

Ambiguity resolution of legal requirements: an exploratory study in the literature

Dorgival Netto¹ and Carla Silva²

¹Universidade Federal do Cariri, Juazeiro do Norte, Brasil
dorgival.netto@ufca.edu.br

²Centro de Informática, Universidade Federal de Pernambuco, Recife, Brazil,
ctlls@cin.ufpe.br

Abstract. **[Background]** Software requirements are usually specified in Natural Language, bringing challenges for Requirements Engineering (RE) as these specifications are inherently ambiguous. These challenges become bigger when dealing with software requirements that must comply with regulations, the so-called legal requirements. **[Goal]** This work investigates how ambiguity in legal requirements specification is addressed in works selected from the literature. **[Method]** We followed a snowballing technique with three backward and forward iterations. **[Results]** We systematically selected 24 papers and classified their proposals in approaches dealing with ambiguous requirements by identifying, reducing or avoiding ambiguity. Moreover, the protocol used to select the papers and the selected papers can be used as starting point to expand this study's results. **[Conclusions]** Our analysis in the selected papers revealed that a considerable amount of works deal with the identification of ambiguous requirements, in opposition to the number of works concerned in avoiding or reducing them.

Keywords: Exploratory Study · Requirements Engineering · Legal Compliance · Ambiguity

1 Introduction

Software development companies must comply with a large number of regulations and ensure that their business and system requirements are in legal compliance [20]. Legal texts, in general, are inflexible, non-negotiable, vague, ambiguous, open to different interpretations, and may change from new legislation [33]. Legal texts are full of ambiguities, often planned, called intentional ambiguity [31], [24]. This type of ambiguity allows laws and regulations to avoid dependence on technologies or practices that may change over time [5, 30].

Ambiguity is a proper attribute of natural language (NL) and a necessary feature that makes NL adaptable in several contexts [4]. Requirements specified in NL tend to be ambiguous [36]. Ambiguity in these specifications may lead different stakeholders, including software designers, regulators, and users, to have different interpretations of system behavior and functionality [6].

Ambiguities in the legal text can make the difference between compliant and non-compliant requirements [24]. Akhigbe et al. [3] define regulatory compliance as ensuring that a software organization, process, or product adheres to laws, guidelines, specifications, and regulations. The cost of non-compliance can also be quite high, and noncompliance can cause critical damage to the organization with fines or reputational harm [18].

Many works were developed to help engineers to address ambiguity in software requirements and to align them with legal constraints [6, 20, 23]. We have identified some Systematic Literature Reviews dealing with Legal Compliance for Requirements Engineering ([31, 3, 28]).

According to Massey et al. [26], many of the approaches developed to mitigate or disambiguate requirements specifications are not appropriate to deal with legal ambiguities. Rewrite legal texts is not an easy task; if ambiguity appears in current law or regulation, it must be clarified through interpretation instead of reformulation. Lawyers and engineers bring different, sometimes conflicting perspectives to the interpretation of legal texts [37].

In Brazil, Law No. 13,709/2018, entitled General Data Protection Law (in Portuguese, Lei Geral de Proteção de Dados (LGPD))[8] sanctioned on August 14, 2018. This Act based on Regulation European Union (EU) 2016/679, called EU General Data Protection Regulation (GDPR) [16]. Therefore, Brazilian companies are in the process of adapting to this legislation that will come into force on August 14, 2020.

Motivated by this scenario, this paper presents a bibliographical study on ambiguity in legal requirements using the snowballing technique [38] to identify works from an initial set and to classify the approaches found according to how they deal with ambiguous requirements when developing legally compliant systems: whether they identify, reduce or avoid ambiguity in legal requirements.

The remaining of this paper is structured as follows. Section 2 describes the research method. as well as data collection and analysis. Section 3 presents our preliminary results. Finally, Section 4 presents the conclusions and future directions of this research.

2 Snowballing - Planning and Execution

This section describes the methodology used to perform this research.

In the planning stage of snowballing, we describe the research questions, procedures, and methods for conducting this study. We used the guidelines proposed by Wohlin [38] to perform snowballing, that refers to using a list of references from a paper or the paper's quotes to identify additional papers [38].

Snowballing aims to respond to Research Question (RQ) *How current approaches deal with ambiguity of legal text in the specification of legal requirements?*

The specific research questions that guided this snowballing were:

RQ1. What approaches deal with ambiguity of legal text in the specification of legal requirements?

RQ2. How the approaches deal with ambiguity of legal text in the specification of legal requirements?

The first challenge is to identify an initial set of papers to use in the snowballing procedure. Wholin [38] mentions that if many papers are found, identifying several relevant and highly cited papers may be an alternative to defining the initial set. One possibility in snowballing is to identify a seminal or highly cited paper in the area of systematic literature study.

After defining the initial set, based on the inclusion and exclusion criteria, we started the process of backward and forward snowballing, performing as many iterations as necessary until new papers are not found.

Backward snowballing means using the reference list to identify new papers to include. The first step is to go through the reference list and exclude papers that do not fulfill the basic criteria. The next step is to remove papers from the list that have already been examined based on being found earlier through either backward or forward snowballing in this or a previous iteration. Once removed, the remaining papers are real candidates for inclusion [38].

Forward snowballing refers to identifying new papers based on those papers citing the paper examined. The citations to the paper examined are obtained from Google Scholar [38].

All papers identified go into data extraction, which should be conducted following the research questions posed in the systematic literature study [38].

It is important to decide on either inclusion or exclusion before starting to use a new paper for snowballing. We defined the inclusion and exclusion criteria, based on the RQ, to achieve consistent results:

Inclusion Criteria: **I1** Primary studies; **I2** Studies that aim to reduce or eliminate ambiguity in Legal Requirements Engineering; **I3** Studies that relate ambiguity and Legal Requirements Specification; **I4** Studies that relate ambiguity and Legal Requirements Elicitation.

Exclusion Criteria: **E1** Duplicated studies (one copy included); **E2** Gray literature (Short papers (less than 3 pages), presentations, reports, dissertations, thesis); **E3** Secondary or tertiary studies; **E4** Papers not written in Portuguese or English; **E5** Not published in peer-reviewed event; **E6** Publications whose text was not available (through search engines or by contacting the authors).

First, the studies have been checked using the exclusion criteria. If a paper could meet any of the exclusion criteria, in turn, if **E1 OR E2 OR E3 OR E4 OR E5 OR E6** is true, then the paper must be removed. Another case for a duplicate **E1** is when a conference paper is followed by a journal article. In such cases, we select the higher-valued publication, i.e., journal over conference [21]. Subsequently, the inclusion criteria were observed. Thus, it was verified if **I1 AND I2 AND I3 AND I4** could meet. If so, papers must be selected, if any criteria are not met, the paper is excluded.

To calibrate the search string and avoid bias, we used Google Scholar, as suggested by Wholin [38]. The search string has the following structure: "*ambiguity*" AND ("*legal requirements*" OR "*regulatory requirements*") AND "*requirements*"

engineering" AND ("*legal compliance*"). We set the display of the papers captured in the search as "*Sort by relevance*" and we did not define a period.

A limitation in our study is that we didn't use the term "law-compliant requirements" in the string to find the seed set of papers, but using snowballing could have contributed to mitigate the impact of this limitation.

3 Data Collection and Analysis

The search and analysis were carried out in the second half of 2017, between June and October. We obtained 110 candidate articles for the initial set.

Analyzing the papers and classifying them according to the inclusion and exclusion criteria, we selected five papers that comprise the initial set to be used as seed for snowballing (Fig. 1). The supplementary material [1] presents the results of each of the three iterations performed in this snowballing process.

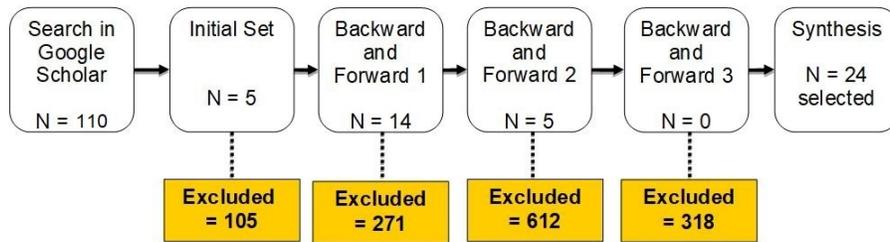


Fig. 1. Selection papers

In Table 1, we present the 24 papers selected and classify them if they aim to identify, reduce or avoid ambiguity. We capture papers that deal with other aspects of legal requirements during the snowballing process, such as cross-reference, requirements extraction, and legal requirements tracking. Therefore, some papers are unclassified in terms of ambiguity. Code [COOO] are part of the seed set. Code [BW - iteration C00] refers to papers discovered in Backward Snowballing iteration number. Code [FW - iteration C00] refers to papers discovered in Forward Snowballing iteration number.

The results of this study indicated that there is a more significant number of studies dealing with the identification of ambiguities rather than avoiding or reducing them.

4 Final Remarks & Future Directions

Ambiguity in legal requirements is a well-known problem both to academic and industry communities. Nevertheless, through the analysis of the papers selected in this study, we could not identify a complete and systematic process able to

Table 1. Classification of the selected papers

Paper	Classification
[C001] [12]; [C023] [34]; [BW - 1 C05] [2]; [BW - 1 C06] [25]	Reduce\Avoid
[C002] [27]; [C003] [13]; [C004] [24]; [FW - 2 -C01] [32]; [BW-2-C02] [22]	Identify
[BW - 1 C01] [14]; [BW - 1 C02] [11]; [BW - 1 C03] [9]; [BW - 1 C04] [19]; [FW-1-C01] [35]; [FW-1-C07] [34]; [FW-1-C08] [33]; [FW-1-C09] [15]; [BW-2-C01] [31]; [FW - 2 -C02] [7]	no classification
[FW-1-C02] [6]; [FW-1-C06] [26]	Identify\Reduce
[FW-1-C03] [10]; [FW-1-C04] [23]; [FW-1-C05] [17]	Reduce

guide the identification and avoidance/reduction of ambiguity in legal requirements. The papers found only make specific advances to solve this problem. In [29], the authors interviewed software practitioners to gather what practices they use to deal with ambiguity in legal requirements. The result was a set of good practices to avoid or solve ambiguity.

As future work, we intend to update this literature study to find new works and compare the results found and presented in this paper, besides comparing them with the results obtained in [3] and [29]. Also, in this update, we intend to add the following research questions:

- RQ3. What are the strategies used in Requirements Engineering to deal with the ambiguity present in legal requirements?
- RQ4. What are the approaches used to specify legal requirements with reduced ambiguity?
- RQ5. What are the challenges and limitations in the approaches analysed in the previous questions?

Acknowledgements

This paper was partially supported by FACEPE.

Referências

1. Supplementary material, <https://anonymous.4open.science/r/wer2023-ID28>
2. Using a security requirements engineering methodology in practice: The compliance with the italian data protection legislation. *Computer Standards Interfaces* **27**(5), 445–455 (2005), formal Methods, Techniques and Tools for Secure and Reliable Applications
3. Akhigbe, O., Amyot, D., Richards, G.: A systematic literature mapping of goal and non-goal modelling methods for legal and regulatory compliance. *Requirements Engineering* **1**(766), 1–23 (2018)

4. Bano, M.: Addressing the challenges of requirements ambiguity: A review of empirical literature. In: 2015 IEEE Fifth International Workshop on Empirical Requirements Engineering (EmpiRE). pp. 21–24 (2015)
5. Berry, D., Kamsties, E.: Ambiguity in requirements specification. In: do P. Leite, J.C.S., Doorn, J.H. (eds.) Perspectives on software requirements, pp. 7–44. Springer US, Boston, MA (2004)
6. Bhatia, J., Breaux, T., Reidenberg, J., Norton, T.: A theory of vagueness and privacy risk perception. In: 24th International Requirements Engineering Conference (RE). pp. 26–35. IEEE, Beijing, China (2016)
7. Boella, G., Humphreys, L., Muthuri, R., Rossi, P., van der Torre, L.: A critical analysis of legal requirements engineering from the perspective of legal practice. In: 2014 IEEE 7th International Workshop on Requirements Engineering and Law (RELAW). pp. 14–21 (2014)
8. BRASIL: Lei geral de protecao de dados n. 13.709, 14/ago. (2018), http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm
9. Breaux, T., Antón, A.: Analyzing regulatory rules for privacy and security requirements. IEEE Transactions on Software Engineering **34**(1), 5–20 (2008)
10. Breaux, T.D., Anton, A.I.: Managing ambiguity and traceability in regulatory requirements: A tool-supported frame-based approach. Tech. rep., North Carolina State University. Dept. of Computer Science (2007)
11. Breaux, T.D., Antón, A.I.: A systematic method for acquiring regulatory requirements: A frame-based approach. RHAS-6), Delhi, India (2007)
12. Breaux, T.D., Antón, A.I., Boucher, K., Dorfman, M.: Legal requirements, compliance and practice: an industry case study in accessibility. In: 16th IEEE Intl. Requirements Engineering Conference. pp. 43–52. IEEE, Catalunya, Spain (2008)
13. Breaux, T.D., Gordon, D.G.: Regulatory requirements traceability and analysis using semi-formal specifications. In: Proceedings of the 19th International Conference on Requirements Engineering: Foundation for Software Quality. p. 141–157. REFSQ’13, Springer-Verlag, Berlin, Heidelberg (2013)
14. Breaux, T.D., Vail, M.W., Anton, A.I.: Towards regulatory compliance: Extracting rights and obligations to align requirements with regulations. In: 14th IEEE International Requirements Engineering Conference (RE’06). pp. 49–58 (2006)
15. Ebad, S.A., Jaha, E.S., Al-Qadhi, M.A.: Analyzing privacy requirements: A case study of healthcare in saudi arabia. Informatics for Health and Social Care **41**(1), 47–63 (2016)
16. EU: The EU general data protection regulation (2018), <https://gdpr.eu/>
17. Ferrari, A., Spoletini, P., Gnesi, S.: Ambiguity and tacit knowledge in requirements elicitation interviews. Requirements Engineering **21**(3), 333–355 (2016)
18. Ghanavati, S.: Legal-Urn Framework for Legal Compliance of Business Processes. Ph.D. thesis, CAN (2013), aAINR97881
19. Ghanavati, S., Amyot, D., Peyton, L.: Towards a framework for tracking legal compliance in healthcare. In: Krogstie, J., Opdahl, A., Sindre, G. (eds.) Advanced Information Systems Engineering. pp. 218–232. Springer Berlin Heidelberg, Berlin, Heidelberg (2007)
20. Ghanavati, S., Amyot, D., Rifaut, A.: Legal goal-oriented requirement language (legal grl) for modeling regulations. In: 6th International Workshop on Modeling in Software Engineering. pp. 1–6. ACM, NY, USA (2014)
21. Giardino, C., Paternoster, N., Unterkalmsteiner, M., Gorschek, T., Abrahamsson, P.: Software development in startup companies: The greenfield startup model. IEEE Transactions on Software Engineering **42**(6), 585–604 (2016)

22. Kamsties, E., Berry, D.M., Paech, B., Kamsties, E., Berry, D., Paech, B.: Detecting ambiguities in requirements documents using inspections. In: Proc of the 1st workshop on inspection in software engineering (WISE'01). vol. 13. Citeseer (2001)
23. Massey, A., Holtgreffe, E., Ghanavati, S.: Modeling regulatory ambiguities for requirements analysis. In: International Conference on Conceptual Modeling (ER). pp. 231–238. Springer Intl. Publishing, Cham (2017)
24. Massey, A., Rutledge, R., Antón, A., Swire, P.: Identifying and classifying ambiguity for regulatory requirements. In: 22nd International Requirements Engineering Conference (RE). pp. 83–92. IEEE, Karlskrona, Sweden (2014)
25. Massey, A.K., Otto, P.N., Hayward, L.J., Antón, A.I.: Evaluating existing security and privacy requirements for legal compliance. *Requir. Eng.* **15**(1), 119–137 (2010)
26. Massey, A.K., Rutledge, R.L., Antón, A.I., Hemmings, J.D., Swire, P.P.: A strategy for addressing ambiguity in regulatory requirements. Tech. rep., Georgia Institute of Technology (2015)
27. Maxwell, J.C., Antón, A.I., Swire, P.: A legal cross-references taxonomy for identifying conflicting software requirements. In: IEEE 19th International Requirements Engineering Conference. pp. 197–206 (2011)
28. Netto, D., Peixoto, M., Silva, C.: Privacy and security in requirements engineering: Results from a systematic literature mapping. *Anais do WER19 - Workshop em Engenharia de Requisitos, Recife-PE, Brasil, Agosto 13-16, 2019* (2019)
29. Netto, D., Silva, C., Araújo, J.a.: Identifying how the brazilian software industry specifies legal requirements. p. 181–186. SBES '19, Association for Computing Machinery, New York, NY, USA (2019)
30. Otto, P.: Reasonableness meets requirements: Regulating security and privacy in software. *Duke LJ* **59**, 309 (2009)
31. Otto, P., Antón, A.: Addressing legal requirements in requirements engineering. In: 15th International Requirements Engineering Conference (RE). pp. 5–14. IEEE, Delhi, India (2007)
32. Otto, P.N., Antón, A.I.: Managing legal texts in requirements engineering. In: Lyytinen, K., Loucopoulos, P., Mylopoulos, J., Robinson, B. (eds.) *Design Requirements Engineering: A Ten-Year Perspective*. pp. 374–393. Springer Berlin Heidelberg, Berlin, Heidelberg (2009)
33. Rabinia, A., Ghanavati, S.: Fol-based approach for improving legal-trl modeling framework: A case for requirements engineering of legal regulations of social media. In: 25th International Requirements Engineering Conference Workshops (REW). pp. 213–218. IEEE, Lisbon, Portugal (2017)
34. Reidenberg, J., Bhatia, J., Breaux, T., Norton, T.: Ambiguity in privacy policies and the impact of regulation. *The Journal of Legal Studies* **45**(S2), S163–S190 (2016)
35. Sannier, N., Baudry, B.: Toward Multilevel Textual Requirements Traceability Using Model-Driven Engineering and Information Retrieval. In: 2nd Intl. Model-Driven Requirements Engineering (MoDRE) Ws. Chicago, United States (2012)
36. Shah, U.S., Jinwala, D.C.: Resolving ambiguities in natural language software requirements: A comprehensive survey. *SIGSOFT Softw. Eng. Notes* **40**(5), 1–7 (sep 2015)
37. Swire, P., Anton, A.: Engineers and lawyers in privacy protection: Can we all just get along. *IAPP Privacy Perspectives* **1**(1), 1–4 (2014)
38. Wohlin, C.: Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: Proceedings of the 18th international conference on evaluation and assessment in software engineering. p. 38. ACM, London, England, United Kingdom (2014)