Scalability of iStar: a Systematic Mapping Study

Paulo Lima[1], Jéssyka Vilela[1], Enyo Gonçalves[12], João Pimentel[3], Ana Holanda[14], Jaelson Castro[1], Fernanda Alencar[1], Maria Lencastre[3]

[1]Universidade Federal de Pernambuco (UFPE)
[2]Universidade Federal do Ceará (UFC)
[3]Universidade de Pernambuco (UPE)
[4]Instituto Federal do Acre (IFAC)

{plc2, jffv, ejtg, jhcp, acah, jbc}@cin.ufpe.br
fernanda.ralencar@ufpe.br
mlpm@ecomp.poli.br

Abstract. Since its first proposal in the nineties, the i* framework has been used to requirements specification in many domains, such as healthcare, telecommunication, and air traffic control. After the modeling of different examples and case studies, it has been observed that i* models become dramatically more difficult to understand and analyze as they grow larger. This issue has led us to investigate scalability in the context of the i* framework, by means of a systematic mapping study. A total of 119 papers were analyzed, in order to understand how scalability is perceived by the i* research community, which proposals have considered this topic, and what open issues still need to be addressed. We found that scalability issues are indeed perceived as relevant and that further work is still required, even though many potential solutions have already been proposed. This study can be a starting point for researchers aiming to further advance the treatment of scalability in social goal models.

Keywords: Systematic Mapping study, Scalability, iStar, i*, Goal Models.

1 Introduction

The i* framework is a modeling language used in Requirements Engineering (RE) to develop models that represent requirements actors and systems [1]. Requirements are represented by dependencies of interest, through goals, softgoals, resources and tasks. The i* language has been used in several situations [1], such as telecommunications, air traffic control, agriculture, e-government, healthcare and business process.

Unfortunately, i* is not suitable for modelling complex cases or involving many parts [3]. The limitation in scaling i* is identified as one of their biggest barriers for industrial adoption [1] [3]. If i* could make possible a good modelling of large and complex cases, there would be greater adoption of i*. Therefore, the i* framework requires solutions and means to address its scalability [2] [3].

This paper aims to map the studies that addressed i* scalability issue. In order to meet our goal, we conducted a systematic mapping study to identify the primary studies on the scalability of i*, following a predefined review protocol. We identified 119 papers about i* scalability and analyzed the distribution of these studies, definitions, mentions for artifacts to address it; how the researchers community perceive the i* ability to scale and, if there are open issues related in this theme.

This goal is important to support researchers in the i* community since i* models can become complex and could be difficult to decompose into smaller pieces [3]. The number of papers that mentioned i* scalability (or near attribute) may serves to future systematic reviews. Moreover, the comparisons between approaches to address i* scalability are necessary.

This paper is structure as follows: Section 2 discusses related works, the study design is presented in Section 3, the results related to each research question are reported in Section 4. Section 5 presents the discussion of the results. Section 6 shows the conclusions and future work of this paper.

2 Related Work

The scalability of i* is an issue addressed by many researchers. However, we identified only two secondary studies that synthesize i* and its scalability: [2] outlined published research on i* in general, commenting also the question of scalability, and [3] present many great challenges and i* problems, including scalability notation.

The first work [2] cites some approaches proposed to improve the scalability. This work reviews the approaches adopted by the i* framework and reports their application in various fields. The authors of this paper explained that models need to reflect software in their social characteristics of complex systems. They argue that i* has a scalability limitation. The second work [2] [3] exposed the scalability as one of the great challenges of i*. It argues that scalability and complexity are two characteristics that i* does not support well and proposed a set of concepts to help improve it.

In order to understand the progress that has been made in integrating goal models with downstream system development, [26] conducted a systematic survey to investigate what approaches exist which map/integrate/transform goal-oriented languages to other software artefacts or languages. They produced a roadmap summarizing 174 publications. They also comment the lack of widespread industrial adoption of goal-oriented models that could be attributed to several factors including the scalability and usability challenges in complex models which are not easily decomposable.

Horkoff and Yu [4] use three available tools implementing seven similar analysis procedures for goal satisfaction to analyze three sample goal models. They performed a comparison to understand the ways in which procedural design choices affect analysis results, and how differences in analysis results could lead to different recommendations over alternatives in the model.

After analyzing these studies, we can notice that scalability is a desirable quality. However, this question is addressed in a broad sense and it becomes difficult to categorize possible evidence since it not possible to replicate those results. Therefore, it is

necessary to know all the works published more systematic way. Accordingly, our work presents the results of systematic mapping study conducted to investigate the scalability of i*. We analyzed the distribution of these studies, definitions, mentions for artifacts to address it; how the researchers community perceive the i* ability to scale and, if there are open issues related in this theme. In the next section, we describe our research protocol.

3 STUDY DESIGN

The purpose of this paper is to map the studies that addressed i* scalability. Being a systematic mapping, it provides a summary of evidence related to the topic, by applying explicit methods and systematic search, critical appraisal and synthesis of selected information [6].

We conducted this systematic mapping study following the guidelines of Kitchenham and Charters [5], while also observing the methodology of other systematic mappings that were recently published in well-known software engineering journals, such as the Journal of Systems and Software, the Empirical Software Engineering journal and the Requirements Engineering journal. From these mappings we obtained insight on creating search queries and on defining inclusion and exclusion criteria.

This research was performed with the following steps: (1) identification of the need for a systematic mapping; (2) formulation of focused research questions; (3) the comprehensive, exhaustive search for primary studies; (4) identification of the data needed to answer the research questions; (5) data extraction; (6) summary and synthesis of study results; (7) interpretation of the results to determine their applicability; and (8) report writing.

The study design described in this section was validated by experienced researchers and adjusted accordingly prior to its execution.

3.1 RESEARCH QUESTIONS

This paper aims to answer the following research questions:

- RQ1: What studies mention the issue of i* scalability?
- RQ2: Which are the scalability definitions in the context of i*?
- RQ3: Which are the types of contributions that have been published to support the i* scalability?
- RQ4: What is the perceived judgment on the scalability of i*?
- RQ5: Are there open issues?

The first research question (RQ1) focuses on identifying the works that address the scalability of the i*. The second research question (RQ2) relates to the conceptualization of scalability. The third research question (RQ3) quantifies which mechanisms have been proposed in order to handle scalability, such as metamodels, formalization, modelling processes, visual constructors, and algorithms. The fourth research ques-

tion (RQ4) summarizes published findings on whether the scalability of i* is considered to be satisfactory or not. The fifth research question (RQ5) summarizes evidences about potential open issues mentioned in the selected studies. Therefore, this paper aims to synthesize information about the i* language regarding its scalability.

3.2 EXCLUSION AND INCLUSION CRITERIA

The selected studies were primary publications that present any discussion or study about the scalability of i*. The exclusion and inclusion criteria adopted are presented in Table 1 and Table 2 respectively.

Table 1. Exclusion criteria.

ID	Definition
EC01	Studies not captured by the keywords in search engines.
EC02	Studies published before 1990.
EC03	Unreadable files (files corrupted).
EC04	Studies not written in English.
EC05	Studies that do not mention i* or variants.
EC06	Studies that do not mention scalability (and similar terms) of i* or (variants)
EC07	Non-Scientific studies (notes, index, editorials, prefaces).

Table 2. Inclusion criteria.

ID	Definition
IC01	Studies that were not eliminated by the exclusion criteria.
IC02	Texts published between 01/01/1990 and 31/12/2014.
IC03	Studies that address some of the study questions
IC04	Theoretical or empirical work will be included.

3.3 SOURCES SELECTION AND SEARCH

The search strategy included only electronic databases and was validated by experts on the requirements engineering area. By using a search string, the following electronic repositories were surveyed: Science Direct, ACM Digital Library, IEEE Xplore, Engineering Village, Scielo and World Scientific. Figure 1 shows the number of papers identified at each search engine and the steps taken to select a subset of papers.

We developed the search string by specifying the main terms used about this topic, the derivation form constructors and synonyms arising from the research questions, previously readings from known studies, consulting from experienced researchers in the field and dictionaries or glossaries.

We performed pilot searches to refine the keywords in the search string using trial and error. After some iterations, we settled on the following search string:

("iStar" OR "i-star" OR "i star" OR "Yu, e" OR "Yu e" OR "GRL" OR "Tropos") AND ("goal-oriented" OR "goal-directed" OR "agent-oriented" OR "requirements engineering" OR "software requirements") AND (model OR diagram*) AND (scal* OR modul* OR complex* OR compreh* OR underst* OR evolu* OR large OR huge OR big OR siz* OR enormous OR immense); Publication Year: 1990 – 2014*

The first set of or-clauses refers to i*, some of its variants (GRL and Tropos), or the i* proponent (Eric Yu). The second or-clause refers to its use: goal-oriented approaches, agent-oriented approaches, or requirements approaches. Then, it is made explicit that we are interested in models or diagrams. The last or-clause defines our focal point: scalability and its related attributes, such as modularity, complexity, and size. The time period was limited to from around the i* creation until the beginning of this mapping. Considering that each search engine has their own syntax to perform automatic searches, the above search string had to be adapted to each one of them. All the search string variations used in this study, as well as other details of the mapping, are available at our website[1].

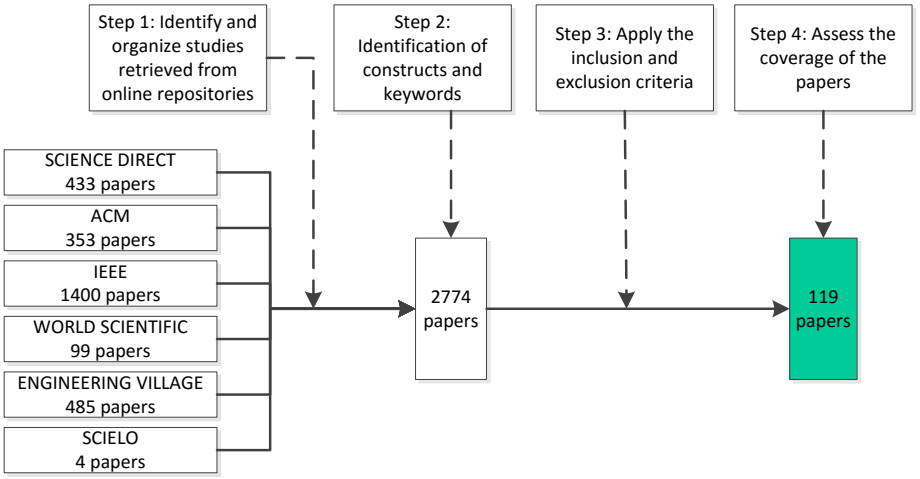

Figure 1. Paper selection.

After applying these steps, a total of 119 studies met the inclusion criteria and their data were extracted. The extraction was performed aiming to answer the research questions described in Section 3.1. The list of selected papers is presented in our website. In the next section, we present the results obtained.

3.4 THREATS TO VALIDITY

This section describes some threats that may be mitigated in future replications of this study and other aspects that must be taken into account when analyzing the re-

[1] www.cin.ufpe.br/~ler/supplement/wer2016

sults described in this paper [8]. We will base our discussion of threats to validity on the categories used by Wohlin et al. [7] but we will only go into details for the categories that we consider important for our study.

Internal validity was enhanced using triangulation in some parts of the method. We consulted experienced researchers to validate the research design and its understanding. Their feedback and the trials contributed to reduce threats.

Threats regarding construct validity required some extra care. This was necessary since the term scalability is an abstract word that has many definitions and the word i* is difficult to insert in the search engines. Hence, to minimize threats of this nature, we discussed synonyms and written alternatives for both terms (for example iStar for i*). We also related potential previous definitions for scalability under the supervision of experienced researchers.

External validity is concerned with establishing the generalizability of the systematic mapping results, which is related to the degree to which the primary studies are representative for the review topic [8]. The external validity (portability, transference) of this study was strengthened by the structure of the extracted data. It was also supported by detailing the research method in order to allow future comparative generalizations.

Regarding the empirical reliability, we tried to run a systematic mapping based on existing guidelines and on similar studies already published and accepted by the academic community.

4 RESULTS

The purpose of this research is to map, in the context of Requirements Engineering, the publications about the i* framework related to its scalability of. In the next sections, we present the answers for our research questions.

4.1 RQ1: What studies mention the issue of i* scalability?

We found 119 studies that mention i* scalability or at least one of its related concepts, such as modularity, complexity, and size. Considering the research types proposed by [9] and also used by [8], we identified that 45 studies can be considered empirical, i.e. they were classified in one the following types: evaluation research, validation research and solution proposals (see Figure 2).

We can notice that the number of studies on i* scalability has increased since 2005, peaking at 19 publications in 2011. Figure 3 shows the relationship between the search sources and the number of excluded and included studies.

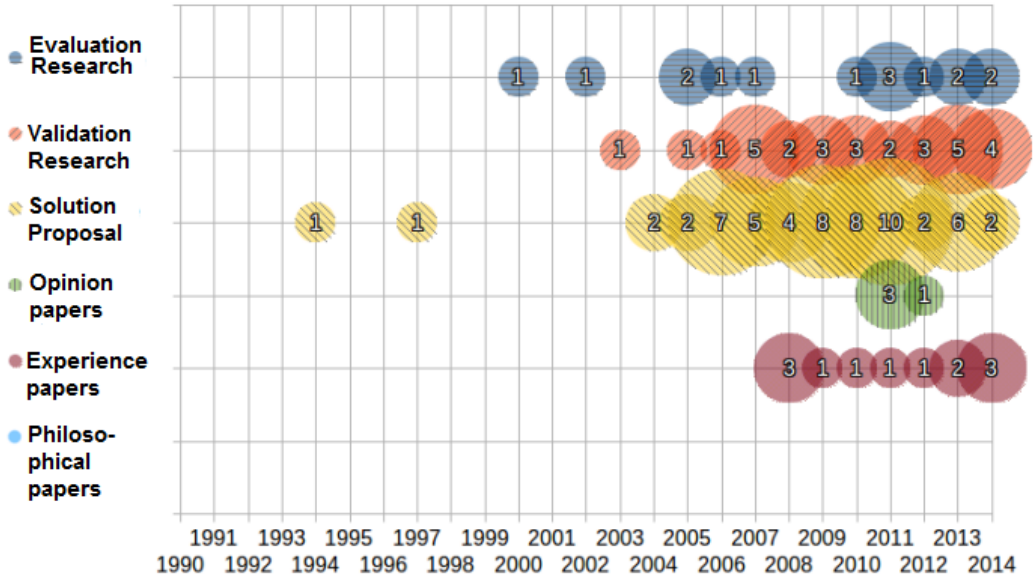

Figure 2. Selected studies distributed over quantity, year of publication and research categories.

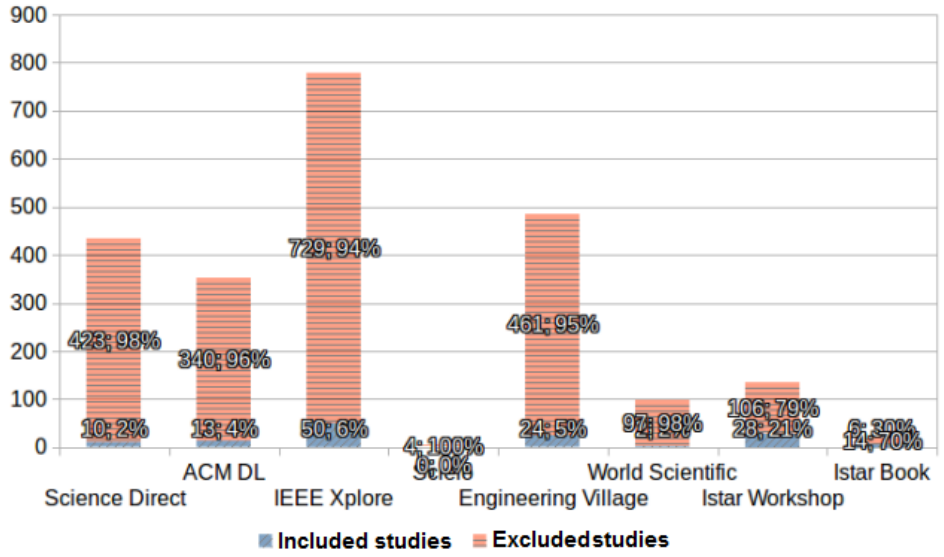

Figure 3. Number of included and excluded studies per search sources.

4.2 RQ2: Which are the scalability definitions in the context of i*?

From the selected studies, we identified 10 works that provide some definition of the term scalability. These definitions are shown in Table 3. In addition, 11 studies

had, in their research objective, the study of scalability or related attributes, such as modularity.

Table 3. Scalability definitions extracted from the selected studies. The full reference of each study is available at www.cin.ufpe.br/~ler/supplement/wer2016

Study Id	Definition or Characterization
[S25]	"Scalability was defined by the number of goal levels and number of variants."
[S61]	"able to have models at different levels of abstraction so that both domain experts and developers alike can get an idea of the overall system behavior or focus on a particular part of the system in more detail if required."
[S66]	"is able to handle numerous Agents in an application."
[S78]	"the property of reducing or increasing the scope of methods, processes, and management according to the problem size (. . .) Inherent in this idea is that software engineering techniques should provide good mechanisms for partitioning, composition, and visibility control. It includes the ability to scale the notation to particular problem needs, contractual requirements, or even to budgetary and business goals and objectives."
[S84]	"the reduced complexity of goal graphs (. . .), the ability to group goal graphs with concerns, the encapsulation provided by concerns, the ability to use parameterized point cut expressions in AoGRL, and the simpler update tasks for AoGRL suggest that AoGRL models are more scalable than GRL models."
[S97]	"measures the methodology's support for designing systems that are scalable. It means that the system should allow the incorporation of additional resources and software components with minimal user disruption."
[S98]	"The degree to which the modeling framework can be used to handle applications of different sizes. Scalability also measures extensibility, the degree to which the inclusion of new modeling elements leaves the understandability of models unaffected. This feature is causally related to refinement and modularity."
[S102]	"features in the technique to scale with the size and complexity of the system under assessment. Examples: Abstraction, refinement, decomposition, different formats, types or versions of technique."
[S104]	"ability of both the approach as well as the specifications to serve for a variety of project sizes and constraints, need to be easily modifiable."
[S111]	"large organizational models (depending on the domain and their description) become complex and inconsistent due to bad labeling and irrelevant information."

4.3 RQ3: Which are the types of contributions that have been published to support the i* scalability?

From the set of selected studies, we identified 150 mentions to different types of contributions to improve the scalability of i* (Figure 4). This includes repetitions to the same technique as well as more than one type of contribution in the same study, i.e., a single paper may be classified in more than one category.

We separate contributions in four categories: metamodels or formalisms (24 mentions), modeling processes (52 mentions), visual constructors (41 mentions), and software or algorithms (33 mentions).

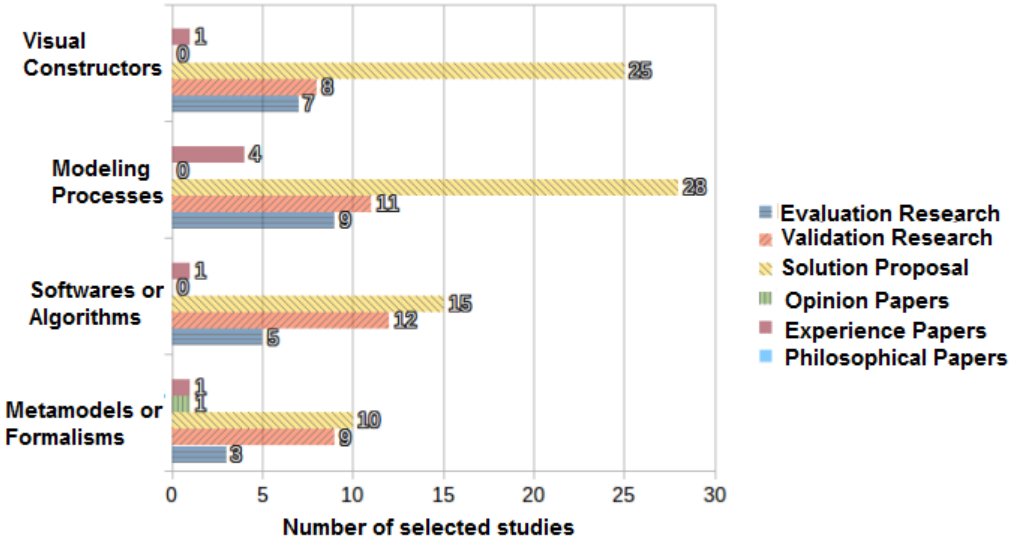

Figure 4. Types of contributions per type of research method.

4.4 RQ4: What is the perceived judgment on the scalability of i*?

According to the selected studies, we identified that i* does not have a good treatment regarding its scalability. In all research types categories (Figure 5), there were quantitatively more studies that reported bad impression on the i* models supporting scalability (62 studies). On the other hand, only eight studies judged the scalability of i*as being well treated. Finally, 49 studies did not mention any information about this question.

In relation to the research types categories (Figure 5), there are five Validation studies, one Evaluation study and two Solution Proposal studies category considering that scalability is well treated. On the other hand, eighteen Validation studies, eight Evaluation studies, twenty nine Solution Proposal studies, three Opinion Papers, and four Experience studies classified the scalability of i* as not being well handled.

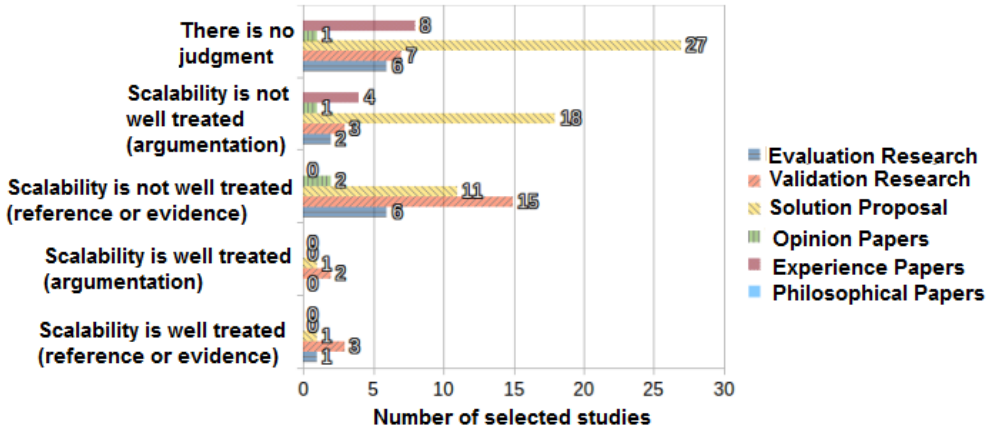

Figure 5. How the i* scalability is perceived.

4.5 RQ5: Are there open issues?

We identified 93 studies that present open issues and related future works about the scalability of i*. Regarding the research types categories (Figure 6), 47 Solution Proposal studies, 24 Validation studies, 11 Evaluation studies, 7 Experience studies, and 4 Opinion papers indicate open issues. On the other hand, 26 studies did not mention open issues. From these results, we concluded that there are indeed open issues on the topic.

4.6 Studies indicated by experienced researchers

After consulting experienced researchers, they pointed out 8 studies that were not captured by our systematic mapping study. The indicated studies were [10], [11], [12], [13], [14], [15], [16], and [17]. Most of these studies are PhD theses, which explains why they have not been captured by our search.

It is important to note that these studies are not inserted in the answers for our research questions since it is outside of our scope to use techniques such as manual inclusion and snowballing. Therefore, the results presented above correspond to the studies exclusively returned by our research protocol.

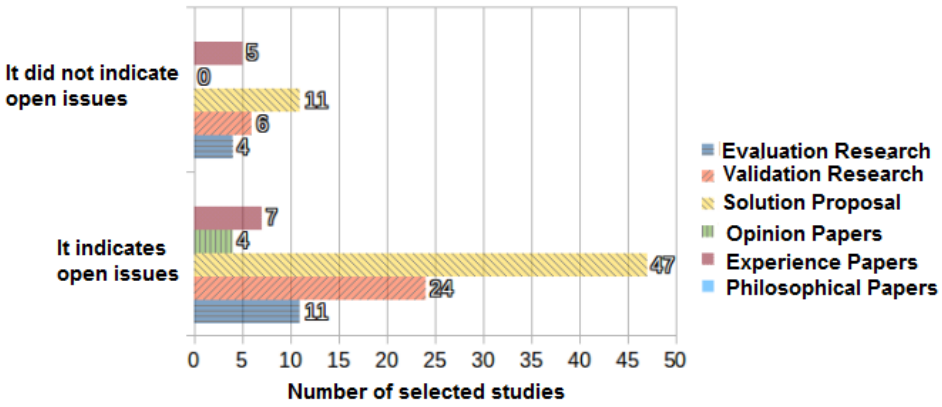

Figure 6. Number of papers that mentions open issues about the i* scalability per research category.

5 DISCUSSION

In this Section, we first discuss the scalability treatments of some of the reviewed papers, in order to provide an overview of approaches that have been proposed. Lastly, some general limitations and recommendations are presented.

Maté et al [25] defined modules for i* models. On an empirical evaluation comparing between regular models and models with modules, the authors observed increased scalability and improved understandability (considering the number of errors) on the latter, even though it took longer to understand them.

Pastor et al [24] presented an empirical evaluation of i* with respect to different concerns, including modularity, complexity management, and scalability. Both modularity and scalability were considered to be not supported, whereas complexity management was considered to be not well supported.

Alencar et al [18] extended i* with aspectual elements aiming to improve the modularity of its models. By distributing repeating elements on aspectual actors it is possible to increase the separation of concerns, with the disadvantage of being necessary to learn new elements in order to use this approach. With increased modularity and separation of concerns, the models are expected to be more scalable.

Previously, Mussbacher et al [22] have also investigated the adoption of aspectual concepts in the context of i*-based models. The proposal was evaluated by comparing three approaches for modeling an example system: regular GRL; monotonic GRL; and aspect-oriented GRL. The latter presented better results in terms of modularity, understandability, reusability, and maintainability, even though it presented an increased vocabulary size.

Oliveira et al [21] defined SDsituations, a modularity construct explicitly aimed at improving the scalability of i* models. SDsituations aggregate different elements of

i* models, and each SDsituation is related to other SDsitutations through logical, temporal or sequential, and physical dependencies.

The approach from Dalpiaz et al [19] allows designing adaptive socio-technical systems. Here, instead of considering the scalability of the models themselves, it was analyzed the scalability of the adaptation algorithms that take i* models as input.

Similarly, Aydemir et al [20] assessed the scalability of its algorithms for model evolution. Moreover, it explicitly takes some precautions in order to improve its visual scalability, such as high visual distance between different kinds of elements and one-to-one correspondence between symbols and concepts.

Horkoff and Yu [23] address the problem of performing an interactive analysis of large goal models by highlighting specific elements depending on the current analysis step.

The following limitations on the overall research about i* scalability were identified: lack of evaluation research on the topic; scarcity of i* models from the industry; there are many papers on the topic, but no clear recommendations on which approach to adopt on different contexts; lastly, some relevant publications are not readily available, preventing wider adoption of the proposed approaches.

In order to improve future research on the topic, the following actions are suggested: make large and complex models publicly available; interact with the industry in order to create and publish models of real systems; make the resources of scalability experiments publicly available; define metrics and exemplars for comparing different approaches.

6 CONCLUSIONS AND FUTURE WORK

In this paper we present a systematic mapping on the scalability of the i* framework. A total of 119 papers were analyzed, resulting on the characterization of the research topic. Based on this study, it was possible to discover the different ways that scalability is considered in the i* community, as well as different mechanisms that have been proposed in order to tackle scalability issues.

We presented 10 concepts related to scalability presented in the selected papers. The contribution types were classified in metamodels or formalisms, processes, visual constructors, and software, and the respective amount of contributions on each type were presented. The existent judgments about the scalability of i* were classified by the following scale: there is no judgment, scalability is not well treated (argumentation); scalability is not well treated (reference or evidence); scalability is well treated (argumentation) and scalability is well treated (reference or evidence).

Regardless of the fact that there are several papers on the topic, it could be observed that the scalability of the i* framework still have many open issues. In future work, we expect to perform a more detailed analysis of the selected papers and to establish mechanisms for comparing different approaches.

Acknowledgments

The following Brazilian institutions have supported this work: FACEPE, CAPES and CNPQ.

References

1. E. Yu. Social modeling and i*. In: Conceptual Modeling: Foundations and Applications, Lecture Notes in Computer Science, vol. 5600, pp. 99–121, 2009.
2. X. Franch. Fostering the adoption of i* by practitioners: Some challenges and research directions. In: Intentional Perspectives on Information Systems Engineering, pp. 177–193, 2010.
3. X. Franch. The i* framework: the way ahead. In: sixth international conference on Research Challenges in Information Science (RCIS), 2012. pp. 1–3.
4. J. Horkoff; E. Yu. Comparison and evaluation of goal-oriented satisfaction analysis techniques. Requirements Engineering, v. 18, n. 3, 2013, pp. 199-222.
5. B. Kitchenham, S. Charters. Guidelines for performing systematic literature reviews in software engineering. In: EBSE Technical Report EBSE-2007-01, 2007
6. K. Petersen, R. Feldt, S. Mujtaba, M. Mattsson. Systematic Mapping Studies in Software Engineering. In: Proceedings of the International Conference on Evaluation and Assessment in Software Engineering (EASE), 2008. pp. 68–77.
7. C. Wohlin et al. Experimentation in software engineering: An introduction. Norwell, MA: Kluwer Academic, 2000.
8. D. Dermeval, J. Vilela, Ig Bittencourt, J. Castro, S. Isotani, P. Brito, A. Silva. Applications of ontologies in requirements engineering: a systematic review of the literature. Requirements Engineering, pp. 1-33, 2015.
9. R. Wieringa; N. Maiden; N. Mead; C. Rolland. Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. In: Requirements Engineering, vol. 11, n. 1, 2006, pp. 102–107.
10. C. Cocca. Towards Improved Visual Support for i* Modelling. diploma thesis, Faculty of Information Studies – University of Toronto, Toronto, Canada (2007), http://www.collectionscanada.gc.ca/obj/thesescanada/vol2/002/ MR27330.PDF
11. H.E Esquivel. A Service-oriented Approach for the i* Framework. Ph.D. thesis, University of Trento, Italy; Universitat Politècnica de Valéncia, Spain (2008), http://hdl.handle.net/10251/3305
12. C. Gralha; M. Goulão; J. Araújo. Identifying modularity improvement opportunities in goal-oriented requirements models. In: Advanced Information Systems Engineering, Lecture Notes in Computer Science, vol. 8484, 2014, pp. 91–104.
13. C. Leica. Scalability Concepts for i* Modelling and Analysis. diploma thesis, Department of Computer Science – University of Toronto, Toronto, Canada (2005), http://www.collectionscanada.gc.ca/obj/thesescanada/vol2/002/mr02471.pdf
14. M. J. N. R Lucena. STREAM: a systematic process to derive architectural models from requirements models. Ph.D. thesis, Centro de Informática, Universidade Federal de Pernambuco (2010).
15. D. Moody; P. Heymans, R. Matulevicius. Visual syntax does matter: improving the cognitive effectiveness of the i* visual notation. In: Requirements Engineering, vol. 15, n. 2, 2010, pp. 141–175.

16. D. Moody. The physics of notations: Toward a scientific basis for constructing visual notations in software engineering. In:IEEE Transactions Software Engineering, vol. 35, n. 6, 2009, pp. 756–779.
17. J. Z. You. Using meta-model-driven views to address scalability in i* models. diploma thesis, Department of Computer Science – University of Toronto, Toronto, Canada (2004), http://www.collectionscanada.gc.ca/obj/s4/f2/dsk4/etd/MQ95286.PDF
18. F. Alencar, J. Castro, M. Lucena, E. Santos, C. Silva, J. Araújo, A. Moreira. Towards modular i* models. Proceedings of the 2010 ACM Symposium on Applied Computing. pp. 292–297, 2010.
19. F. Dalpiaz, P. Giorgini, J. Mylopoulos. Adaptive socio-technical systems: a requirements-based approach. In: Requirements Engineering, vol. 18, 2013, pp. 1–24.
20. F. Aydemir, P. Giorgini, J. Mylopoulos, F. Dalpiaz. Exploring alternative designs for socio-otechnical systems. Eighth International Conference on Research Challenges in Information Science, pp. 1–12, 2014.
21. A. Oliveira, J. Leite, L. Cysneiros, C. Cappelli. Eliciting multi-agent systems intentionality: from language extended lexicon to i* models. XXVI International Conference of the Chilean Society of Computer Science, pp. 40–49, 2007.
22. G. Mussbacher, D. Amyot, J. Araujo, A. Moreira, M. Weiss. Visualizing aspect-oriented goal models with aogrl. Second International Workshop on Requirements Engineering Visualization, 2007.
23. J. Horkoff, E. Yu. Visualizations to support interactive goal model analysis. Fifth International Workshop on Requirements Engineering Visualization, pp. 1–10, 2010.
24. O. Pastor, H. Estrada, A. Martínez. Strengths and weaknesses of the i* framework: An empirical evaluation. Social Modeling for Requirements Engineering, chapter 18, pp. 607–643, 2011.
25. A. Maté, J. Trujillo, X. Franch. Adding semantic modules to improve goal-oriented analysis of data warehouses using i-star. Journal of Systems and Software 88, pp. 102 – 111, 2014.
26. J. Horkoff, T. Li, F. Li; M. Salnitri, E. Cardoso, P. Giorgini, J. Mylopoulos, J. Pimentel. Taking goal models downstream: A systematic roadmap. In: Eighth International Conference on Research Challenges in Information Science (RCIS), 2014, pp.1-12, 2014.