feature toggles is a technique that supports rollback, canary testing, and A/B testing.

- **Input information**: the output of the previous activities S[60, 89]; workflows design and service interface S[111]; legacy scripts S[106].
- **Output information**: integrated services S[106, 111], Rest API S[112], API Gateway S[98].

7. Verify and validate the microservices: this activity focuses on verifying and validating whether the microservices were developed properly and the integration with the legacy was performed accordingly. This activity of Verify and validate the microservices can use the legacy system as oracle for testing tasks.

Considering a continuous integration environment, unit tests are expected to be created during the implementation of microservices, but in addition, black box integration tests are also required S[109]. According to a primary source S[100], automated tests must be performed to verify/validate the interoperability of the message exchanges, and check whether microservices are in accordance with the specified architecture. In the case of inadequate definition and implementation of microservices, or other bugs, developers can go back to the previously proposed activities and refine or redefine the problematic part. As the modernization is an incremental process, regression testing and continuous integration should be used to ensure that new bugs are not introduced S[83]. Furthermore, when adopting continuous deployment, practitioners must consider user’s feedback as part of the modernization verification and validation S[83].

- **Input information**: the output of the previous activities; usability constraints S[100]; improvements and new features S[100]; non-functional requirements S[79]; and feedback from users S[100].
- **Output information**: unit and integration tests S[69, 73, 100, 110] and black-box/functional testing S[109] reports; metrics about satisfaction of non-functional requirements S[79, 83]; and microservices ready to deploy S[72].

8. Monitor the microservices / infrastructure: this activity is devoted to constantly check the behavior of the deployed microservices. Monitor the microservices / infrastructure should focus on service availability, bottlenecks, performance, use of infrastructure resources, and so on. Here, practitioners are able to evaluate and analyze information of the modernized system, in comparison to the driving forces that motivated the modernization (see Activity 1 - Analyze the driving forces and Section 3.1). For example, if we observed that scalability is the most common driving force for modernization, then, in this monitoring activity, the performance of the microservices can be observed. In addition to checking system health, monitoring can also support decision-making S[59]. For example, which microservices are the most used by customers, and should receive more attention from the development experts.

In the recent survey with practitioners working with microservice architectures, all 58 participants agreed to the importance of a robustness logging and monitoring framework, to achieve a mature level of microservices development S[50]. In addition, 90% of the participants believe that logging and monitoring should be set up as early as the project starts. In fact, “early” in our proposed roadmap means to define, configure, and execute automated frameworks and tools also in the previous activities, such as Activity 4. Define the microservice architecture and Activity 5. Execute the modernization S[50]. However, these actions can be postponed to a later phase when the project shows failures and delays S[50]. To support practitioners on conducting the activity of Monitor the microservices / infrastructure, Dave and Degioanni describe five microservice monitoring principles S[9]: (i) monitor containers and what runs inside them, (ii) use orchestration systems, (iii) prepare for elastic, multi-country services, (iv) monitor APIs, and (v) map monitoring to the organizational structure.

- **Input information**: the output of the previous activities; microservices deployed S[72], team structures, development and operation processes S[59].
- **Output information**: logs and metrics S[72]; change in team structures S[59].

4 THREATS TO VALIDITY

In this section we present the threats to validity of our study, and how we mitigated/reduced them.

**Internal Validity**: The first threat to internal validity concerns the terms used in the search string. To address this threat we composed a set of keywords based on secondary studies S[11, 14, 19, 22, 33, 41, 46, 48], to make sure of covering the important terms. The second internal threat is the selection of the digital libraries to search for primary sources. In addition to selecting well-known libraries, we also validated the returned results comparing to primary sources of existing secondary study S[11, 18, 20, 41], which confirmed the reliability of the libraries. The third threat is related to incorrect classification of primary sources. To reduce this threat, the authors in charge of the classification received training on the topic and extensively discussed the extracted data and performed cross-validation of the data classification. To further mitigate internal threats, we compared our results with existing secondary studies, describing similarities and differences found. In addition,
we performed a validation of the proposed roadmap in comparison to a seminal modernization planning process.

**External Validity:** Our roadmap is a set of high-level activities designed to serve as a starting point for the modernization with microservices. In our perspective, this roadmap can be easily adopted and aid the conduction of the modernization in different contexts, proving its external validity. Once we performed a comprehensive analysis of existing primary studies that describe the modernization process, the external validity should be appropriate. However, we believe that validation of the roadmap in real-world scenarios and with the participation of practitioners are needed.

## 5 RELATED WORK

This section describes related work on the topic of modernizing monolithic legacy systems with microservices and highlights existing differences to our study.

Ponce et al.[41] performed a study relying on 20 primary studies on modernization with microservices. They conducted a rapid review and identified three techniques, namely model-driven, static analysis, and dynamic analysis. The most of the techniques use elements of design as input to identify microservice candidates. This correlates to some studies identified for the Activity 3. Decompose the legacy system of our roadmap. Fritzsch et al. [19] performed a survey with experts that adopted microservice to identify migration strategies. For example, they identify different ways to decompose monolithic systems into microservices. They mentioned the use of the *strangler pattern* and *functional decomposition* as strategies. Carvalho et al. [4] performed a survey with experienced practitioners that adopted microservices. They investigated the criteria used for supporting decision-making along microservice extraction. The practitioners mentioned that they commonly consider at least four criteria to identify microservices in legacy systems. Wang et al. [50] also performed a survey and interviews with practitioners. In this study they investigated the best practices learned by practitioners that adopted microservices. These authors also investigated the challenges faced by practitioners and how these challenges were overcome. These four studies differ from ours, as they focus on specific topics and issues of the migration, rather than how the modernization process is conducted as a whole.

Fritzsch et al. [20] describes a guide for decision making that classifies the decomposition techniques to aid the architects during the migration from monolithic to microservices. They discuss refactoring activities to support the modernization with microservices. They mention that decomposition of monolithic into services with appropriate granularity is the main challenge during the refactoring process. Similarly, Francesco et al. [10] identify and classify approaches to modernize with microservices. They found that most existing approaches focus on handling the trade-off between flexibility and complexity in the microservice architecture. Differently from our study, the goal of these studies is not the definition of a roadmap for aiding the conducting of the modernization process.

The work most similar to ours is presented by Silva et al. [8], in which the authors present a modernization process based on lessons learned from two studies (a pilot and a case study). The main drawback of their process is its validity, since it relies only on authors’ experience. Our roadmap presents a more robust process based on a systematic mapping of the literature and on the results obtained by several researchers. In addition, we present two additional phases/activities compared to Silva et al. [8].

There are other secondary studies that explore the microservice architectural style and its relation to important concepts/technologies such as, devops [49], granularity [22], SOA [14], microservice gains and pains [33], and microservice smells [46]. However, their motivations are different from our study. Although these studies provide useful insights to understand the migration process, they do not focus on defining a generic roadmap. In this context, to the best of our knowledge, our paper is the first to describe recommendations based on 62 studies about migration.

## 6 CONCLUSION

Modernization of monolithic systems with microservices is a trend, which has called the attention of the software engineering research community. However, despite the existence of many studies on this topic, most of them focus on specific points of the modernization. We observed that there is a lack of a comprehensive study investigating the modernization of legacy systems with microservices as a whole. To fill this gap, this paper presented a study based on a systematic mapping method to answer the research question: Why and how are monolithic legacy systems migrated to microservices?

The answer for this question is that, based on 62 studies, we find 11 different driving forces that motivate the modernization. These driving forces are related to technical, operational, and organizational aspects. By an extensive analysis of 54 studies, we could observe that the modernization is composed of eight activities, covering the phases of initiation, planning, execution, and monitoring, which can be accomplished in an incremental manner during the migration process. The identified activities were the basis to define a comprehensive roadmap to contribute to practitioners, researchers and tool builders.

As future work, we consider the validation of our roadmap in industrial settings. More studies focusing the activities less investigated, namely execution of the modernization, integration of microservices and legacy, verification and validation, and monitoring of the modernized system and infrastructure. We also see the opportunity for the development of tools to support the modernization.

## ACKNOWLEDGMENTS

This work is supported by the FAPERJ PDR-10 program under Grant No. 202073/2020 and CNPq under Grant No. 408356/2018-9.

## REFERENCES


Modernizing Legacy Systems with Microservices: A Roadmap


A PRIMARY SOURCES


