UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

DANIEL BRUNO DE CASTRO REIS

Como as leis de privacidade e de proteção de dados impactam as atividades de engenharia de software: uma revisão sistemática da literatura

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Marcelo Soares Pimenta

Porto Alegre
2023

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Reis, Daniel Bruno de Castro

Como as leis de privacidade e de proteção de dados impactam as atividades de engenharia de software: uma revisão sistemática da literatura / Daniel Bruno de Castro Reis – Porto Alegre: PPGC da UFRGS. – 2023.

49 f.:il.

Orientador: Marcelo Soares Pimenta.

Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação. Porto Alegre, BR – RS, 2023.

1.Leis de Privacidade. 2. Leis de Proteção de Dados. 3. GDPR. 4. LGPD. 5. Atividades de Engenharia de Software. I. Pimenta, Marcelo Soares. III. Como as leis de privacidade e de proteção de dados impactam as atividades de engenharia de software: uma revisão sistemática da literatura.

AGRADECIMENTOS

Gostaria de agradecer aos meus pais, parentes e amigos por toda compreensão, companheirismo e cumplicidade nesse período de realização do mestrado acadêmico.

Também gostaria de deixar meu agradecimento aos professores, funcionários e colegas do Instituto de Informática da UFRGS; Em especial, gostaria de agradecer meus colegas de setor, por esse período e pela compreensão.

Não poderia de deixar de agradecer a dedicação de meu orientador, professor Dr. Marcelo Soares Pimenta, que se empenhou bastante para que eu pudesse realizar a apresentação deste trabalho em caráter excepcional.

Agradeço também ao professor Alberto, coordenador do PPGC, e aos colegas da Secretaria do Programa de Pós-Graduação em Computação pela oportunidade dada para a conclusão deste trabalho.

Enfim, deixo meu agradecimento a todos que, em algum momento, passaram ou estiveram presentes nesta fase de minha vida.

RESUMO

Garantir a proteção de dados pessoais nos sistemas de software contemporâneos é um desafio multifacetado, exigindo a implementação de regulamentos e restrições legais para a gestão de dados pessoais, bem como um suporte metodológico para o desenvolvimento de sistemas de software que preservem a privacidade dos dados dos seus utilizadores. Esta pesquisa investiga as ramificações das leis de privacidade e proteção de dados nas atividades de engenharia de software e busca determinar as modificações ou ampliações necessárias para garantir a adesão às leis de privacidade. Foi realizada uma revisão sistemática para descobrir literatura pertinente relativa às questões de privacidade e proteção de dados nas práticas de engenharia de software. Além disso, foram explorados os métodos e atividades concebidos para facilitar o desenvolvimento de software sensível à privacidade. As descobertas ressaltam a necessidade de (i) integrar a abordagem da Engenharia de Privacidade (PE) e da Privacidade desde o Design (PbD) na engenharia de software, abrangendo todas as fases do ciclo de vida de desenvolvimento de software, e (ii) investigar estratégias para aumentar a maturidade e a consciência da privacidade dentro das equipes de desenvolvimento de software.

Palavras-chave: Leis de Privacidade, Leis de Proteção de Dados, GDPR, LGPD, Atividades de Engenharia de Software.

How privacy and data protection laws impact software engineering activities: a systematic literature review

ABSTRACT

Ensuring the protection of personal data within contemporary software systems is a multifaceted challenge, demanding the implementation of legal regulations and constraints for managing personal data as well as a methodological support for the development of software systems that uphold the privacy of their users' data. This research delves into the ramifications of privacy and data protection laws on software engineering activities and seeks to ascertain the requisite modifications or augmentations within the software engineering process to ensure adherence to privacy laws. Was Conducted a systematic literature review to unearth pertinent literature pertaining to privacy and data protection concerns in software engineering practices. Furthermore, was explored the methods and activities designed to facilitate privacy-aware software development. Thefindings underscore the need to (i) integrate both Privacy Engineering (PE) and the Privacy by Design (PbD) approach into software engineering, spanning all phases of the software development lifecycle, and (ii) investigate strategies for augmenting the maturity and privacy awareness within software development teams.

Keywords: Privacy Laws, Data Protection Laws, GDPR, LGPD, Software Engineering Activities.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E DE SIGLAS

ER	Engenharia de Requisitos
EUA	Estados Unidos da América
GDPR	General Data Protection Regulation
LGPD	Lei Geral de Proteção de Dados Pessoais
NDPR	Nigeria Data Protection Regulation
PbD	Privacy by Design
QP	Questões de Pesquisa
RNP	Rede Nacional de Ensino e Pesquisa
RSL	Revisão Sistemática da Literatura
SRL	Systematic Literature Review
UE	União Europeia
UFRGS	Universidade Federal do Rio Grande do Sul
UML	Unified Modeling Language

SUMÁRIO

1 INTRODUÇÃO

A cada dia, torna-se maior a quantidade de dados produzidos por pessoas e por sistemas ao redor do mundo (FRANÇA et al., 2014). Entre esses dados, estão informações sensíveis, como dados pessoais, cadastros, fotos, números de documentos, entre outros. Ao mesmo tempo, com o acesso à rede mundial de computadores cada vez mais democratizado, tornou-se frequente – muitas vezes sem que o usuário tenha conhecimento – a divulgação, transferência, acesso indevido, vazamento e/ou venda desses dados sem a explícita ou prévia autorização dos seus titulares, gerando um problema de privacidade (ESTRADA-JIMÉNEZ et al, 2017).

Com o objetivo de coibir o acesso a informações sensíveis, sem autorização do seu proprietário, entrou em vigência em 2018 a *General Data Protection Regulation* (GDPR) (ALBRECHT et al., 2016; DUGGINENI et al., 2023) na União Europeia, uma lei que tem por objetivo proteger os dados usuários. No Brasil, a privacidade já é uma garantia constitucional reafirmada em mecanismos legais de proteção, como o Marco Civil da Internet - Lei nº 12.965, de 23 de abril de 2014 (Brasil, 2014). Contudo, é importante frisar que privacidade se diferencia de proteção de dados, e que mesmo um dado público deve ser protegido (GARCIA et al., 2020). Nesse contexto, e com forte inspiração no GDPR, em 18 de setembro de 2020, entrou em vigor no Brasil a LGPD – Lei Geral de Proteção de Dados Pessoais – Lei nº 13.709, de 14 de agosto 2018 (Brasil, 2018), representando um avanço importante para o país. Assim, o Brasil passou a fazer parte de um grupo de nações que possuem uma legislação específica com o objetivo de proteger os dados dos seus cidadãos. Diante de casos de uso indevido, comercialização e vazamento de dados, os dispositivos da lei visam garantir a privacidade dos brasileiros.

A Engenharia de Software refere-se à aplicação disciplinada de princípios e métodos de engenharia visando a produção econômica de software de qualidade (HUMPHREY et al., 1988), a partir de atividades e tarefas pré-definidas. Contudo, não existe uma “bala de prata” no processo de desenvolvimento de software (BROOKS et al., 1987) e, por isso, diferentes abordagens de desenvolvimento são adotadas pelos engenheiros de software em cada projeto. Não somente em novos projetos, mas também como forma de adequar-se às leis de privacidade, diversos sistemas de informação existentes tendem a necessitar de manutenção ou de ajustes, seja na forma como os dados são tratados, ou no modo como o uso e o acesso às informações pessoais é autorizado pelo usuário (DUGGINENI et al., 2023).

Por essa razão, considerando as necessidades de adaptação de softwares em produção, e as etapas de desenvolvimento e de implementação de novos sistemas, o objetivo deste trabalho é investigar como as leis de privacidade e de proteção de dados impactam as atividades de engenharia de software e quais mudanças foram necessárias incluir ou adicionar ao processo para se garantir o cumprimento das regulamentações de privacidade. Para isso, optou-se por realizar uma revisão sistemática da literatura (RSL) a partir de estudos primários sobre esse tema.

O restante deste trabalho está organizado da seguinte maneira: O capítulo 2 descreve a base teórica sobre os tipos de revisões sistemáticas da literatura, definições acerca de leis de privacidade e de proteção de dados – como GDPR e da LGPD –, e uma breve revisão sobre atividades de engenharia de software. O capítulo 3 descreve outras revisões sistemáticas relacionadas a desenvolvimento, implementação e/ou adequação de sistemas a leis de privacidade. O capítulo 4 descreve o protocolo estipulado para a realização da RSL, incluindo a motivação, questões de pesquisa, string de busca e avaliação da qualidade. O capítulo 5 relata os resultados com base na seleção de estudos primários, o processo de avaliação da qualidade e responde às questões de pesquisa definidas no capítulo 4. No capítulo 6, são resumidos os principais achados e tendências identificadas. Ao final, o capítulo 7 apresenta as conclusões e possibilidades de trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo aborda aspectos e características relacionadas ao processo de elaboração de uma revisão sistemática. Também são descritos conceitos relacionados a leis de privacidade e, por fim, destaca tópicos relevantes referentes às leis de privacidade na Europa (GDPR) e no Brasil (LGPD). Ao final, também é realizada uma análise referente aos termos relacionados a atividades de engenharia de software.

2.1 Revisão Sistemática da Literatura

Revisões da literatura são estudos que podem ser definidos como um meio de identificar, avaliar e/ou interpretar toda pesquisa disponível relevante a uma questão, área ou fenômeno de interesse de uma pesquisa particular (KITCHENHAM et al., 2004). Seu objetivo é sintetizar os principais conceitos e terminologias de um determinado tópico em um único estudo, permitindo identificar tendências, lacunas e áreas desse tema em que ainda é preciso aprofundar pesquisas (ROWLEY et al., 2004). Por isso, uma revisão de literatura é um formato de estudo que pode ser utilizado para resumir o estado da arte de um tópico de pesquisa, facilitando a sua compreensão e o conhecimento que se tem acerca do mesmo.

Através de uma revisão da literatura também é possível avaliar e comparar teorias por meio do exame das relações entre as variáveis, o que fornece uma base conceitual para a criação de novas teorias conceituais. O método ou tipo de revisão pode variar de acordo com o objetivo da revisão. Os estudos individuais (pesquisas, questionários, estudos de caso, pré-experimentos) que contribuem para a produção de uma RSL são denominados de estudos primários (WOHLIN et al., 2012). Uma RSL é uma forma de estudo secundário que possui uma metodologia bem definida para identificar, avaliar e interpretar todas as evidências disponíveis relacionadas a uma questão de pesquisa específica (KITCHENHAM et al., 2007) e que é, até certo ponto, repetível. Os seguintes objetivos servem como justificativa para o desenvolvimento de uma RSL (KITCHENHAM et al., 2007): (i) resumir evidências existentes sobre um método, tratamento, conceito, teoria ou abordagem expondo limitações e benefícios; (ii) identificar eventuais lacunas de pesquisa mostrando à academia quais temas precisam de mais pesquisas e de novas soluções; e (iii) fornecer uma base para novas atividades de pesquisa, além de examinar até que ponto as evidências empíricas são evidenciadas ou contrariadas.

Ao contrário de uma revisão de literatura tradicional, uma RSL possui um protocolo explícito a ser seguido para identificar estudos primários e analisá-los de maneira completa e

imparcial (MACDONELL et al., 2010). Esse protocolo costuma ser dividido geralmente em três fases principais: planejamento, condução e relatoria (KITCHENHAM et al., 2007). A primeira fase, de planejamento, consiste em identificar a necessidade da realização da RSL, definir o escopo através das questões de pesquisa e desenvolver um protocolo de revisão para seleção dos estudos, extração de dados necessários para permita responder às questões e sumarizar os dados para que as questões possam ser respondidas. Na fase de condução, as atividades são realizadas conforme planejadas na fase anterior, de protocolo. A fase final, de relatoria, tem por objetivo elaborar um documento que apresente e relate os resultados de forma eficiente.

2.2 Leis de Privacidade

Conforme (VOSS et al., 2019), os termos “informações pessoais”, também denominados de “informações de identificação pessoal” e de “dados pessoais” são fundamentais para a compreensão de leis de privacidade de dados, pois esses termos delimitam o escopo das leis. A seguir, são apresentadas as principais características da GDPR (UE) e da LGPD (Brasil):

2.2.1 GDPR – General Data Protection Regulamentation (UE)

O GDPR é um projeto que visa proteger os dados e as identidades dos cidadãos da União Europeia (VIEIRA et al., 2021). O projeto foi concebido em 2012, aprovado em 2016 e está em vigor desde maio de 2018 (ZAEEM et al., 2020). Ela possui 11 capítulos e 99 artigos, sendo que os capítulos principais regem sobre: princípios; direitos do titular dos dados; controlador e processador; transferência de dados para países ou organizações internacionais; recursos, responsabilidades e sanções; situações específicas de tratamento. Para a União Europeia a proteção dos dados pessoais é um direito dos seus cidadãos e, portanto, todas as empresas e organizações devem seguir regras rígidas para coletar, processar, compartilhar e proteger dados pessoais – independentemente da escala ou ramo de atividade (VIEIRA et al., 2021). Além disso, este regulamento se aplica a qualquer tipo de serviço oferecido para cidadãos que chegam a um dos países da UE, isto é, qualquer país que desejar ter relações comerciais – ou enviar produtos para clientes da UE – deve também se adaptar ao GDPR para não violar a lei 2016/679 (União Europeia, 2016).

2.2.2 LGPD – Lei Geral de Proteção de Dados Pessoais (Brasil)

A Lei Geral de Proteção de Dados Pessoais (LGPD) – Lei nº 13.709, de 14 de agosto de 2018 (Brasil, 2018) – é a lei brasileira para proteção de dados pessoais e foi inspirada fortemente na GDPR (Albrecht et al.,2016; DUGGINENI et al., 2023). Ela possui 10 capítulos e 65 artigos, sendo que os capítulos mais relevantes regem sobre: Tratamento de Dados Pessoais; Direitos do Titular; Transferência Internacional de Dados; Segurança e Boas práticas. Nesse último capítulo citado, são estabelecidas diretrizes para a segurança e para o sigilo dos dados, bem como a governança e boas práticas em relação ao tratamento e aos dados, à natureza, ao escopo, à finalidade e à probabilidade e a gravidade dos riscos e dos benefícios decorrentes de tratamento de dados do titular. Conforme o texto da lei, essas medidas deverão ser observadas desde a fase de concepção do produto ou do serviço até a sua fase de execução.

A publicação da LGPD ocorreu em agosto de 2018 e está em vigor desde agosto de 2020. Em relação à privacidade de dados, (DIAS et al., 2020) elenca que ela compreende os dados do usuário – criados por ele mesmo ou por terceiros – e sua utilização por indivíduos através de observações, análises, entre outros. De acordo com a lei, o dado pessoal é considerado uma “informação relacionada à pessoa natural identificada ou identificável”, ou seja, dados como nome, endereço, sexo, RG e CPF (GARCIA et al., 2020; VIEIRA et al., 2021). Assim, dado pessoal pode ser definido como toda e qualquer informação, seja ela direta ou indireta, que possa permitir a identificação de uma pessoa (MENDES et al., 2020).

2.3 Atividades de Engenharia de Software

Engenharia de software se refere à aplicação disciplinada de princípios e métodos de engenharia, científicos e matemáticos para a produção econômica de software de qualidade. (HUMPHREY et al., 1988). Embora cada termo nesta definição possa ser discutido mais detalhadamente, o foco para este trabalho está apenas nas palavras “disciplina” e/ou “atividades” de engenharia de software. Por usar disciplinas estabelecidas, como gerenciamento de configuração, padrões de codificação ou nomenclatura de convenções, uma organização de software pode evitar repetir o processo de resolução e/ou de análise de problemas encontrados anteriormente (HUMPHREY et al., 1988).

Como o processo de engenharia de software utilizado para um projeto específico deve refletir suas necessidades específicas, é necessária uma estrutura para fornecer consistência entre os projetos. Uma arquitetura de processo de software fornece os principais pontos de verificação, atividades e definições necessárias para permitir a adoção de tecnologias,

métodos e medições comuns (HUMPHREY et al., 1988), visto que o desenvolvimento de software é diversificado e as empresas precisam adotar rapidamente novas tecnologias e mercados (KUHRMANN et al., 2017).

Não existe “bala de prata” no processo de desenvolvimento de software (BROOKS et al., 1987), por isso engenheiros de software sempre realizam diferentes abordagens de desenvolvimento, que sejam mais adequadas. Mesmo assim, enfrentam uma enorme variedade de desafios e fatores contextuais que influenciam na definição de processos de desenvolvimento que sejam apropriados (KUHRMANN et al., 2017). Para abordar essas diferenças contextuais, foram propostos diferentes métodos de desenvolvimento de software e de sistemas (HIJAZI et al., 2012), sendo os principais *Waterfall*, Incremental, *V-Model*, Espiral e metodologias ágeis.

Cada metodologia de desenvolvimento citada possui um conjunto de tarefas e/ou atividades, sendo, em essência: elicitação de requisitos, análise, design, codificação, teste e operação (HIJAZI et al., 2012). As nomenclaturas podem variar de acordo com o processo de desenvolvimento, mas realizam, em geral, as mesmas atividades. Por exemplo, as histórias de usuário, nos métodos ágeis, são equivalentes às elicitações de requisitos de outros processos de desenvolvimento (LEE et al., 2003).

A etapa de elicitação de requisitos é tão importante que possui uma disciplina especial, denominada Engenharia de Requisitos (ER) (ZOWGHI et al., 2005), que é o processo pelo qual os requisitos para desenvolvimento de uma aplicação são reunidos, analisados, documentados e gerenciados durante todo o ciclo de vida do software. A principal função da ER é interpretar e compreender os objetivos, necessidades e crenças das partes interessadas. Em geral, desenvolvedores de software percebem que um forte processo de gerenciamento de requisitos é essencial para a conclusão bem-sucedida dos projetos de software (ZOWGHI et al., 2005). Os requisitos podem, ainda, ser classificados como funcionais e não funcionais (DABBAGH et al., 2014). Os requisitos funcionais descrevem o comportamento funcional do sistema, enquanto os requisitos não funcionais expressam o quão bem um sistema deve funcionar (DABBAGH et al., 2014). Por essa razão, requisitos não funcionais também chamados de requisitos de qualidade (WAGNER et al., 2013), cujo objetivo é descrever como o sistema deve performar (BEHUTIYE et al., 2017).

De acordo com (FERRÃO et al., 2022), as normas contidas em leis de privacidade são geralmente definidas como requisitos de privacidade e categorizadas como requisitos não funcionais. Por sua natureza, os requisitos de privacidade podem ser utilizados para registros de regras fundamentadas em bases legais.

3 TRABALHOS RELACIONADOS

Este capítulo descreve revisões existentes no contexto de como leis de privacidade impactam as atividades de engenharia de software. Na sequência, também é realizada uma comparação com este trabalho, justificando a necessidade da sua realização.

3.1 Revisões da Literatura sobre o impacto de leis de privacidade nas atividades de engenharia de software

A primeira etapa na fase de planejamento de uma RSL é identificar a sua necessidade, comparando-a com outras revisões da literatura existentes. Nesta seção, são resumidos estudos secundários (RSL, pesquisas, revisões) no contexto de como leis de privacidade impactam as atividades de engenharia de software ou tratamento de dados. Os estudos que listados a seguir, em ordem cronológica, foram coletados durante o processo de condução dessa RSL (detalhes do protocolo, critérios de inclusão e critérios de exclusão são apresentados no capítulo 4), no entanto, foram descartados do conjunto de trabalhos selecionados por se tratarem de estudos secundários. O objetivo e as questões de pesquisa abordadas por cada trabalho estão resumidos na Tabela 3.1. A Tabela 3.2 resume a fonte de informação, a sequência de pesquisa e o número de estudos selecionados.

O estudo (CANEDO et al., 2021) elenca que a privacidade se tornou uma grande preocupação no processo de desenvolvimento de software devido às exigências de leis de privacidade, citando tanto a LGPD (Brasil) como a GDPR (UE). Os autores realizaram revisão da literatura e aplicação de um questionário para identificar as metodologias e técnicas utilizadas para a fase de elicitação de requisitos no contexto do desenvolvimento ágil de software. Os resultados dessa pesquisa apontam que as equipes ágeis estão cientes dos impactos provocados pelas leis de privacidade e que essas equipes já trabalham com alguns princípios de privacidade. Como principais desafios, os autores identificaram que as equipes enfrentam problemas com especificações de requisitos de software desatualizadas e falta de conhecimento das partes interessadas sobre privacidade de dados; que geralmente não usam as técnicas propostas na literatura para atender aos requisitos de privacidade. Esse é um passo inicial, no entanto, sente-se a falta de um estudo mais abrangente que possa elencar impactos causados por regulamentações de privacidade com outras metodologias de desenvolvimento de software.

Já (GEORGIADIS et al., 2022) aborda a questão da privacidade e violação de dados pessoais no contexto de grandes volumes de dados. Por isso, o objetivo dos autores é identificar riscos de privacidade e proteção de dados específicos do contexto de *Big Data Analytics* que poderiam impactar negativamente os direitos e liberdades dos usuários. Os resultados elencaram que nenhuma metodologia de avaliação do impacto na privacidade parece ser capaz de abordar todas as questões específicas de riscos de privacidade e de proteção de dados de grandes volumes de dados. Embora esse estudo tenha por objetivo avaliar riscos de privacidade e de proteção de dados, o contexto é muito específico e não há atividades de engenharia de software envolvidas. O estudo considera apenas os dados brutos em grandes volumes.

Em (CANEDO et al., 2022) os autores realizaram uma RSL para investigar e entender como as equipes de desenvolvimento de software que utilizam metodologias ágeis estão realizando a elicitação de requisitos de privacidade após a entrada em vigor da LGPD. Os autores realizaram também questionários e entrevistas semiestruturadas com integrantes de equipes ágeis para compreender o que as organizações modificaram no processo de desenvolvimento para atender aos requisitos de privacidade da LGPD. Os resultados dessa pesquisa identificaram que as áreas mais impactadas pela LGPD são as de Elicitação de Requisitos e de Implementação de Software; que mais de 50% das equipes ágeis não contam com nenhuma ferramenta para eliciar requisitos de privacidade, que a falta de conhecimento das equipes ainda é um desafio para identificar corretamente os princípios de privacidade e que nenhuma das organizações utiliza software ou guias para garantir a conformidade com a LGPD. Comparado a (CANEDO et al., 2021), esse estudo aborda mais profundamente impactos causados por uma regulação de privacidade que equipes ágeis elencam. No entanto, o estudo não aborda apenas impactos para outras metodologias ou atividades tradicionais de engenharia de software.

3.2 Considerações Finais e Comparação

As Tabelas 3.1 e 3.2 resumem as revisões da literatura existentes e fazem um comparativo com este trabalho. Com base nos dados apresentados na Tabela 3.1, é possível identificar o quão diferentes são os objetivos e as questões de pesquisa de cada estudo. A revisão sistemática apresentada neste trabalho visa compreender melhor os impactos causados por leis de privacidade às atividades de engenharia de software durante o processo de desenvolvimento, com questões de pesquisa que não foram abordadas na literatura.

Tabela 3.1 – Questões de Pesquisa das revisões de literatura relacionadas

Estudo	Questões de Pesquisa	Objetivos
(CANEDO et al., 2021)	(I) Como as equipes ágeis implementam o conceito de privacidade em seu trabalho diário? (II) Como as equipes ágeis interpretam o conceito de privacidade em relação à Lei Geral de Proteção de Dados Pessoais implementada em 2020? (III) Como as equipes ágeis percebem as ações que devem ser adotadas para reduzir o impacto da privacidade no desenvolvimento ágil de software?	investigar se as equipes ágeis interpretam os princípios de privacidade corretamente e se implementam esses conceitos e princípios durante o desenvolvimento de software;
(GEORGIADIS et al., 2022)	(I) Quais são os riscos específicos de privacidade e proteção de dados para Análise de Big Data? (II) Até que ponto as metodologias de Avaliação de Impacto na Privacidade que receberam atenção na literatura revisada cobrem os riscos identificados em (I)?	identificar riscos de privacidade e proteção de dados específicos do contexto de Análise de Big Data
(CANEDO et al., 2022)	(I) Quais práticas/técnicas e ferramentas são utilizadas por equipes ágeis para eliciar requisitos de privacidade que atendam à Lei Geral de Proteção de Dados do Brasil? (II) Quais ações/mudanças as organizações brasileiras têm feito para desenvolver software compatível com a LGPD?	investigar o nível de conhecimento das equipes ágeis sobre a LGPD e seus princípios e quais soluções de privacidade as equipes estão adotando
Este Trabalho	(I) Como as leis de privacidade e de proteção de dados impactam as atividades de engenharia de software? (II) Quais aspectos foram necessários alterar ou incluir ao processo para se garantir o cumprimento das leis de privacidade? (III) Uma vez identificados os processos que necessitaram ser alterados ou aprimorados, como a solução proposta foi implementada ou validada no contexto de desenvolvimento de software em produção? (IV) Quais desafios, problemas, limitações foram identificados em alguma atividade de engenharia de software, decorrente das leis de privacidade, que esteja sem solução?	investigar o impacto que leis de privacidade e de proteção de dados geram em uma ou mais atividades de engenharia de software para garantir a conformidade de sistemas com regulamentações dessa natureza

Fonte: o autor

A Tabela 3.2 complementa a anterior, com mais detalhes sobre as revisões da literatura: as bibliotecas digitais utilizadas para busca de estudos primários, as strings de busca e o número de estudos selecionados. É possível perceber que todos os estudos variam enormemente nas estratégias de busca e seleção, resultando em um conjunto completamente diferente de resultados primários.

Tabela 3.2 – Comparação das revisões de literatura relacionadas

Estudo	Fontes	String de Busca	Estudos selecionados
(CANEDO et al., 2021)	IEEE, DBLP, Scopus, ACM DL	(“requirements engineering” OR “requirements approach” OR “requirements methodology” OR “requirements process” AND (“elicitation” OR “requirements elicitation” OR “requirements specification” OR “requirements gathering” OR “requirements capture”) AND (“technique” OR “method” OR “tool”) AND (“agile software development” OR “agile development”)) AND (“privacy” or “security”))	23
(GEORGIADIS et al., 2022)	ACM, EBSCOhost, Science Direct, HeinOnline, IEEE Xplore, International Data Privacy Law, Web of Sci- ence, ProQuest, Scopus and Taylor&Francis Online	(‘privacy impact assessment’ OR pia OR ‘privacy impact state- ment’ OR ‘data protection impact assessment’ OR dpia OR ‘im- pact assessment’ OR ‘privacy risk assessment’ OR ‘privacy risk’ OR ‘privacy evaluation’ OR ‘data protection risk assessment’ OR ‘data protection risk’ OR ‘data protection evaluation’) AND ‘big data’	159
(CANEDO et al., 2022)	IEEE, DBLP, Scopus, ACM DL	((“requirements engineering” OR “requirements approach” OR “requirements methodology” OR “requirements process” AND (“elicitation” OR “requirements elicitation” OR “requirements specification” OR “requirements gathering” OR “requirements capture”) AND (“technique” OR “method” OR “tool”) AND (“agile software development” OR “agile development”) AND (“privacy” or “security”) or “ Brazilian General Data Protection Law” or “LGPD”)).	36
Este trabalho	IEEE, Science Direct, Springer, ACM	(“privacy by design” OR “data protection law” OR “privacy issues”) AND (“software engineering” OR “Software Systems”)	11

Fonte: o autor

A RSL que é apresentada neste trabalho é a única que investiga como as atividades de software são impactadas pelas leis de proteção de dados e de privacidade. Essa constatação justifica a necessidade do desenvolvimento de uma RSL como esta, que busca examinar não somente quais artefatos de software são afetados, como também quais as propostas para minimizar esse impacto e quais limitações são relatadas. Um aspecto importante que deve ser considerado é que a metodologia apresentada neste trabalho prezou por realizar um trabalho iterativo, sistemático e repetível, isto é, sem inclusão de trabalhos manuais ou outras formas de coleta de dados. (CANEDO et al., 2021; CANEDO et al., 2022) além da RSL, também aplicaram um questionário a um conjunto de participantes e realizaram entrevistas semi-esteuturadas. (GEORGIADIS et al., 2022), além dos estudos coletados pela metodologia de revisão sistemática, realizou a inclusão de estudos de maneira manual, o que afeta a reprodução da metodologia por terceiros futuramente. Portanto, considera-se que essa RSL complementa e inova esse panorama de revisões de literatura, no contexto de quais impactos são causados por leis de privacidade nas atividades de engenharia de software.

4 REVISÃO SISTEMÁTICA DA LITERATURA: DEFINIÇÃO DO PROTOCOLO

Neste capítulo, é apresentado o protocolo definido para a realização desta RSL A metodologia adotada segue as diretrizes de revisões sistemáticas da literatura propostas por (WOHLIN et al., 2012).

4.1 Objetivo e Questões de Pesquisa

O papel principal de uma RSL consiste em sumarizar todas as informações do estado da arte acerca de um determinado campo de pesquisa ou fenômeno, de forma completa e imparcial. Partindo dessa premissa, o objetivo deste trabalho é investigar o impacto que as leis de privacidade e de proteção de dados geram em uma ou mais atividades de engenharia de software para garantir a conformidade de sistemas com regulamentações dessa natureza. Também se deseja coletar os principais problemas detectados e as soluções propostas. De acordo com esse objetivo, foram definidas quatro questões de pesquisa (QP), apresentadas na tabela 4.1, com a respectiva motivação.

Tabela 4.1 – Questões de pesquisa e respectivas motivações

ID	Questão de Pesquisa (QP) e Motivação (M)
QP1	QP: Como as leis de privacidade e de proteção de dados impactam as atividades de engenharia de software? M: O objetivo desta questão é investigar quais necessidades de adaptações as leis de privacidade geraram durante as atividades do processo de desenvolvimento engenharia de software tradicional (elicitação de requisitos, implementação, testes, manutenção, suporte, etc)
QP2	QP: Quais aspectos foram necessários alterar ou incluir ao processo para se garantir o cumprimento das leis de privacidade? M: O objetivo desta questão é investigar quais aspectos ou artefatos precisaram de mudanças para se adequar às leis de privacidade durante o processo de desenvolvimento; se foi necessário adotar assinatura de novos contratos ou termos de confidencialidade.
QP3	QP: Uma vez identificados os processos que necessitaram ser alterados ou aprimorados, como a solução proposta foi implementada ou validada no contexto de desenvolvimento de software em produção? M: O objetivo desta questão é investigar, uma vez identificado os aspectos impactados pelas leis de privacidade, se alguma solução foi implementada ou, em caso negativo, se há alguma proposta de como implementá-la, de forma a adequar o processo de desenvolvimento de software às leis de privacidade.
QP4	QP: Quais desafios, problemas, limitações foram identificados em alguma atividade de engenharia de software, decorrente das leis de privacidade, que esteja sem solução? M: O objetivo desta questão é investigar se foram identificados problemas para atender às leis de privacidade, cujas soluções ainda permanecem pendentes.

Fonte: o autor

4.2 Estratégia de busca para estudos primários

A estratégia de busca de estudos primários é um processo que visa encontrar todos os estudos primários relevantes, sem obter um número esmagador de estudos falsos positivos, isto é, um resultado que é erroneamente positivo quando devia não ser.

Para cobrir toda a literatura de interesse, foram utilizadas bibliotecas digitais, conforme descrito na tabela 4.2. Estudos duplicados identificados foram removidos, mantendo-se a versão mais recente. Para esse processo, foi preciso definir e validar uma string de busca a partir da área a ser abordada e das questões de pesquisa. Em um primeiro momento, uma amostra de artigos foi levantada, o que ajudou a montar a primeira versão da string de busca. Após, foram aplicados refinamentos de forma iterada e de acordo com a qualidade e a quantidade de estudos resultantes das buscas nas bibliotecas digitais. A string de termos de busca final foi composta por (I) sinônimos para leis de privacidade ou leis de proteção de dados e (II) engenharia de software. Foram analisadas cada alteração dos termos de busca de forma a avaliar a obtenção de um conjunto adequado de estudos, verificando se os resultados incluíam artigos estabelecidos como "de controle". A figura 4.1 apresenta a string de busca final utilizada para execução deste trabalho.

Quando disponibilizado pela biblioteca digital que estava sendo consultada, foram aplicados filtros para as áreas: (a) privacidade de dados (ou *Data Privacy*); (b) proteção de dados (ou *Data Protection*); (c) engenharia de software (ou *Software Engineering*). Quando possível, também foi realizado um filtro para estudos da área de ciência da computação (ou C*omputer Science*).

Figura 4.1 – String de Busca final utilizada na busca de estudos primários em bibliotecas digitais

```
("privacy by design" OR "data protection law" OR "privacy
issues") AND ("software engineering" OR "Software Systems")
```

Fonte: o autor

Tabela 4.2 – Relação de Bibliotecas Digitais

ID	**Biblioteca Digital**	**URL**
ACM	ACM Digital Library	<http://dl.acm.org>
IEEE	IEEE Digital Library	<http://ieeexplore.ieee.org>
SD	Science Direct	<http://www.sciencedirect.com>
Springer	Springer	< https://link.springer.com/ >

Fonte: o autor

4.3 Seleção de Estudos

Para este trabalho, foram estabelecidos os seguintes critérios de inclusão: (I) estudos escritos em inglês ou português; (II) estudos que estabelecem relação entre leis de privacidade e/ou proteção de dados e atividades de engenharia de software; (III) estudos que relatam mudanças decorrentes de leis de privacidade e/ou proteção de dados.

Foram definidos, também, os seguintes critérios de exclusão: (I) publicações escritas em outros idiomas, que não sejam inglês ou português; (II) não ser um full-paper (isto é, são descartados *short-papers*, editoriais, livros, capítulos de livros, tutoriais, resumos, *worshops*, revisões, estudos secundários, teses, dissertações, pôsteres, comentários etc.); (III) artigos focados somente em leis de privacidade e/ou de proteção de dados, sem relação alguma com engenharia de software); (IV) estudos relacionados à privacidade de dados focados em outras áreas não relacionadas às atividades de engenharia de software; (V) estudos cujo texto completo esteja indisponível.

A seleção de estudos ocorreu em três etapas: (I) uma primeira seleção por análise do título, resumo e palavras-chave do estudo; (II) uma seleção a partir da leitura da introdução e da conclusão, uma leitura superficial do estudo e observação de figuras; (III) seleção final com base na leitura completa do estudo.

Não foi estabelecido um período para seleção de estudos para o protocolo dessa RSL, ou seja, qualquer estudo relacionado a leis de privacidade e à engenharia de software pode ser selecionado independente do ano de sua publicação. No entanto, a seleção de trabalhos foi realizada entre outubro e dezembro de 2022, assim, os estudos selecionados incluídos neste trabalho têm como período máximo o ano citado.

De forma a evitar o viés do autor, dois autores (orientando e orientador deste trabalho) realizaram a fase de triagem de forma independente, sem contato. Para que houvesse convergência na interpretação dos critérios de inclusão/exclusão, esses dois autores também fizeram individualmente a leitura todos os estudos selecionados. Em casos em que houve dúvida ou discordância, os autores conversaram pessoalmente e expuseram suas observações relacionadas ao devido estudo, chegando a um consenso de acordo com argumentos ou trechos do texto dos estudos que davam sustentação ao ponto de vista.

4.4 Avaliação da Qualidade

Em uma RSL é importante avaliar a qualidade dos estudos primários é importante, especialmente quando os estudos relatam resultados contraditórios. O protocolo de avaliação da qualidade adotado neste trabalho inclui identificar a contribuição individual de cada estudo

e avaliar se diferenças de qualidade ajudam a explicar os resultados. Foram estabelecidos nove critérios para avaliar cada estudo sob a ótica da qualidade metodológica e acerca da relação das leis de privacidade e/ou proteção de dados com as atividades de engenharia de software. As questões de qualidade são apresentadas na Tabela 4.3.

Tabela 4.3 – Critérios de avaliação de qualidade

ID	Questão de Qualidade
QQ1	Há uma declaração clara dos objetivos da pesquisa (DERMEVAL et al., 2016)?
QQ2	O problema a ser resolvido pela técnica/método/abordagem/framework está claramente explicado (TIWARI et al., 2015)?
QQ03	Há discussão suficiente de trabalhos relacionados (TIWARI et al., 2015)?
QQ4	A técnica/método/abordagem/framework proposto está claramente descrito (DERMEVAL et al., 2016)?
QQ5	Existe descrição adequada do contexto (indústria, ambiente laboratorial, produtos utilizados etc.) em que a pesquisa foi realizada (DERMEVAL et al., 2016)?
QQ6	Existe uma discussão sobre os resultados do estudo (DERMEVAL et al., 2016)?
QQ7	As limitações deste estudo são explicitamente discutidas (DERMEVAL et al., 2016)?
QQ8	A relevância mais ampla do trabalho é discutida (TIWARI et al., 2015)?
QQ9	O estudo aumenta o conhecimento sobre o impacto das leis de privacidade e/ou de proteção de dados em atividades de engenharia de software?

Fonte: o autor

Cada questão de qualidade (QQ) foi avaliada e respondida em relação a três possíveis cenários: “Sim” (pontuação = 1), “Parcialmente” (pontuação = 0.5) e “Não” (pontuação = 0). Assim, o escore de qualidade de cada estudo é calculado pela soma dos escores das respostas às questões e pode chegar a até 9 pontos. Com base nesse valor, definimos um limite mínimo de qualidade de 5 pontos (ou seja, mínimo de cinco). Caso este valor não seja atingido, o estudo é descartado.

4.5 Extração e Monitoramento de Dados

A estratégia de extração de dados tem por objetivo auxiliar a responder às questões de pesquisa, permitindo que os pesquisadores possam resumir e categorizar os estudos primários, melhorando assim a compreensão do domínio. Os dados extraídos de cada artigo selecionado foram incluídos em um formulário com os campos predefinidos descritos na Tabela 4.4. Visando facilitar a análise dos dados, foram utilizadas planilhas para tabular os dados extraídos, resumir os resultados e gerar tabelas e gráficos.

Tabela 4.4 – Formulário de Extração de Dados

#	Dados do Estudo	Descrição	QP relacionada
1	Identificador	Identificador único para cada estudo	visão geral
2	Ano, Autores, País	Ano de publicação, nome dos autores, instituições de ensino, e respectivos países	visão geral
3	Local Científico	Periódico ou Conferência	visão geral
4	Tipo de Instituição	Universidade, Instituição, Indústria	visão geral
5	Atividade(s) de ES	Qual(is) atividade(s) de Engenharia de software é(são) abordada(s)	QP1
6	Impacto / Problema	Qual o impacto ou problema foi detectado para atender às leis de privacidade e/ou de proteção de dados	QP1
7	Solução Proposta	Quais mudanças foram necessárias para atender à legislação	QP2
8	Implementação / Validação	A solução proposta chegou a ser validada, implementada ou testada em produção?	QP3
9	Limitação Existente	Descrição de limitações que não puderam ser resolvidas para atender à legislação	QP4

Fonte: o autor

4.6 Ameaças à validade do estudo

Neste trabalho teve-se uma preocupação com a sua validade. Uma das ameaças identificadas ao estudo é a string de busca que foi definida. Por essa razão, de forma a minimizar os riscos de não se coletar o máximo de estudos possíveis, foram feitos pequenos exercícios e testes até se obter a string de busca apresentada na Figura 4.1. Outra ameaça que pode ser considerada é a avaliação da qualidade dos estudos. Embora a definição de qualidade possa ser subjetiva, a avaliação da qualidade é uma ferramenta essencial para minimizar vieses e erros sistemáticos quanto à validade interna e externa. Cada estudo precisa ser avaliado de acordo com a lista de verificação do protocolo e de acordo com a motivação de tal avaliação. Por fim, outra ameaça à validade deste trabalho a ser considerado é o número de autores que realizaram a condução do processo de seleção dos estudos. Esse processo foi realizado somente por dois autores devido a eles estarem envolvidos com esta pesquisa. No entanto, houve poucas situações de divergência durante a análise dos estudos selecionados, e, quando houve, essa divergência pôde ser resolvida com uma leitura em pares, considerando o conteúdo apresentado no texto dos próprios estudos.

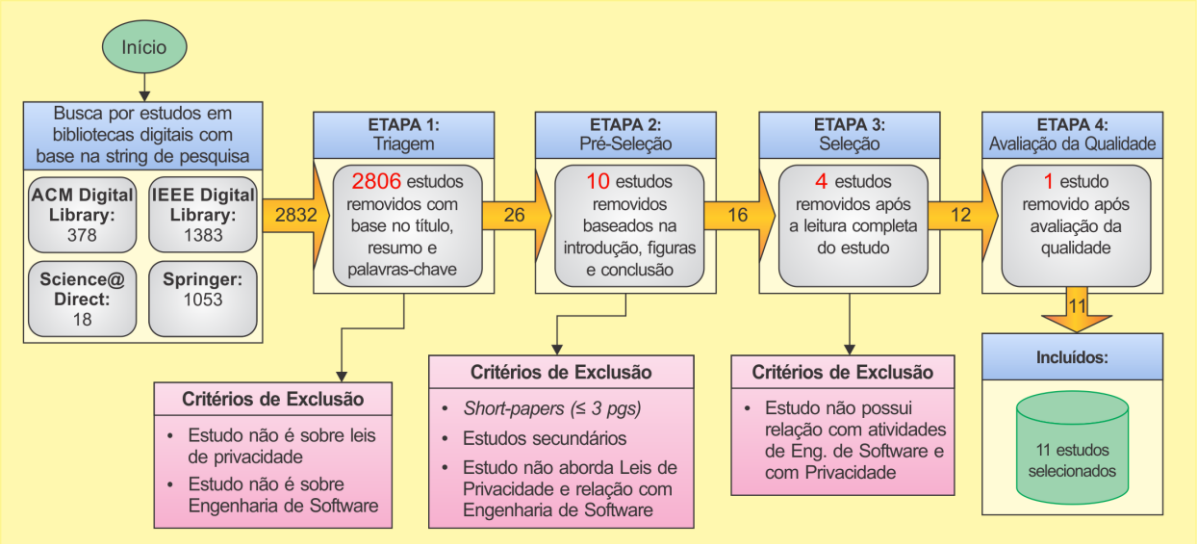
5 RESULTADOS

Neste capítulo, são apresentados os resultados e achados obtidos de acordo com as questões de pesquisa definidas para esta revisão sistemática.

5.1 Seleção de Estudos

A Figura 5.1 descreve o fluxo completo de como os estudos primários para este trabalho foram selecionados. A busca nas quatro bibliotecas digitais utilizadas identificou 2.832 estudos. Com base no critério de exclusão (IV), foram removidos 2.806 estudos por não estarem relacionados diretamente à privacidade de dados e à engenharia de software. Estes estudos descartados estavam relacionados à privacidade de dados em outras áreas de pesquisa, como: inteligência artificial; drones; aprendizado de máquina; comunicação veicular; *blockchain*, *cloud computing*, internet das coisas e redes wireless. Após, pelo segundo critério de exclusão definido, um estudo foi removido por se tratar de um *short-paper*, e outros dois por se tratar de estudos secundários. Com base na leitura do título, do resumo, das figuras e da conclusão, outros sete artigos foram descartados pelo critério de exclusão (III) por focarem somente em leis de privacidade, mas sem vínculo com engenharia de software. Por fim, após a leitura completa dos estudos e pelo mesmo critério de exclusão, foram descartados outros quatro estudos, relacionados com privacidade de dados, no entanto, com foco em outras aplicações, sem vínculo com as atividades de engenharia de software. A partir da leitura completa, foram selecionados 12 estudos primários, que foram submetidos à avaliação de qualidade.

Figura 5.1 – Seleção de estudos e avaliação da qualidade



Fonte: o autor

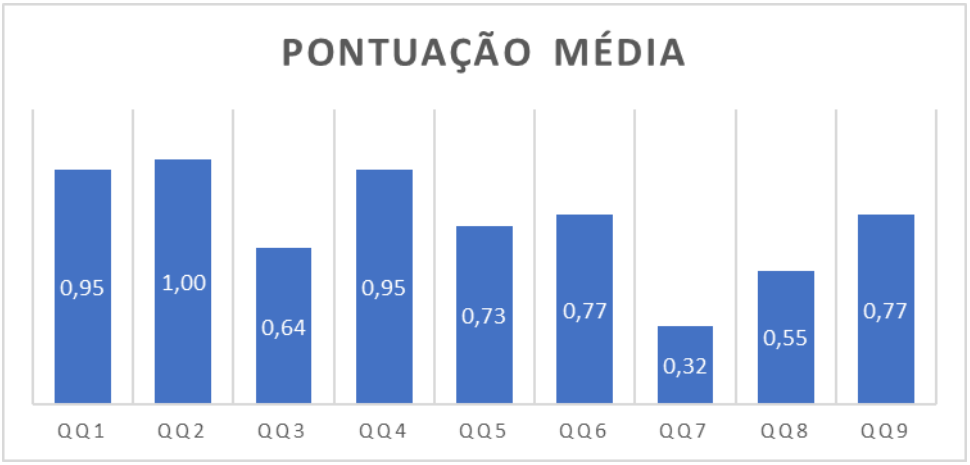
5.2 Avaliação da Qualidade

Os 12 estudos primários selecionados foram submetidos à avaliação da Qualidade, com base nos critérios definidos na tabela 4.3. As pontuações de qualidade atribuídas para cada estudo estão disponíveis no Apêndice A, tabela A-1.

Com base na avaliação realizada, foi removido mais um estudo, tendo em vista que não apresentava relevância e completude para o tema de pesquisa deste trabalho. Assim, o número final de estudos incluídos foi 11 (onze).

A pontuação média dos 11 estudos restantes foi 6,69. A Figura 5.2 apresenta detalhes de média de pontuação para cada questão de qualidade.

Figura 5.2 – Pontuação média de cada questão de qualidade



Fonte: o autor

A avaliação da qualidade permitiu identificar que três critérios (QQ3, QQ7, QQ8) apresentaram a menor pontuação em relação aos demais. A pontuação média do critério QQ3 foi de apenas 0,58, o que demonstra que incluir aspectos de leis de privacidade nas atividades de engenharia de software ainda não é um tópico maduro, uma vez que, através dessa questão, o objetivo era avaliar se os estudos apresentam discussão acerca de trabalhos relacionados suficientes, e se há uma comparação sobre técnicas diferentes adotadas. A pontuação média do critério QQ7 foi de apenas 0,32, sendo a pontuação mais baixa de todos os critérios. Através desta questão, se constatou que os estudos discutem não discutem explicitamente suas limitações. Esse resultado demonstra não só a falta de domínio do problema, como também a confiabilidade das soluções propostas. Por fim, a pontuação média do critério QQ8 foi de apenas 0,5. Através dessa questão, tínhamos por objetivo avaliar os resultados obtidos pelos estudos e o impacto da solução proposta para o problema detectado. No entanto, muitos estudos não chegaram a ter uma validade da solução, implementaram pequenos projetos

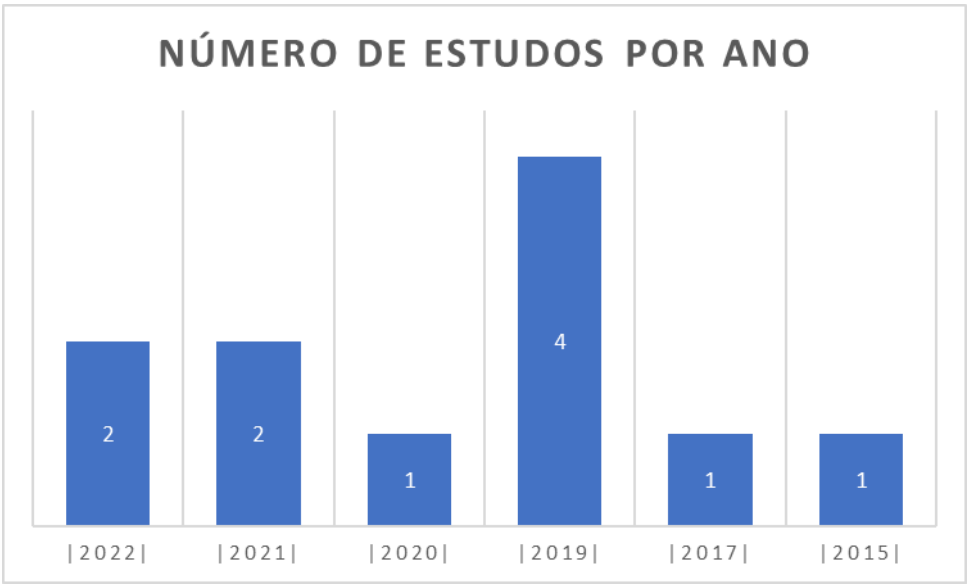
pilotos ou apenas especificaram quais as implementações precisariam ser feitas, sem realizá-las de fato.

5.3 Extração dos Resultados

Foram extraídos os dados dos 11 estudos selecionados, conforme o protocolo definido na tabela 4.4. Esses são apresentados na tabela 5.2. Antes de discutir os resultados para cada questão de pesquisa, são apresentados alguns dados estatísticos dos estudos selecionados.

A Figura 5.3 apresenta a distribuição das publicações por ano. Pode-se observar que o número de publicações que abordam o impacto das leis de privacidade nas atividades de engenharia de software aumenta a partir do ano de 2019. A hipótese mais provável para isso é que a partir do final de 2018 a GDPR entrou em vigor na Europa e então academia e indústria precisaram avaliar como adequar processos de desenvolvimento de software para atender às leis de privacidade.

Figura 5.3 – Número de estudos por ano



Fonte: o autor

A Figura 5.4 apresenta a frequência de autores que desenvolvem pesquisas nesta área por país, baseado nas organizações envolvidas, conforme os estudos selecionados. Foram identificados 13 países no total, sendo que os mais ativos são Brasil, Bélgica e Espanha. Esses países possuem ou fazem parte de um bloco que possuem leis de proteção de dados e/ou de privacidade em vigor, o que auxilia a explicar esses resultados.

Figura 5.4 – Número de pesquisadores por país

Fonte: o autor

Os estudos selecionados também foram categorizados por tipo de veículo científico e tipo de instituição, conforme a Tabela 5.1. Os resultados demonstram que a maioria dos estudos foram realizados pela academia (72,7%), ao passo que o restante foi através de parcerias entre academia e indústria. Não houve a seleção de estudos efetuados apenas pela indústria. Pode-se perceber que, dado o maior número de estudos acadêmicos, que este ainda é um tema com baixa maturidade, além de ainda ser pouco abordado em implementações de novos sistemas e em softwares em produção na indústria.

Tabela 5.1 – Estudos agrupados com base no veículo científico e tipo de instituição

Veículo Científico	**Academia**	**Indústria**	**Ambas**
Periódicos	2	0	0
Conferências	6	0	3

Fonte: o autor

Tabela 5.2 – Estudos Selecionados

Referência	Atividade de ES	Impacto / Problema	Solução Proposta	Implementação ou Validação	Desafio, Problema, Limitação
(CAMPANILE et al., 2022)	Especificação	projeto, manutenção	Novo ciclo de desenvolvimento com ênfase a testes de privacidade	-	-
(OLUKOYA et al., 2022)	Requisitos	Extração sistemática de requisitos de acordo com a lei	Metodologias para obter requisitos	Estudo de Caso	Dificuldade em atingir consistência, integridade e utilidade; leis podem evoluir rapidamente e inviabilizar métodos.
(MENDES et al., 2021)	Requisitos	Baixo número de engenheiros de software com conhecimento de leis de privacidade; Interpretação de textos legais, ambiguidade, definições específicas de domínio	Checklist de inspeção e adequação de sistemas	Prova de Conceito	-
(TORRE et al., 2021)	Requisitos	Auditorias manuais e caras para garantir conformidade com a GDPR	Modelo UML para projetar metodologias automatizadas para verificar conformidade com GDPR.	-	Generalização, Resiliência, Extensibilidade
(VALENÇA et al., 2020)	-	Garantir a conformidade com a GDPR em ecossistemas de software	Funções de proteção de dados em ecossistemas de software	-	Coleta, controle, consentimento, transparência, adequação à diferentes legislações.
(SILVA et al., 2019)	-	Violações de dados em aplicativos web descentralizados	Framework Esfinge Guardian	-	-
(SION et al., 2019)	-	Dicotomia entre raciocínio legal e abordagens de engenharia de software; atividades são executadas de forma isolada.	Uma arquitetura de software de uma perspectiva de proteção de dados, vinculada aos Diagramas de Fluxo de Dados	Protótipo	-
(NETTO et al., 2019)	Requisitos	Especificações de requisitos legais são inerentemente ambíguos	Através de entrevistas, detectar como a indústria trata a ambiguidade na especificação de requisitos e como é feita a conformidade legal	-	-
(HJERPPE et al., 2019)	Requisitos	Implicações práticas da GDPR para engenharia de requisitos	Elaboração de restrições práticas, extração de requisitos a partir da GDPR	Estudo de caso	Generalização
(MOUGIAKOU et al., 2017)	Requisitos	Falta de teoria legislativa para implementação de requisitos de privacidade	Analisar privacidade de informações por meio de diagramas de casos de uso UML	-	-
(NOTARIO et al., 2015)	Requisitos	Dificuldade de compreensão, por parte dos engenheiros de software, de como adequar sistemas às leis de privacidade	Metodologia sistemática para engenharia de privacidade	-	-

Fonte: o autor

5.4 QP1: Como as leis de privacidade e de proteção de dados impactam as atividades de engenharia de software?

A maioria dos estudos citam impactos na fase de especificação e/ou de requisitos, (CAMPANILE et al., 2022; OLUKOYA et al., 2022; MENDES et al., 2021; TORRE et al., 2021; NETTO et al., 2019; HJERPPE et al., 2019; MOUGIAKOU et al., 2017; NOTARIO et al., 2015). (NETTO et al., 2019) cita o problema de, geralmente, os requisitos de software serem especificados em linguagem natural, e que muitas vezes essas especificações possuem ambiguidade, trazendo desafios para a Engenharia de Requisitos. Citam também que esses problemas se tornam maiores quando se trata de requisitos baseados em leis, visto que estas possuem trechos que podem ser interpretados com mais de um sentido. Isso também foi

relatado por (MOUGIAKOU et al., 2017), que detecta que falta uma teoria legislativa para implementação de requisitos de privacidade. (CAMPANILE et al., 2022) observa que a regulamentação de privacidade teve um impacto significativo no modo como software, sistemas e empresas devem ser projetados ou mantidos para cumprir regras e que, inicialmente, questões de privacidade eram consideradas requisitos externos (não funcionais), onde a indústria operava de forma reativa para atendê-los, adaptando artefatos de software existentes. Na sequência, com base na experiência adquirida, se passou a fazer auditorias internas para se gerenciar questões relacionadas à privacidade. (TORRE et al., 2021) cita o fato de essas auditorias serem custosas e manuais, e que, no momento, não há uma solução automatizada, de aplicabilidade e escala industrial, para substituir essa situação. Em (OLUKOYA et al., 2022), os autores contextualizam que metodologias e ferramentas foram propostas para lidar com requisitos de privacidade, mas pouco se estudou sobre isso fora da região da UE e dos EUA. Em (NOTARIO et al., 2015), os autores diagnosticam que engenheiros de software possuem dificuldade em compreender o que deve ser feito para que seus sistemas estejam de acordo com leis de privacidade desde a fase de projeto, o que dificulta a adoção de práticas da engenharia de privacidade. Corroborando com isso, no contexto de Brasil, (MENDES et al., 2021) menciona que, embora a LGPD tenha se tornado aplicável em 2020, várias equipes de desenvolvimento de software ainda não possuem conhecimento sobre quais atributos de qualidade são necessários para que um sistema esteja em conformidade com a referida lei e evite penalidades legais e monetárias. Por fim, com viés industrial, (HJERPPE et al., 2019) considera que pouca pesquisa foi realizada para entender melhor as implicações práticas do GDPR para engenharia de requisitos e arquiteturas de software.

Outros trabalhos não citam uma atividade específica de engenharia de software (VALENÇA et al., 2020; SILVA et al., 2019; SION et al., 2019). Nestes estudos, os autores demonstraram uma preocupação mais ampla do impacto das leis de privacidade, não só na fase de requisitos, como toda a vida útil do software. Em (VALENÇA et al., 2020), os autores tratam de ecossistemas de software, que consistem em vários negócios e atividades econômicas atuarem em conjunto, como uma unidade, e interagirem como um mercado compartilhado de softwares e serviços e, nesse contexto, consideram que a privacidade é um ponto crítico, visto que os dados dos usuários podem ser acessados por desenvolvedores terceiros e porque pode haver vazamento de dados pessoais. Em (SILVA et al., 2019), os autores abordam a autorização de acesso a dados pessoais em aplicativos web descentralizados e citam uma plataforma em que os usuários são os únicos responsáveis pelo

controle de acesso aos seus dados, o que, no entanto, não garante o cumprimento da LGPD, tendo em vista que é preciso que, pela lei, as empresas também devem proteger o acesso aos dados dos usuários. O principal problema relatado é que, mesmo que um usuário marque uma aplicação como confiável, ele não teria como impedir o acesso não autorizado aos seus dados pessoais. Os autores do estudo (SION et al., 2019) detectaram que há uma bipartição entre o raciocínio legal, conduzido em avaliações de impacto, e as atividades de engenharia de software, responsável pela modelagem de ameaças identificadas, fazendo com que essas atividades, geralmente, sejam executadas sem conexão uma com a outra, afetando especialmente o exercício da conformidade e a capacidade de evolução de um sistema. Algo na mesma direção dos estudos (MENDES et al., 2021; NOTARIO et al., 2015).

A tabela 5.3 a seguir apresenta um quadro resumo das principais respostas referentes à questão de pesquisa deste capítulo:

Tabela 5.3 – Quadro Resumo referente à QP1

ID	Respostas
1º	72% dos estudos citam impactos na fase de especificação de requisitos.
2º	Requisitos especificados em linguagem natural possuem ambiguidade. Esta, por sua vez, aumenta quando requisitos são baseados em textos legais.
3º	O processo de lidar com questões de leis de proteção de dados, em geral, possui três fases de maturidade: Fase 1: Reativa – onde se adapta sistemas existentes. Fase 2: Auditorias – são custosas e manuais. Falta de solução automatizada. Fase 3: Privacy by Design (PbD) – requisitos de privacidade devem ser integrados desde as primeiras fases de projeto.
4º	Há uma dicotomia entre o raciocínio legal e as atividades de software.
5º	Ecossistemas de software possuem grandes desafios, como dados de usuários, que podem ser acessados sem consentimento por desenvolvedores terceiros; pode haver haver vazamento de dados.

Fonte: o autor

5.5 QP2: Quais aspectos foram necessários alterar ou incluir ao processo para se garantir o cumprimento das leis de privacidade?

Os estudos (TORRE et al., 2021; SION et al., 2019; MOUGIAKOU et al., 2017) propõem soluções utilizando UML. Em (TORRE et al., 2021), os autores, em conjunto com especialistas jurídicos, elaboraram um modelo utilizando UML e OCL para representar os principais conceitos e relacionamentos da GDPR, isto é, que não possuem relação com leis nacionais ou jurisprudência de países membros da UE. (MOUGIAKOU et al., 2017) utiliza diagramas de casos de uso UML para modelar a estrutura arquitetônica de um software de forma a tratar a privacidade desde a fase de projeto, estabelecendo atores e requisitos de privacidade e inventário de dados. (SION et al., 2019) propõe a criação de uma arquitetura de

software, a partir de uma perspectiva de proteção de dados utilizando Diagrama de Fluxo de Dados.

Em (OLUKOYA et al., 2022), os autores propõem avaliar metodologias visando obter requisitos de privacidade e segurança de leis e de regulamentos, a partir de métricas de medição, como consistência, abrangência, utilidade e evolução na legislação de privacidade. Para isso, os autores utilizam trabalhos anteriores revisados, onde estruturas codificam regras e artigos de textos legais em requisitos de software, extraindo direitos de acesso, obrigações e restrições Segundo os autores, a partir disso, as organizações podem ser auxiliadas em seus programas de conformidade com diferentes regulamentações, visto que um software desenvolvido na União Europeia que precise operar em outro país, precisa atender tanto à GDPR como às leis de privacidade deste outro país. Em (NOTARIO et al., 2015), os autores propõem uma metodologia sistemática, visando integrar a engenharia de privacidade com as melhores práticas existentes para os estágios de análise e de ciclo de vida de desenvolvimento de um sistema, potencializando as melhores práticas de privacidade complementares disponíveis. Em (MENDES et al., 2021), os autores propõem o uso de um *checklist* de inspeção, visando avaliar a aderência de sistemas à lei brasileira de privacidade. Esse *checklist* foi elaborado a partir da transformação de requisitos legais em requisitos técnicos, onde foram analisados os textos da LGPD e de artigos científicos relacionados ao tema. Por fim, a seleção dos melhores descritores da lei brasileira, dos artigos e da revisão informal foram transformados em 52 itens presentes no *checklist*. (CAMPANILE et al., 2022) propõe um novo ciclo de desenvolvimento, onde seja dada uma atenção especial à fase de testes de privacidade. Os autores também elaboram orientações para que processos de desenvolvimento de software possam ser adaptados para essa finalidade.

Focado no ambiente industrial, (HJERPPE et al., 2019) propõe nove restrições práticas com as quais pequenas e médias empresas precisam lidar para atender a restrições legais. Também extrai requisitos a partir da GDPR que precisam ser atendidos para propor uma arquitetura de software que atenda tanto às restrições práticas levantadas como os requisitos legais elencados para o estudo de caso analisado. Focado em ecossistemas de software, (VALENÇA et al., 2020) analisa funções para promover a proteção de dados e a privacidade dos usuários, que são o ativo mais importante dessas plataformas, segundo os autores. Ainda, com base nessa análise e na especificação da GDPR, elencam uma série de desafios e restrições visando a conformidade com a privacidade. Para tentar resolver o problema de acesso irrestrito aos dados pessoais de um usuário que marcou uma aplicação web descentralizada como confiável, (SILVA et al., 2019) propõe utilizar um framework

denominado Esfinge Guardian, cujo papel é interceptar chamadas para operações protegidas, visando é aumentar a conformidade com os requisitos de governança de dados da LGPD. Ainda, segundo os autores, com o uso deste framework é possível anonimizar dados pessoais, filtrando informações que possam levar à identificação dos usuários.

Em (NETTO et al., 2019), a partir da metodologia de realizar entrevistas semiestruturadas, os autores citam que não foi possível identificar uma metodologia sistemática para detectar e resolver a ambiguidade em requisitos baseados em leis. Também não foi possível detectar um processo para garantir que requisitos estejam em conformidade com leis de privacidade.

A tabela 5.4 a seguir apresenta um quadro resumo das principais respostas referentes à questão de pesquisa deste capítulo:

Tabela 5.4 – Quadro Resumo referente à QP2

ID	Respostas
1º	São feitas propostas utilizando modelagem UML, especialmente casos de uso e diagramas de fluxo de dados.
2º	Checklist de inspeção para avaliar a aderência de sistemas às leis de privacidade.
3º	Atenção à fase de testes – para que sejam incluídos testes referentes à privacidade.
4º	Na indústria, são adotadas restrições práticas para pequenas e médias empresas; Processos para anomização de dados e filtragem de informações.

Fonte: o autor

5.6 QP3: Uma vez identificados os processos que necessitaram ser alterados ou aprimorados, a solução proposta chegou a ser implementada ou validada no contexto de desenvolvimento de software em produção?

Poucos estudos citam uma etapa de validação ou de implementação de suas propostas. Com base nos estudos selecionados para este trabalho, foi possível verificar que somente o estudo (SION et al., 2019) produziu um protótipo. Os autores desse estudo validaram a proposta no contexto de um sistema de tratamento de doenças cardiovasculares, através da implementação do metamodelo e do ponto de vista de proteção de dados alinhado ao diagrama de fluxo de dados propostos. O protótipo destaca quais tipos de dados são usados em quais fluxos de dados, e documenta o envolvimento do titular dos dados com o médico (como destinatário legal). Segundo os autores, as correspondências entre essas duas visões permitem uma ampla variedade de verificações de conformidade e mecanismos de *feedback* para garantir consistência e conformidade técnica com os requisitos impostos pela lei de proteção de dados (GDPR).

Para validar a sua proposta, (HJERPPE et al., 2019) realizou um estudo de caso, uma empresa bem estabelecida no ramo de desenvolvimento de software, onde foram extraídos nove requisitos, agrupados em cinco categorias: privacidade e segurança; minimização de dados; direitos dos usuários; capacidade de rastreamento; e capacidade de rastreamento. A partir desse levantamento, os requisitos foram endereçados para mudanças na arquitetura de software. Segundo os autores, as principais mudanças necessárias foram em relação a interfaces de usuário, isolamento de dados pessoais, mecanismos de controle de acesso, pseudoanonimização, registro e anotações. O estudo de caso apresentado por (OLUKOYA et al., 2022) tem por objetivo avaliar a proposta para a lei de privacidade da Nigéria (NDPR). Os autores modelaram e analisaram a NDPR, e transformaram artigos da lei de privacidade nigeriana em requisitos. No entanto, segundo os autores, elaborar estruturas com diferentes leis de segurança da informação, que lidam com artefatos que são tecnicamente diferentes, gerenciados por diferentes organizações e partes interessadas, é um desafio que impacta a generalização almejada.

(MENDES et al., 2021) realiza uma prova de conceito para o *checklist* proposto, aplicando-o durante o uso de um portal de uma instituição federal brasileira. No resultado, foi identificado que 23 atributos foram feridos ou não foram atendidos. Não foi possível avaliar 13 itens, uma vez que, para realizar a sua análise, seria preciso possuir um alto nível de acesso ao sistema da instituição. Conforme concluem os autores, o restante dos itens do *checklist* não foram violados.

A tabela 5.5 a seguir apresenta um quadro resu pseudoanonimização mo das principais respostas referentes à questão de pesquisa deste capítulo:

Tabela 5.5 – Quadro Resumo referente à QP3

ID	Respostas
1º	Poucos estudos citaram uma fase de validação ou de implementação.
2º	Apenas um protótipo – um sistema de tratamento de doenças cardiovasculares
3º	Dois estudos de caso – NDPR e indústria
4º	Uma prova de conceito – checklist.

Fonte: o autor

5.7 QP4: Existe algum desafio, problema ou limitação em alguma atividade de engenharia de software, decorrente das leis de privacidade, que esteja sem solução?

Poucos estudos dedicaram uma seção ou capítulo para abordar limitações dos estudos. (HJERPPE et al., 2019) cita que o enquadramento explícito da GDPR para as pequenas e médias empresas limita a generalização, e que o enquadramento para arquiteturas de software aumenta ainda mais essa limitação. Os autores citam desafios ainda para

pseudoanonimização, análise estática, software distribuído e rastreamento por todo ciclo de vida do software. (TORRE et al., 2021) também menciona o aspecto da generalização como uma limitação, tendo em vista que a abordagem construída no estudo para a GDPR pode não ser aplicada a outras leis de privacidade. (TORRE et al., 2021) trata como fatores limitantes, ainda, a questão da extensibilidade – futuros trabalhos poderão considerar aspectos e detalhes que não foram considerados neste estudo – e da resiliência, que aborda o fato de que, se houver mudanças na legislação da GDPR, como inclusão de novas disposições, será necessário evoluir e alinhar também o modelo proposto. (OLUKOYA et al., 2022) elenca, como desafios do estudo, que é difícil atingir consistência, integridade e utilidade das abordagens propostas. Embora a mesma metodologia fosse aplicada para diferentes atores-chave de TI envolvidos no ciclo de vida de um software, o número de direitos, obrigações e restrições não foi o mesmo para todos, o que afeta a consistência. Os autores também citam que até os *frameworks* mais recentes não incluem a avaliação por engenheiros de privacidade. Assim como elencado por (TORRE et al., 2021), outra limitação, é a rápida mudança ou alteração da legislação de privacidade, e que as estruturas propostas devem considerar essa evolução. (VALENÇA et al., 2020) elenca sete desafios para que ecossistemas de software possam estar em conformidade com leis e regulamentações de privacidade, sendo os mais relevantes o aspecto de identificar quais dados pessoais são coletados dentro do ecossistema; garantir que apenas os dados realmente necessários são coletados; garantir que empresas do ecossistema tenham acesso somente aos dados aos quais elas têm (legalmente) direito; gerenciar o consentimento do usuário em todo o ecossistema, com seus diferentes atores e artefatos; assegurar a transparência relativa ao tratamento realizado; e garantir que os requisitos legais de diferentes legislações possam ser atendidos.

A tabela 5.6 a seguir apresenta um quadro resumo das principais respostas referentes à questão de pesquisa deste capítulo:

Tabela 5.6 – Quadro Resumo referente à QP4

ID	Respostas
1º	Poucos estudos citaram limitações.
2º	Generalização – Estruturas para GDPR pode não ser aplicáveis a outras leis de privacidade
3º	Extensibilidade e Resiliência – estruturas e modelos podem precisar evoluir de acordo com mudanças na lei (que ocorrem rapidamente).
4º	Uma prova de conceito – checklist.

Fonte: o autor

6 DISCUSSÃO

No capítulo anterior, foram resumidas as principais contribuições dos trabalhos selecionados que tratam do impacto das leis de privacidade nas atividades de engenharia de software e que propõe soluções para um ou mais problemas identificados. Este capítulo resume as conclusões relativas a estas questões e discute as tendências e oportunidades identificadas durante este trabalho.

6.1 Impactos gerados no desenvolvimento de software, devido a leis de privacidade

Considerando a alta quantidade de acesso indevido aos dados pessoais de usuários, é de suma importância que leis de privacidade sejam estabelecidas. No entanto, ao passo que leis desta natureza são promulgadas em determinados países ou blocos econômicos, é preciso que sistemas e aplicações, desde os mais simples, aos mais complexos, atendam a essas legislações, sob risco de receber penalidades não só financeiras como também de operação de serviços (TORRE et al., 2021). Por essa razão, é preciso garantir que os softwares em operação estejam de acordo com a legislação de privacidade vigente no seu país de desenvolvimento e/ou no seu país de operação, conforme avaliado por (OLUKOYA et al., 2022).

Ao passo que é preciso analisar sistemas visando garantir sua conformidade, o primeiro problema que é detectado é o fato de engenheiros de software não terem familiaridade com textos jurídicos, a fim de extrair requisitos de privacidade. Há profissionais que, através de entrevistas, expressam que sequer tiveram contato ou ouviram falar de leis de privacidade, como GDPR ou LGPD (NETTO et al., 2019), e desconhecem a sua importância não só para o setor de TI, mas para a empresa como um todo. (NOTARIO et al., 2015) reconhece que desenvolvedores de software possuem dificuldade em compreender o que deve ser ajustado para que softwares estejam em conformidade com à privacidade desde o a fase de projeto. Na mesma direção, (CAMPANILE et al., 2022) propõe um novo ciclo de desenvolvimento com ênfase em testes de privacidade, mas destaca que o cenário atual está longe do desejado para implementar tais soluções, visto que os testes são realizados por pessoas que não recebem treinamentos específicos e que também não possuem experiência com privacidade. (TORRE et al., 2021) sugere que as equipes de desenvolvimento necessitam de apoio de pessoas especializadas em interpretar textos legais para que os requisitos de privacidade possam ser extraídos corretamente. No mesmo sentido, (SION et al., 2019) percebe que existe uma separação entre o raciocínio legal e as atividades de engenharia de

software, que deveriam ser realizadas de forma integrada, mas acabam sendo executadas de forma isolada. (MENDES et al., 2021; NETTO et al., 2019; NOTARIO et al., 2015) ainda elencam o problema da ambiguidade existente em trechos de leis de privacidade – o que corrobora com a necessidade desses especialistas –, e que somente uma empresa, onde um entrevistado do estudo (NETTO et al., 2019) trabalha, possui um departamento especializado em ambiguidade. Mais além, (MOUGIAKOU et al., 2017) ainda sugere uma integração Jurídica-TI, que possa traduzir a crescente pesquisa do assunto em soluções práticas, sugerindo padrões, paradigmas e ferramentas.

Outro aspecto crítico, que pode ser percebido a partir dos estudos selecionados, é a questão da generalização das propostas (TORRE et al., 2021; HJERPPE et al., 2019), onde as soluções elencadas podem não ser válidas para toda e qualquer lei de privacidade, isto é, arquiteturas ou modelos sugeridos por estudos, com base na legislação de privacidade de um país, podem não ser aplicáveis a leis de outra nação, em virtude de trechos específicos das leis, ou de políticas sobre como a privacidade é tratada. Há, por essa razão, um problema de se implementar em larga escala as soluções propostas. (TORRE et al., 2021) cita que a evolução de software também é um dos itens impactados pelas leis de privacidade, uma vez que, os modelos e requisitos que representam trechos de regulações podem perder a eficácia ou sentido, caso leis de privacidade alterem artigos ou incluam novas estruturas. Além disso, sistemas podem deixar de estar em conformidade com legislações de privacidade, caso as organizações desenvolvedoras de software não estejam atentas a modificações na legislação. Por isso é importante, como citado anteriormente, que haja pessoas integradas à organização, com experiência jurídica, capazes de detectar essas modificações, a fim de manter as aplicações desenvolvidas em conformidade com as leis de privacidade.

A tabela 6.1 a seguir apresenta um quadro resumo das principais descobertas (ou achados) deste estudo:

Tabela 6.1 – Quadro Resumo com principais descobertas (achados)

ID	Descobertas
1º	É preciso que sistemas atendam às disposições de leis de privacidade, sob risco de receber sanções financeiras e de operação; Os softwares desenvolvidos precisam também atender à legislação do país de operação.
2º	Há uma baixa familiaridade com textos jurídicos por parte dos engenheiros de software, o que resulta em uma dificuldade de compreender o que deve ser feito para que sistemas estejam em conformidade com leis de privacidade vigentes.
3º	Especialista jurídico – Pessoas com experiência em textos legais podem auxiliar na redução de ambiguidade e na especificação de requisitos de privacidade.
4º	Testes de privacidade ainda estão precários, visto que são realizados por pessoas que não possuem treinamentos específicos ou experiência com privacidade.

Fonte: o autor

6.2 Tendências e oportunidades

Sendo o estudo mais recente incluído, (CAMPANILE et al., 2022) relata que as empresas desenvolvedoras de software se viram provocadas a incluir proativamente garantias adequadas conforme leis e regulamentações em seus produtos, e sintetiza três fases pelas quais as empresas e indústrias de software tendem a passar: a primeira, que já era adotada por algumas empresas que consideravam questões de privacidade como algo relevante, é a que se operava de maneira reativa, estimulada por métodos de avaliação da privacidade, com as devidas adaptações de artefatos de software existentes em seus próprios produtos; com base na experiência adquirida na etapa anterior, a segunda fase adota procedimentos de auditoria (internas e externas), cujo impacto é percebido nas últimas etapas do ciclo de vida de desenvolvimento. (TORRE et al., 2021) cita, como pontos negativos das auditorias, o fato de elas serem custosas e manuais; e, por fim, a terceira fase conta com a estratégia de PbD, que assume que os requisitos relacionados à privacidade devem ser integrados desde as primeiras fases, cujo impacto é sentido em todas as etapas do ciclo de vida de desenvolvimento de um software. A abordagem de PbD é uma tendência também discutida por (MOUGIAKOU et al., 2017) – onde um dos modelos propostos já utiliza essa abordagem, colocando como ator o controlador de dados – e (NOTARIO et al., 2015), que relata que os seus princípios estão se tornando cada vez mais reconhecidos. Neste estudo, os autores elencam que a sua adoção por desenvolvedores de software ainda era prejudicada pela falta de maturidade da abordagem em termos práticos, no entanto, através de (CAMPANILE et al., 2022) pode-se perceber como ela evoluiu, sendo a fase mais avançada sugerida para o tratamento de privacidade dentro do ciclo de vida de um software.

Ainda, de forma não madura, pode-se perceber que os estudos (OLUKOYA et al., 2022; TORRE et al., 2021; NOTARIO et al., 2015) abordam propostas visando, de alguma maneira, automatizar metodologias e práticas, com o objetivo de minimizar o trabalho e a

análise manual e individual de cada caso de lei de privacidade existente. Contudo, essas abordagens ainda barram no aspecto de generalização dessas legislações – identificada por (TORRE et al., 2021; HJERPPE et al., 2019) –, uma vez que leis de determinados países podem ter artigos específicos que podem não ser capturados por seus modelos propostos. Há também o problema da evolução (ESTRADA-JIMÉNEZ et al., 2017), visto que leis de privacidade podem sofrer alterações e os modelos propostos ficarem defasados, sendo necessário revisá-los. (HJERPPE et al., 2019) tenta extrair requisitos a partir da GDPR, no entanto, contudo, está amarrado a um estudo de caso. (OLUKOYA et al., 2022) tenta aplicar estruturas propostas na avaliação de uma lei de privacidade fora da área de abrangência da UE e dos EUA, mas barra nas diferenças específicas das leis. (NOTARIO et al., 2015), ainda antes da GDPR entrar em vigor, discute metodologias sistemáticas com o objetivo de integrar as etapas de análise e projeto do ciclo de vida de um software com a engenharia de privacidade.

Assim como diversas atividades básicas da engenharia de software – como Engenharia de Requisitos, Projeto, Implementação etc. –, (OLUKOYA et al., 2022; (SION et al., 2019; NOTARIO et al., 2015) propõem adotar a atividade específica de Engenharia de Privacidade, um campo embrionário de pesquisa que busca abordagens sistemáticas para o início e aplicação de soluções orientadas à privacidade em sistemas e em processos de desenvolvimento de software. (OLUKOYA et al., 2022) elenca que a engenharia de privacidade pode auxiliar no desenvolvimento de sistemas de maneira segura e, ao mesmo tempo, na preservação da privacidade e da conformidade de leis. Para (SION et al., 2019), trabalhar com questões de proteção de dados desde a fase de projeto de um software – ao invés de se adicionar uma camada desajeitada de conformidade legal na etapa final de desenvolvimento de um sistema – é cada vez mais reconhecido como uma abordagem correta para a engenharia de privacidade. (NOTARIO et al., 2015), contudo, diagnostica que os desenvolvedores têm dificuldade em entender o que deve ser feito para que sistemas obedeçam à privacidade desde o a fase de projeto, dificultando a adoção de práticas de engenharia de privacidade na prática.

Outro aspecto relevante, e que são cada vez mais tendência atualmente, são os ecossistemas de software, onde vários negócios podem atuar em conjunto, como uma unidade, e podem interagir com um mercado compartilhado de software e de serviços (VALENÇA et al., 2020). Com a definição de legislações de privacidade, é preciso garantir que esses ecossistemas funcionem de acordo com essas regulamentações, evitando vazamento ou acesso indevido de dados pessoais de usuários por desenvolvedores ou revendedores. Por

essa razão, mais estudos são necessários a temas como: coleta de dados; controle de acesso; consentimento do usuário; transparência relativa ao tratamento; adequação a leis de diferentes regiões – assim como elencado também por (OLUKOYA et al., 2022).

A tabela 6.2 a seguir apresenta um quadro resumo das principais tendências detectadas a partir deste estudo:

Tabela 6.2 – Quadro Resumo com principais tendências

ID	Tendências
1º	Privacy by Design (PbD) – Princípios cada vez mais são reconhecidos e aplicáveis durante todo o ciclo de vida do software.
2º	Busca por ferramentas de automatização do processo de extração de requisitos legais
3º	Engenharia de Privacidade como uma Disciplina – Embrionária ainda, e baseada em abordagens sistemáticas, pode auxiliar no desenvolvimento de sistemas de forma segura e na preservação da privacidade e da conformidade.

Fonte: o autor

Para empresas e engenheiros de software que desejam investigar e adotar estratégias de privacidade e de proteção de dados em suas aplicações, a indicação deste trabalho é que sejam estudadas as leis vigentes de proteção de dados pessoais (de países ou de blocos econômicos) onde o software ou produto estará em operação. Essa análise é de suma importância para evitar sanções financeiras e de operação pelo fato de o software não atender a algum requisito legal. Não somente por isso, mas também para evitar desperdício de recursos no processo de desenvolvimento. Também é importante adotar, o mais cedo possível, estratégias para que questões relacionadas às leis privacidade possam ser atendidas. Para esse processo de *compliance* (de estar em conformidade com determinadas leis), uma sugestão é estudar e adotar o processo de PbD, onde requisitos de privacidade podem ser analisados desde a fase de projeto, além de poderem ser monitorados e acompanhados durante todo o ciclo de vida do software. Caso haja dúvidas a respeito de alguma legislação de privacidade ou de proteção de dados vigente, é importante realizar uma consultoria com especialistas jurídicos a este respeito, evitando que a implementação de requisitos de privacidade seja realizada de forma incorreta e desconforme a leis e normas.

7 CONCLUSÃO

Este trabalho teve como objetivo investigar o estado da arte de como as leis de privacidade e de proteção de dados impactam as atividades de engenharia de software através de um processo sistemático de revisão da literatura. Com base no conhecimento produzido através dos estudos primários selecionados, foram respondidas as quatro questões de pesquisa definidas, o que permitiu sintetizar o conhecimento sobre os impactos causados, quais soluções foram propostas para mitigar esses problemas e foi possível identificar desafios e oportunidades de pesquisa.

Esta RSL foi motivada pela ausência de levantamento ou revisão sobre os impactos e desafios que leis de privacidade trazem às atividades de engenharia de software no contexto de desenvolvimento, manutenção e evolução de sistemas e aplicações. Enquanto (CANEDO et al., 2022) e (CANEDO et al., 2021) investigam se os princípios de privacidade são corretamente interpretados por equipes de metodologia ágil, qual o conhecimento dessas equipes acerca da LGPD e quais soluções de privacidade estão sendo adotadas, (GEORGIADIS et al., 2022) está mais focado em identificar riscos de privacidade e proteção de dados específicos do contexto de análise de grandes volumes de dados, que poderiam impactar negativamente os direitos e liberdades dos usuários. Portanto, essa RSL apresenta uma perspectiva mais ampla ao considerar todas as atividades de engenharia de software e metodologias de desenvolvimento.

De forma sintetizada, as respostas às questões de pesquisa podem ser sintetizadas da seguinte maneira:

QP1: Como as leis de privacidade e de proteção de dados impactam as atividades de engenharia de software? A maioria dos estudos (72%) citam impactos na fase de especificação e/ou de requisitos, dado que os requisitos de software geralmente baseados em leis costumam possuir ambiguidade, fazendo com que trechos possam ser interpretados com mais de um sentido. Outro problema é que várias equipes de desenvolvimento de software ainda não possuem conhecimento maduro sobre leis de privacidade e como garantir que sistemas estejam em conformidade com a referida lei e evite penalidades legais e monetárias.

QP2: Quais aspectos foram necessários alterar ou incluir ao processo para se garantir o cumprimento das leis de privacidade? Nesta questão, foi possível perceber uma gama de soluções heterogêneas dos estudos. Para garantir a conformidade com as leis de privacidade, 27,27% dos estudos propuseram soluções utilizando UML, com o objetivo de representar conceitos e relacionamentos da GDPR. 18,18% propuseram o uso de

metodologias visando obter requisitos de privacidade com base em textos de leis de privacidade. Como forma de checar a conformidade, um estudo elaborou um *checklist* de inspeção de forma a verificar se sistemas estão em conformidade com a legislação de privacidade, enquanto outro sugeriu a elaboração de um novo ciclo de desenvolvimento, onde seja dada uma atenção especial à fase de testes de privacidade. Um estudo focado no ambiente industrial elencou restrições práticas que pequenas e médias empresas precisam lidar para atender as restrições da GDPR. Outro estudo com foco em ecossistemas de software analisou funções de modo a proteger a privacidade de seus usuários.

QP3: Uma vez identificados os processos que necessitaram ser alterados ou aprimorados, a solução proposta chegou a ser implementada ou validada no contexto de desenvolvimento de software em produção? Poucos estudos citaram uma etapa de validação ou de implementação de suas propostas. Apenas um estudo produziu um protótipo. Os autores desse estudo validaram a proposta no contexto de um sistema de tratamento de doenças cardiovasculares, onde o protótipo destaca quais tipos de dados são usados em quais fluxos de dados, e documenta o envolvimento do titular dos dados com o médico (como destinatário legal). Para validar suas propostas, 18,18% dos estudos utilizaram estudo de caso. Em um deles, os autores avaliaram a estrutura proposta com base uma lei de privacidade fora da União Europeia e dos EUA, mas perceberam que lidar com artefatos que são tecnicamente diferentes, gerenciados por diferentes organizações e partes interessadas, é um desafio que impacta a generalização desejada. Por fim, um estudo realizou uma prova de conceito para validar o *checklist* proposto durante o uso de um portal de uma instituição federal brasileira.

QP4: Existe algum desafio, problema ou limitação em alguma atividade de engenharia de software, decorrente das leis de privacidade, que esteja sem solução? Poucos estudos dedicaram uma seção ou capítulo para abordar limitações dos estudos. 18,18% dos estudos citam a generalização como uma limitação tendo em vista que a abordagem construída no estudo para a GDPR pode não ser aplicada a outras leis de privacidade. Um desses estudos trata como fatores limitantes, ainda, a questão da extensibilidade e da resiliência. Outra limitação levantada pelos estudos é a rápida mudança ou alteração da legislação de privacidade, e que as estruturas propostas devem considerar essa evolução, caso contrário, os modelos e propostas podem ficar defasados. Um dos estudos elenca que é difícil atingir consistência, integridade e utilidade. No âmbito de ecossistemas de software, são levantados vários desafios em aberto, como questões de quais dados pessoais são coletados dentro do ecossistema; como garantir que apenas os dados realmente necessários sejam coletados; de qual forma assegurar que empresas do ecossistema tenham acesso somente aos

dados aos quais elas têm (legalmente) direito, como garantir que os requisitos legais de diferentes legislações possam ser atendidos, entre outros.

A avaliação da qualidade permitiu refinar a seleção dos estudos e explicar os resultados pelas diferenças de qualidade. Os estudos selecionados apresentam pontuação média de 6,69, o que é razoável. Porém, a expectativa deste trabalho era encontrar soluções mais robustas no que diz respeito a como sistemas são evoluídos de forma a garantir a conformidade com leis de privacidade. Pode-se concluir que incluir aspectos de privacidade dentro de softwares existentes é um tópico de estudo relativamente novo, principalmente porque a comparação com os trabalhos relacionados e a discussão das limitações das proposições dos estudos foram questões de qualidade com pontuações médias baixas. Por essa razão, é preciso realizar estudos mais profundos, tanto em termos teóricos quanto práticos, para avançar o conhecimento sobre como minimizar o impacto das leis de privacidade ao longo do ciclo de vida de desenvolvimento de sistemas.

O processo de seleção de estudos foi o maior desafio para este trabalho. Foi realizada uma tentativa de coletar o máximo possível de artigos relevantes para o assunto deste trabalho nas principais bibliotecas digitais, mantendo uma metodologia sistemática, que possa ser reproduzida numa eventual continuidade desse estudo, isso é, sem inclusão de estudos de forma manual ou com base na opinião pessoal do autor acerca da qualidade dos estudos primários coletados. No entanto, encontrar todos os trabalhos relevantes já publicados é impossível (VOM BROCKE et al., 2015). Para dar conta do enorme volume de artigos encontrados nas bibliotecas digitais (2.832 estudos no caso), foi realizado um processo de três etapas para triagem e seleção dos estudos que dariam conta desse volume. Assim, a seleção final foi refinada iterativamente com base no título, resumo, palavras-chave, numa segunda etapa na visão geral do artigo (com atenção especial à introdução, figuras dos estudos e conclusão) e, finalmente, pela leitura completa dos artigos. Esse processo permitiu minimizar o viés na triagem dos artigos por dois leitores independentes e consolidar os critérios de seleção/exclusão dos estudos posteriormente, com base na leitura completa deles. Assim, acredita-se de que a estratégia e processo metodológico adotados foram bem definidos e executados, o que permitiu maximizar o escopo e a qualidade da revisão efetuada.

Os resultados relatados neste trabalho abrangem estudos coletados entre os anos de 2015 e de 2022. Como as leis de privacidade se tornaram um tópico de tendência, dado o lançamento – pode-se dizer recente – das leis de privacidade da União Europeia (GDPR) e do Brasil (LGPD), espera-se que o número de estudos aumente significativamente nos próximos

anos. É importante coletar novos estudos como trabalhos futuros, inclusive para medir os avanços e o preenchimento de lacunas detectadas.

Como principais descobertas, a partir da pesquisa realizada, podemos destacar que é cada vez mais é preciso que sistemas atendam às disposições de leis de privacidade, sob o risco de receber sanções financeiras e de operação; Outro aspecto relevante é que os softwares desenvolvidos precisam atender à legislação do país ou bloco econômico de operação. Também foi possível descobrir que há uma baixa familiaridade de engenheiros de software com textos jurídicos, o que resulta em uma dificuldade de compreender o que deve ou precisa ser realizado para que sistemas estejam ou fiquem em conformidade com leis de privacidade vigentes. Outro achado é a necessidade de um haver – em organizações ou em equipes/times – a presença de uma pessoa experiência em análise e interpretação de textos legais, para que haja uma redução de ambiguidade e melhoria no processo de especificação de requisitos de privacidade. Também se pôde verificar que testes de privacidade, se existem, ainda são precários, visto que estes são geralmente realizados por pessoas que não possuem treinamentos específicos ou experiência com privacidade e proteção de dados.

Para preencher as lacunas em aberto, há uma diversidade de trabalhos futuros que podem ser realizados: como primeira sugestão, essa RSL poderia ser complementada, considerando a literatura mais recente, além de outras bibliotecas digitais. No entanto, outros tópicos também poderiam ser investigados, como formas de mapear estruturas das leis em requisitos de forma a maximizar a generalização e a evolução de softwares em produção. Também podem ser feitos estudos mais aprofundados de como a engenharia de privacidade pode auxiliar no processo de desenvolvimento de sistemas. De uma perspectiva mais prática, poderia ser investigado como garantir a conformidade de softwares legados com as leis de privacidade. Outra sugestão é avaliar se a presença de especialistas jurídicos na equipe pode, de fato, garantir que sistemas ou artefatos fiquem em conformidade com leis de privacidade, minimizando aspectos de ambiguidade, por exemplo. Outro tópico relevante é analisar o amadurecimento da abordagem de PbD. Também podem ser realizados estudos para investigar quais soluções ou estratégias estão sendo adotadas pela indústria para minimizar o impacto da legislação de privacidade no processo de desenvolvimento de software.

REFERÊNCIAS

ALBRECHT, Jan Philipp. How the GDPR will change the world. **European Data Protection Law Review**. v. 2, p. 287, 2016.

BEHUTIYE, Woubshet et al. Non-functional requirements documentation in agile software development: challenges and solution proposal. In: **International conference on product-focused software process improvement**. Cham: Springer International Publishing, 2017. p. 515-522.

BRASIL. Lei nº 12.965, de 23 de abril de 2014. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Brasília, DF: 2014. Disponível em < https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm>. Acesso em: 14 de jan. de 2023.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: 2018. Disponível em < https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em: 14 de jan. de 2023.

BROOKS, Frederik P.; BULLET, No Silver. Essence and accidents of software engineering. **IEEE computer**, v. 20, n. 4, p. 10-19, 1987.

CAMPANILE, Lelio; IACONO, Mauro; MASTROIANNI, Michele. Towards privacy-aware software design in small and medium enterprises. In: **2022 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)**. IEEE, 2022. p. 1-8.

DIAS CANEDO, Edna et al. Perceptions of ICT practitioners regarding software privacy. **Entropy**, v. 22, n. 4, p. 429, 2020.

CANEDO, Edna Dias et al. Agile teams' perception in privacy requirements elicitation: Lgpd's compliance in brazil. In: **2021 IEEE 29th International Requirements Engineering Conference (RE)**. IEEE, 2021. p. 58-69.

CANEDO, Edna Dias et al. Guidelines adopted by agile teams in privacy requirements elicitation after the Brazilian general data protection law (LGPD) implementation. **Requirements Engineering**, v. 27, n. 4, p. 545-567, 2022.

DABBAGH, Mohammad et al. An approach for integrating the prioritization of functional and nonfunctional requirements. **The Scientific World Journal**, v. 2014, 2014.

DERMEVAL, Diego et al. Applications of ontologies in requirements engineering: a systematic review of the literature. **Requirements engineering**, v. 21, p. 405-437, 2016.

DUGGINENI, Sasidhar. Impact of Controls on Data Integrity and Information Systems. **Science and Technology**, v. 13, n. 2, p. 29-35, 2023.

ESTRADA-JIMÉNEZ, José et al. Online advertising: Analysis of privacy threats and protection approaches. **Computer Communications**, v. 100, p. 32-51, 2017.

FERRÃO, Sâmmara Éllen Renner; CANEDO, Edna Dias. Uma taxonomia para requisitos de privacidade e sua aplicação no Open Banking Brasil. In: **WER 2022**. 2022.

FRANÇA, Tiago Cruz et al. Big Social Data: princípios sobre coleta, tratamento e análise de dados sociais. **XXIX Simpósio Brasileiro de Banco de Dados–SBBD**, v. 14, 2014.

GARCIA, Lara Rocha et al. **Lei Geral de Proteção de Dados (LGPD): guia de implantação**. Editora Blucher, 2020.

GEORGIADIS, Georgios; POELS, Geert. Towards a privacy impact assessment methodology to support the requirements of the general data protection regulation in a big data analytics context: A systematic literature review. **Computer Law & Security Review**, v. 44, p. 105640, 2022.

HIJAZI, Haneen; KHDOUR, Thair; ALARABEYYAT, Abdulsalam. A review of risk management in different software development methodologies. **International Journal of Computer Applications**, v. 45, n. 7, p. 8-12, 2012.

HJERPPE, Kalle; RUOHONEN, Jukka; LEPPÄNEN, Ville. The general data protection regulation: requirements, architectures, and constraints. In: **2019 IEEE 27th International Requirements Engineering Conference (RE)**. IEEE, 2019. p. 265-275.

HUMPHREY, Watts S. The software engineering process: definition and scope. In: **Proceedings of the 4th international software process workshop on Representing and enacting the software process**. 1988. p. 82-83.

KITCHENHAM, Barbara. Procedures for performing systematic reviews. **Keele, UK, Keele University**, v. 33, n. 2004, p. 1-26, 2004.

KEELE, Staffs et al. Guidelines for performing systematic literature reviews in software engineering. 2007.

KUHRMANN, Marco et al. Hybrid software and system development in practice: waterfall, scrum, and beyond. In: **Proceedings of the 2017 international conference on software and system process**. 2017. p. 30-39.

LEE, Christopher; GUADAGNO, Luigi; JIA, Xiaoping. An agile approach to capturing requirements and traceability. In: **Proceedings of the 2nd International Workshop on Traceability in Emerging Forms of Software Engineering (TEFSE 2003)**. 2003.

MACDONELL, Stephen et al. How reliable are systematic reviews in empirical software engineering?. **IEEE Transactions on Software Engineering**, v. 36, n. 5, p. 676-687, 2010.

MENDES, João; VIANA, Davi; RIVERO, Luis. Developing an inspection checklist for the adequacy assessment of software systems to quality attributes of the brazilian general data protection law: An initial proposal. In: **Proceedings of the XXXV Brazilian Symposium on Software Engineering**. 2021. p. 263-268.

MENDES, Laura Schertel et al., Tratado de Proteção de Dados Pessoais. Vol. 1. GEN - Grupo Editorial Nacional. 2020.

MOUGIAKOU, Eirini; VIRVOU, Maria. Based on GDPR privacy in UML: Case of e-learning program. In: **2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA)**. IEEE, 2017. p. 1-8.

NETTO, Dorgival; SILVA, Carla; ARAÚJO, João. Identifying how the brazilian software industry specifies legal requirements. In: **Proceedings of the XXXIII Brazilian Symposium on Software Engineering**. 2019. p. 181-186.

NOTARIO, Nicolás et al. PRIPARE: integrating privacy best practices into a privacy engineering methodology. In: **2015 IEEE Security and Privacy Workshops**. IEEE, 2015. p. 151-158.

OLUKOYA, Oluwafemi. Assessing frameworks for eliciting privacy & security requirements from laws and regulations. **Computers & Security**, v. 117, p. 102697, 2022.

PARLIAMENT, European e Council of the European Union: General data protection regulation. **European Commission**, 1:99, 2018. Disponível em <https://gdpr-info.eu/. 1, 2, 3, 9, 10, 11, 14, 17, 18, 19, 20, 21, 22>. Acesso em 16 jan. 2023.

ROWLEY, Jennifer; SLACK, Frances. Conducting a literature review. **Management research news**, v. 27, n. 6, p. 31-39, 2004.

SILVA, Jefferson; CALEGARI, Newton; GOMES, Eduardo. After Brazil's general data protection law: Authorization in decentralized web applications. In: **Companion proceedings of the 2019 World Wide Web conference**. 2019. p. 819-822.

SION, Laurens et al. An architectural view for data protection by design. In: **2019 IEEE International Conference on Software Architecture (ICSA)**. IEEE, 2019. p. 11-20.

TIWARI, Saurabh; GUPTA, Atul. A systematic literature review of use case specifications research. **Information and Software Technology**, v. 67, p. 128-158, 2015.

TORRE, Damiano et al. Modeling data protection and privacy: application and experience with GDPR. **Software and Systems Modeling**, v. 20, p. 2071-2087, 2021.

VALENÇA, George; KNEUPER, Ralf; REBELO, Maria Eduarda. Privacy in software ecosystems-an initial analysis of data protection roles and challenges. In: **2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)**. IEEE, 2020. p. 120-123.

VIEIRA, Iuri Sousa. Aplicações de software desenvolvidas no contexto da inteligência artificial (IA), Machine Learning e big data e o direito dos cidadãos de acordo com a lei geral de proteção de dados (LGPD). 2021.

VOM BROCKE, Jan et al. Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. **Communications of the association for information systems**, v. 37, n. 1, p. 9, 2015.

VOSS, W. Gregory; HOUSER, Kimberly A. Personal data and the GDPR: providing a competitive advantage for US companies. **American Business Law Journal**, v. 56, n. 2, p. 287-344, 2019.

WAGNER, Stefan. "Software product quality control." Springer, Heidelberg. 2013:210.

WOHLIN, Claes et al. **Experimentation in software engineering**. Springer Science & Business Media, 2012.

ZAEEM, Razieh Nokhbeh; BARBER, K. Suzanne. The effect of the GDPR on privacy policies: Recent progress and future promise. **ACM Transactions on Management Information Systems (TMIS)**, v. 12, n. 1, p. 1-20, 2020.

ZOWGHI, Didar; COULIN, Chad. Requirements elicitation: A survey of techniques, approaches, and tools. **Engineering and managing software requirements**, p. 19-46, 2005.

APÊNDICE A – TABELA DE AVALIAÇÃO DA QUALIDADE DOS ESTUDOS SELECIONADOS

Tabela A-1 – Pontuação da avaliação da qualidade dos estudos selecionados

ID Estudo	QQ1	QQ2	QQ3	QQ4	QQ5	QQ6	QQ7	QQ8	QQ9	Pontuação Total
S1	1	1	1	1	0,5	1	0	0,5	0,5	6,5
S2	1	1	0	1	1	1	1	0,5	1	7,5
S3	1	1	1	0,5	1	0,5	0	0,5	1	6,5
S4	1	1	1	1	0,5	1	1	0,5	1	8
S5	0,5	1	0	1	0,5	1	0,5	0,5	0,5	5,5
S6	1	1	1	1	0,5	0,5	0	0,5	0,5	6
S7	1	1	1	1	0,5	0,5	0	0,5	1	6,5
S8	1	1	0	1	1	1	0	1	1	7
S9	1	1	1	1	1	1	1	0,5	1	8,5
S10	1	1	1	1	0,5	0,5	0	0,5	0,5	6
S11	1	1	0	1	1	0,5	0	0,5	0,5	5,5
Média	0,95	1	0,64	0,95	0,73	0,77	0,32	0,55	0,77	6,69

Fonte: o autor