Automatic Identification of Privacy and Security Requirements: A Systematic Literature Review

FRANCESCO CASILLO, Department of Computer Science, University of Salerno, Italy
VINCENZO DEUFEMIA, Department of Computer Science, University of Salerno, Italy
CARMINE GRAVINO, Department of Computer Science, University of Salerno, Italy

The utmost importance of privacy and security requirements in software development calls for adopting methods that enable the identification and proactive mitigation of these issues during the system development. Our survey of 45 primary studies provides a comprehensive overview of the methods, document types, and datasets employed in tackling this critical challenge, along with an analysis of approaches demonstrating superior performance based on document types and specific identification problems. Analysis reveals a wide adoption of AI-based systems on diverse datasets, showcasing the effectiveness of leveraging various source of information to identify privacy and security requirements in software development.

CCS Concepts: • **Software and its engineering** → **Requirements analysis**; • **Security and privacy** → **Software security engineering**.

Additional Key Words and Phrases: Software requirements, privacy & security, automatic identification

ACM Reference Format:
Francesco Casillo, Vincenzo Deufemia, and Carmine Gravino. 2023. Automatic Identification of Privacy and Security Requirements: A Systematic Literature Review. 1, 1 (August 2023), 31 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Requirements Engineering (RE) plays a pivotal role in software development by encompassing vital activities that focus on understanding and fulfilling the capabilities and characteristics demanded by a system [61]. Through stages like elicitation, analysis, specification, and validation, RE aims to comprehend customer needs and translate them into precisely defined requirements [55], which include non-functional requirements (NFRs). The latter specify system qualities that extend beyond its core functionality, encompassing various attributes essential for its overall performance and success. Among NFRs, privacy and security have emerged as prominent concerns in software development. Incidents involving unauthorized data exploration, misuse of information, and unauthorized disclosure of personal data have raised awareness regarding privacy risks [52]. Users may be unaware of when and for what purposes their sensitive information is collected, analyzed, or

Authors' addresses: Francesco Casillo, fcasillo@unisa.it, Department of Computer Science, University of Salerno, Via Giovanni Paolo II, 132, Fisciano(SA), Italy, 84084; Vincenzo Deufemia, deufemia@unisa.it, Department of Computer Science, University of Salerno, Via Giovanni Paolo II, 132, Fisciano(SA), Italy, 84084; Carmine Gravino, gravino@unisa.it, Department of Computer Science, University of Salerno, Via Giovanni Paolo II, 132, Fisciano(SA), Italy, 84084.

Manuscript submitted to ACM