3 RELATED WORK

This Chapter presents the related work to this Ph.D. thesis, which serves as a basis for understanding the research gaps pointed out in this work. Therefore, this chapter answer the RQ1 - *What are the existing NFRs catalogs and how they are defined?* Firstly, Section 3.1 presents the research method used to collect and analyze related studies. Then, Section 3.2 presents the results. In Section 3.3, the results are discussed in terms of key findings, research opportunities and threats to validity. Finally, Section 3.4 summarizes this chapter.

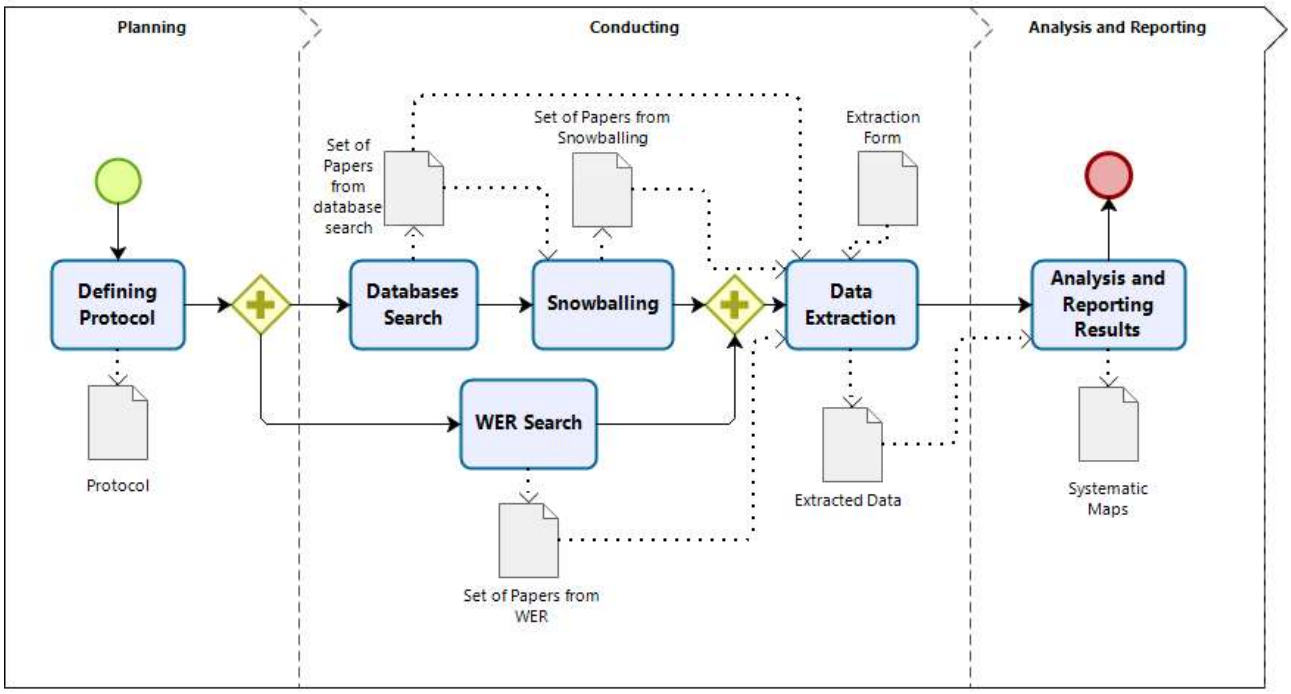
3.1 Research Method

There are several catalogs available in the literature. Many are scattered in several online libraries of scientific studies. Despite their importance, the requirements community lacks a synthesized study about NFRs catalogs in order to understand how existing catalogs are constructed, represented and evaluated. Therefore, to collect an overview of the state of the art and then understand better the related studies, this work performed a systematic mapping (SM) study. The aim is to identify what NFRs catalogs have been proposed in the literature, how they have been represented, defined and evaluated. SM studies are a research method that provides an overview of an area and allows to discover research gaps and trends.

To perform this SM study, a process with three main phases proposed by (KITCHENHAM; CHARTERS, 2007) for systematic studies: *(i)* Planning; *(ii)* Conducting, and *(iii)* Analysis and Reporting (see Figure 18). These phases are usually performed in systematic studies.

In Planning phase, the research protocol should be defined (KITCHENHAM; CHARTERS, 2007). This document combines all information necessary to perform the study. In conducting phase, the researcher should execute what was planned in the protocol. In this work, the conduction was performed through a databases' search followed by a snowballing and, additionally, a search in a specific research event (Workshop on Requirements Engineering - WER) was performed. Then, data extraction took place, where papers from databases, snowballing and WER were used as inputs together with the extraction form. By having all the data, the analysis and report was performed. The next subsections better explain how and why each activity was done.

Figure 18 – Systematic Mapping Process



Source: Author.

3.1.1 Planning

A mapping protocol is usually composed of the following information (KITCHENHAM; CHARTERS, 2007): research goal, questions, databases, selection criteria, screening process, and extraction strategy. All this information is explained as follows.

3.1.1.1 Research Goal and Questions

The goal of this SM is to provide an overview of the literature regarding NFRs catalogs. To achieve our goal, this work investigates the following questions (Systematic Mapping Research Question (SM-RQ)):

- **SM-RQ1 - What NFR catalogs have been proposed in the literature?** This question aims to understand what the existing catalogs are and more specifically what NFRs are covered by these catalogs and how they are classified.
- **SM-RQ2 - How is the information represented in the NFR catalogs?** This question seeks to identify the forms of representation of the extracted catalogs (such as Softgoals Interdependency Graph, Matrices, and others).
- **SM-RQ3 - How are the NFR catalogs defined?** This question aims to check what the approaches are used to define all information presented in these catalogs. In this way,

researchers who want to build their own catalogs can use the answers to this question to guide themselves.

- **SM-RQ4 - How are the NFR catalogs evaluated?** This question aims to know how the proposed catalogs have been evaluated regarding their usefulness and effectiveness.

3.1.1.2 Search Terms, String and Databases

The strategy PICO (Population, Intervention, Comparison, and Outcomes) suggested by (PETERSEN *et al.*, 2015) was considered to identify keywords and formulate the search string. However, only P and I were used. According to (PETERSEN *et al.*, 2015), the other dimensions (C and O) could restrict the search too much and remove important papers. After testing several combinations of terms, the final set is presented in Table 8.

Table 8 – Search Terms

Population	non-functional requirements
	NFR OR "quality characteristic" OR "quality attribute" OR "non-functional property" OR "extra-functional requirement" OR "non-behavioural requirement" OR "quality requirement" OR "quality factor"
Intervention	catalog
	catalogue OR SIG OR "softgoal interdependency graph"

Source: Author.

The population refers to specific roles of software engineering, an application area or an industrial group. In the context of this work, the population is non-functional requirement. There are several synonyms for this term (*e.g.*, quality attribute, quality requirement), which were used in the set of search terms. Intervention refers to a methodology, tool, software procedure or something to be investigated regarding the population. In this work, the intervention is the catalogs. As synonyms, "catalogues" and "softgoal interdependency graphs" were used since this notation has been widely used as NFRs catalogs (CHUNG *et al.*, 2000).

The final search string is presented bellow. To evaluate the quality of it, three control papers that were knew before the execution of the search were selected (MAIRIZA; ZOWGHI, 2011) (NIXON, 2000) (CYSNEIROS, 2007). If this string would find the control papers, then the search string could be considered good.

Search String

(*"quality characteristic"* OR *"non-functional requirement"* OR NFR OR *"quality attribute"* OR *"non-functional property"* OR *"extra-functional requirement"* OR *"non-behavioural requirement"* OR *"quality requirement"* OR *"quality factor"*) **AND** (*catalog* OR *catalogue* OR SIG OR *"softgoal interdependency graph"*)

Once the string has been defined, it was applied in two general databases: Web of Science and Scopus. They were chosen because, according to (SANTOS *et al.*, 2017), they have good coverage and stability, as well as Scopus cover other bases, such as IEEE.

In addition to the database searches, a backward snowballing was performed. This procedure was done because four papers regarding NFRs' Catalogs (MAIA *et al.*, 2009) (FEITOSA *et al.*, 2015) (EGYED; GRUNBACHER, 2004) (GARCIA-MIRELES *et al.*, 2015), known before the execution of this study, would not be found out by the database search, since they do not call what they were presenting as "catalog" or even "Softgoal Interdependency Graph", even though one of them used SIG to catalog requirements (MAIA *et al.*, 2009).

Furthermore, one specific database from the requirements area was also added: Workshop on Requirements Engineering (WER). This workshop was added for three reasons: it represents an important event in the area where researchers usually publish NFRs catalogs, publications from there are not all indexed in the databases this study previously choose, and it provides a search engine of its own, making the search for catalogs easier.

Snowballing was not performed in the selected papers from WER because the stop criteria for snowballing was at least found out the four papers previously known.

3.1.1.3 Selection Criteria

The following Inclusion Criteria (IC) and Exclusion Criteria (EC) to select the most suitable studies were defined: IC1 - the study presents a NFR catalog; EC1 - the study is not written in English; EC2 - the study is not from Computer Science or Engineering related areas; and EC3 - the paper does not present a NFR catalog. It is important to notice that short papers or books were not excluded. In this systematic mapping, all kind of study was accepted if they meet the criteria.

3.1.1.4 Data Extraction

An extraction form was elaborated to be used for each selected paper in this phase. This form is organized into four parts. The first one is about the papers that were accepted during the conduction phase: title, authors, year, source (Scopus, Web of Science, Snowballing or WER), location venue, type of publication (*e.g.*, conference, journal) and main contribution of the paper.

The other parts are related to data to answer the research questions. For SM-RQ1, the following data was defined: catalog ID, type of the catalog, NFRs considered in the catalog, subcharacteristics, development strategies, quantity of negative and positive correlations, level of correlation and domain area by which the catalog was developed. Also, a data item called "Original Authorship" was defined because many studies did not by themselves proposed a catalog, but it used a catalog from another study. Thus, it was possible to obtain the information about the origin of the catalog to get more information and even more catalogs.

Regarding SM-RQ2, the following data were defined to be extracted: general form of representation for each information in the catalog extracted by SM-RQ1 and form of representation for correlations. For answering SM-RQ3, data about how each component in the catalog (subcharacteristics, strategies, correlations) was defined. Finally, regarding SM-RQ4, data about how the catalog was evaluated should be extracted.

All information extracted was managed in a online collaborative spreadsheet editor. At the end of this phase, the data to draw conclusions was analyzed and then it was possible to give an overview about this topic.

3.1.1.5 Data Analysis

In this work, the results should be analyzed according to the type of data extracted. Table 9 summarizes the analysis strategy and type planned to be used. For SM-RQ1 and SM-RQ2, a quantitative analysis should be performed due to the nature of the data collected, which would be numerical in a nominal scale. Measures like central tendency (mode) and dispersion (frequency) could be used to present the results. Regarding SM-RQ3 and SM-RQ4, to better explore the data, it was decided to use a qualitative method and, consequently, answer more appropriately the question. This work used Content Analysis (CA) (BARDLN, 1977).

CA is seen as a research method to classify any kind of communication material

Table 9 – Systematic Mapping Data Analysis

Question	Types of Analysis	Strategy of Analysis
SM-RQ1, SM-RQ2	Quantitative	Descriptive Statistic: Mode and Frequency
SM-RQ3, SM-RQ4	Quantitative and Qualitative	Descriptive Statistic: Mode and Frequency, and Content Analysis Method

Source: Author.

into identified categories of similar meanings (CHO; LEE, 2014). It is suitable for subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns (HSIEH; SHANNON, 2005). Therefore, this method comes as a strategy to better analyze the extracted data in order to create categories of different approaches regarding how the catalogs have been defined and how they have been evaluated.

The same procedures from (CHO; LEE, 2014) was planned to be followed to perform the inductive approach for the qualitative content analysis. The unit of analysis is all extracted data for answering both questions SM-RQ3 and SM-RQ4.

3.1.2 Conducting

As presented in Figure 18, this phase was performed in four steps. The first one is the conduction of a search in two digital databases. The second step is the snowballing based on the papers selected in databases search, and, then, the third step consisted of searching papers at all proceedings of the Workshop on Requirements Engineering (WER). Finally, the last step is the Data Extraction. All these steps are described in the next subsections.

3.1.2.1 Database Search

This step is about selecting studies in Scopus and Web of Science libraries. This selection started by applying the search string into the search machines of those libraries. To select the most suitable set of papers after applying the search string, a screening step was performed using the following five filters: (1) Applying EC1 and EC 2 in the found studies; (2) Excluding duplicate studies; (3) Applying the exclusion criteria in the abstract and title reading; (4) Applying the exclusion criteria in the introduction and conclusion reading, also apply the exclusion criteria by checking if there is an image of a catalog in the study (*e.g.*, graph, table); and (5) Applying the exclusion criteria in the entire paper reading.

Filter 1 was performed in the own search machines (Web of Science and Scopus)

whereas the other filters (2, 3, 4 and 5) were performed with support of Parsif.al tool[1]. Filters 3, 4 and 5 were executed by peers - an undergraduate student and the author of this thesis. First, the student performed the filter alone. Then, the author reviewed the papers excluded by the student. For example, if the student had excluded a paper by reading the abstract and title, the author would review them to check whether she agrees with the exclusion. If not, the paper would be included in the study again.

3.1.2.2 Snowballing

This work also performed the backward snowballing procedures defined by (WOHLIN *et al.*, 2013) to complement the set of papers found by the database search. The backward procedure consists of checking the references list of a set of papers. In the case of this work, the set of papers were the ones selected in the databases searches.

The procedure from (WOHLIN *et al.*, 2013) was adapted and used to conduct this search, consisting of four filters. These filters are not totally similar to the filters used earlier since the selection through backward snowballing is slightly different. First, all references of each accepted paper from databases searches were manually reviewed. Then, for each reference the following filters were applied: (1) Applying all exclusion criteria in the reference. In this case, it was necessary to be more specific since only titles, authors, venue and year were being read. Thus, papers who presented any keyword from the search string were accepted; (2) Applying all exclusion criteria in the abstract reading; (3) Applying the exclusion criteria in the most relevant part of the papers (introduction, conclusion, and images of catalogs); and (4) Applying the exclusion criteria in the entire paper reading.

Then, the Data Extraction step with the same extraction strategy was followed to collect data from the papers selected by the snowballing procedure.

3.1.2.3 WER Search

The Workshop on Requirements Engineering (WER) exists since 1998 and has been an advance for the Ibero-American community of researchers. This workshop provides a Google search engine that explores papers in all WER editions. Therefore, the same search string was applied on it and the same filters were used in the databases searches. Also, just like the other papers, studies selected from WER search also used the same extraction strategy.

[1] https://parsif.al/ - Free online collaborative tool for systematic studies

3.1.2.4 Data Extraction

After performed all searches, the data extraction took place. In this way, first, the extraction of data was done in the papers found in the electronic databases. Then the extraction was done in the papers of snowballing and, finally, the extraction was carried out in the papers from the workshop.

Regarding to the papers obtained by the databases, the same student who performed the filters during the filtering phase also performed the data extraction. Each extraction was reviewed by the author of this thesis. The extractions of the papers obtained by the snowballing and the workshop were made by the author of this thesis and reviewed by the student.

During extraction, the extraction form was updated to include a data not considered during planning: type of correlation. Through one of the obtained papers, it was possible to note that the correlations can appear between requirements (*e.g.*, Usability and Performance), which is called INTER-NFR, or within a same requirement (*e.g.*, Performance subcharacteristics conflict with one another), called INTRA-NFR. Thus, all papers that had already been extracted were revised again to extract this specific data.

3.1.3 Analysis and Reporting Results

As presented in Figure 18, this phase was about analysis and reporting results. SM-RQ1 and SM-RQ2 were indeed analyzed with the measures Mode and Frequency, generating bar and line graphs, as well as a word cloud and a frequency table. SM-RQ3 and SM-RQ4 were also analyzed using quantitative measures, but it was also analyzed through a qualitative method, Context Analysis.

According to (CHO; LEE, 2014), there are two approaches of conducting qualitative content analysis: inductive approach and deductive approach. The inductive approach is suitable when prior knowledge regarding the topic under investigation is limited or fragmented. Therefore, codes, categories, or themes are directly drawn from the data. The deductive approach starts with preconceived codes or categories derived from prior relevant theory, research, or literature. In this work, an inductive approach was used to analyze data from both SM-RQ3 and SM-RQ4 (See Figure 19).

The main procedure for qualitative content analysis is the coding. According to (CHARMAZ, 2006), coding means that labels are attached to segments of data that depict what

Figure 19 – Procedures used in Content Analysis - Inductive Approach

Source: Adapted from (CHO; LEE, 2014).

each segment is about. In open coding, each extracted data form was read. Then, the preliminary codes that emerged from the text were formulated. After that, all texts were codified with those codes. Every time a new information was found out but it did not fit into an existing code, a new one was created. The next steps were to review these codes in order to group the similar ones into categories, always with the care that the categories are mutually exclusive (a rule in CA method). The MAXQDA12 tool (GODAU *et al.*, 2004) was used to support all these procedures.

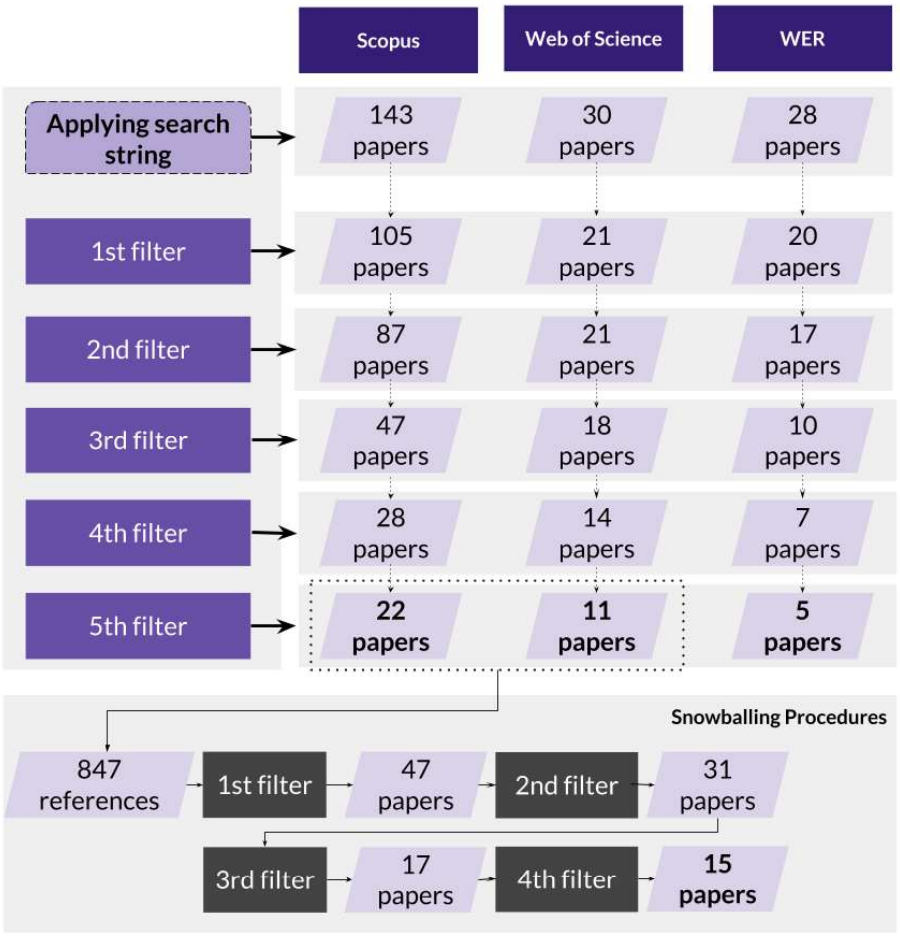
In the case of data from SM-RQ3, although some existing techniques to define knowledge were already knew, the knowledge about this topic was still minimal and very fragmented. In the case of data from SM-RQ4, which were data about evaluation, although some approaches such as experiment, case study were knew, it was clear that the extracted data could bring more information on the evaluation of this specific artifact: NFRs catalogs.

3.2 Results

Figure 20 presents the results from all searches during the conducting phase. In summary, the databases search resulted in 173 papers (143 from Scopus and 30 from Web of Science). After filtering them according to the selection criteria, 33 papers were accepted. This set of papers was defined as the start set for the backward snowballing, which after screening them, the result was a set of 15 additional papers.

Besides, the WER search, which initially brought 28 papers after applying the search string in their google engine, resulted in 5 papers. Therefore, this systematic mapping study resulted in 53 papers (33 from databases, 15 from snowballing and 5 from WER). They are listed in Appendix A. Next subsections present the results for each research question.

Figure 20 – Results from Filtering



Source: Author.

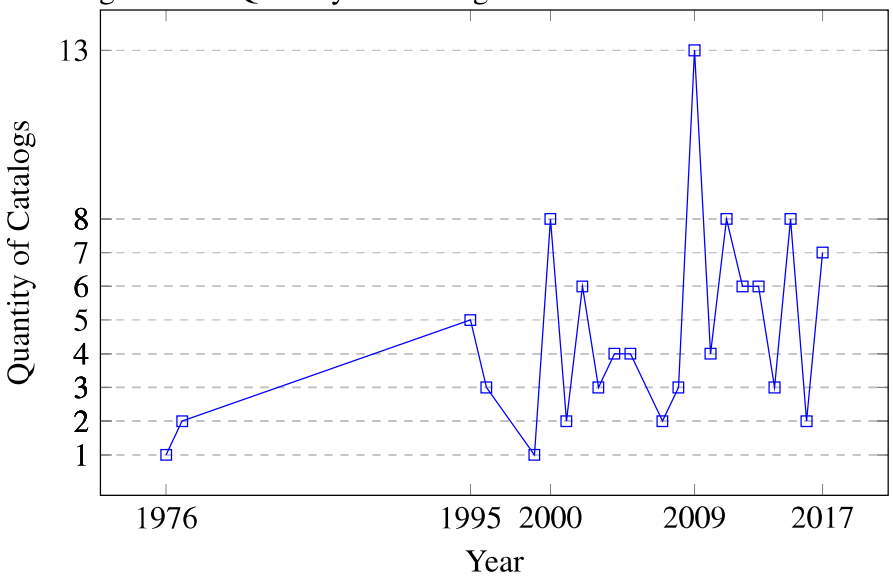
3.2.1 SM-RQ1 - What catalogs have been proposed in the literature?

102 catalogs were found out, ranging from 1976 to 2017 by year of publication, being 2009 the year with most published catalogs (13), as illustrated in Figure 21.

During extraction, it was realized that many catalogs cannot be classified exclusively in one of the three kinds of catalogs: catalog of subcharacteristics, catalog of implementation or design strategies and catalog of correlations. The found catalogs present a mix of stored knowledge. Many of the found catalogs (30) present three types of information: Subcharacteristics, Strategies, and Correlations (See Figure 22). Others present only subcharacteristics (28) or present subcharacteristics and strategies (22). Therefore, this work came up with a different classification from the previous one presented in Background Section for types of catalogs:

- T1. Subcharacteristics - this kind of catalog stores only decomposition of NFRs into more specific NFRs, which this work calls as subcharacteristics;
- T2. Subcharacteristics and Strategies - this catalog stores not only decompositions of

Figure 21 – Quantity of Catalogs vs. Year of Publication



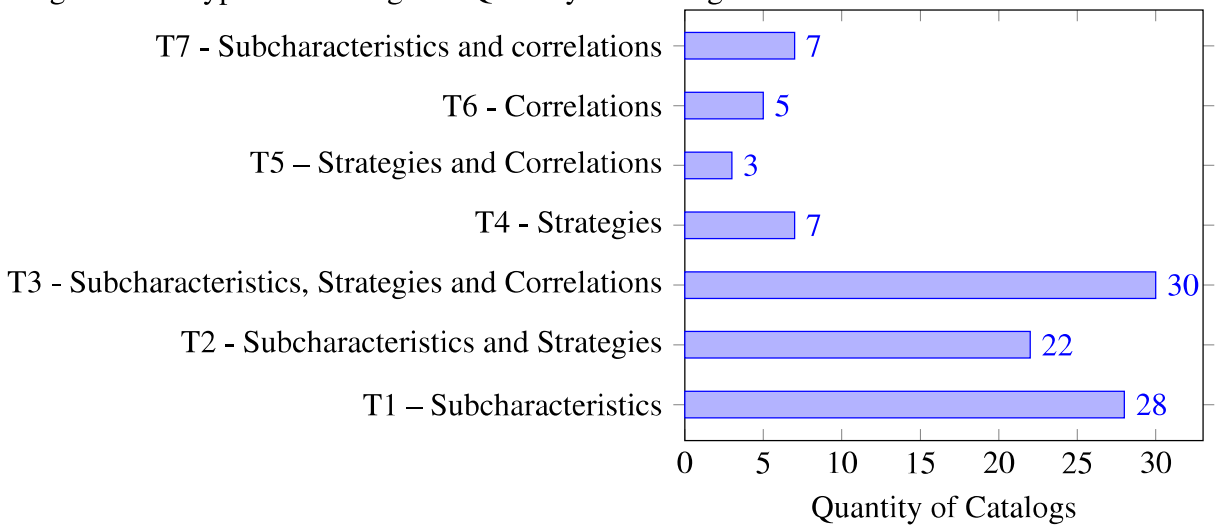
Source: Author.

NFRs, but also development strategies that help achieve subcharacteristics;

- T3. Subcharacteristics, Strategies and Correlations - this kind of catalog contains all the three knowledge;
- T4. Strategies - this catalog stores only strategies to operationalize NFRs. Therefore, this kind of catalog does not present subcharacteristics;
- T5. Strategies and Correlations - this kind of catalogs store correlations between strategies of NFR;
- T6. Correlations - these catalogs store correlations between NFRs, not considering subcharacteristics. That means the correlations are defined in the level of characteristics;
- T7. Subcharacteristics and correlations - this catalog stores correlations between subcharacteristics of NFR. Therefore, the catalog classified in this type should present subcharacteristics and correlations;

The NFRs supported by these catalogs were also analyzed. In total, 348 NFRs were extracted. However, this number includes duplicated NFRs, which means they had the same name or were synonyms. Therefore, they were analyzed in order to establish a set with different NFRs. The analysis was performed in two steps. The first was the removal of NFRs that had exactly the same name. The second step consisted of grouping NFRs with similar names, such as traceability and traceable, functional suitability and functionality, among others. At the end of this analysis, it was concluded that these catalogs support 86 different NFRs. Figure 23 presents all NFRs and Table 10 presents the frequency of the five most cited NFRs. Performance was

Figure 22 – Types of Catalogs vs. Quantity of Catalogs



Source: Author.

the most cited (34), followed by Security (29), Usability (23) and Reliability (14). Regarding subcharacteristics and strategies, 1269 and 1113 were found out, respectively.

Figure 23 – Word Cloud of NFRs identified in the SM Study

Source: Author.

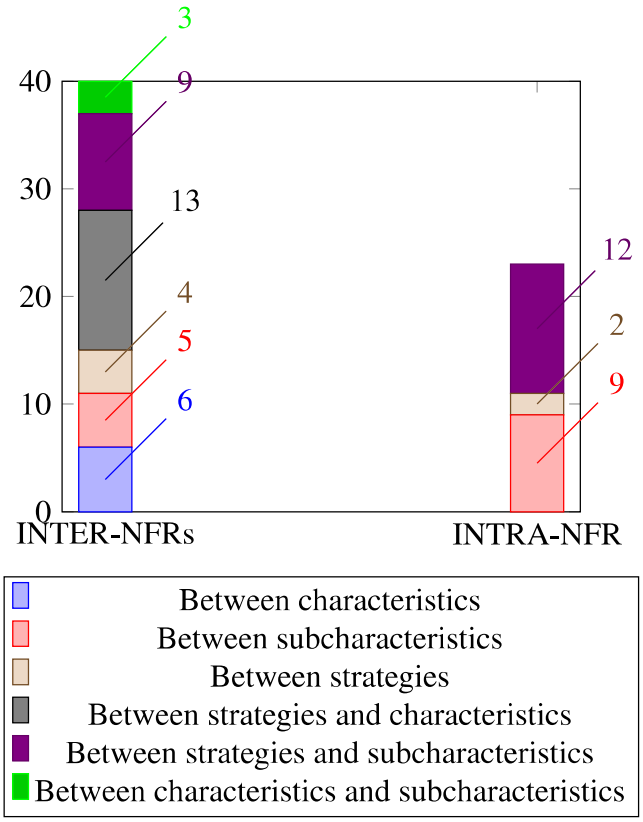
Table 10 – Frequency of NFRs in the SM Study

NFR	Frequency	Relative Frequency
Performance	34	9,7%
Security	30	8,6%
Usability	23	6,6%
Reliability	14	4%
Reusability	14	4%

Source: Author.

With respect to the correlation catalogs, they present 473 positive and 395 negative correlations in total. They are mostly classified as INTER-NFRs (24 out of 45 catalogs that present correlations). Furthermore, there are also catalogs presenting INTRA-NFR correlations (11) and catalogs presenting both (10). Figure 24 shows the type of correlation regarding the level of correlation. Catalogs that are INTER-NFRs type have most correlations in the level "between strategies and characteristics", 13 catalogs in total. While for INTRA-NFR catalogs, the level of correlation is "between strategies and subcharacteristics". This information is interesting because it gave the idea that no matter what type of correlation is, the catalogs show them in a more specific level, which is considered good for developers who wants to make a quick decision.

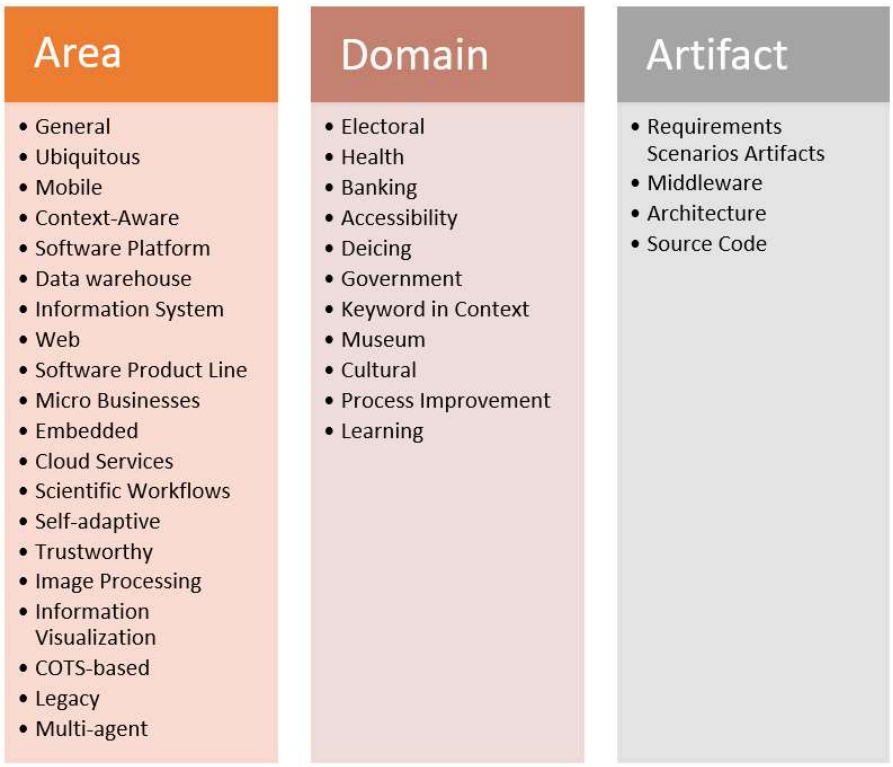
Figure 24 – Type of Correlation vs Level of Correlation



Source: Author.

Furthermore, while analyzing the domain by which the catalog was proposed, it was found out that there are catalogs specific not only for a domain, but also for areas and artifacts. Figure 25 presents them, ordered by highest frequency. Area represents the kind of system the catalog is dealing with. For example, there are catalogs specific for mobile applications, web application or embedded applications. Domain represents the kind of problem the catalog is proposed to. For example, health, learning, banking are domains. Artifact represents the part of

Figure 25 – Domain, Area and Artifacts of Catalogs identified by the SM Study



Source: Author.

the system for which the catalog is being proposed. For example, there are catalogs specific for source code or specific for requirements scenarios.

Looking to the NFR catalogs proposed for UbiComp, Context-Aware, Self-Adaptive, Mobile areas, eight catalogs that contain one or more characteristics from AMICCaS were found out. Table 11 presents them. However, it was possible to see that there are few correlations and they are very specific for one kind of system: health domain.

Seven catalogs are related to Context-Awareness characteristics, six are related to Mobility and four are related to Invisibility. Most of the catalogs (5) do not present correlations. The ones that have correlations use to present them in a specific level: between strategies and subcharacteristics.

3.2.2 SM-RQ2 - How is information represented in the catalogs?

In the set of identified catalogs, eight types of representing the knowledge about NFRs were identified: SIG, matrices, i* notation, tables, hierarchical structures, SIG adaptations, list, template, and pattern. To better describe these representations, the types of catalogs that are most stored in each of them were analyzed. Figure 26 presents a bubble plot crossing data

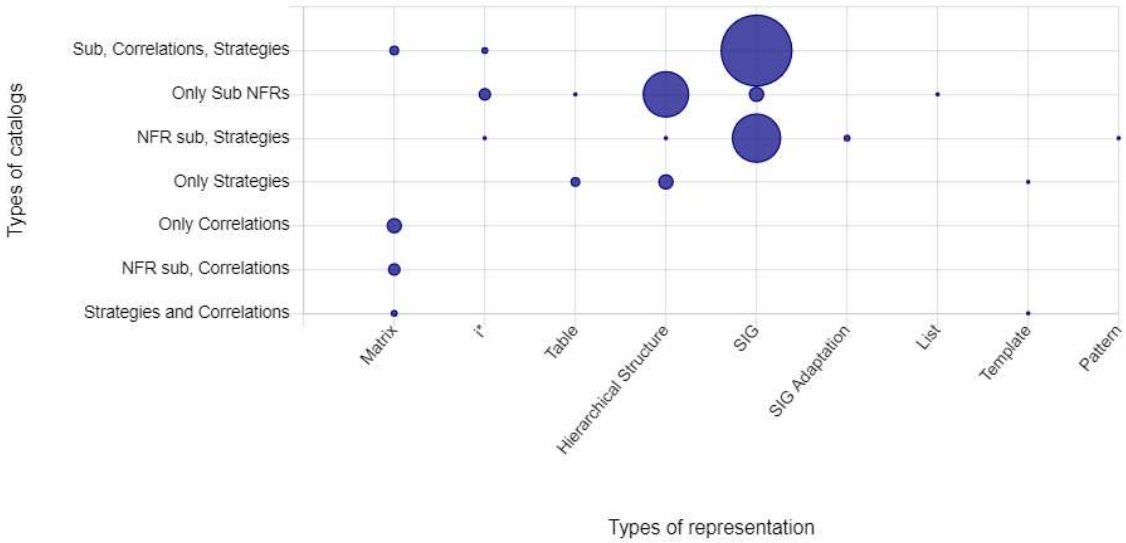
Table 11 – NFRs Catalogs related to AMICCaS identified by the SM Study

CAT ID	*Presence of AMICCaS*	*# Positive Correlation*	*# Negative Correlation*	*Level of Correlation*
18	Context-Awareness	-	-	-
22	Mobility, Context-Awareness	-	-	-
26	Invisibility, Context-Awareness, Mobility	-	-	-
27	Context-Awareness	-	-	-
28	Invisibility, Mobility, Context-Awareness	2	1	Between strategies and subcharacteristics / Between strategies
29	Invisibility, Mobility, Context-Awareness	1	5	Between strategies and subcharacteristics
30	Invisibility, Mobility, Context-Awareness	0	1	Between strategies and subcharacteristics / Between strategies
67	Mobility	-	-	-

Source: Author.

between the type of catalog and the type of representation. The size of a bubble is proportional to the number of catalogs that are in a pair of categories corresponding to the bubble coordinates. Therefore, it is possible to see that SIGs and hierarchical structures are the most used.

Figure 26 – Types of Catalogs vs Type of Representation

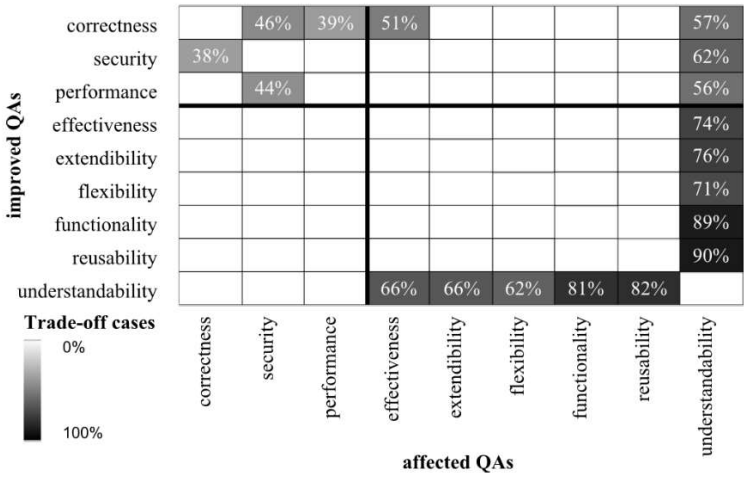


Source: Author.

The first representation is Matrix, which is a rectangular array organized in columns and rows to store information such as numbers and symbols. In case of NFRs catalogs, matrices are usually used to store correlations among pairs of NFRs. Figure 26 shows that regarding

"Matrix", there are only catalogs that contain correlations (T7 - Subcharacteristics and Correlations, T6 - Correlations, T5 - Strategies and Correlations, T3 - Subcharacteristics, Strategies and Correlations). Furthermore, Matrix is most used to store correlations directly between NFRs, *e.g.*, Usability hurts Security. Figure 27 presents an example of a matrix (FEITOSA *et al.*, 2015), where the correlations are correlations between NFRs (INTER-NFRs), each cell represents the effect of improving one NFR (vertical axis) over another (horizontal axis).

Figure 27 – Example of NFR Catalog represented as a Matrix



Source: (FEITOSA *et al.*, 2015)

Framework i* is a goal-oriented modeling language which consists of two main modeling components (YU, 1997). In this notation, it is possible to describe relationships among actors in an organizational context, stakeholder interests and concerns, and how several configurations of systems and environments can address them. Figure 26 shows that it is possible to represent all three knowledge (subcharacteristics, strategies, and correlations).

Hierarchical Structures are typically used to display a hierarchy of NFRs (ISO/IEC 25010, 2011). Figure 26 shows that this representation is more used in catalogs that present only subcharacteristics of NFRs. However, it also has been adopted to represent a hierarchy of strategies.

Softgoal Interdependency Graph (SIG) is a notation proposed by the NFR Framework to analyze and rationale about NFRs (CHUNG *et al.*, 2000). SIGs use NFRs catalogs to support the analysis and help developers. However, this notation has been largely used to not only analyze but also store information that can be reused in other opportunities. This notation was one of the most cited, unsurprisingly given the defined search string in this work. Furthermore,

with this notation, it is possible to represent all kinds of knowledge: subcharacteristics, strategies and correlations.

There are also some Adaptations to SIGs, such as the Softgoal Interdependency Ruleset Graphs (SIRG) (BURGESS *et al.*, 2009), which includes an automated propagation of labels by introducing a new node type: Interdependency Rulesets (IR's). It creates the possibility of analysis without developer input. Another adaptation is from (CYSNEIROS; LEITE, 1999), which is the integration of NFRs into data models, aiming to be more helpful in identifying conflicts. Both adaptations were used to catalog subcharacteristics and strategies.

The other types of representation are List, similar to hierarchical structures but only used in one catalog to store subcharacteristics (FREITAS *et al.*, 2013), Template (BOEHM; IN, 1996) and Pattern (CARVALLO, 2015). The last two are a more formal structure to organize a piece of information.

3.2.3 SM-RQ3 - How the catalogs are defined?

As previously mentioned in Section 3, the Content Analysis (CA) methodology (BARDLN, 1977) was used to analyze the data to answer this question. Therefore, all extracted data of SM-RQ3 was used to find out how catalogs are defined. Thus, the coding activity was performed in the material exploration step. This activity means extracting and relating codes from raw data through inspection. Codes are conceptual names that represent the understanding of the researcher about a text. A set of codes can be grouped to form a category, which is a higher-level concept. Table 12 presents examples of text segments, their codes and category.

Each information that was being read was compared with other information along the set of data for similarities and differences. Every time a similar understanding was found, the existing code could be used. Then, codes with the same characteristics or purpose could be grouped in a category. This was the case for the codes presented in Table 12: Expert, Literature and Author are all sources for extracting knowledge to define a NFR catalog. Therefore, a category called "Source of Knowledge Extraction" was created.

At the end of CA steps, 173 texts segments were obtained and they were represented by 20 codes. These codes were analyzed and organized into 3 identified categories: *(i)* Sources of Knowledge Extraction; *(ii)* Extraction Techniques; and *(iii)* Analysis Techniques. These categories were created because during the analysis it was realized that there is a variation among where to collect knowledge, how to collect knowledge and how to analyze knowledge (*i.e.*,

Table 12 – Examples of Texts, Codes and Category for SM-RQ3

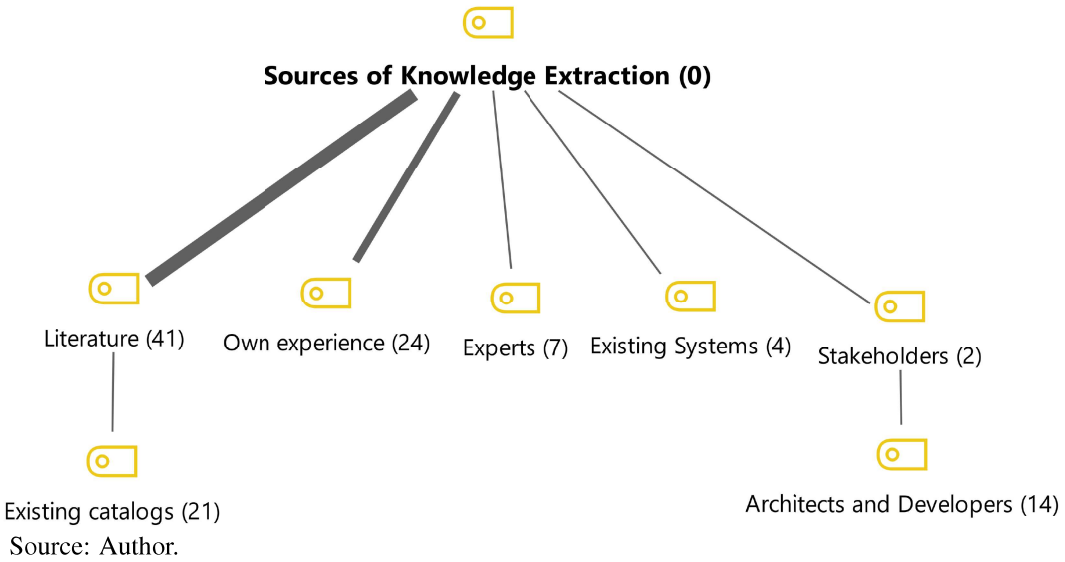
Text Segment	Code	Category
"The positive or negative impact of the alternative on fulfilling the related alternatives is assessed. The assessment is done based on expert knowledge from the design alternatives..." (SADI; YU, 2017)	Expert	Source of Knowledge Extraction
"Based on our bibliographical research, together with our Software Engineering experience" (SILVA et al., 2003)	Literature, Own Experience	
The content of this group of catalogues is extracted from a set of Software Engineering and Business literature discussing problems, concerns, and requirements in opening up software platforms." (SADI; YU, 2017)	Literature	

Source: Author.

subcharacteristics, strategies and correlations).

Figures 28, 29 and 30 present them organized in trees, where the categories are at the highest level and codes are below them. Also, each code presents a number that represents the frequency that this code appeared in data. For example, Figure 28 presents a code "Experts", where the number next to it is 7, meaning this code appeared seven times in data.

Figure 28 – "Sources of Knowledge" Category



Source: Author.

The first category illustrated in Figure 28 refers to the source of extraction, *i.e.*, from where the knowledge of the catalogs was extracted. Six sources were identified: stakeholders; architects and developers; experts; literature; existing catalogs and own experience of the authors.

The source with most citations is Literature, which means that knowledge may have come from books, papers, technical reports, among others. Below literature, there is the source "existing catalogs", these are also part of the literature, but are a particular class because they represent sources of literature with more organized content. Some studies adapt the knowledge of existing catalogs to structure a new catalog (SILVA *et al.*, 2003) (TOTIYA; SENIVONGSE, 2017) (SADI; YU, 2017).

The second source most cited is the own experience of the authors. Many studies are based on the authors' own knowledge and experience (SILVA *et al.*, 2003) (CHUNG *et al.*, 1995). Therefore they do not use an external source, but catalog knowledge based on their experience. Experts also appear as a source of extraction (GARCIA-MIRELES *et al.*, 2015), *i.e.*, experienced people that the authors of the studies consulted to extract knowledge.

The last source identified is the "Stakeholders", which relates to Architects and Developers. They appeared because many catalogs were built for specific systems and thus were made from the experience of the stakeholders who were working on the development of that system.

The relative frequency of each source regarding the type of NFR knowledge that the source was used to collect was analysed. Table 13 presents these frequencies regarding subcharacteristics (S), development strategies (DS) and correlations (C) and the catalogs where these sources were used.

Table 13 – Frequency of Extraction Sources

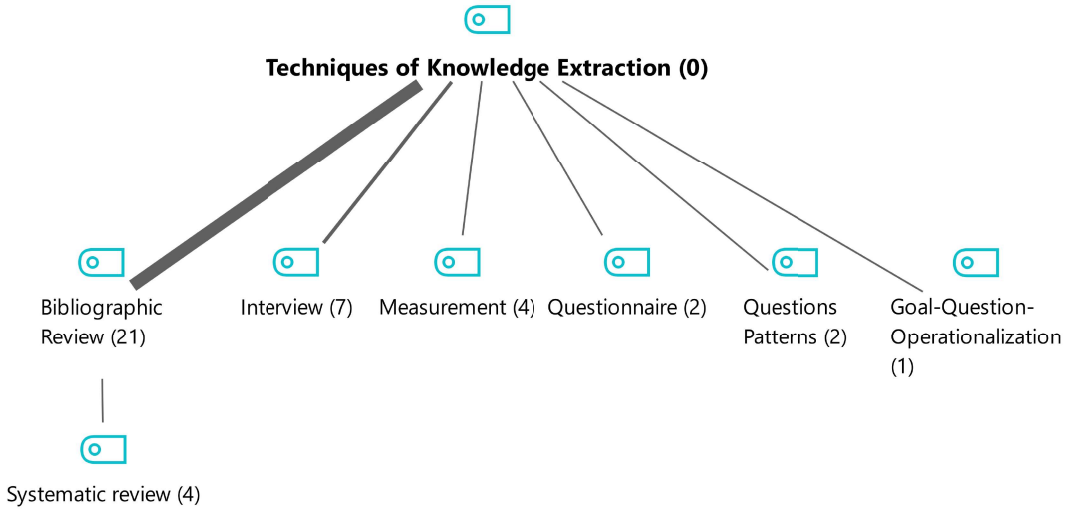
Source	NFR Knowledge **S**	**DS**	**C**	**Catalogs**
Literature	56%	27%	17%	C2, C3, C5, C6, C9, C10, C11, C14, C15, C18, C21, C22, C25, C26, C27, C28, C29, C30, C31, C44, C58, C61, C63, C64, C86, C88, C89, C93, C101
Existing Catalogs	71%	14%	14%	C5, C6, C7, C8, C13, C15, C16, C19, C22, C43, C54, C61, C74, C97, C90, C79
Own Experience	33%	29%	38%	C14, C17, C19, C28, C29, C30, C42, C43, C44, C85, C86, C92, C101
Experts	43%	0%	57%	C4, C25, C74, C75, C76, C88, C89
Stakeholders	100%	0%	0%	C20, C21
Architects and Developers	43%	36%	21%	C12, C40, C41, C87, C88, C89
Existing Systems	0%	0%	100%	C72, C73, C98, C99

Source: Author.

In this way, it is possible to highlight that Literature is a source more used to extract subcharacteristics as well as Existing Catalogs; Own Experience is a common source for correlations; Experts are sources for only subcharacteristics and correlations; Stakeholders were cited only for subcharacteristics; and Architects and Developers are almost equally balanced among the three types of knowledge.

The second category (see Figure 29) refers to the technique of extraction, which means how the knowledge of the catalogs was extracted. Seven techniques were identified. The most used technique is Bibliographic Review, which can be done through a Systematic Review of the literature. It makes sense that this technique is the most cited because the source most used to extract is the literature itself. Interview is another found technique, which can be used together with two sources: stakeholders and experts. Questionnaires were also used to extract knowledge. Another one is Measurement (FEITOSA *et al.*, 2015) (ANDREOPOULOS, 2004), which is about using existing systems to find possible correlations between NFRs. Unlike other techniques, this is specific for extracting correlations.

Figure 29 – "Techniques of Extraction" Category



Source: Author.

Different techniques to extract knowledge are: Goal-Question-Operationalization (SERRANO; LEITE, 2011) and Question Patterns (LEAL *et al.*, 2015). Both help to obtain "strategy" knowledge because they refine subcharacteristics in questions that help to identify developing strategies. These techniques may support interviews or questionnaires for experts.

The relative frequency of each technique regarding the type of knowledge and the catalogs where these techniques were used are presented in Table VII.

It is possible to observe that **Bibliographic Review** and **Systematic Reviews** are

Table 14 – Frequency of Extraction Techniques

Source	NFR Knowledge S	DS	C	Catalogs
Bibliographic Review	62%	24%	14%	C15, C18, C22, C25, C28, C29, C30, C31, C44, C58, C63, C64, C93, C101
Systematic Review	50%	25%	25%	C14, C78
Interview	43%	14%	43%	C20, C21, C74, C88, C89, C91
Measurement	0%	0%	100%	C72, C73, C98, C99
Questionnaire	50%	0%	50%	C15
Question Patterns	0%	100%	0%	C13, C18
Goal-Question-Operationalization	0%	100%	0%	C79

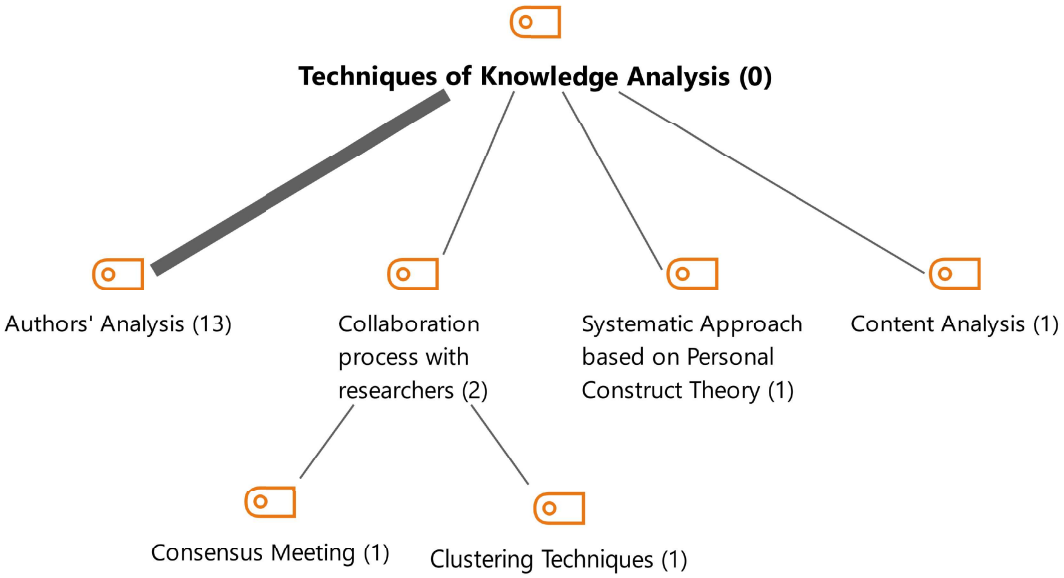
Source: Author.

techniques used to extract all kinds of knowledge: subcharacteristics, strategies and correlations; **Interviews** are used to extract all types of knowledge, being most cited for subcharacteristics and correlations; **Questionnaires** are cited for subcharacteristics and correlations, but it could be used for strategies as well; **Goal - Question - Operationalization** and **Question Patterns** are specific techniques for extracting strategies; and **Measurement** is specific for extracting correlations.

The third category (Techniques of Knowledge Analysis) refers to approaches to analyze data to arrive at an accurate result of knowledge (see Figure 30). In general, the studies present analyzes by the **authors** themselves, not specifying a technique. However, there are also approaches that involve authors in a **collaboration process with experts' researchers**, where they discuss data through **consensus meetings** and **grouping techniques** (LEITE; CAPPELLI, 2010). Also there is one catalog in which correlations were defined through **Content Analysis** (MAIRIZA; ZOWGHI, 2011). In another catalog, the correlations were defined through on an approach based on **Personal Construct Theory** (CAPPELLI *et al.*, 2010).

The relative frequency of each analysis technique regarding the type of knowledge is presented in Table 15 as well as the catalogs where these techniques were used. From this data, it is possible to conclude that many catalogs do not present a technique for the analysis of the extracted data, which is done by the authors themselves; **Consensus Meeting** and **Clustering Techniques** appeared as alternatives to analyze data for defining subcharacteristics; and **Content Analysis** and a proposed technique based on **Personal Construct Theory** are specifically used to identify correlations.

Figure 30 – "Techniques of Analysis" Category



Source: Author.

Table 15 – Frequency of Analysis Techniques

Source	NFR Knowledge S	DS	C	Catalogs
Author's Analysis	38%	23%	38%	C15, C17, C28, C29, C30, C42, C43, C44, C63, C64
Collaboration Process with researchers	50%	50%	0%	C15, C79
Consensus Meeting	100%	0%	0%	C15
Clustering Techniques	100%	0%	0%	C15
Content Analysis	0%	0%	100%	C31
Technique based on Personal Construct Theory	0%	0%	100%	C6

Source: Author.

3.2.4 SM-RQ4 - How the catalogs are evaluated?

The data extracted to answer this question provided a great variety of information to be analyzed, different ideas and concepts related to evaluation appeared in the texts. Thus, the same methodology used to analyze the SM-RQ3 data was used to analyze the SM-RQ4 data. In summary, 88 segments were codified with 18 codes. These codes were analyzed and organized into three categories (see Figures 31, 32 and 33). Table 16 presents examples of text segments, their codes and categories related to this question.

During the analysis, it was observed that some catalogs were not directly evaluated, but rather used to support a proposal. This situation occurred when the focus of the paper was on

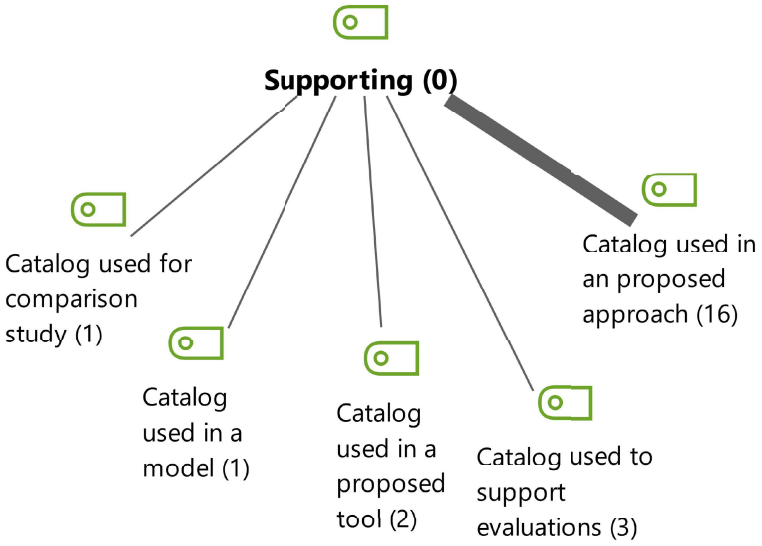
Table 16 – Examples of Texts, Codes and Category for Quest4

Text Segment	Code	Category
"we have carried out a survey with several stakeholders..." (SUBRAMANIAN *et al.*, 2014)	Survey	Evaluation Approaches
"...continued refining the SIG, using questionnaire responses submitted to 16 international modeling experts" (LEITE; CAPPELLI, 2010)	Questionnaire	
"In order to exemplify the use of Provenance SIG, this section describes an usage scenario where the scientific software SWf is modelled" (LEAL *et al.*, 2015)	Design	Evaluation Purpose
"We use the proposed approach to revisit the high-level architectural design of data pro vision service in two real-world open software platform" (SADI; YU, 2017)	Redesign	
"this model can calculate all NFRs contribution values by which developers could make tradeoff decisions among NFRs competing alternatives" (ZHU *et al.*, 2012)	Catalog used in a model	Supporting

Source: Author.

some solution proposal that uses catalogs instead of the catalog itself. In this way, a category "Supporting" was created to represent this concept (See Figure 31).

Figure 31 – "Supporting" Category



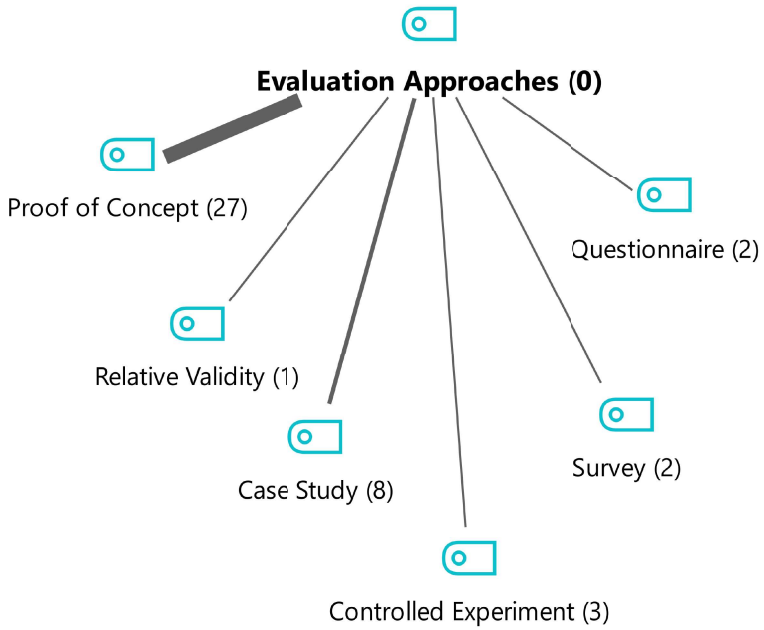
Source: Author.

Five codes were identified in this category: *(i)* Catalog used to support evaluations –

catalogs used as a base to create verification checklists or software measures (ISO/IEC 25010, 2011) (SOAD; BARBOSA, 2016); *(ii)* Catalog used in a model – catalog to support a model that can calculate all NFRs contribution values by which developers could make decisions (ZHU *et al.*, 2012); *(iii)* Catalog used for comparison study – catalog used to compare methodologies (SILVA *et al.*, 2003); *(iv)* Catalog used in a tool – catalogs used inside tools, for example, to help modeling properties of NFRs (UCHÔA *et al.*, 2017); and *(v)* Catalog used in proposed approaches – an example of approach is one from (EGYED; GRUNBACHER, 2004) which aims to identify more precise conflicts through requirements traceability. This last category was the one most cited, 16 texts segments were codified in it.

Besides, a category called "Evaluation Approaches" was created to represent all approaches found out in the studies (see Figure 32). Although, at the beginning, some of the well known approaches (such as case studies and experiments) were expected, this information was scattered in the papers and not explicitly defined, so it was essential to perform the content analysis. Figure 32 shows that Proof of Concept (PoC) is the approach that most appears.

Figure 32 – "Evaluation Approaches" Category

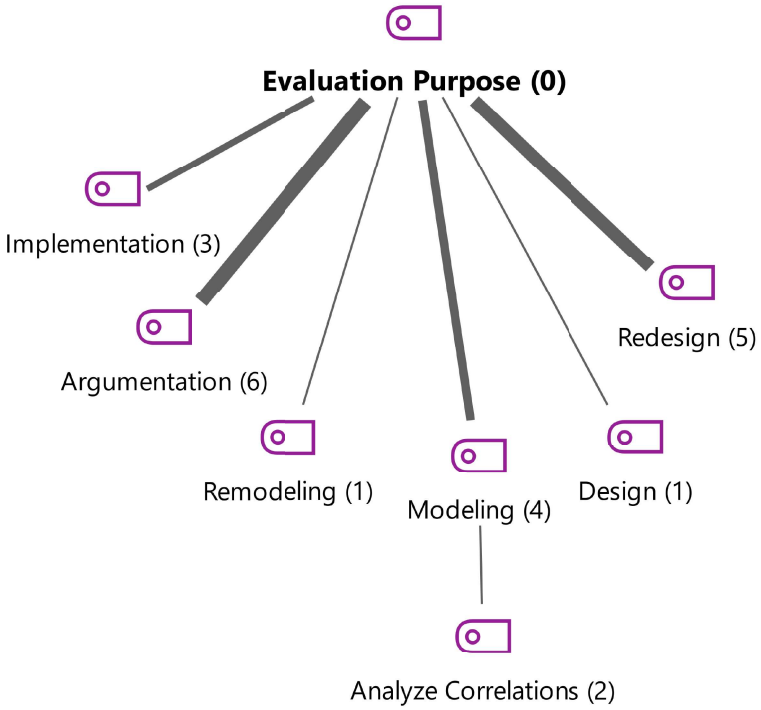


Source: Author.

The third category created is called "Evaluation Purpose". This category was created because, besides the approach, there are still different purposes. Also, many papers did not clarify exactly which approach (Case Study, Experiment or other) they used but have explained about the purpose of the evaluation. As can be seen in Figure 33, some evaluations had the following purpose: *(i)* designing a specific system considering a NFR; *(ii)* redesigning an existing system

to show improvement with the use of the catalog; *(iii)* building a new model, often reusing knowledge from the existing catalog; *(iv)* remodeling an existing catalog; *(v)* arguing about the effectiveness of the catalog; and *(vi)* Support a system's implementation.

Figure 33 – "Evaluation Purposes" Category



Source: Author.

3.3 Discussion

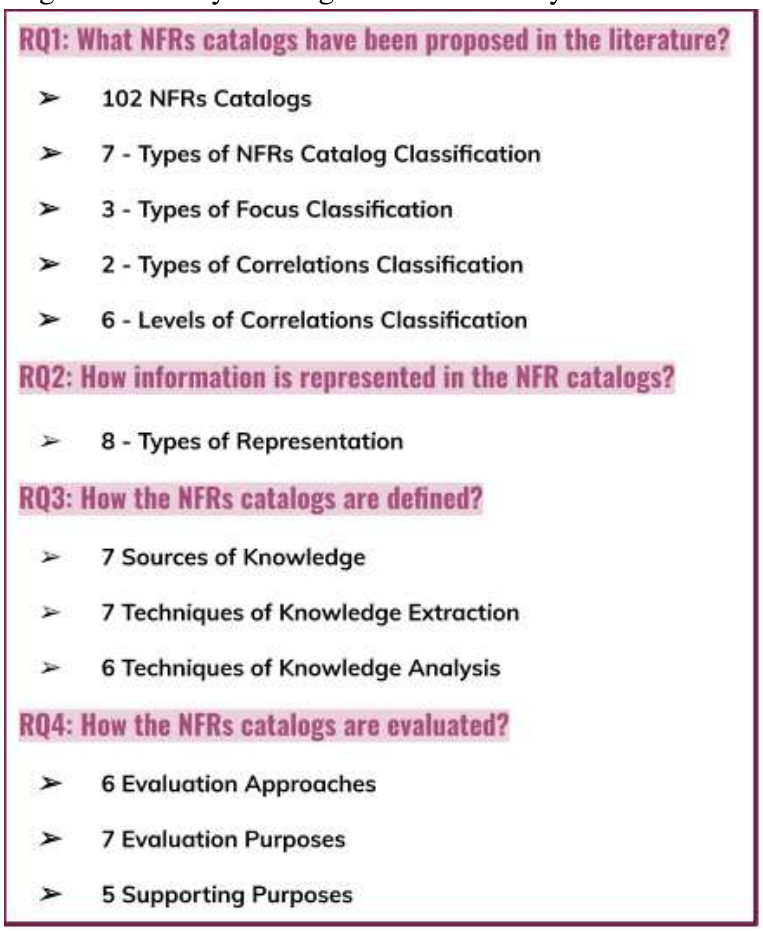
Through this mapping study, it was possible to obtain interesting findings about NFRs catalogs, which was presented through answers for each research question (SM-RQs). A synthesis of these findings is presented and discussed in Subsection 3.3.1. Then, research opportunities are presented in Subsection 3.3.2 and, finally, the threats to validity in Subsection 3.3.3.

3.3.1 Synthesis of the Results

This work consolidates the key findings into a single comprehensive view presented in Figure 34 and each one of these findings are discussed in this section.

The primary focus with this mapping study was first to collect as many catalogs as possible to understand them more deeply. Through SM-RQ1, 102 catalogs were obtained and

Figure 34 – Key Findings of the SM Study



Source: Author.

then it was possible to better understand how they are characterized. One of the main findings is that catalogs cannot be classified as mutually exclusive in three types of catalogs, as previously proposed by (CHUNG *et al.*, 2000). In fact, a catalog can be in more than one classification. In this way, the initial classification of (CHUNG *et al.*, 2000) was extended to include seven types: T1 – Subcharacteristics, T2 - Subcharacteristics and strategies, T3 - Subcharacteristics, Strategies and Correlations, T4 – Strategies, T5 - Strategies and Correlations, T6 – Correlations and T7 - Subcharacteristics and correlations.

Another interesting finding is that initially it was expected that the catalogs could be for a specific area, but in fact, they can be proposed to particular areas, domains or artifacts. These different views are named as the "focus" of the catalog. Catalogs specific to artifacts of the system (*e.g.*, middleware) were found out, but also to a domain (*e.g.*, health) and to a area (*e.g.*, mobile). It was also possible to find combinations of these foci, such as catalogs for mobile applications focused on health. Or even catalogs for ubiquitous applications with a focus on middleware. The more specific the more the catalog can help developers in decision making.

As seen previously, many catalogs have correlations. This kind of knowledge is more complex than subcharacteristics and strategies and deserves a separate analysis. This SM study found out that correlations can occur not only between NFRs but also within the same NFR because their subcharacteristics or strategies may conflict with one another. In this way, a classification of correlation types was found out: INTER-NFRs and INTRA-NFR. This is interesting because it shows that even a single NFR cannot be wholly achieved in a system, thus demonstrating the high complexity that a NFR can present. Besides, this SM study found out that there are six levels of correlation. This level varies from the most generic, which are correlations directly between NFRs, to more specific levels, which are correlations between strategies and strategies and characteristics. Correlations with strategies are also more useful to developers because they can help more accurate decision making. It is difficult for the developer to understand or decide only with a more general level correlation, *e.g.*, Performance Hurts Security. Even because this relationship may be relative, in one type of system, this may be true, in another, it does not. Also, it depends on the strategy used.

The catalogs found are represented in eight different ways (SM-RQ2). Some representations are specific notations that the developer can use to analyze the satisfaction of NFRs, as is the case of SIGs, their adaptations and i *. Other representations are more informal, such as hierarchical structures, matrices, tables, and lists. Some are not notations, but they better organize knowledge, such as patterns and templates. SIG is the representation that was knew before starting this work, which is why it was part of the search string and was obviously the representation with more catalogs. Despite this, other representations have been found, and the interesting fact is to see that some stand out for a specific type of knowledge. For example, hierarchical structures are widely used to store a hierarchy of NFRs and their subcharacteristics. Matrices are widely used to represent correlations. Understanding these representations is important because a researcher or practitioner who wants to catalog knowledge about NFRs can use one of these representations.

Additionally, a mapping of approaches to define a catalog was investigated(SM-RQ3). Although there are catalogs defined by the authors themselves based on their experience, this study realized that the definition of a catalog can be done in two steps. First, it is necessary to collect the information, and then it is necessary to analyze this information to arrive at a more organized knowledge. Concerning the collection, there are six external sources by which the catalog creator can search for information: literature, existing catalogs, existing systems,

experts, stakeholders and architects/developers. Also, there are seven techniques to extract the information from these sources: bibliographic review, systematic review, interview, questionnaire, questions patterns, goal-question-operationalization and measurement.

Table 17 presents a mapping among these sources, techniques and the type of knowledge that can be extracted from them. For sources like "literature" and "existing catalogs", the technique is the "bibliographic review", which can be done through a "systematic review". Although this study did not identified it, this literature review could also be done through systematic mapping or snowballing. For sources "experts" and "stakeholders", "questionnaire" and "interview" can be used to extract all kinds of knowledge (subcharacteristics, strategies and correlations). Additionally, the techniques "question patterns" and "goal-question-operationalization" can be used to extract strategies. From the source "existing systems", "measurements" can be performed to find out correlations.

Table 17 – Mapping among Source, Technique and Kind of Knowledge

Source of Knowledge	***Technique of Knowledge Extraction***	***Kind of Knowledge***
Literature, Existing Catalogs	Bibliographic Review, Systematic Review	Subcharacteristics, Strategies, Correlations
Experts, Stakeholders	Questionnaire, Interview	Subcharacteristics, Strategies, Correlations
	Question Patterns	Strategies
	Goal-Question-Operationalization	
Existing Systems	Measurement	Correlations

Source: Author.

In addition to the sources and extraction techniques, few researchers use techniques to better understand the data that has been extracted and thus define the knowledge more reliably. The techniques found were: Collaboration Process with researchers, Consensus Meeting, Clustering Techniques, Content Analysis and a Technique based on Personal Construct Theory. Collaboration Process with researchers together with consensus Meeting and clustering techniques were used to define subcharacteristics. Content analysis was used to analyze correlations in the level "between characteristics", which means correlations between NFRs *e.g.*, Security hurts Usability. The technique based on Personal Construct Theory was used to define correlations in the level "between subcharacteristics", but using strategies to analyze conflicts and cooperations among NFRs.

Thus, different studies use different combinations of approaches to arrive at a proposed catalog. Some papers cite that they used well-known research strategies, others use their own knowledge and experience. However, a systematic and reusable process that organizes a step by step with inputs, outputs, and approaches on how to create a complete NFR catalog, including steps for refinement of NFRs and strategies, especially for innovative NFRs, was not found in this systematic mapping.

The last findings were regarding the evaluation of a catalog. Six evaluation approaches were found out: proof of concept, case study, survey, questionnaire, relative validity and controlled experiment. Seven evaluation purposes: designing a specific system considering a NFR; redesigning an existing system to show improvement with the use of the catalog; building a new model, often reusing knowledge from the existing catalog; remodeling an existing catalog; arguing about the effectiveness of the catalog; and support a system's implementation. Also, many catalogs are not evaluate themselves but used to support another proposal. Five supporting purposes were found out: to support evaluations, to be used in a model, to be used for comparison study, to be used in a tool, to be used in proposed approaches.

3.3.2 Research Opportunities

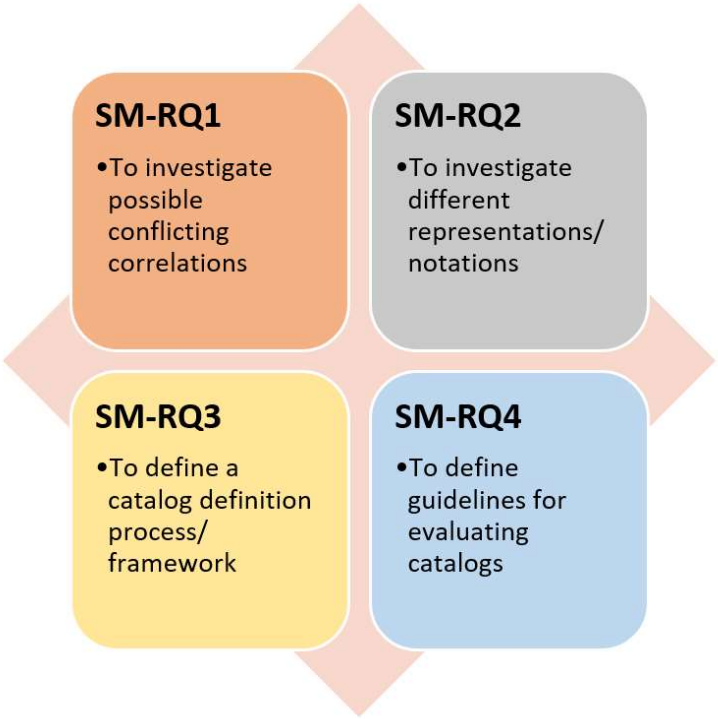
For each research question, it is possible to identify a research opportunity. Figure 35 presents an overview of these four opportunities.

The first opportunity is related to the investigation of possible conflicting (positive or negative) correlations. 473 positive correlations and 395 negative correlations were found. However, they have not been analyzed so that there may be opposing correlations of different catalogs. This research is interesting to show that correlations between NFRs are relative, not always it is true that a pair of NFRs will be in conflict or harmony. Then, a general catalog of these catalogs can be built to support developers.

The second opportunity is related to the catalog's representation. Future research could investigate these representations in a way that could indicate which would be the most appropriate to deal with each knowledge of a catalog of NFRs: subcharacteristics, strategies and correlations. Or even if there is one that prevails in all these senses.

The third opportunity is related to the definition of a complete catalog. Although some papers explain their ways of constructing the proposed catalogs, a generic framework or process that provides a detailed guideline to create NFRs catalog was not found.

Figure 35 – Research Opportunities



Source: Author.

The fourth and last opportunity is related to the evaluation of a proposed catalog. Few catalogs presented detailed evaluation procedures. Thus, there is a need to create a guideline to guide NFRs catalog evaluations.

This thesis deals with the third research opportunity, which is the definition of NFR catalog. The idea is to group a source of knowledge; and techniques of extraction and analysis into a single comprehensive process to help researchers in defining NFRs catalogs.

3.3.3 Threats to Validity

Following the suggestion of (PETERSEN *et al.*, 2015), which is also a systematic study, the following types of validity were considered: descriptive validity, theoretical validity, generalizability, interpretive validity.

Descriptive validity means the extent to which observations are described accurately. This kind of threat usually has more risk in qualitative studies. In the case of this study, there are two questions by which the extracted data was expressed as textual paragraphs. Therefore, to reduce this threat, a qualitative methodology was used to avoid an informal analysis and bias during the analysis. Hence, this threat is considered as being under control.

Theoretical validity is determined by the ability to being able to capture what the

researchers intend to capture. Therefore, study identification, study selection and data extraction are two important threats to be considered. In this study, two databases (Scopus and Web of Science) were initially selected, but to reduce the threat regarding study identification, a snowballing for the selected papers from these databases were performed. Also, a requirement engineering workshop (WER) database was added, because it represents an important event in the requirements area where researchers usually publish their NFRs catalogs, but publications are not all indexed in the databases. However, it is worth to say that only a small number (5) of new studies was obtained from WER, indicating that the overall conclusions of this mapping would not change.

Regarding the study selection, valid papers could have been rejected. To avoid this threat, during the studies selection from Web of Science and Scopus, the review was performed by peers, the author reviewed the selection every time a paper was rejected by the undergraduate student. Therefore, a paper who was rejected could be considered again after the review. Also, whenever an excluded paper returned to the set of selected papers, an explanation was given to the student who was making the selection to achieve a common understanding of the studies.

The suggestions from the literature say that it is also appropriate to have one researcher extracting data and another reviewing the extraction (PETERSEN *et al.*, 2015). Therefore, extractions were also performed by peers in studies from Web of Science and Scopus.

Generalizability is determined by the degree that researchers can generalize results and it can be classified between external and internal. A possible threat is that data interpretation could be different for different researchers. To mitigate this threat, most of the extraction was performed in peers. Thus, this work considers that internal generalizability is not a major threat. Regarding external generalizability, this threat was minimized by performing this work with researchers from three organizations and two countries, increasing the possibilities of generalization. However, this study cannot generalize completely so this risk is accepted.

Interpretive validity is concerned about if the conclusions were based on the data, whether objective or subjective. In this work, most of the extractions were reviewed by the author. Also, the data analysis, both objective and subjective, following a methodology, not being conducted informally. The quantitative analysis followed Wholin's suggestions for presenting the data in graphs. Qualitative analysis was performed through Content Analysis.

Repeatability requires detailed reporting of the research process. To achieve this kind of validity, guidelines from literature were followed to perform several steps of this research.

For performing the SM study in general, instructions proposed by (PETERSEN *et al.*, 2015) were used, and also this work considered other SM processes, such as (SANTOS *et al.*, 2017), (MONTAGUD *et al.*, 2012) (NETO *et al.*, 2011) (CARVALHO *et al.*, 2017). Besides, guidelines for performing snowballing proposed by (JALALI; WOHLIN, 2012), (BADAMPUDI *et al.*, 2015) were followed. Finally, all the methodology from (HSIEH; SHANNON, 2005) to conduct content analysis was followed as well.

3.4 Chapter Summary

This Chapter has presented an exploratory study about NFRs' catalogs through a systematic mapping[2]. The purpose of this study is to investigate the existing catalogs, specially if there are existing catalogs dealing with AMICCaS, and also understand how they are defined and evaluated before starting the construction of a catalog for UbiComp and IoT systems.

Through this study, 102 NFRs catalogs were found out. Most of them present the three kinds of knowledge: subcharacteristics, strategies and correlations. There are catalogs for a general focus and for a specific focus. Regarding the area of this work, Internet of Things and UbiComp, few catalogs present characteristics related to AMICCaS (8), and only three of them present few correlations, which is still very scarce.

Furthermore, several techniques that could potentially help identify correlations in UbiComp and IoT systems were found. However, it was possible to see that there is not a systematic process that groups techniques to help researchers and developers to define correlations by looking to the development strategies, especially for quality characteristics that are new and no taxonomy is available, which is the case of AMICCaS.

This lack of such approach makes the definition of catalogs arduous. Thus, part of this Ph.D. work is first dedicated to define a process capable to define NFRs catalog composed of subcharacteristics, development strategies and correlations (answering RQ2).

[2] All materials and results from this SM study are available at https://github.com/great-ufc/SM-NFRsCatalogs