Past and Future of Software Architectural Decisions – a Systematic Mapping Study

Dan Tofan $^a$, Matthias Galster $^b$, Paris Avgeriou $^a$, Wes Schuitema $^a$

$^a$ Department of Mathematics and Computing Science, University of Groningen, Nijenborgh 9, 9747 AG Groningen, Netherlands

$^b$ Department of Computer Science and Software Engineering, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

Abstract

Context: The software architecture of a system is the result of a set of architectural decisions. The topic of architectural decisions in software engineering has received significant attention in recent years. However, no systematic overview exists on the state of research on architectural decisions.

Objective: The goal of this study is to provide a systematic overview of the state of research on architectural decisions. Such an overview helps researchers reflect on previous research and plan future research. Furthermore, such an overview helps practitioners understand the state of research, and how research results can help practitioners in their architectural decision-making.

Method: We conducted a systematic mapping study, covering studies published between January 2002 and January 2012. We defined six research questions. We queried six reference databases and obtained an initial result set of 28,895 papers. We followed a search and filtering process that resulted in 144 relevant papers.

Results: After classifying the 144 relevant papers for each research question, we found that current research focuses on documenting architectural decisions. We found that only several studies describe architectural decisions from the industry. We identified potential future research topics: domain-specific architectural decisions (such as mobile), achieving specific quality attributes (such as reliability or scalability), uncertainty in decision-making, and group architectural decisions. Regarding empirical evaluations of the papers, around half of the papers use systematic empirical evaluation approaches (such as surveys, or case studies). Still, few papers on architectural decisions use experiments.

Conclusion: Our study confirms the increasing interest in the topic of architectural decisions. This study helps the community reflect on the past ten years of research on architectural decisions. Researchers are offered a number of promising future research directions, while practitioners learn what existing papers offer.

Keywords: software architecture, architectural decisions, systematic mapping study

1 Corresponding author. Tel.: +31 503 633 968
E-mail address: d.c.tofan@rug.nl
# Table of Contents

1 Introduction ......................................................................................................................... 3

2 Research Methodology ......................................................................................................... 4

2.1 Research Directives ......................................................................................................... 5

2.1.1 Protocol Definition .................................................................................................... 5

2.1.2 Generic Decision Literature Survey ........................................................................... 6

2.1.3 Research Questions Definition .................................................................................... 7

3 Data Collection .................................................................................................................. 9

3.1 Source Selection and Search String .................................................................................. 9

3.2 Inclusion and Exclusion Criteria ....................................................................................... 10

3.3 Search Process ................................................................................................................ 11

4 Results .................................................................................................................................. 12

4.1 Overview of Selected Papers ........................................................................................... 12

4.1.1 Empirical Evaluation Approaches ............................................................................. 13

4.1.2 Publication Venues and Years ................................................................................... 14

4.2 RQ1 – Documenting Architectural Decisions .................................................................. 15

4.3 RQ2 – Functional Requirements and Quality Attributes ............................................... 17

4.4 RQ3 – Domain-specific Architectural Decisions ............................................................. 18

4.5 RQ4 – Descriptive and Normative Papers ...................................................................... 19

4.6 RQ5 - Addressing Uncertainty in Architectural Decisions ............................................. 21

4.7 RQ6 - Group Architectural Decisions .............................................................................. 21

5 Discussion ........................................................................................................................... 22

5.1 Analysis and Synthesis of Results .................................................................................... 22

5.1.1 Empirical Evaluation Approaches, Publication Venues and Years ............................ 22

5.1.2 Documenting Architectural Decisions ...................................................................... 24

5.1.3 Functional Requirements and Quality Attributes ....................................................... 24

5.1.4 Domain-specific Architectural Decisions .................................................................. 25

5.1.5 Descriptive and Normative Papers .......................................................................... 25

5.1.6 Uncertainty in Architectural Decisions ..................................................................... 28

5.1.7 Group Architectural Decisions .................................................................................. 28

5.2 Implications for Researchers and Practitioners ............................................................. 30
1 Introduction

The software architecture of a software-intensive system is created in the early development phases and evolves throughout the whole development cycle. The architecture influences the extent to which a system achieves its functional and quality requirements [1]. The academic software architecture community accepts that a software architecture is the result of a set of architectural decisions [2]. Architectural decisions are a subset of design decisions that are costly to change and hard to make [3]. Also, architectural decisions involve important choices on core components or connectors, and the overall software-intensive system, to satisfy and balance stakeholders’ concerns [3]. Capturing architectural decisions promises significant practical benefits: lower costs of change, increased reuse of architectures, and less design erosion. These benefits are obtained by offering architects access to the decisions that lead to the architecture, rather than only the resulting outcomes and artifacts from the decisions [2]. These practical benefits encouraged more and more researchers to work on the topic of architectural decisions.

Given the increasing importance of architectural decisions in software engineering, there is a need to provide a systematic overview of the current state of research on architectural decisions. To achieve this goal, we present a systematic mapping study. The overview in this systematic mapping study provides value by offering the list of papers on architectural decisions, clusters of papers based on various research topics (i.e. detailed in sub-section 2.1.3) relevant to architectural decisions, as well as findings and future research directions derived from the clusters of papers.

Such an overview benefits two types of audiences. First, the overview enables researchers to reflect critically on the current state of research on architectural decisions. Moreover, the overview offers researchers promising future research directions, such gaps in current research and topics that need further attention. Second, the overview enables practitioners to learn about the state of existing research on architectural decisions, and potentially adopt state of the art approaches (e.g. methodologies or tools) and use other insights on architectural decisions (e.g. empirical evidence about the importance of architectural decisions) in their own architectural decision-making practices. We provide a more detailed discussion on the benefits of our study when introducing the rationales of our research questions in sub-section 2.1.3.
Previously, only partial overviews of the state of research on architectural decisions have been presented: Falesti et al. [1] survey fifteen architectural decision-making techniques proposed in the literature, with the goal of helping architects choose among decision-making techniques for their practice. Shahin et al. [4] compare features of nine tools for documenting architectural decisions from the literature. Bu et al. [5] analyze various aspects of design reasoning (such as rationale reuse) in nine decision-centric architectural design approaches from the literature. Overall, these three studies only offer partial overviews of research on architectural decisions. In contrast, we present a systematic overview following a rigorous approach to identify all relevant papers. Furthermore, rather than comparing existing tools or approaches, we aim to map the whole domain of architectural decisions.

Architectural decisions are part of architectural knowledge (in addition to other aspects, such as context and assumptions [6]). Thus, we also identified overviews of the state of research on the topic of architectural knowledge. De Boer and Farenhorst [7] present a systematic literature review on definitions of architectural knowledge, and notice that many definitions of architectural knowledge in fact refer to architectural decisions. Tang et al. [8] compare several tools for architectural knowledge management. Li et al. [9] present a systematic mapping study on applying knowledge-based approaches in software architecture. Overall, studies on architectural knowledge are on a different abstraction level, compared to studies on architectural decisions. Furthermore, this study has a different goal compared to other literature reviews [7-9], and a broader scope than existing partial overviews [1, 4, 5].

The rest of this paper is structured as follows. Section 2 presents the mapping study methodology, research questions, study search strategy, data extraction and analysis. Section 3 presents the data collection efforts. Section 4 presents the answers to the research questions. Section 5 discusses the study results and limitations. Section 6 concludes the paper.

2 Research Methodology

A systematic mapping study is an evidence-based form of secondary study that provides a comprehensive overview of a research area, identifying common publication venue types (e.g. conference or journal), quantitative analyses (e.g. number of published studies per year), and research findings in the investigated research area [10]. Systematic mapping studies offer multiple benefits. First, mapping studies identify gaps and clusters of papers based on frequently occurring themes in current research, including the nature and extent of empirical data on a topic, using a systematic and objective procedure [11]. Second, mapping studies help plan new research, avoiding effort duplication [12]. Third, they identify topics and areas for future systematic literature reviews, a more in-depth form of secondary studies with focus on smaller research areas, compared to mapping studies.

We also considered performing a systematic literature review, rather than a systematic mapping study. These two types of literature reviews differ as follows. First, mapping studies are suitable for analyzing broad research areas, compared to systematic literature reviews which are suitable for answering in-depth research questions [13]. Typically, systematic literature reviews cover fewer papers than mapping studies [13]. Second, mapping studies focus on broad analysis of the literature (e.g. by classifying and summarizing literature), rather than in-depth analysis of the literature (e.g. in terms of outcomes and quality assessments of the papers) [10, 13].
From our previous work on architectural decisions [14, 15], we noticed that much work exists on architectural decisions. However, no systematic overview of this area exists. This prevents an in-depth analysis of the literature in this research area (e.g. analysis of a specific type of architectural decisions). Therefore, given the large amount of work and architectural decisions and lack of existing systematic overviews, we concluded that a systematic mapping study benefits the community more than a systematic literature review, by offering a broad perspective of existing work on architectural decision. Future systematic literature reviews could start from the results of this mapping study and define research questions for sub-topics from literature on architectural decisions.

To conduct this study, we extended the mapping study process proposed by Petersen et al. [10] and used a process similar to [12, 16], that included the definition of a data collection form and a study protocol. As shown in Figure 1, the authors of [12] split the process proposed by [10] into three phases: 1) Research directives, 2) Data collection, and 3) Results. In addition to the process in [12, 16], we added an extra step to the first phase: we surveyed generic decision literature (detailed in Section 2.1.2), after defining the protocol. The survey of generic decision literature resulted in extra dimensions for classifying papers, from which we derived additional research questions to the preliminary questions defined in the protocol. In the second phase (detailed in Section 3), we identified relevant papers according to inclusion and exclusion criteria defined in the protocol. In the third phase (detailed in Section 4), we created the classification scheme to classify existing papers on architectural decisions using the instructions from [10, 17] and performed the actual mapping of current literature.

![Systematic mapping study process used in this paper.](image)

2.1 Research Directives

In this sub-section, we present the first phase of the mapping study process. In this phase, we define the research protocol, survey generic decision-making literature, and define the research questions.

2.1.1 Protocol Definition

We developed the research protocol based on the template from [18]. In the protocol, we specified the study topic, its justification, and preliminary research questions. Also, we specified the search strategy (detailed in sub-section 3.1), selection criteria (detailed in sub-section 3.2), and a data extraction form. We reviewed and updated the protocol in several iterations.

In the protocol, we also indicated offering an overview of the selected papers in terms of their empirical evaluation approaches. The overview on empirical evaluation approaches indicated if existing research
used empirical evidence, and what kind of empirical evidence. The rationale of such an overview was to understand what empirical evidence supports existing work. For example, a paper might validate an approach using a case study with practitioners. Synthesizing such information can indicate what approaches are most used and least used in research on architectural decisions. Results on most used approaches offer researchers examples for conducting future empirical studies. Results on least used approaches encourage researchers to use such approaches, since evidence from multiple empirical evaluation approaches is stronger than evidence from one approach. Also, these results help practitioners judge the validity and applicability of research results on architectural decisions. For example, practitioners might consider that research results from evaluations with other practitioners are more applicable than results from evaluations with students.

Other mapping studies also present aspects related to empirical evaluation approaches, such as research type in [12, 19], which use a classification that includes opinions, evaluations, and solution proposals. We used the classification from [9, 16, 17] that includes experiments, case studies, and surveys.

In the protocol, we also indicated offering an overview of the selected papers in terms of their publication venues and years. This overview helps researchers and practitioners understand the leading venues to publish or read about research results on architectural decisions. Although there is clear interest in the community on architectural decisions, there is no empirical evidence about the trend with the number of publications on architectural decisions over the years. Such trend helps researchers and practitioners understand if there is a growing or decreasing in the interest on the topic of architectural decisions, and if there are gaps in the interest on the topic of architectural decisions. Moreover, other mapping studies also present publication venues and years for the selected papers [9, 12, 16, 19].

We defined research questions (detailed in sub-section 2.1.3) derived from two sources:

1. **Software architecture decisions literature** helps classify relevant papers using research questions emerging from the relevant papers themselves (such as the role of non-functional requirements in architectural decision-making).
2. **Generic decision literature** helps classify relevant papers using research questions derived from a larger body of work, beyond software architecture. We further motivate and explain the use of generic decision literature in sub-section 2.1.2).

### 2.1.2 Generic Decision Literature Survey

Leveraging generic decision literature enables us to position research on architectural decisions in a larger body of work. Generic decision literature refers to literature on decisions independent of a domain and independent from software architectural decisions (e.g. strategic decisions in organizations).

Architectural decisions are a particular type of decisions (i.e. decisions made by software architects, while architecting software systems). Since research on generic decisions is more mature than research on architectural decisions, we consider that this perspective allows us to reuse results, and to better transfer research results from the generic decisions literature to the software architecture community. For example, if the generic decision body of work offers results for compensating bias in decision-making, then software architecture researchers can investigate such results to address bias in architectural decision-making. Therefore, we conducted a lightweight literature survey to identify possible items for the research questions and the classification scheme. We used the following criteria for identifying relevant generic decision literature:
1. We chose to use books instead of articles, because books offer more comprehensive and broad content on an existing body of knowledge in an area. In contrast, research articles tend to present novel, in-depth contributions.
2. The books should target an academic audience, because an academic audience has high expectations on content quality (e.g. providing detailed references and avoiding speculations).
3. The authors of the books should have a publication track record on decision-related topics in peer-reviewed venues to indicate their expertise on decisions. We ensured this by checking the background of authors and their publication record.
4. The books should focus on generic decision-making field.

We identified six books (i.e. [20-25]) that comply with the above criteria. We are aware that other books on decision-making exist that comply with the criteria. However, we consider that the six books offer sufficient content for identifying representative ideas from generic decision-making literature. We read each book and extracted major ideas from them. We considered a major idea relevant for our study if it referred to major established concepts (e.g. estimating probabilities is a relevant major idea, but specific approaches for how to do this are out of scope). Next, we consolidated the major ideas and discussed the potential applicability of each idea as a mapping dimension for literature on decision-making in software architecture. Finally, following discussions among researchers, we kept the mapping dimensions that we found most relevant for the software architecture field, and formulated four additional research questions based on them. In the next sub-section, we present all research questions and their rationale, including the four research questions derived from this literature survey.

2.1.3 Research Questions Definition
We present the six research questions for this study and their rationale. We categorize the research questions in two groups, based on their source (see sub-section 2.1.1).

2.1.3.1 Research Questions Derived from Software Architecture Literature

RQ1. What are the papers on documenting architectural decisions?
Rationale: Documenting decisions reduces architectural knowledge vaporization, which, in turn, reduces maintenance costs [26]. Various approaches have been proposed for documenting architectural decisions. However, an overview of papers on documenting architectural decisions is currently missing. Such an overview helps researchers analyze existing approaches on documenting architectural decisions and identify gaps in existing work. In addition, practitioners can use the proposed approaches to improve their decision documentation practices.

RQ2. Does current research on architectural decisions consider functional requirements and quality attributes?
Rationale: In their activities, architects need to consider the functional requirements and quality attributes (or non-functional requirements) of software systems. In addition, quality attributes play an important role in the decision-making process, since architects must make tradeoffs between quality attributes (e.g. security versus usability). However, in some decision-making situations, specific quality attributes become architectural key drivers and therefore receive more attention. For example, the quality attribute security would be an architectural key driver when architecting a security-intensive system. Answering RQ2 helps researchers identify quality attributes that are rarely addressed in current decision-making approaches and therefore might need more attention in future research. Also, practitioners can use the
answer to RQ2 to select approaches from the literature to make decisions related to specific quality attributes.

**RQ3. What specific domains for architectural decisions are investigated?**
Rationale: Architectural decisions are made in various domains. Domains include application domain (e.g. healthcare) and technology domain (e.g. service-oriented architectures). A paper may belong to more domains. Different domains bring different challenges for architects, so architects in industry can benefit from choosing approaches that are geared towards the challenges of a particular domain. Furthermore, answering RQ3 helps researchers identify domains that may need more attention. Practitioners can use the answer to RQ3 to select approaches from the literature to help them in their domain-specific architectural decision-making. This offers them better targeted approaches, compared to generic approaches for architectural decision-making. Also, practitioners learn whether different domains require different approaches or if the domain has only limited influence on how decisions are made.

**2.1.3.2 Research Questions Derived from Generic Decision Literature**

**RQ4. What are the normative and descriptive papers?**
Rationale: Generic decision literature distinguishes between normative and descriptive theories about decisions [21, 22]. Descriptive theories aim at explaining and predicting how decisions are actually made in the real world. In contrast, normative theories aim at prescribing how decisions should be made in a rational manner. However, normative and descriptive theories complement each other. Developing a normative theory (i.e. on how architectural decisions should be made) benefits from understanding how architectural decisions are actually made. Thus, we use the normative/descriptive classification for papers on architectural decisions. Answering RQ4 helps researchers understand the existing descriptive papers, and plan future descriptive studies. For example, current descriptive studies might present only parts of the lifecycle of real-world architectural decisions. Thus, future descriptive studies can present the full lifecycle of such decisions, from initial need to make the decision to its actual implementation and results. The answer to RQ4 includes a list of descriptive papers that helps researchers uncover such missing aspects. Also, researchers can use the existing descriptive papers to propose better normative approaches. Practitioners can use descriptive work on real-world architectural decisions to understand how other practitioners deal with architectural decisions. In addition, practitioners can use approaches from normative work to improve their decision-making activities.

**RQ5. What are the papers on addressing uncertainty in architectural decisions?**
Rationale: Addressing uncertainty is a major issue in generic decision literature [21, 23, 24], because most decisions involve uncertainty about the future consequences of choosing a certain alternative. If a decision does not involve uncertainty, then such decision is trivial to make, because the decision maker can simply choose the alternative with the highest benefits. Generic decision literature proposes various approaches to address uncertainty, such as Bayesian theory.

Uncertainty increases the difficulty of architectural decisions [27]. For example, an architect might design a new software system, and when deciding on a piece of technology, the architect might be confronted with uncertainty regarding the future of that technology, or on how well the technology satisfies scalability requirements. Therefore, addressing uncertainty is important for architectural decisions. Answering RQ5 enables researchers to understand the existing approaches for addressing uncertainty, and propose improved approaches. Also, practitioners learn which papers help them address uncertainty.
RQ6. What are the papers on group architectural decisions?
Rationale: Group decisions is an important topic in generic decision literature [20-23, 25], because many important decisions are made by groups, rather than individuals. Group decisions entail different challenges compared to individual decisions (e.g. ‘group thinking’ is a bias occurring in groups, not individuals). The majority of architectural decisions are group, rather than individual decisions [27]. Therefore, answering RQ6 helps researchers understand and improve approaches for group architectural decisions. In addition, practitioners can improve their group decision-making skills by learning from existing approaches.

Next, we present the steps for answering these six research questions.

3 Data Collection
The study search strategy must lead to inclusion of relevant papers and exclusion of irrelevant papers. The search strategy of this study involves querying reference databases with customized search strings, followed by manual filtering of the query results, using predefined inclusion and exclusion criteria. Three researchers were involved in executing the search strategy.

3.1 Source Selection and Search String
Since using only one reference database might miss some of the relevant papers on architectural decisions, we queried six reference databases, all of which index software engineering papers:

1. ACM Digital Library
2. IEEE Xplore
3. ScienceDirect
4. Scopus
5. SpringerLink
6. Web of Science

We searched for papers published between the 1st of January 2002 and the 1st of January 2012. We chose 2002 as a starting date, because in our previous work on architectural decisions, we observed that many papers on architectural decisions refer to an influential position paper [26] from 2004, and very little work on architectural decisions existed before that. In addition, the first conference dedicated to software architecture (Working IEEE/IFIP Conference on Software Architecture) took place in 1999. Therefore, papers from early 2002 bring a comprehensive overview of existing work on architectural decisions. We chose 1st of January 2012 as end date, because this mapping study started in April 2012.

To evaluate the results of the queries on the reference databases, we developed a quasi-gold standard, as recommended by [28]. The quasi-gold standard is a manually collected set of relevant papers from a small number of venues, which are well-known for publishing work on the relevant topic. The results of the queries on the reference databases must include the quasi-gold standard. Otherwise, the search strategy needs to be revised. The quasi-gold standard for this study consisted of 40 papers that we collected manually from three venues: the European Conference on Software Architecture, the Working IEEE/IFIP Conference on Software Architecture, and the Journal of Systems and Software. Working on the quasi-gold standard helped us validate and refine the search string, as well as the inclusion and exclusion
criteria. Finally, all items in the quasi-gold standard appeared in the results of the queries on the reference databases.

Starting from our research questions, we identified the keywords for the search string. Furthermore, we refined the search string using the papers from the quasi-gold standard. For example, in addition to ‘architectural decision’, we added ‘architectural knowledge’, because many authors consider architectural decisions as part of architectural knowledge. Moreover, we identified relevant synonyms for the initial keywords, such as ‘design decision’. Also, we considered variations of keywords, such as ‘architecture’ and ‘architectural’, as well as singular and plural forms, such as ‘decision’ and ‘decisions’. Finally, all the items in our search string (Table 1) are connected with OR statements, to make sure that all relevant papers are retrieved.

Table 1. Search string. Similar items are on the same line.

<table>
<thead>
<tr>
<th>Search string items</th>
</tr>
</thead>
<tbody>
<tr>
<td>('architecture decision' OR 'architectural decision')</td>
</tr>
<tr>
<td>OR ‘architecture choice’ OR ‘architectural choice’</td>
</tr>
<tr>
<td>OR ‘architecture decisions’ OR ‘architectural decisions’</td>
</tr>
<tr>
<td>OR ‘architecture choices’ OR ‘architectural choices’</td>
</tr>
<tr>
<td>OR ‘architecture rationale’ OR ‘architectural rationale’</td>
</tr>
<tr>
<td>OR ‘architecture knowledge’ OR ‘architectural knowledge’</td>
</tr>
<tr>
<td>OR ‘design decision’ OR ‘design decisions’</td>
</tr>
<tr>
<td>OR ‘design choice’ OR ‘design choices’</td>
</tr>
<tr>
<td>OR ‘design rationale’</td>
</tr>
<tr>
<td>OR ‘design knowledge’)</td>
</tr>
</tbody>
</table>

In each reference database, we searched the above string in the title, abstract, and keywords fields. Depending on the options offered by each reference database, we refined results by the selected time interval and by the relevant topic (e.g. software engineering).

3.2 Inclusion and Exclusion Criteria
We formulated inclusion and exclusion criteria for filtering papers. If a paper met an exclusion criterion, then the paper was removed from the study. Only if a paper met all inclusion criteria, the paper was kept. When a researcher was not sure about including or excluding a paper, the other researchers were asked to discuss and decide.

The inclusion criteria are:

I1. The study refers to the software architecture of software-intensive systems (e.g. not hardware or buildings architecture)
I2. The paper focuses on the topic of software architectural decisions

Regarding I2, the focus on architectural decisions means that the paper uses an ‘architectural’ perspective (or level of abstraction), rather than a more generic one such as ‘design’. Architectural decisions are a sub-category of design decisions [3]. However, papers on design decisions are excluded, because design decisions are out of scope for this study. Only if a paper focuses on ‘early design decisions’ (i.e. architectural design decisions), rather than design decisions in general, such paper is included, after discussions among researchers.
Some papers mention architectural decisions incidentally (e.g. a paper presents a tool implementation and some decisions for implementing this tool with a clear focus on the tool rather than related decisions). Such papers do not meet I2 (i.e. no focus on architectural decisions). Typical examples of papers with a focus on architectural decisions are papers on 1) documenting architectural decisions (e.g. templates, viewpoints, or tool support), 2) making architectural decisions (e.g. approaches or tool support for decision-making), or 3) describing real-world architectural decisions (e.g. exploratory studies to investigate how architectural decisions are made in industrial projects).

The exclusion criteria are:

E1. Papers in a language other than English
E2. Papers published before 1st of January 2002 or after 1st of January 2012. Even though the search interval was defined in the search scope, we defined this as an exclusion criterion since some databases did not allow filtering the search results based on time. Thus, this criterion was applied on results provided by database searches when no time filtering option was available
E3. Gray literature, because of their unclear peer review process, as recommended by [29]: editorials, extended abstracts, tutorials, tool demos, doctoral symposium papers, research abstracts, book chapters (other than proceedings), keynote talks, workshop reports, and technical reports
E4. Secondary studies on architectural decisions, because such papers are related work to this study
E5. Papers that focus on hardware topics, computer-aided non-software design (e.g. industrial), non-software design (e.g. products, systems, or where the main product is not software)
E6. Papers from other venues than software engineering (e.g. physics, law)
E7. Papers with a distinct body of work, which focus on specialized related topics, such as components selection or software release decisions. Although these topics can be considered architectural decisions, they have a distinct and mature body of work (e.g. much literature exists on components selection)

Two researchers reviewed the criteria for each study, using title, keywords, and abstract. If needed, the researchers reviewed also the paper content. Differences between researchers were discussed and reconciled, sometimes involving a third researcher.

### 3.3 Search Process

We followed the search process in Figure 2. We used the search string on each of the six reference databases and we obtained various numbers of papers for each venue. We used EndNote to manage the references. The tool allowed us to remove some duplicates from the initial set of 28,895 references automatically. We split the remaining 28,201 references in batches of up to 3,000 references, for easier handling.

Next, two researchers filtered the articles in each batch by title, by removing articles with titles that were clearly out of scope. When not sure about keeping a paper or not, we chose to keep the paper, to avoid the risk of filtering relevant papers. Researchers shared their experiences and discussed examples of papers, to make sure that they had a common understanding about the filtering process. Following the filter by title step, we obtained 2,283 references.

Two researchers filtered the papers by reading the abstract and the introduction sections of each paper. In addition, if a researcher was not sure about keeping a specific paper, the researcher read the conclusion
section. If still not sure, the researcher kept the paper, to decide about it in the final step. This step resulted in 262 references.

Finally, we read the full content of each paper, to make the final decision about each paper. When a researcher was not sure about keeping or removing a paper, the other two researchers analyzed the paper, and made a final decision. During discussions, the two researchers used a dialectical decision-making approach, inspired from [30]: initially, each researcher argued either for keeping or removing the paper, and then a consensus was reached after debates. This approach helped us evaluate systematically the arguments for keeping or removing a paper, and avoid the decision-making bias of relying heavily on the initial impression about a paper. Overall, this step produced the final set of 144 references.

4 Results
We prepared a data extraction form with fields corresponding to each research question, and two researchers extracted data for the 144 papers. If the two researchers needed help on data extraction, then a third researcher was involved. Next, we used descriptive statistics and frequency analysis to answer the research questions. We present an overview of the papers and the answers to the six research questions from sub-section 2.1.3.

4.1 Overview of Selected Papers
We present an overview of the selected papers in terms of their empirical evaluation approaches and publication venues/years.
4.1.1 Empirical Evaluation Approaches
We classified each paper in one of the following empirical evaluation approaches: experiment, survey, case study, or example. Easterbrook et al. [33] consider experiments, surveys and case studies as the established research methods that are most relevant to software engineering. By examples, we refer to early validation attempts (such as using a hypothetical situation or a toy example in the paper), which can be a preliminary step for using later an established research method.

To classify papers, we checked if papers referenced established guidelines for conducting experiments (i.e. [34, 35]), surveys (i.e. [36-39]), or case studies (i.e. [40-42]). If a paper did not reference any guideline, we could still classify it as an experiment, survey, or case study, provided that the paper included important elements from the guidelines. For example, a paper can include one or more of the following elements: a protocol, research questions, hypotheses, validity threats, and interpretation of results. If a paper lacked such elements, it was classified as using an example for its empirical evaluation. Table 2 summarizes the numbers of papers for each category. This table also shows the type of publication venue at which a paper has been published.

<table>
<thead>
<tr>
<th></th>
<th>Examples</th>
<th>Case studies</th>
<th>Surveys</th>
<th>Experiments</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal</td>
<td>25</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>35</td>
</tr>
<tr>
<td>Conference</td>
<td>41</td>
<td>23</td>
<td>11</td>
<td>2</td>
<td>77</td>
</tr>
<tr>
<td>Workshop</td>
<td>25</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>32</td>
</tr>
<tr>
<td>Total</td>
<td>91</td>
<td>31</td>
<td>16</td>
<td>6</td>
<td>144</td>
</tr>
</tbody>
</table>

To zoom into the results in Table 2, we created the bubble chart in Figure 3, which classifies all papers on the venue type and empirical evaluation approach. Based on Table 2 and Figure 3, we notice the following points.
First, examples are the dominant evaluation approach: 63% of the papers on architectural decisions use examples as their evaluation approach. Examples are used by 78% of the workshop papers, which is not surprising, since workshops tend to present work in progress. Thus, the evaluation is rather an illustration of a new approach. As much as 71% of the journal papers rely on examples. One explanation is that some papers are published in journals that target practitioners, who are interested in the core findings, rather than details from a thorough evaluation approach. This becomes clearer when we filter the seven papers from the IEEE Software and the one paper from the IBM Systems journals (both journals are geared towards practitioners). The remaining papers make up for 50%, which is slightly less than the percentage for conferences (i.e. 53%).

Second, experiments are, by far, the least used evaluation approach: only 4% of all papers use experiments. Out of the six papers with experiments, only P106 reports an experiment with fourteen architects. Three of the papers (P45, P46, and P47) report experiments with 25-50 students each. The remaining two papers (P62 and P102) use a mix of students and architects, in total 10 (for P62) and 16 participants (for P102).

Third, surveys make for 11% of all papers. Most surveys are published in conferences. Only one survey is published in a workshop. Overall, surveys have similar distributions to case studies, but the number of surveys is about half the number of case studies, which make for 21% of all papers.

### 4.1.2 Publication Venues and Years

Table 3 shows the top ten most popular publication venues for papers on architectural decisions. We notice that the top ten venues include around 57% of the 144 selected papers. The SHARK workshop series, ECSA/WICSA conferences, JSS and IEEE Software journals are the most popular venues for
publications on architectural decisions. In the Appendix (sub-section 7.2), we present all venues for the 144 selected papers. In total, there are 59 different venues: 13 workshops, 29 conferences, and 17 journals.

Table 3. Number of papers published in the top ten most popular publication venues.

<table>
<thead>
<tr>
<th>Venue</th>
<th>Type</th>
<th>No</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workshop on SHAring and Reusing Architectural Knowledge</td>
<td>Workshop</td>
<td>17</td>
<td>11.81</td>
</tr>
<tr>
<td>European Conference on Software Architecture</td>
<td>Conference</td>
<td>16</td>
<td>11.11</td>
</tr>
<tr>
<td>Working IEEE/IFIP Conference on Software Architecture</td>
<td>Conference</td>
<td>10</td>
<td>6.94</td>
</tr>
<tr>
<td>Journal of Systems and Software</td>
<td>Journal</td>
<td>7</td>
<td>4.86</td>
</tr>
<tr>
<td>Working IEEE/IFIP Conference on Software Architecture/European</td>
<td>Conference</td>
<td>7</td>
<td>4.86</td>
</tr>
<tr>
<td>Conference on Software Architecture</td>
<td>Journal</td>
<td>7</td>
<td>4.86</td>
</tr>
<tr>
<td>IEEE software</td>
<td>Journal</td>
<td>7</td>
<td>4.86</td>
</tr>
<tr>
<td>International Conference on Software Engineering</td>
<td>Conference</td>
<td>6</td>
<td>4.17</td>
</tr>
<tr>
<td>Quality of Software Architectures</td>
<td>Conference</td>
<td>5</td>
<td>3.47</td>
</tr>
<tr>
<td>Information and Software Technology</td>
<td>Journal</td>
<td>4</td>
<td>2.78</td>
</tr>
<tr>
<td>Workshop on Traceability, Dependencies and Software Architecture</td>
<td>Workshop</td>
<td>3</td>
<td>2.08</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td></td>
<td>82</td>
<td>56.94</td>
</tr>
</tbody>
</table>

We found evidence on the increasing interest in the community on the topic of architectural decisions. Figure 4 shows the distribution of the 144 papers over the past ten years (i.e. 2002 – 2011). We plotted the numbers of papers for each year, and added trend lines for the intervals 2002 – 2004 and 2005 – 2011. We notice that between 2002 and 2004, there were very few papers on architectural decisions. However, the number of papers on architectural decisions has grown steadily since 2005. Overall, we see that since 2005 research on architectural decisions started to gain much traction in the community.

Next, we answer the six research questions.

4.2 RQ1 – Documenting Architectural Decisions
Much interest exists in documenting architectural decisions for reducing architectural knowledge vaporization [26]. However, no overview exists on the papers that contribute to the topic of documenting architectural decisions. To provide such an overview, we checked each of the 144 papers to see if they refer to documentation, or include a process for documentation, or tool support for documentation of
architectural decisions. To answer RQ1, we provide the overview in Figure 5 and offer examples of papers for each sub-category.

Figure 5 shows that out of the 144 selected papers, 83% (or 120 papers) propose some form of documentation approach for architectural decisions. The remaining 24 papers do not refer to documentation, processes or tool support for documentation of architectural decisions. Most of these 24 papers (i.e. 18) are either case studies or surveys that took place in the industry, describing various aspects of architectural decision-making. For example, P129 presents a case study on real-world architectural decisions, and P9 presents a survey on organizational factors that influence architectural decisions in data warehouse projects.

Out of the 120 papers on documentation, 24 papers present no process and no tool support for documenting architectural decisions. Instead, they focus on other topics related to architectural decisions. For example, P19 explores ideas for embedding rationale of decisions in architecting activities, and P60 focuses on business aspects of architectural decisions.

Seventy-six papers on documentation also propose a process that results in documentation for architectural decisions. These 76 papers consist of 44 papers with no tool support, 26 papers with custom-made tool support, and 6 papers with off-the-shelf tool support. For example, P2 proposes an approach for facilitating the architectural decision-making process, but without tool support. Also, P144 presents an approach to help the architectural decision-making process, with a custom made wiki-based tool support. Finally, P111 proposes a process for assisting change impact analysis for architectural analysis, using off-the-shelf tools such as an UML editor and a Bayesian Belief Network tool.

\[\begin{array}{c|c|c}
\text{144 papers: all} & \text{120 papers: documentation} & \text{52 papers: tools} \\
\hline
24 papers - no documentation, no process, no tool support: P9, P17, P33, P35, P38, P42, P48, P59, P61, P67, P85, P91, P93, P97, P98, P109, P119, P120, P129, P130, P131, P132, P133, P141 \\
\hline
24 papers on documentation, no process, no tool support: P12, P19, P27, P32, P39, P40, P41, P43, P49, P53, P54, P55, P60, P68, P70, P71, P89, P103, P115, P118, P125, P126, P139, P142 \\
\hline
\hline
26 custom tools: P4, P6, P18, P22, P26, P62, P74, P76, P77, P79, P80, P81, P83, P86, P87, P88, P92, P95, P104, P114, P116, P122, P123, P127, P137, P144 \\
\hline
6 off-the-shelf tools: P3, P8, P36, P58, P101, P111 \\
\hline
16 custom tools: P7, P10, P11, P20, P21, P23, P24, P25, P31, P37, P56, P65, P69, P75, P136, P143 \\
\hline
4 off-the-shelf tools: P1, P44, P72, P102
\end{array}\]

As visible in Figure 5, 52 out of the 120 papers include tool support for documenting architectural decisions. From these 52 papers, 32 papers (26 custom tools and 6 off-the-shelf tools) overlap with the
category of papers that also include a process. The remaining 20 papers use custom (16 papers) and off-the-shelf (4 papers) tools for documenting architectural decisions. For example, P7 proposes an ontology for documenting architectural decisions, and custom tool support for it. Also, P102 presents an experiment for visualization of architectural decisions using the off-the-shelf, open source tool Compendium. Regarding open-source tool support, four papers (i.e. P1, P44, P101, P102) use off-the-shelf open-source tools (i.e. Protégé in P1, OSATE in P44, Compendium in P101 and P102), and one paper (i.e. P136) presents a custom open-source tool (i.e. Frag). Overall, many papers include tool support, most of the tools are custom made, rather than off-the-shelf tools, and very few papers include open-source tool support.

4.3 RQ2 – Functional Requirements and Quality Attributes
Architectural decisions must help satisfy functional requirements and achieve quality attributes for a software system. We investigated if the collected papers mention addressing functional requirements or quality attributes. Moreover, we investigated whether the collected papers focus on specific functional requirements or specific quality attributes.

We mapped the papers by analyzing them for mentions of functional requirements and/or quality attributes. If we could not find such explicit mention, then we marked the paper as having an unclear treatment of functional requirements and/or quality attributes. Thus, each paper was classified in one of the four possible states, as follows:

- S1. Unclear functional requirements, and explicit quality attributes
- S2. Explicit functional requirements, and unclear quality attributes
- S3. Unclear functional requirements, and unclear quality attributes
- S4. Explicit functional requirements, and explicit quality attributes

Figure 6 summarizes the results. Out of the 144 selected papers, thirteen papers have unclear treatment of functional requirements but discuss quality attributes (i.e. S1), ten have unclear treatment of quality attributes but discuss functional requirements (i.e. S2), and seven have unclear treatment of both (i.e. S3). For example, P42 is a conceptual paper that compares architectural decisions with design decisions, and addressing functional requirements and quality attributes is implicit in the paper. The remaining 114 papers (i.e. the 144 collected papers minus the 30 papers in Figure 6) clearly address both functional requirements and quality attributes (i.e. S4) and are not shown in Figure 6.

<table>
<thead>
<tr>
<th>Unclear Functional Requirements</th>
<th>S1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P4, P6, P13, P60, P61, P81, P82, P107, P108, P112, P113, P126, P137</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unclear Quality Attributes</th>
<th>S2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P9, P30, P31, P34, P47, P56, P97, P127, P128, P129</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unclear Functional Requirements</th>
<th>S3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P42, P68, P69, P130, P131, P132, P133</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Out of the 124 papers that explicitly treat functional requirements, we found no specific functional requirements that are recurring among papers, since such requirements vary much with the actual software architecture project.
Out of the 127 papers that explicitly treat quality attributes, seven papers focus on specific quality attributes. P8 refers to achieving both **security** and **reliability**, through a decision-making framework. The other six papers focus on only one quality attribute, as follows:

- Achieving **usability** is the focus of P103 and P104. Both publications refer to decisions on usability patterns (e.g. fixed or requested menus).
- Achieving **reliability** is the focus of P82, through a decision-making framework for achieving reliability.
- Achieving **scalability** is the focus of P58, through goal-oriented analysis and simulation of different decision scenarios.
- Achieving **evolvability** in P16 is proposed through a quality model for supporting evolvability, for evaluating decisions that may affect evolvability, such as choice of architectural patterns.
- Achieving **safety** is proposed by P124 through a framework for eliciting and formulating negative requirements (e.g. unplanned or unwanted events).

Overall, most collected papers address explicitly both functional requirements and quality attributes.

### 4.4 RQ3 – Domain-specific Architectural Decisions

Out of the 144 collected papers, 22% of the papers (or 32 papers) refer to domain-specific architectural decisions. The other papers refer to architectural decisions in general. Figure 7 presents the domains and the IDs of the papers. Some domains refer to specific application domains (e.g. telecommunication, web applications, healthcare, defence), while other domains refer to generic domains (e.g. SOA, enterprise architecture, software product lines). We notice that the domains are not strictly orthogonal, for example, a paper from the healthcare domain can also be a SOA paper. However, due to the low number of papers, we were not able to separate the domains. Only three domains have three or more papers: software product lines, enterprise architecture, and service-oriented architecture (SOA). We discuss briefly the papers in each of the three domains.

**Service-oriented architecture domain:**
• P49 presents an approach to support service design decisions.
• P52 proposes modeling SOA process decisions.
• P53 presents a template for documenting SOA decisions.
• P54 proposes a set of factors that affect SOA decision-making, such as types of service consumers, and perspectives of service providers and consumers.
• P93 analyzes in-depth the selection between two types of services (i.e. RESTful and WSDL/SOAP web services).
• P116 proposes a decision model for SOA projects.
• P140 presents a decision-modeling framework for SOA systems.
• P141 describes the rationale for various SOA architectural decisions from a real-world project.
• P142 proposes a design approach for decisions on SOA transactional workflows.

Enterprise architecture domain:

• P2 and P4 consider the architectural design as a search problem, and propose approaches for searching the design space.
• P9 investigates organizational factors (e.g. resource constraints, perceived skills of IT staff) for decisions on enterprise data warehouses.
• P85 presents an approach for collaborative decision-making for the design of enterprise architecture.
• P94 presents a process model for architectural decisions management.
• P97 reports the role of enterprise architecture in group decisions in e-Government.
• P143 proposes a conceptual framework for collaborative decision-making, identifying and enforcing decisions in enterprise architectures.

Software product lines domain:

• P20 presents tool support for capturing architectural decisions for software product lines.
• P103 discusses decisions on usability patterns for product lines.
• P114 uses ideas from software product lines to increase reusability of documented architectural decisions.

4.5 RQ4 – Descriptive and Normative Papers
We classified each paper as either a normative, or a descriptive paper. We identified 20 descriptive papers, representing 13.8% of all collected papers: P9, P32, P33, P35, P43, P48, P59, P61, P91, P97, P98, P117, P119, P120, P129, P130, P131, P132, P133, and P141. The remaining 124 papers are normative.

The number of descriptive papers is much lower than the number of normative papers. However, descriptive papers are very important for understanding real-world architectural decisions. Furthermore, descriptive papers can be used to propose more targeted approaches for normative papers. For example, a descriptive paper can present real-world challenges (such as how to increase consensus for group architectural decisions), and such challenges can be addressed in normative papers that propose approaches for group architectural decisions. If few descriptive papers exist, then researchers have little idea on real-world architectural decision-making, and might risk focusing on a subset of challenges.
Given the importance of descriptive papers, we need to understand to what extent these papers cover real-world architectural decisions and decision-making.

To understand the descriptive papers, we define four factors for descriptive papers. The factors help researchers understand the real-world decisions from the literature, and plan future research to cover existing gaps in the literature. These factors are not used to assess the quality of descriptive papers. Next, we present the factors, their rationale, and characterize the 20 descriptive papers using the factors.

**Number of described architectural decisions:** This factor indicates the breadth of a paper, with regard to the paper’s contribution on describing architectural decisions: a paper that describes many decisions has a broad contribution. By describing many architectural decisions, researchers can synthesize findings based on multiple decisions. Overall, researchers benefit from papers that present a clear number of described decisions.

Out of the 20 descriptive publications that we identified, sixteen refer to an unclear number of architectural decisions. The remaining four publications refer to one (P9), fifteen (P33), twenty-eight (P132), and eighty (P43) architectural decisions from real-world projects.

**Time spent for making architectural decisions:** We chose this factor because saving time for architects is critical, due to their busy schedules. To propose timesaving approaches for architects, researchers need to understand how architects spend time in making architectural decisions. Papers that describe decisions over a longer time might present new challenges of real-world decisions.

Out of the 20 descriptive papers, only P33 refers to the time spent for making architectural decisions (i.e. the number of minutes for each architectural decision), in the context of a study on observing and analyzing the design process of practitioners.

**Number of participants:** This factor is applicable to both types of papers. First, for descriptive papers, it refers to the number of persons that were involved in the decisions in the paper. Second, for normative papers, it refers to the number of persons that participated in an empirical evaluation. This factor is important because more participants indicate papers with stronger empirical evidence, for both descriptive and normative papers.

Table 5 in sub-section 5.1.5 (in which the results are further discussed) summarizes the results: four papers do not describe clearly the number of participants involved in the decisions described in the empirical evaluation. The number of participants in the other papers varies from 3 to 436 (see Table 5). All participants are from industry, except for P119.

**Classes of architectural decisions:** We chose this factor because understanding the kind of decisions described in the literature helps researchers understand which classes of decisions need further attention, including better descriptions of the actual decisions. According to [31], architectural decisions can be classified in three classes:

- **existence** decisions (i.e. indicating existence of some artifacts, such as that the system will have three layers).
- **property** decisions (i.e. stating an enduring trait of the system, such as that the system will use open-source libraries).
out of the 20 descriptive papers, two papers (P32 and P98) describe existence decisions. In addition, only P9 has an executive decision (i.e. data warehouse selection). No paper describes property decisions. However, four papers (P35, P117, P129, and P141) describe multiple decisions, from all classes of decisions. The remaining thirteen papers describe unclear classes of decisions.

We discuss the results for descriptive and normative papers in sub-section 5.1.5.

4.6 RQ5 - Addressing Uncertainty in Architectural Decisions

Out of the 144 papers, only nine papers (or 6%) address uncertainty in architectural decision-making. Given the low number, we summarize below the approaches in the papers.

- P5 proposes a process for identifying risks in architectural decisions, and the use of Bayesian networks to quantify and manage the risks.
- P111 proposes quantifying the relationship between architectural decisions and design artefacts using Bayesian belief networks.
- P36 analyzes ‘build vs. buy’ architectural decisions, and uses probabilities to check reliability constraints for the decisions.
- P46 proposes an approach for documenting the rationale of architectural decisions, which includes documenting the occurrence probabilities of the various scenarios for decisions.
- P61 mentions the need for probabilities of market changes and technology risks.
- P77 considers architectural decisions as solving a multi-attribute decision problem, which uses the occurrence probabilities for possible combinations of alternatives.
- P82 uses probabilities for making decisions that result in more reliable software systems.
- P96 proposes an approach for evaluation design alternatives that uses the probabilities for evaluating consequences of the alternatives.
- P109 discusses the influence of cognitive biases on evaluating probabilities that influence architectural decisions.

4.7 RQ6 - Group Architectural Decisions

We found out that 15% (or 22 papers) from the selected papers refer to group architectural decision-making. We summarize below the approaches in the papers.

- P3 uses the Analytical Hierarchy Process with three stakeholders who evaluate the importance of various quality attributes for an architectural decision.
- P32 describes the importance of drawings for focusing the group discussions on various decisions.
- P33 analyzes the dynamics of decision-making of three teams of two practitioners.
- P43 describes consensus for decisions made by the architecting team at Volvo Cars.
- P45 presents an experiment on team architectural decision-making.
- P48 and P59 present a survey with architects portraying them as lonesome, rather than team decision makers.
• P53 presents a template for documenting SOA architectural decisions. The template was used by teams of three students to document their SOA decisions.
• P66 proposes an extension to the CBAM [32] framework, which considers explicitly stakeholders’ preferences in group decision-making.
• P83 proposes a traceability framework for group decisions, based on integrating knowledge from various sources.
• P84 describes a framework to quantify economically the value of architectural design decisions.
• P85 proposes an approach for supporting collaborative decision-making for enterprise architectures.
• P97 describes an instance of group architectural decision-making in a Finnish e-Government project.
• P101 analyzes the group collaborative features of various tools for capturing architectural decisions.
• P103 mentions involving team members in architectural decisions.
• P104 describes an approach for group decision-making, which includes a facilitator.
• P106 describes an industry study about consensus in group decision-making.
• P129, P130, P132, and P133 describe the impact of group interactions on decision-making.
• P143 presents a framework that facilitates group decision-making.

5 Discussion
In this section, we present analyses and syntheses of results for all research questions, so that the numbers in the systematic overview in Section 4 are used to propose future research directions. Afterwards, we present implications for researchers and practitioners, and discuss validity threats.

5.1 Analysis and Synthesis of Results

5.1.1 Empirical Evaluation Approaches, Publication Venues and Years
To understand further the empirical evaluation approaches for papers on architectural decisions, we present in Figure 8 two charts with the evolution of validation approaches over time. Systematic evaluations are especially relevant for conferences and journals, but matter less for workshops, since workshops include early ideas, for which systematic evaluations might be less feasible to conduct. Overall, we present two charts with the evolution of systematic evaluations and examples for conferences and journals.

The left chart shows the number of studies that use examples, and the number of systematic studies (i.e. total number of experiments, surveys, and case studies) for conferences and journals. The right chart shows the ratio of number of systematic studies on the number of all studies (e.g. in 2003 the ratio is 1, as the only paper for 2003 had a systematic evaluation). We notice that the ratio was particularly low in 2009, when 4 systematic studies and 19 papers with examples were published.
We make several observations on the charts in Figure 8. First, similar to the trend in Figure 4, the numbers of conference and journal papers are very low (i.e. less than three) for 2002, 2003, and 2004, which explains the extreme values for their corresponding ratios. Second, for 2005 and 2006, we notice that about a third of all conference and journal papers had systematic evaluations. Third, since 2007 about half of all conference and journal papers had systematic evaluations. The exception is 2009, for which the ratio is much lower, due mostly to only 4 papers with systematic evaluations, 6 papers with examples at joint WICSA/ECSA conference, 3 papers with examples at the Journal of Systems and Software, and 2 papers at IEEE Software.

Here are basic statistics (i.e. median, mode, and average) about the citation count of the 144 papers. The median citation count for the 144 selected papers was six, and the mode citation count was zero (i.e. 17 papers had no citations). Since most cited papers have more influence, we present the top 10% (i.e. 14) most cited papers out of the 144 selected papers in Table 4, and their empirical evaluation, venue type, publication year, and average citation count.

Table 4 shows some interesting results. Regarding the citation count, P93 has the highest number of citations, with most citations from papers on service-oriented systems. Regarding publication year, we notice three papers (i.e. P17, P70, and P118) published during 2002 – 2004, which influenced much subsequent work on architectural decisions.

Figure 8. Number of conference and journal papers with systematic evaluations (case studies, experiments, surveys) and examples (left chart), their ratio (right chart).
Table 4. Overview of the top 10% most cited papers on architectural decisions.

<table>
<thead>
<tr>
<th>Paper ID</th>
<th>Empirical evaluation</th>
<th>Venue type</th>
<th>Year</th>
<th>Citation count</th>
<th>Average citation count</th>
</tr>
</thead>
<tbody>
<tr>
<td>P93</td>
<td>Example</td>
<td>Conference</td>
<td>2008</td>
<td>440</td>
<td>88</td>
</tr>
<tr>
<td>P115</td>
<td>Example</td>
<td>Journal</td>
<td>2005</td>
<td>281</td>
<td>35.13</td>
</tr>
<tr>
<td>P63</td>
<td>Example</td>
<td>Conference</td>
<td>2005</td>
<td>264</td>
<td>33</td>
</tr>
<tr>
<td>P72</td>
<td>Example</td>
<td>Conference</td>
<td>2006</td>
<td>196</td>
<td>28</td>
</tr>
<tr>
<td>P17</td>
<td>Example</td>
<td>Workshop</td>
<td>2004</td>
<td>214</td>
<td>23.78</td>
</tr>
<tr>
<td>P70</td>
<td>Example</td>
<td>Workshop</td>
<td>2004</td>
<td>158</td>
<td>17.56</td>
</tr>
<tr>
<td>P32</td>
<td>Survey</td>
<td>Conference</td>
<td>2007</td>
<td>102</td>
<td>17</td>
</tr>
<tr>
<td>P65</td>
<td>Example</td>
<td>Conference</td>
<td>2007</td>
<td>95</td>
<td>15.83</td>
</tr>
<tr>
<td>P71</td>
<td>Example</td>
<td>Journal</td>
<td>2009</td>
<td>55</td>
<td>13.75</td>
</tr>
<tr>
<td>P118</td>
<td>Example</td>
<td>Journal</td>
<td>2002</td>
<td>148</td>
<td>13.45</td>
</tr>
<tr>
<td>P55</td>
<td>Example</td>
<td>Journal</td>
<td>2007</td>
<td>80</td>
<td>13.33</td>
</tr>
<tr>
<td>P10</td>
<td>Example</td>
<td>Workshop</td>
<td>2007</td>
<td>79</td>
<td>13.17</td>
</tr>
<tr>
<td>P141</td>
<td>Case study</td>
<td>Conference</td>
<td>2005</td>
<td>94</td>
<td>11.75</td>
</tr>
<tr>
<td>P116</td>
<td>Example</td>
<td>Journal</td>
<td>2009</td>
<td>38</td>
<td>9.5</td>
</tr>
</tbody>
</table>

5.1.2 Documenting Architectural Decisions

The key facts are that 83% of all papers on architectural decisions propose some form of documentation approach, and that 53% of all papers propose processes that result in documentation of architectural decisions. These facts indicate that documenting architectural decisions is a well-covered research topic.

Given the large number of studies on documenting decisions, we conclude that there is a need to consolidate research on documenting architectural decisions. This includes obtaining evidence on the real-world benefits for documenting architectural decisions. Such evidence can be obtained from success stories from early adopters of documentation approaches. The core benefit for documenting architectural decisions lower knowledge vaporization and therefore maintenance costs of software systems (as in [26]). The cost reduction takes place by capturing, sharing and reusing architectural knowledge, thus reducing the time that new developers need to familiarize with an existing software system and potentially increasing the quality of the software system. However, documenting decisions is a cost by itself. Therefore, the main goal of consolidating research on documenting architectural decisions is to get insights on the right amount of documentation that offers most benefits, at acceptable costs.

Regarding tool support for documenting architectural decisions, we notice that very few tools are open-source and easily accessible for practitioners (i.e. four off-the-shelf open-source tools, and one custom open-source tool). We consider that the community can benefit from more open-source tools for documenting architectural decisions since they would facilitate the adoption of documentation approaches by practitioners. Thus, we encourage future efforts in this direction.

5.1.3 Functional Requirements and Quality Attributes

The results in Section 4.3 suggest that papers on architectural decisions have considered functional requirements and quality attributes. We think these trends are due to the wide acceptance of the concepts of functional requirements and quality attributes in the community. For the future, we expect this acceptance to persist.
Given the wide acceptanc of quality attributes, we observe the low number of papers on addressing specific quality attributes, such as scalability and reliability. Satisfying specific quality attributes requires dedicated approaches, which offer better-targeted approaches, compared to generic approaches for making architectural decisions. For example, P58 indicates the use of simulations to compare scalability in various workloads and processing resources (such as adding more processing resources or more powerful processing resources). As a future direction, we encourage more work on approaches for satisfying specific quality attributes.

Regarding papers for satisfying functional requirements, although such requirements vary much across projects, we think that there is much potential for distilling architectural decisions that can be reused to satisfy common functional requirements across software projects and products. For example, architectural decisions for some functional requirements (e.g. from open source projects) can be reused by other architects. An encouraging piece of work in this direction is [43], which presents the architectures of various popular open source projects. Researchers can collect reusable decisions in decision repositories for satisfying functional requirements from these projects, and practitioners can reuse such decisions in their own architectural decision-making.

5.1.4 Domain-specific Architectural Decisions
The results in Section 4.4 show an interesting trend: domain-specific architectural decisions refer mostly to the service-oriented and enterprise domains. These domains are well-established in the sense that there is considerable research and practice on these two domains. In contrast, other domains need more attention.

Mobile computing is a domain with huge recent adoption, as part of the post-PC era (i.e. continuous decline of personal computers sales in favor of mobile devices, such as tablets and smartphones). Much development effort exists for creating mobile applications (i.e. apps) for mobile operating systems (e.g. iOS, Android). Architecting apps involves making mobile-specific architectural decisions. Research on mobile-specific architectural decisions is missing from our results. Still, such research has much potential to help practitioners architect apps. Similar to mobile computing, we consider that other neglected domains are cloud computing and internet of things.

5.1.5 Descriptive and Normative Papers
Descriptive papers help researchers understand real-world architectural decisions and decision-making. Furthermore, descriptive papers help propose better approaches for normative papers. In Table 5, we present the 20 descriptive papers, which consist of five workshop papers, eleven conference papers, and four journal papers. These numbers suggest that conferences are the most popular venue type for publishing descriptive papers.

For each paper, we indicate its average citation count. Citation count indicates the impact of a paper in the community. We gathered the citation count for each of the 144 selected papers as indicated by Google Scholar (including self-citations), at the start of April 2013. Since papers might receive more citations over the years, we calculate the average citation count by dividing the number of citations with the number of elapsed years since publication.
Table 5. Summary of descriptive papers, sorted by average citation count.

<table>
<thead>
<tr>
<th>Paper ID</th>
<th>Empirical evaluation</th>
<th>Number of participants</th>
<th>Venue type</th>
<th>Year</th>
<th>Average citation count</th>
</tr>
</thead>
<tbody>
<tr>
<td>P32</td>
<td>Survey</td>
<td>436</td>
<td>Conference</td>
<td>2007</td>
<td>17</td>
</tr>
<tr>
<td>P141</td>
<td>Case study</td>
<td>Unclear</td>
<td>Conference</td>
<td>2005</td>
<td>11.75</td>
</tr>
<tr>
<td>P9</td>
<td>Case study</td>
<td>Unclear</td>
<td>Conference</td>
<td>2005</td>
<td>11.75</td>
</tr>
<tr>
<td>P129</td>
<td>Case study</td>
<td>25</td>
<td>Journal</td>
<td>2007</td>
<td>5.83</td>
</tr>
<tr>
<td>P35</td>
<td>Survey</td>
<td>107</td>
<td>Conference</td>
<td>2007</td>
<td>4.5</td>
</tr>
<tr>
<td>P120</td>
<td>Survey</td>
<td>53</td>
<td>Conference</td>
<td>2011</td>
<td>3.5</td>
</tr>
<tr>
<td>P48</td>
<td>Survey</td>
<td>142</td>
<td>Conference</td>
<td>2009</td>
<td>3</td>
</tr>
<tr>
<td>P59</td>
<td>Survey</td>
<td>142</td>
<td>Journal</td>
<td>2011</td>
<td>3</td>
</tr>
<tr>
<td>P33</td>
<td>Case study</td>
<td>6</td>
<td>Journal</td>
<td>2010</td>
<td>2</td>
</tr>
<tr>
<td>P61</td>
<td>Survey</td>
<td>21</td>
<td>Workshop</td>
<td>2010</td>
<td>2</td>
</tr>
<tr>
<td>P130</td>
<td>Case study</td>
<td>27</td>
<td>Workshop</td>
<td>2005</td>
<td>1.75</td>
</tr>
<tr>
<td>P119</td>
<td>Survey</td>
<td>22</td>
<td>Conference</td>
<td>2010</td>
<td>1.33</td>
</tr>
<tr>
<td>P43</td>
<td>Case study</td>
<td>Unclear</td>
<td>Conference</td>
<td>2010</td>
<td>1</td>
</tr>
<tr>
<td>P91</td>
<td>Survey</td>
<td>9</td>
<td>Workshop</td>
<td>2010</td>
<td>1</td>
</tr>
<tr>
<td>P132</td>
<td>Case study</td>
<td>3</td>
<td>Conference</td>
<td>2007</td>
<td>1</td>
</tr>
<tr>
<td>P117</td>
<td>Case study</td>
<td>Unclear</td>
<td>Conference</td>
<td>2009</td>
<td>0.75</td>
</tr>
<tr>
<td>P131</td>
<td>Case study</td>
<td>12</td>
<td>Conference</td>
<td>2006</td>
<td>0.71</td>
</tr>
<tr>
<td>P133</td>
<td>Case study</td>
<td>3</td>
<td>Workshop</td>
<td>2007</td>
<td>0.17</td>
</tr>
<tr>
<td>P97</td>
<td>Case study</td>
<td>16</td>
<td>Workshop</td>
<td>2011</td>
<td>0</td>
</tr>
<tr>
<td>P98</td>
<td>Case study</td>
<td>Unclear</td>
<td>Conference</td>
<td>2010</td>
<td>0</td>
</tr>
</tbody>
</table>

Regarding empirical evaluation (RQ7), we notice that the descriptive papers consist of eleven case studies and nine surveys. The number of participants in the case studies and surveys varies from 3 to 436 (in Table 5). We observed a high correlation coefficient (i.e. 82.3%) between the number of participants (i.e. one of the factors in sub-section 4.5) in a descriptive paper and the average citation count of a paper. This suggests that descriptive studies with more participants receive more citations.

Based on the results for the other three factors in sub-section 4.5, we propose several recommendations for researchers who plan to conduct descriptive work on architectural decisions.

- Since most papers refer to an unclear number of decisions, we recommend more clarity on the number of decisions in descriptive papers. This would allow researchers to synthesize findings based on more decisions and therefore increase the validity of studies.

- Since only one paper discusses the time effort spent on making architectural decisions, we recommend paying attention towards understanding real-world time effort for making architectural decisions. We have already collected some data on the time effort spent on making architectural decisions in the real world [27], but more work is needed: Understanding the required time effort is a necessary step for researchers to propose approaches that help architects reduce the needed effort and to perform a cost-benefit analysis of systematic architectural decision making.

- Most papers refer to unclear classes of decisions. We recommend more clarity on the classes of decisions in descriptive work. This is because different classes of decisions may be treated differently, require different effort, etc. In addition, we recommend more descriptions of property...
decisions, since no paper describes property decisions. We consider that insights about property decisions are particularly useful for practitioners as property decisions include design rules, design guidelines, and design constraints, which influence many elements of a software system [31].

As an example of descriptive work from a related field, during the decision literature survey (in subsection 2.1.2), we found a thorough piece of work in the field of strategic organizational decisions: researchers analyzed hundreds of interviews on 150 real-world decisions in 30 organizations [25]. This descriptive study offered in-depth insights on real-world decision-making in various organizations (e.g. identifying and explaining various types of real-world decision-making processes [25]). These insights lead to over 20 publications on organizational decisions, including improved approaches for organizational decisions [25], such as a list of key factors that influence the success of implementing decisions [44]. Similarly, in-depth research on real-world architectural decisions would be very valuable.

Normative papers propose various approaches on architectural decisions. We notice that the 22 normative papers in Table 6 validate their proposed approaches using practitioners and students, who were asked to use the approaches on realistic architectural decisions. These normative papers offer empirical evidence for their proposed approaches, so that practitioners get a better idea of their value. We found a weak correlation between the number of participants in normative papers and their average citation count.

Table 6. Summary of normative papers with validations, sorted by average citation count.

<table>
<thead>
<tr>
<th>Paper ID</th>
<th>Empirical evaluation</th>
<th>Number of participants involved in empirical evaluation</th>
<th>Venue type</th>
<th>Year</th>
<th>Average citation count</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3</td>
<td>Case study</td>
<td>3</td>
<td>Conference</td>
<td>2005</td>
<td>8.88</td>
</tr>
<tr>
<td>P62</td>
<td>Experiment</td>
<td>16</td>
<td>Journal</td>
<td>2009</td>
<td>7.75</td>
</tr>
<tr>
<td>P94</td>
<td>Case study</td>
<td>3</td>
<td>Conference</td>
<td>2006</td>
<td>7.29</td>
</tr>
<tr>
<td>P45</td>
<td>Experiment</td>
<td>50</td>
<td>Conference</td>
<td>2006</td>
<td>3.29</td>
</tr>
<tr>
<td>P74</td>
<td>Survey</td>
<td>3</td>
<td>Conference</td>
<td>2008</td>
<td>3.2</td>
</tr>
<tr>
<td>P107</td>
<td>Survey</td>
<td>8</td>
<td>Journal</td>
<td>2005</td>
<td>3.13</td>
</tr>
<tr>
<td>P46</td>
<td>Experiment</td>
<td>50</td>
<td>Conference</td>
<td>2008</td>
<td>3</td>
</tr>
<tr>
<td>P47</td>
<td>Experiment</td>
<td>25</td>
<td>Workshop</td>
<td>2008</td>
<td>2.4</td>
</tr>
<tr>
<td>P66</td>
<td>Example</td>
<td>4</td>
<td>Journal</td>
<td>2005</td>
<td>2</td>
</tr>
<tr>
<td>P123</td>
<td>Survey</td>
<td>3</td>
<td>Conference</td>
<td>2010</td>
<td>2</td>
</tr>
<tr>
<td>P85</td>
<td>Survey</td>
<td>70</td>
<td>Conference</td>
<td>2010</td>
<td>1.67</td>
</tr>
<tr>
<td>P106</td>
<td>Experiment</td>
<td>14</td>
<td>Journal</td>
<td>2004</td>
<td>1.67</td>
</tr>
<tr>
<td>P23</td>
<td>Case study</td>
<td>45</td>
<td>Conference</td>
<td>2008</td>
<td>1.4</td>
</tr>
<tr>
<td>P112</td>
<td>Survey</td>
<td>7</td>
<td>Conference</td>
<td>2011</td>
<td>1</td>
</tr>
<tr>
<td>P21</td>
<td>Case study</td>
<td>45</td>
<td>Journal</td>
<td>2010</td>
<td>0.67</td>
</tr>
<tr>
<td>P31</td>
<td>Case study</td>
<td>8</td>
<td>Conference</td>
<td>2010</td>
<td>0.67</td>
</tr>
<tr>
<td>P2</td>
<td>Case study</td>
<td>1</td>
<td>Conference</td>
<td>2005</td>
<td>0.63</td>
</tr>
<tr>
<td>P6</td>
<td>Example</td>
<td>4</td>
<td>Workshop</td>
<td>2009</td>
<td>0.5</td>
</tr>
<tr>
<td>P5</td>
<td>Case study</td>
<td>17</td>
<td>Journal</td>
<td>2009</td>
<td>0</td>
</tr>
<tr>
<td>P53</td>
<td>Case study</td>
<td>105</td>
<td>Conference</td>
<td>2010</td>
<td>0</td>
</tr>
<tr>
<td>P102</td>
<td>Experiment</td>
<td>10</td>
<td>Workshop</td>
<td>2011</td>
<td>0</td>
</tr>
<tr>
<td>P113</td>
<td>Survey</td>
<td>20</td>
<td>Conference</td>
<td>2011</td>
<td>0</td>
</tr>
</tbody>
</table>
As future work, we consider that there is a need to investigate how to increase the practitioners’ acceptance of approaches on architectural decisions, so that practitioners can benefit from researchers’ efforts. For example, the Technology Acceptance Model [45] offers a powerful model on how users accept and use a new technology.

The remaining 102 normative papers propose approaches on architectural decisions, mostly for documenting architectural decisions. The answer to RQ1 and the corresponding discussion present in detail the documentation approaches in the normative papers.

5.1.6 Uncertainty in Architectural Decisions
We synthesize the nine papers (see sub-section 4.6) by identifying their core contributions towards addressing uncertainty in architectural decisions, and grouping the papers by how they address uncertainty, which is made explicit through probabilities. We identified two main categories of papers:

1. Basic addressing of uncertainty. Four papers (P46, P61, P82, and P109) recognize the importance of probabilities (i.e. of specific scenarios to occur), but do not propose concrete approaches for addressing uncertainty in architectural decisions.

2. Advanced addressing of uncertainty. Two papers (P77 and P96) use decision theory for making architectural decisions. In both papers, probabilities are considered first-class entities in decision-making that help evaluate systematically the consequences of each decision alternative. Other three papers (P5, P36, and P111) also regard probabilities as first-class entities in decision-making. In addition, these papers use Bayesian theory for impact analysis of each decision alternative.

Overall, although there is some work on addressing uncertainty in architectural decision-making, we consider that the topic is under-represented, given its practical importance. For example, we did not find work on how architects can improve their estimation of probabilities, so that architect can better address uncertainty. As a future trend, we consider that approaches for decision documentation should help architects quantify uncertainties, so that architects can present stakeholders how uncertainty levels influence the architectural decisions.

5.1.7 Group Architectural Decisions
Table 7 summarizes the 22 papers (i.e. ten conference, seven journal, and only four workshop papers) on group decisions. Regarding empirical evaluation (RQ8), four papers use surveys, eleven papers use case studies, two papers use experiments, and five papers use examples. Similar to the papers in Table 5, we observe a high correlation (75%) between the numbers of participants in the empirical evaluations and the average citation count of the papers. This suggests a link between using more participants in the empirical evaluation of a paper and the subsequent impact of that paper.

Table 7. Summary of papers on group decisions, sorted by average citation count.

<table>
<thead>
<tr>
<th>Paper ID</th>
<th>Empirical evaluation</th>
<th>Descriptive</th>
<th>Number of participants involved in empirical evaluation</th>
<th>Venue type</th>
<th>Year</th>
<th>Average citation count</th>
</tr>
</thead>
<tbody>
<tr>
<td>P32</td>
<td>Survey</td>
<td>Yes</td>
<td>436</td>
<td>Conference</td>
<td>2007</td>
<td>17</td>
</tr>
<tr>
<td>P3</td>
<td>Case study</td>
<td>No</td>
<td>3</td>
<td>Conference</td>
<td>2005</td>
<td>8.88</td>
</tr>
<tr>
<td>P143</td>
<td>Example</td>
<td>No</td>
<td>NA</td>
<td>Conference</td>
<td>2007</td>
<td>8.5</td>
</tr>
<tr>
<td>Paper</td>
<td>Type</td>
<td>Approach</td>
<td>No.</td>
<td>Source</td>
<td>Year</td>
<td>Score</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>----------</td>
<td>-----</td>
<td>---------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>P129</td>
<td>Case study</td>
<td>Yes</td>
<td>25</td>
<td>Journal</td>
<td>2007</td>
<td>5.83</td>
</tr>
<tr>
<td>P84</td>
<td>Case study</td>
<td>No</td>
<td>Unclear</td>
<td>Conference</td>
<td>2003</td>
<td>3.8</td>
</tr>
<tr>
<td>P45</td>
<td>Experiment</td>
<td>No</td>
<td>50</td>
<td>Conference</td>
<td>2006</td>
<td>3.29</td>
</tr>
<tr>
<td>P48</td>
<td>Survey</td>
<td>Yes</td>
<td>142</td>
<td>Conference</td>
<td>2009</td>
<td>3</td>
</tr>
<tr>
<td>P59</td>
<td>Survey</td>
<td>Yes</td>
<td>142</td>
<td>Journal</td>
<td>2011</td>
<td>3</td>
</tr>
<tr>
<td>P83</td>
<td>Case study</td>
<td>No</td>
<td>Unclear</td>
<td>Journal</td>
<td>2007</td>
<td>2.83</td>
</tr>
<tr>
<td>P101</td>
<td>Example</td>
<td>No</td>
<td>NA</td>
<td>Workshop</td>
<td>2010</td>
<td>2.67</td>
</tr>
<tr>
<td>P33</td>
<td>Case study</td>
<td>Yes</td>
<td>6</td>
<td>Journal</td>
<td>2010</td>
<td>2</td>
</tr>
<tr>
<td>P66</td>
<td>Example</td>
<td>No</td>
<td>4</td>
<td>Journal</td>
<td>2005</td>
<td>2</td>
</tr>
<tr>
<td>P130</td>
<td>Case study</td>
<td>Yes</td>
<td>27</td>
<td>Workshop</td>
<td>2005</td>
<td>1.75</td>
</tr>
<tr>
<td>P85</td>
<td>Survey</td>
<td>No</td>
<td>70</td>
<td>Conference</td>
<td>2010</td>
<td>1.67</td>
</tr>
<tr>
<td>P106</td>
<td>Experiment</td>
<td>No</td>
<td>14</td>
<td>Journal</td>
<td>2004</td>
<td>1.67</td>
</tr>
<tr>
<td>P43</td>
<td>Case study</td>
<td>Yes</td>
<td>Unclear</td>
<td>Journal</td>
<td>2010</td>
<td>1</td>
</tr>
<tr>
<td>P132</td>
<td>Case study</td>
<td>Yes</td>
<td>3</td>
<td>Conference</td>
<td>2007</td>
<td>1</td>
</tr>
<tr>
<td>P104</td>
<td>Example</td>
<td>No</td>
<td>Unclear</td>
<td>Conference</td>
<td>2006</td>
<td>0.71</td>
</tr>
<tr>
<td>P103</td>
<td>Example</td>
<td>No</td>
<td>NA</td>
<td>Journal</td>
<td>2011</td>
<td>0.5</td>
</tr>
<tr>
<td>P133</td>
<td>Case study</td>
<td>Yes</td>
<td>3</td>
<td>Workshop</td>
<td>2007</td>
<td>0.17</td>
</tr>
<tr>
<td>P53</td>
<td>Case study</td>
<td>No</td>
<td>105</td>
<td>Conference</td>
<td>2010</td>
<td>0</td>
</tr>
<tr>
<td>P97</td>
<td>Case study</td>
<td>Yes</td>
<td>16</td>
<td>Workshop</td>
<td>2011</td>
<td>0</td>
</tr>
</tbody>
</table>

There are relatively few normative papers with approaches for practitioners on group architectural decisions. Table 7 also indicates that 10 out of the 22 papers are descriptive work (RQ4), meaning that they offer evidence on group architectural decisions in the industry. However, the ratio of descriptive per normative papers is much higher for the papers in Table 7, compared to the similar ratio for papers on architectural decisions in general. This suggests an insufficient number of normative papers on group architectural decisions. The results from a recent survey on architectural decisions [27] which reports that 86% of architectural decisions in the industry are made by groups, rather than individual decision makers. Therefore, we suggest that future work is needed on approaches for group architectural decisions. To propose such approaches, researchers need to understand the key findings of existing papers on group architectural decisions.

The 10 descriptive papers contribute at increasing our understanding of group dynamics in architectural decision-making (e.g. P33, P97). However, more work needs to be done for understanding how to reach consensus among decision makers for architectural decisions, as indicated by P32. Additionally, Zannier’s papers (e.g. P129, P130) call for more descriptive work, to increase community’s understanding of group dynamics in architectural decision-making.

The 12 normative papers provide contributions for improving group architectural decision-making. We categorize these contributions in three types: CBAM [32] extensions, documentation approaches, and processes. P66 and P84 propose CBAM extensions that help decision makers elicit and evaluate alternatives for architectural decisions. The documentation approaches (in P45, P53, P83, and P143) help capture the perspectives of the multiple decision makers. Finally, the processes (in P3, P85, P103, and P104) help decision-makers structure their interactions for faster decision-making.
Furthermore, the normative papers indicate several areas that need future work. First, P3 and P66 call for more work on treating judgment uncertainty of the decision-makers. Second, other papers (i.e. P53, P66, P83, and P101) ask for better tool support for group decision-making. Third, P66 and P101 call for further empirical validations. Forth, P66 and P143 ask for further work on dependency analysis of decision makers’ perspectives.

5.2 Implications for Researchers and Practitioners

This mapping study confirms that architectural decisions are an increasingly important topic in software architecture research. In the recent decade, the number of papers on architectural decisions has increased steadily since 2005, as shown in Figure 4. Three seminal papers (P17, P70, and P118 – see sub-section 5.1.1) that were published before 2005 influenced much subsequent research, as indicated by their high citation count. The key message of the three seminal papers is to reduce maintenance costs by fighting the vaporization of architectural knowledge, through documentation of architectural decisions.

This study shows much effort has been invested in improving documenting architectural decisions. We agree with the importance of documenting architectural decisions. However, following this study, we speculate that critics who look at the efforts for documenting architectural decisions might demand evidence on how documenting architectural decisions actually reduces maintenance costs for the industry, as envisioned by the three seminal papers (i.e. P17, P70, and P118). As future work, researchers can use the 144 selected papers in this study to search for such evidence.

Following this study, we see value in diversifying research efforts to cover three additional topics on architectural decisions: descriptive work, uncertainty, and group architectural decisions. These topics receive much attention in generic decision-making literature (e.g. [7-12]). We do not see any reason for neglecting these topics in research on architectural decisions.

Addressing specific quality attributes is often needed in practice. For example, Poort et al. [46] bring evidence on the importance of modifiability in practice, but we could not identify any paper on architectural decisions that focuses on achieving modifiability. Also, Ameller et al. [47] investigate the role of specific quality attributes in architecting service-based systems. In this mapping study, we found only 7 papers that address specific quality attributes. Therefore, more work is needed for architectural decisions that address specific quality attributes.

Regarding domain-specific architectural decisions, we notice that most papers address architectural decisions in the SOA and enterprise architecture domains. There are very few papers for other domains, so more work is needed to cover other domains, such as mobile. This will lead to better architectural decisions by using targeted approaches for the specific domains.

Similar to other mapping studies [9, 19], we found improvement opportunities for the empirical evaluation of future studies, by moving away from examples to surveys, case studies or experiments. However, we found no clear trend (see Figure 8) suggesting that, over time, papers on architectural decisions improved their empirical evaluation.

When analyzing the results on empirical evaluation, we were surprised to find only six experiments on architectural decisions. Experiments use 49 participants on average [48]. Recruiting professional architects is very challenging, due to their busy schedules. Using students as participants raises validity
threats, although some authors indicate concrete steps towards reduce validity threats from using students [49]. However, experiments allow researchers to test hypotheses on causes and effects in a systematic manner. Therefore, despite the difficulties, conducting more experiments on architectural decisions is critical for advancing research on architectural decisions.

Regarding implications for practitioners, we notice the following points. First, this study shows that current work on architectural decisions offers a rich set of approaches for documenting architectural decisions (see sub-section 4.2), so practitioners can incorporate documentation approaches in their activities. Second, this study indicates papers that help practitioners achieve specific non-functional requirements (see sub-section 4.3) and in certain specific domains (see sub-section 4.4) for their architectural decisions. Third, this study helps practitioners understand the existing approaches for addressing uncertainty in architectural decisions (see sub-section 4.6), and for improving group decision-making (see sub-section 4.7). However, practitioners must be aware that the maturity of these topics varies: most mature research results are on documenting architectural decisions, and more work is needed for the other topics. Therefore, we encourage practitioners to document their architectural decisions using approaches from the literature, and offer researchers feedback about the approaches.

5.3 Validity threats
In this mapping study we addressed conclusion, construct, internal and external validity threats.

5.3.1 Conclusion Validity
Conclusion validity refers to obtaining the same study results, if other researchers replicate the study [33]. By making explicit the criteria for including and excluding papers, we believe that our conclusions are valid and can be replicated using the same research questions for three reasons. First, we followed a systematic mapping study process (in Section 2), which helps other researchers replicate the study. Second, we made explicit our data collection efforts (in Section 3), including details on the number of papers in each step of the search process. Third, three researchers were involved in data collection and classification, thus reducing the threat of overlooking relevant papers or misinterpreting the data.

5.3.2 Construct Validity
Construct validity refers to the correct interpretation and measurement of the theoretical concepts [33]. In our mapping study, ‘architectural decision’ is the key theoretical concept. However, we spent extra efforts for distinguishing architectural decisions from other theoretical concepts that might be confounded with, such as ‘architectural knowledge’ and ‘design decisions’. For example, we noticed that in some studies the line between these theoretical concepts is blurred, therefore we had to discuss such studies in detail among all researchers to make sure that we refer correctly to architectural decisions. Also, architectural decisions might be made by professionals who do not have the official role of software architect, such as software developers. Thus, we had to discuss if the decisions in the papers were architectural or design decisions, because the distinction between them is sometimes fuzzy.

To achieve construct validity, the final set of papers had to be complete. To achieve this, we created the search string systematically (see sub-section 3.1), making sure to include not only the most relevant keywords (i.e. ‘architectural decision’) in the automated search, but also variations of these keywords (e.g. ‘architecture choice’). Moreover, we searched for items similar theoretical concepts such as ‘design rationale’, ‘design knowledge’, or ‘design decisions’. On the one hand, by extending the search to similar theoretical concepts, we increased confidence in retrieving all relevant papers. On the other hand, the
extended search required massive efforts from us to filter manually the initially more than 28,000 references (see Figure 2). Also, to further ensure the completeness of included papers, we checked several randomly selected papers for references to other relevant papers, and we found that all relevant papers were included. Furthermore, we verified the results of the automated search against a quasi-gold standard.

5.3.3 Internal Validity
Internal validity refers to how well the data supports the study results, and a typical mistake is to misuse statistical analysis [33]. In this study, we used basic statistics for analyzing the data, so internal validity threats are minimal.

5.3.4 External Validity
External validity refers to the strength of generalizability claims of the study results [33]. The results of this mapping study refer to the state-of-research on architectural decisions in the software architect field, from the perspective of researchers. Since we identified a comprehensive list of 144 papers on architectural decisions using a broad search, we consider that our study results have strong generalizability claims, within the selected timeframe (i.e. 2002-2011). Given the detailed presentation of the protocol, other researchers can extend this study beyond our selected timeframe.

6 Conclusions
We conducted a systematic mapping study with the goal of providing a systematic overview of literature on architectural decisions. As part of the overview, we identified gaps in existing research, and promising future research directions. We obtained the overview by querying six search engines that returned 28,895 papers, covering a decade of research (i.e. from the 1st of January 2002 until the 1st of January 2012). After multiple filtering steps, we obtained a set of 144 relevant papers.

For each of these papers, we extracted data to answer six research questions. The six research questions belonged to two groups. The first group covered software architecture-specific topics: documenting decisions, (non-)functional requirements, and domain-specific architectural decisions. The second group covered topics inspired from generic decision-making literature: normative and descriptive papers, uncertainty in architectural decisions, and group architectural decisions. In addition, we presented an overview of empirical evaluation approaches, publication venues, and average citation count of papers.

Our analysis of existing research on architectural decisions found the following. Much work exists on documenting architectural decisions. However, other topics are not studied in detail: domain-specific architectural decisions, and decisions for achieving specific quality attributes. Also, we found that little descriptive work exists. Little work exists on addressing uncertainty in architectural decisions, and group decision-making. Finally, we found that the number of papers validated with surveys and case studies increased over time, but few papers used experiments.

7 Appendix

7.1 Selected Papers
Here is the list with the 144 selected papers. The papers that were included in the quasi-gold standard (see sub-section 3.1) are marked with a star (e.g. P12*).
P27* Carignano, M.C., S. Gonnet, and H. Leone. A model to represent architectural design rationale. in 2009 Joint Working IEEE/IFIP Conference on Software Architecture (WICSA) & 3rd European Conference on Software Architecture


Clerc, V., P. Lago, and H. Vliet. The Architect’s Mindset. in 3rd international conference on Quality of Software architectures, components, and applications. 2007: Springer Berlin Heidelberg.


Falessi, D., R. Capilla, and G. Cantone. A value-based approach for documenting design decisions rationale: a replicated experiment. in Proceedings of the 3rd international workshop on Sharing and reusing architectural knowledge. 2008: ACM.


Grunske, L. Identifying "good" architectural design alternatives with multi-objective optimization strategies. in 28th International Conference on Software Engineering. 2006: ACM.

Gu, Q. and P. Lago. SOA process decisions: new challenges in architectural knowledge modeling. in Proceedings of
the 3rd international workshop on Sharing and reusing architectural knowledge. 2008: ACM.


P76 MacDonald, S., et al., Deferring Design Pattern Decisions and Automating Structural Pattern Changes Using a Design-Pattern-Based Programming System. ACM Transactions on Programming Languages and Systems, 2009. 31(3).


Shahin, M., P. Liang, and M.R. Khayyambashi, Improving understandability of architecture design through visualization of architectural design decision. in 2010 ICSE Workshop on Sharing and Reusing Architectural Knowledge. 2010.


Zannier, C. and F. Maurer. A qualitative empirical evaluation of design decisions. in 2005 workshop on Human and social factors of software engineering, 2005: ACM.


Zannier, C. and F. Maurer. Comparing decision making in agile and non-agile software organizations. in 8th international conference on Agile processes in software engineering and extreme programming. 2007: Springer-Verlag.

Zannier, C. and F. Maurer. Social Factors Relevant to Capturing Design Decisions. in Proceedings of the Second


### 7.2 Publication Venues

Table 8. Distribution of the 144 papers over all publication venues.

<table>
<thead>
<tr>
<th>Venue</th>
<th>Type</th>
<th>No</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workshop on SHAring and Reusing Architectural Knowledge</td>
<td>Workshop</td>
<td>17</td>
<td>11.81</td>
</tr>
<tr>
<td>European Conference on Software Architecture</td>
<td>Conference</td>
<td>16</td>
<td>11.11</td>
</tr>
<tr>
<td>Working IEEE/IFIP Conference on Software Architecture</td>
<td>Conference</td>
<td>10</td>
<td>6.94</td>
</tr>
<tr>
<td>Journal of Systems and Software</td>
<td>Journal</td>
<td>7</td>
<td>4.86</td>
</tr>
<tr>
<td>Working IEEE/IFIP Conference on Software Architecture</td>
<td>Conference</td>
<td>7</td>
<td>4.86</td>
</tr>
<tr>
<td>IEEE software</td>
<td>Journal</td>
<td>7</td>
<td>4.86</td>
</tr>
<tr>
<td>International Conference on Software Engineering</td>
<td>Conference</td>
<td>6</td>
<td>4.17</td>
</tr>
<tr>
<td>Quality of Software Architectures</td>
<td>Conference</td>
<td>5</td>
<td>3.47</td>
</tr>
<tr>
<td>Information and Software Technology</td>
<td>Journal</td>
<td>4</td>
<td>2.78</td>
</tr>
<tr>
<td>Workshop on Traceability, Dependencies and Software Architecture</td>
<td>Workshop</td>
<td>3</td>
<td>2.08</td>
</tr>
<tr>
<td>Engineering of Computer Based Systems</td>
<td>Conference</td>
<td>3</td>
<td>2.08</td>
</tr>
<tr>
<td>Hawaii International Conference on System Sciences</td>
<td>Conference</td>
<td>3</td>
<td>2.08</td>
</tr>
<tr>
<td>Decision Support Systems</td>
<td>Journal</td>
<td>3</td>
<td>2.08</td>
</tr>
<tr>
<td>European Workshop on Software Architecture</td>
<td>Workshop</td>
<td>2</td>
<td>1.39</td>
</tr>
<tr>
<td>Euromicro Conference on Software Engineering and Advanced Applications</td>
<td>Conference</td>
<td>2</td>
<td>1.39</td>
</tr>
<tr>
<td>Journal of Computer Science</td>
<td>Journal</td>
<td>2</td>
<td>1.39</td>
</tr>
<tr>
<td>Extreme Programming and Agile Processes in Software Engineering</td>
<td>Conference</td>
<td>2</td>
<td>1.39</td>
</tr>
<tr>
<td>Software Engineering and Knowledge Engineering</td>
<td>Conference</td>
<td>2</td>
<td>1.39</td>
</tr>
<tr>
<td>Canadian Conference on Electrical and Computer Engineering</td>
<td>Conference</td>
<td>2</td>
<td>1.39</td>
</tr>
<tr>
<td>Australian Software Engineering Conference</td>
<td>Conference</td>
<td>2</td>
<td>1.39</td>
</tr>
<tr>
<td>Symposium on Human Factors in Computing Systems</td>
<td>Conference</td>
<td>1</td>
<td>0.69</td>
</tr>
<tr>
<td>Product Line Approaches in Software Engineering (In conjunction with ISCE)</td>
<td>Workshop</td>
<td>1</td>
<td>0.69</td>
</tr>
<tr>
<td>Empirical Software Engineering Journal</td>
<td>Journal</td>
<td>1</td>
<td>0.69</td>
</tr>
<tr>
<td>IEEE/ACM International Conference on Automated Software Engineering</td>
<td>Conference</td>
<td>1</td>
<td>0.69</td>
</tr>
<tr>
<td>Software Engineering for Automotive Systems</td>
<td>Workshop</td>
<td>1</td>
<td>0.69</td>
</tr>
<tr>
<td>IFIP TC 2 Central and East European Conference on Software Engineering Techniques</td>
<td>Conference</td>
<td>1</td>
<td>0.69</td>
</tr>
<tr>
<td>International Workshop on The Economics of Software and Computation</td>
<td>Workshop</td>
<td>1</td>
<td>0.69</td>
</tr>
<tr>
<td>Journal/Magazine/Conference/Workshop</td>
<td>Type</td>
<td>Impact Factor</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------------------------------------</td>
<td>---------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Design Studies Journal</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Journal of Software Maintenance and Evolution: Research and Practice</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Conference of Advanced Communication Technology</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>SIGCHI Conference on Human Factors in Computing Systems</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Conference on Design Science Research in Information Systems and Technology</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Software Quality Journal</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Conference on Intelligent Agent &amp; Multi-Agent Systems</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Workshop on Models and Processes for the Evaluation of off-the-shelf Components</td>
<td>Workshop</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Conference on Quality Software</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Workshops and Posters on the Move to Meaningful Internet Systems</td>
<td>Workshop</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Conference on Service-Oriented Computing</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Journal of Software</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages, and Applications</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Asia Pacific Software Engineering Conference</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Conference on Software Engineering Advances</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>IBM Systems Journal</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Conference on World Wide Web</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Advanced Information Systems Engineering Workshops</td>
<td>Workshop</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Ubiquitous, Autonomic and Trusted Computing Symposia and Workshops</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Software Quality Control Journal</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Formal Methods for Components and Objects</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Software: Practice and Experience</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Computers &amp; Operations Research</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>IEEE Workshop on Software Evolvability</td>
<td>Workshop</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Groningen Workshop on Software Variability</td>
<td>Workshop</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>WSEAS Transactions on Information Science and Applications</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Workshop on Human and Social Factors of Software Engineering</td>
<td>Workshop</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Journal of Advanced Computer Science and Applications</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>ACM Transactions on Programming Languages and Systems</td>
<td>Journal</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Symposium on Architecting Critical Systems</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Symposium on Empirical Software Engineering</td>
<td>Conference</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>International Enterprise Distributed Object Computing Conference Workshops</td>
<td>Workshop</td>
<td>0.69</td>
<td></td>
</tr>
</tbody>
</table>
8 References


