Um Modelo de Conceitos Relacionados à Privacidade de Dados Pessoais

Arthur Santos, Mariana Peixoto and Carla Silva

Centro de Informática, Universidade Federal de Pernambuco, Recife, Brasil
arthursantosvs@gmail.com, {mmp2, ctlls}@cin.ufpe.br

Abstract. Muitas informações disponíveis no dia a dia têm sido digitalizadas para tornar o seu acesso mais rápido e fácil. Esses dados muitas vezes apresentam informações pessoais, o que pode acarretar em problemas e preocupações com a privacidade. Porém, para tratar com eficiência esses problemas e preocupações no ambiente de desenvolvimento de software, os engenheiros de software precisam compreender melhor o conceito de privacidade. Dada a relevância do tópico nos últimos anos, especialmente após regulamentações de proteção de dados, o presente trabalho atualiza, por meio de uma Revisão Sistemática de Literatura, uma pesquisa prévia que define conceitos relacionados à privacidade na forma de um modelo conceitual. A atualização do modelo conceitual resulta em uma visão clara e atualizada dos conceitos relacionados à privacidade, tanto para profissionais da indústria como para acadêmicos da área.

Keywords: Engenharia de Software, Privacidade, Linguagens de Modelagem, Modelos Conceituais.

1 Introdução

Muitas informações disponíveis no dia a dia estão sendo digitalizadas para facilitar o acesso rápido e fácil. Os serviços digitais, por exemplo, dependem de dados armazenados para identificar clientes, suas preferências e transações [4, 9]. Esses dados geralmente revelam grandes quantidades de informações pessoais, e a exposição dessas informações de forma não regulamentada pode ameaçar a privacidade do usuário. Assim, a privacidade deve ser um fator considerado desde o início do processo de desenvolvimento de sistemas de software [11, 19, 20].

Gharib et al. [7] afirma que violações de privacidade podem ser evitadas se os requisitos de privacidade forem descobertos adequadamente durante a fase de Engenharia de Requisitos (ER) ao desenvolver um sistema sensível à privacidade. No entanto, apesar de vários esforços feitos para esclarecer conceitos relacionados à privacidade, vinculando-os a outros mais refinados e bem-definidos, não há consenso sobre a definição destes conceitos ou qual deles deve ser usado para analisar os requisitos de privacidade ao desenvolver um sistema que lida com à privacidade [7, 8]. Essa situação resultou em muita confusão entre designers e partes interessadas e levou, por sua vez, a decisões de design erradas [7].

O problema aumenta porque muitos desenvolvedores não têm conhecimento e compreensão suficientes sobre privacidade, nem sabem suficientemente como desenvolver software com privacidade [10,26]. Mesmo assim, a maioria dos trabalhos

Proceedings of the 26th Workshop on Requirements Engineering (WER23), August 15-17, 2023, Porto Alegre, RS, Brazil.
DOI: 10.29327/1298356.26-6

existentes sobre requisitos de privacidade muitas vezes lidam com eles como requisitos não funcionais (RNFs) sem técnicas específicas sobre como tais requisitos podem ser atendidos [5], ou como requisitos de segurança (por exemplo, [4,6], etc.), ou seja, focando principalmente na confidencialidade e negligenciando importantes aspectos como anonimato, pseudonimato, desvinculabilidade, inobservância, etc.

Neste contexto, uma melhor compreensão dos conceitos relacionados à privacidade (ou seja, conceitos considerados quando estiver desenvolvendo um sistema que precisa de privacidade), e suas inter-relações, seria um passo importante para fornecer aos desenvolvedores mais conhecimento para elicitar requisitos de privacidade e, consequentemente, aprimorar a qualidade do sistemas que precisam de privacidade [7].

Dada então a importância da privacidade no desenvolvimento de software, e a falta de definição clara e ampla do conceito, representando muitas condições relacionadas com que se deseja manter privado [4, 7], surgiu a necessidade de se ter um guia de conceitos relacionados à privacidade. Motivado por este cenário, Peixoto et al. [3] realizou uma Revisão Sistemática da Literatura (RSL) com o objetivo de encontrar linguagens de modelagem que tratassem de conceitos relacionados à privacidade para que esses conceitos pudessem ser extraídos e usados na criação de um modelo conceitual e uma descrição dos conceitos relacionados à privacidade e como eles se relacionam.

Percebida então a importância da pesquisa realizada em Peixoto et al. [3], e o número exponencialmente crescente de sistemas que lidam com informações pessoais (e.g. informação sobre cidadãos, clientes, etc.) [1], notou-se a importância de manter um modelo de conceitos relacionados à privacidade atualizado. Neste contexto, surgiu então a necessidade de se observar o atual estado da pesquisa, executar validações para medir a real necessidade de uma atualização da RSL [13] e, se confirmada a necessidade, realiza-la usando a metodologia proposta por Mendes et al. [12]. Com a confirmação da necessidade de atualizar a RSL, procurou-se por novos trabalhos sobre linguagens de modelagem que lidam com conceitos relacionados à privacidade, de forma a executar uma nova extração de dados e atualizar o modelo conceitual. Desta forma, será possível fornecer uma assistência mais completa e útil para os engenheiros de sistemas que lidam com a privacidade, independente da lei de proteção de dados em vigor no seu país.

Este trabalho está estruturado da seguinte forma: Na Seção 2 encontra-se a apresentação da fundamentação teórica. Na Seção 3 são apresentados os procedimentos metodológicos utilizados para atualizar a RSL originalmente apresentada em [3]. Na Seção 4. Os trabalhos selecionados na RSL, bem como o novo modelo conceitual obtido são apresentados e explicados. Na Seção 5 as limitações e ameaças à validade da pesquisa são pontuadas. E, finalmente, na Seção 6 são apresentadas as conclusões finais e os trabalhos futuros.

2 Fundamentação Teórica

A privacidade tornou-se uma preocupação para os designers de sistema. Em outras palavras, lidar com questões relacionadas à privacidade é uma obrigação porque violações de privacidade podem resultar em custos enormes, bem como consequências a longo prazo [2, 21, 22].

Violações de privacidade podem acontecer devido a políticas de privacidade inadequadas [23,24] e más práticas de segurança que levam a ataques, furtos de dados, etc. [1,15]. No entanto, a maioria dessas violações podem ser evitadas se os requisitos de privacidade do sistema foram capturados corretamente durante o design do sistema (por exemplo, Privacidade por Projeto (Privacy by Design -PbD) [15,4].

Por outro lado, a privacidade é um conceito vago e elusivo [4,16]. Apesar de vários esforços feitos para esclarecer o conceito, ligando-o a conceitos mais refinados como "secrecy", "personhood", controle de informações pessoais, etc., não há consenso sobre a definição desses conceitos ou quais deles deveriam ser usados para analisar privacidade [16]. Isso resulta em muita confusão entre os engenheiros de software e as demais partes interessadas no software e, por sua vez, leva a decisões de design erradas.

Dentre os estudos que se esforçam para definir esse conceito, algumas características são recorrentes, como requisitos de privacidade, "goals" (objetivos de privacidade), informação pessoal e privada, "threats" (ameaças) de privacidade, dimensões de privacidade e atores de diferentes papéis no processo de manuseio de dados pessoais.

As definições de privacidade também foram influenciadas pelas novas definições legais e leis de proteção de dados, especialmente a GDPR (General Data Protection Regulation ou, em Português, Regulamento Geral sobre a Proteção de Dados) e a LGPD (Lei Geral de Proteção de Dados Pessoais). A GDPR, lei de proteção de dados europeia, teve grande impacto no desenvolvimento de software que lida com a privacidade. Modelos que capturam conceitos de privacidade, como o proposto por Agostinelli et al. em [14] e Vilela et al. [25], são usados para elaboração modelos de desenvolvimento de software em conformidade com a GDPR e LGPD (lei brasileira), respectivamente.

3 Protocolo para atualização da Revisão Sistemática da Literatura

Para atingir os objetivos deste trabalho, uma Questão de Pesquisa (QP) adaptada de Peixoto et al. [3] foi elaborada: QP1: Quais são os novos conceitos relacionados à privacidade capturados pelas linguagens usadas para modelagem e análise de requisitos de privacidade? (Esta questão pretende identificar quais conceitos são suportados pelas linguagens de modelagem selecionadas para capturar necessidades de privacidade e como eles se relacionam entre si. Essas entradas apoiarão a criação de um modelo conceitual sobre privacidade).

Este trabalho leva em consideração o método descrito por Wohlin et al. [13] para atualização de RSLs. Resumidamente, as suas diretrizes sugerem o uso de uma foward snowballing com uma única iteração usando o Google Scholar (https://scholar.google.com.br/), empregando como seed set os estudos primários da RSL original.

Segundo Kitchenham e Charters [17], os critérios de seleção dos estudos servem para identificar quais são os estudos primários que fornecem evidências diretas para responder as perguntas de pesquisa, logo, critérios de Inclusão e Exclusão devem ser definidos tomando como base as perguntas de pesquisa. Esses critérios de seleção, assim

como as questões de pesquisa, de acordo com Wohlin et al. [13], são herdados e adaptados da pesquisa original [3].

Os critérios elaborados na pesquisa original [3] seguiram Kitchenham e Charters [17]. Além dos critérios herdados, foram adicionados os critérios de que os artigos devem ter sido publicados a partir de 2017, já que a RSL original de Peixoto et al. [3] foi feita neste ano.

A seleção dos estudos da presente pesquisa foi realizada utilizando o modelo de Cruzes e Dyba [18], assim como no estudo original. Porém, a pesquisa também diferiu na fase 1 do modelo padrão de Cruzes e Dyba devido à abordagem específica à atualização de RSL de Wohlin et al. [13]. A seguir, são detalhadas as fases da pesquisa:

- Fase 1: Foi realizado o forward snowballing no Google Scholar usando os artigos selecionados em Peixoto et al. [3] como seed set, e os resultados foram exportados para uma planilha Excel[1] categorizada com informações de contexto, formando assim a base de dados inicial do estudo.
- Fase 2: Cada estudo selecionado na fase anterior foi então analisado. Os títulos e resumos foram lidos e os critérios de inclusão e exclusão foram aplicados. Em caso de incerteza quanto à sua aceitação, o artigo era incluído para análise nas etapas seguintes.
- Fase 3: Os critérios de inclusão e exclusão foram aplicados sobre os estudos com base na leitura da introdução e conclusão dos estudos. Quando ainda assim houver incerteza quanto à sua aceitação, o artigo era lido por completo.
- Fase 4: Os estudos resultantes da terceira fase foram então lidos por completo e a avaliação da qualidade (omitida por questão de espaço) foi aplicada sobre eles.

Após feita a seleção dos artigos, a pesquisa avançou para a fase de análise e síntese dos dados, assim como a atualização do modelo conceitual de privacidade. Além do formulário de extração de dados, foi criado um documento de análise[2] que detalha todas as linguagens de modelagem apresentadas nos artigos da seleção final e os conceitos de privacidade presentes nessas linguagens. O intuito de criar este documento foi centralizar as informações sobre conceitos de privacidade e suas relações para facilitar a atualização e síntese do novo modelo conceitual fornecido por esta pesquisa e entender de quais linguagens eles vieram, assim como quais são as linguagens mais relevantes para modelagem de privacidade.

Para a extração dos conceitos dos trabalhos selecionados, três etapas foram seguidas para obter o modelo conceitual original apresentado em Peixoto et al [3], as quais serão repetidas neste trabalho, conforme Wohlin et al [13] recomenda. Na primeira etapa, extraímos dos artigos selecionados os conceitos relacionados à privacidade (por exemplo, Conscientização / Necessidade de saber / Saber). De um conjunto de conceitos correlacionados, uma única categoria foi derivada (por exemplo, apenas “Consciência (Awareness)”). Na segunda etapa, observamos quais são as relações entre as categorias (por exemplo, “Consciência (Awareness)” é um "Mecanismo de Privacidade (Privacy

[1] Planilha usada: https://docs.google.com/spreadsheets/d/1hZ3hCliJWTlk14pXGVNCOrq9ELjH-Cx9_WrWhlmUk-g/edit?usp=sharing

[2] Documento de análise das linguagens de modelagem: https://docs.google.com/spreadsheets/d/16DxbhzXoPCRvtanQqtVkzIp5uLZQSNCn4j7LMiaxkHU/edit#gid=0

Mechanism)"). Finalmente, na terceira etapa, criamos a relação entre as categorias (por exemplo, a relação entre “Consciência/Awareness” e "Mecanismo de Privacidade (Privacy Mechanism)" é uma relação de generalização).

4 Resultados da atualização da Revisão Sistemática da Literatura

Esta RSL foi realizada entre Agosto e Novembro de 2020. Na fase 1, o resultado das buscas por repositório foi de 811 estudos. Durante a Fase 2 do processo de seleção, 623 estudos foram descartados após a investigação dos títulos e resumos, restando 188 estudos potencialmente relevantes. Na Fase 3 do processo de seleção, os estudos foram analisados por meio da leitura de sua introdução e conclusão, sendo descartados 118 estudos, resultando em 70 estudos.

Foi realizada a leitura integral e a avaliação de qualidade dos 70 estudos resultantes da fase anterior, resultando na exclusão de 44 estudos. Restaram então 26 estudos considerados relevantes para serem utilizados na pesquisa (Tabela 1). Destes, apenas 10 contribuíram diretamente para encontrar novo conceitos e relacionamentos para o modelo conceitual de privacidade: ARXIV1, IEEE4, SS1, SPRINGER8, WILEY1, SPRINGER12, SPRINGER9, EMINS2, IEEE1, SPRINGER1. Eles estão destacados em negrito na Tabela 1. No documento de análise[2] encontra-se mais detalhes sobre a análise feita nas linguagens de modelagem encontradas nos artigos selecionados a fim de identificar os novos conceitos de privacidade.

Tabela 1. Estudos Primários Selecionados.

ID	Trabalho Selecionado
ARXIV1	Gharib, Mohamad, and John Mylopoulos. "A Core Ontology for Privacy Requirements Engineering." arXiv preprint arXiv:1811.12621 (2018).
IEEE1	Alkubaisy, Duaa. "A framework managing conflicts between security and privacy requirements." 2017 11th International Conference on Research Challenges in Information Science (RCIS). IEEE, 2017.
EMINS1	Alshammari, Majed, and Andrew Simpson. "A model-based approach to support privacy compliance." Information & Computer Security (2018).
SPRINGER1	Ramadan, Q., Strüber, D., Salnitri, M., Jürjens, J., Riediger, V., & Staab, S. "A semi-automated BPMN-based framework for detecting conflicts between security, data-minimization, and fairness requirements." Software and Systems Modeling (2020): 1-37.
SS1	Argyropoulos, N., Shei, S., Kalloniatis, C., Mouratidis, H., Delaney, A., Fish, A., & Gritzalis, S. "A semi-automatic approach for eliciting cloud security and privacy requirements." Proceedings of the 50th Hawaii international conference on system sciences. 2017.
SPRINGER2	Chergui, Mohamed El Amine, and Sidi Mohamed Benslimane. "A valid

	BPMN extension for supporting security requirements based on cyber security ontology." International Conference on Model and Data Engineering. 2018.
SPRINGER3	Agostinelli, S., Maggi, F. M., Marrella, A., & Sapio, F."Achieving GDPR compliance of BPMN process models." International Conference on Advanced Information Systems Engineering. Springer, Cham, 2019.
SP1	Xia, T., Washizaki, H., Kato, T., Kaiya, H., Ogata, S., Fernandez, E. B., ... & Hazeyama, A. "Cloud Security and Privacy Metamodel." Proceedings of the 6th International Conference on Model-Driven Engineering and Software Development. SCITEPRESS - Science and Tech. Publications, Lda, 2018.
SPRINGER4	Gonçalves, António, Anacleto Correia, and Luis Cavique. "Data protection risk modeling into business process analysis." International Conference on Computational Science and Its Applications. Springer, Cham, 2017.
SPRINGER5	Ramadan, Qusai, et al. "Detecting conflicts between data-minimization and security requirements in business process models." European Conference on Modelling Foundations and Applications. Springer, Cham, 2018.
ACM1	Ahmadian, A. S., Jürjens, J., & Strüber, D. "Extending model-based privacy analysis for the industrial data space by exploiting privacy level agreements." Proceedings of the 33rd Annual ACM Symposium on Applied Computing. 2018.
EMINS2	Kalloniatis, Christos. "Incorporating privacy in the design of cloud-based systems: a conceptual meta-model." Information & Computer Security (2017).
IEEE2	Wuyts, Kim, Laurens Sion, and Wouter Joosen. "LINDDUN GO: A Lightweight Approach to Privacy Threat Modeling." 2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). IEEE, 2020.
ACM2	Ahmadian, Amir Shayan, et al. "Model-based privacy and security analysis with CARiSMA." Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering. 2017.
SPRINGER6	Robol, Marco, et al. "Modeling and Reasoning About Privacy-Consent Requirements." IFIP Working Conference on The Practice of Enterprise Modeling. Springer, Cham, 2018.
SCIENCE1	Mai, Phu X., et al. "Modeling security and privacy requirements: a use case-driven approach." Information and Software Tech. 100 (2018): 165-182.
IEEE3	Antignac, Thibaud, Riccardo Scandariato, and Gerardo Schneider. "Privacy compliance via model transformations." 2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). IEEE, 2018.
SPRINGER7	Pullonen, Pille, et al. "Privacy-enhanced BPMN: Enabling data privacy analysis in business processes models." Software and Systems Modeling 18.6 (2019): 3235-3264.

WILEY1	Basso, T., Montecchi, L., Moraes, R., Jino, M., & Bondavalli, A "PrivAPP: An integrated approach for the design of privacy-aware applications." Software: Practice and Experience 48.3 (2018): 499-527.
SPRINGER8	Caramujo, J., Rodrigues da Silva, A., Monfared, S., Ribeiro, A., Calado, P., & Breaux, T. "RSL-IL4Privacy: a domain-specific language for the rigorous specification of privacy policies." Requirements Engineering 24.1 (2019): 1-26.
SPRINGER9	Diamantopoulou, V., Kalloniatis, C., Gritzalis, S., & Mouratidis, H. "Supporting privacy by design using privacy process patterns." IFIP International Conference on ICT Systems Security and Privacy Protection. Springer, Cham, 2017.
IEEE4	Diamantopoulou, V., Argyropoulos, N., Kalloniatis, C., & Gritzalis, S. "Supporting the design of privacy-aware business processes via privacy process patterns." 2017 11th International Conf. on Research Challenges in Information Science (RCIS). IEEE, 2017.
SPRINGER10	Mavroeidi, Aikaterini-Georgia, Angeliki Kitsiou, and Christos Kalloniatis. "The interrelation of game elements and privacy requirements for the design of a system: A metamodel." International Conference on Trust and Privacy in Digital Business. Springer, Cham, 2019.
IEEE5	Coles, Joshua, Shamal Faily, and Duncan Ki-Aries. "Tool-supporting data protection impact assessments with CAIRIS." 2018 IEEE 5th International Workshop on Evolving Security & Privacy Requirements Engineering (ESPRE). 2018.
SPRINGER11	Robol, Marco, Mattia Salnitri, and Paolo Giorgini. "Toward GDPR-compliant socio-technical systems: modeling language and reasoning framework." IFIP Working Conf. on The Practice of Enterprise Modeling. Springer, Cham, 2017.
SPRINGER12	Gharib, Mohamad, Paolo Giorgini, and John Mylopoulos. "Towards an ontology for privacy requirements via a systematic literature review." International conference on conceptual modeling. Springer, Cham, 2017.

4.1 Descrição de Conceitos e Relações Atualizada

Nesta seção se encontram as descrições dos conceitos do modelo conceitual de privacidade, incluindo os achados em Peixoto et al. [3] e os levantados por essa pesquisa, os quais estão destacados nas cores rosa, cinza, amarelo, azul, verde e rosa claro nas Fig. 1 e Fig. 2. Alguns conceitos do modelo conceitual original também são detalhados seguir para melhor compreensão dos conceitos novos adicionados.

No fragmento do modelo conceitual na Fig.1, existe um "Proprietário/controlador (Owner/controller)" que possui uma associação a um "Terceiro (Third Party)" e a um "Processador (Processor)". Segundo o estudo original, um "Proprietário/controlador" pode ser uma entidade ativa com relações entre "Processador" e "Terceiros". O

"Proprietário/controlador" tem zero ou mais "Informação Pessoal (Personal Information)". O estudo original afirma que o tratamento da proteção da privacidade requer o tratamento da propriedade dos dados, que no caso foi indicada como "Informação Pessoal" (conceitos obtidos do modelo conceitual original [3]).

Esta "Informação Pessoal" pode ser especializada aos tipos "Privado (Private)", "Público (Public)" e "Semi-Público (Semi-Public)". Segundo o estudo original, os utilizadores (aqui apresentados como "Proprietários/controladores") podem decidir o que querem partilhar com os outros e existem tipos de "Informação Pessoal" de três graus. Esta informação pode ser privada quando apenas o "Proprietário/controlador" tem acesso a ela. "Semi-Público" quando o "Proprietário/controlador" decide partilhar a informação com determinado "Terceiro". Ou pública quando alguém tem acesso à informação (conceitos obtidos do modelo conceitual original [3]).

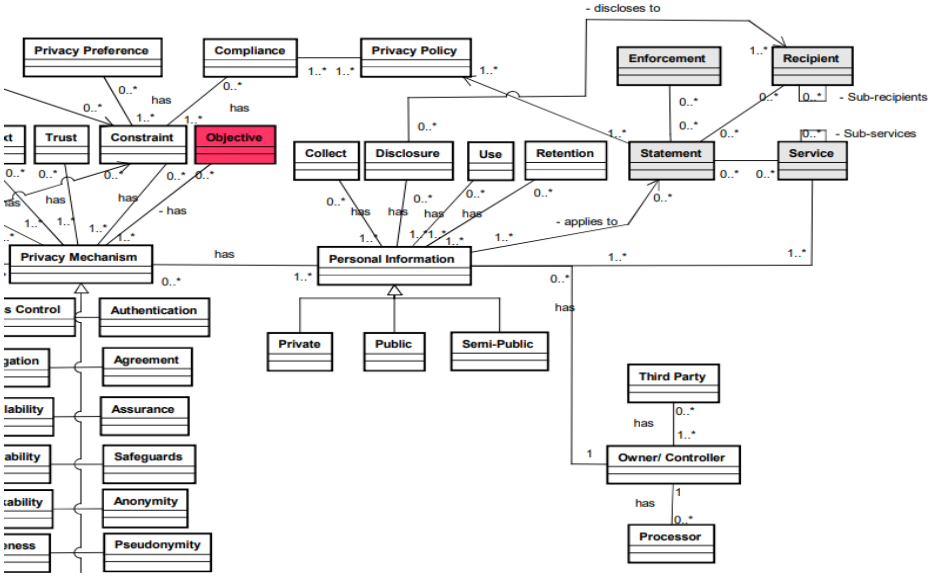

Fig.1. Modelo Conceitual Atualizado.

"Informação Pessoal" tem zero ou mais associações com "Coleta (Collect)", "Utilização (Use)", "Divulgação (Disclosure)" e "Mecanismo de Privacidade (Privacy Mechanism)". Por exemplo, de acordo com o estudo original, os "Mecanismos de Privacidade" podem concentrar-se na proteção da privacidade dos dados pessoais garantindo o seu "Anonimato (Anonimity)" ao preservar a revelação dos seus dados pessoais a terceiros maliciosos conceitos obtidos do modelo conceitual original [3].

Segundo SPRINGER8, as "Políticas de Privacidade" são definidas por um ou mais "Declarações (Statement)" legais, que por sua vez são relacionadas com formas de "Execução (Enforcement)", "Serviços (Service)" e seus sub-serviços, "Informações Pessoais" às quais as "Declarações" se aplicam (WILEY1) e "Recipientes (Recipient)" e seus sub-recipientes, aos quais as "Políticas de Privacidade" interessam, ou seja, os stakeholders da aplicação. WILEY1 também define a importância da relação entre a "Divulgação" e os "Recipientes" das "Políticas de Privacidade". Ou seja, é de

responsabilidade da “Política de Privacidade” definir a quem interessa e quem vai receber (“Recipient”) a “Divulgação” da “Informação Pessoal” (os novos conceitos estão representados na cor cinza na Fig. 1 e os trabalhos citados de onde eles foram obtidos estão na Tabela 1).

No fragmento do modelo conceitual apresentado na Fig. 2, “Mecanismo de Privacidade” tem uma associação com “Risco (Risk)”, “Confiança (Trust)”, “Contexto (Context)”, “Restrição (Constraint)” (conceitos obtidos do modelo conceitual original [3]) e “Objetivo (Objective) que, se revelada, pode resultar numa violação da privacidade (conceito representado na cor rosa na Fig. 2 e obtido do estudo SS1 da Tabela 1).

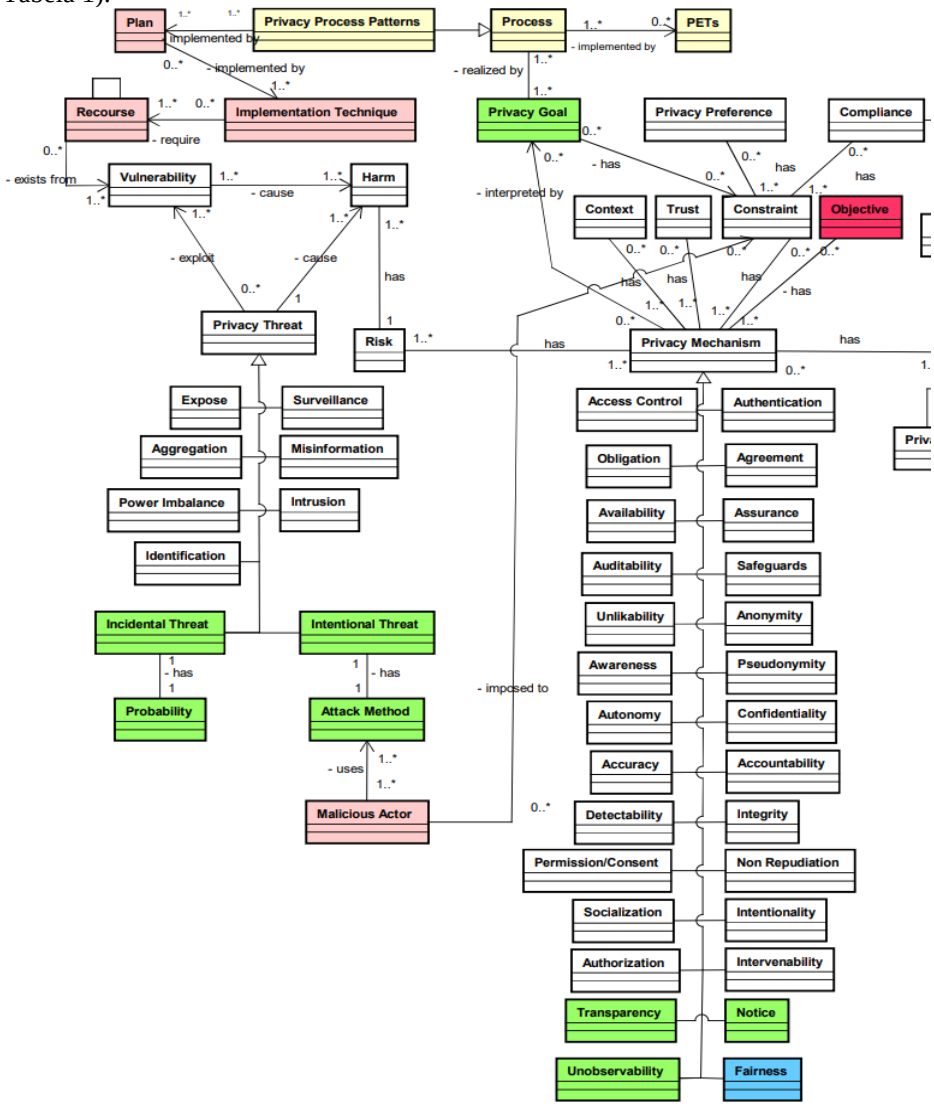

Fig.2. Modelo Conceitual Atualizado.

“Restrição” tem zero ou mais "Preferência de Privacidade (Privacy Preference)" ou "Conformidade de Privacidade (Compliance)" como restrições. A “Conformidade de Privacidade” exige zero ou mais "Políticas de Privacidade" para sua validação (conceitos obtidos do modelo conceitual original [3]). As “Restrições” também são limitadoras opcionais de 0 ou mais “Metas de Privacidade (Privacy Goals) (destacada em verde da parte superior da Fig. 2 e obtida de ARXIV1/SPRINGER12 da Tabela 1). As “Metas de Privacidade” podem ser interpretadas por zero ou mais “Mecanismos de Privacidade”, que podem interpretar zero ou mais “Metas de Privacidade” (o relacionamento novo também foi obtido de ARXIV1/SPRINGER12 da Tabela 1).

“Metas de Privacidade” são realizadas por um ou mais processos “Processos (Process)”, que realizam um ou mais “Metas de Privacidade” e são implementados por “PETs” (Privacy Enhancing Technologies ou Tecnologias de Privacidade). Os “Processos” são especializados por “Padrões de Processo de Privacidade (Privacy Process Patterns)”, que são usados na implementação de aplicações de privacidade (os novos conceitos estão destacados em amarelo na Fig. 2 e os trabalhos de onde eles foram obtidos foram IEEE4/SPRINGER9/EMINS2, os quais estão na Tabela 1).

Ademais, IEEE1 afirma que os “Padrões de Processo de Privacidade (Privacy Process Patterns)” são necessariamente operacionalizados por pelo menos um “Plano (Plan)”, que são implementados por pelo menos uma “Técnica de Implementação (Implementation Technique)”, que por sua vez, requerem pelo menos um “Recurso (Resource)” de privacidade para sua construção (os novos conceitos estão destacados em rosa claro na Fig. 2 e foram obtidos de IEEE1, descrito na Tabela 1).

O modelo conceitual original [3] estabelece que as ameaças à privacidade são violações de "Risco" de privacidade que são susceptíveis de acontecer. Quando as ameaças à privacidade são analisadas em conjunto com o sistema de software existente, as suas falhas podem ser capturadas como preocupações de privacidade. De acordo com o estudo original, a preocupação com a privacidade (indicada como um "Mecanismo de Privacidade") tem de lidar com ameaças que um sistema deve ter em conta para apoiar a privacidade do utilizador final. Por conseguinte, um ou mais "Mecanismos de Privacidade" têm zero ou mais "Ameaças à Privacidade (Privacy Threats)".

Também é possível dividir as “Ameaças à Privacidade” em “Ameaças Incidentais (Incidental Threats)”, possuem uma probabilidade de acontecer sozinhas e “Ameaças Intencionais (Intentional Threats)” que são feitas a partir de um “Método de Ataque (Attack Method)” (conceitos novos em verde na parte esquerda da Fig. 2 e obtidos de ARXIV1/ SPRINGER12, descritos na Tabela 1), executado por um “Ator Malicioso (Malicious Actor)” (conceito destacado em rosa claro e obtido do artigo IEEE1 destalhado na Tabela 1). Segundo IEEE1, “Atores Maliciosos” são limitados apenas por uma ou mais restrições impostas pelo sistema, que podem limitar um ou mais atores.

Segundo o modelo conceitual original [3], "Ameaças de Privacidade" pode ainda ter uma associação com "Vulnerabilidade (Vulnerability)" e "Danos (Harms)", por exemplo, a “Identificação” pode causar danos financeiros. Segundo IEEE1, “Vulnerabilidades” existem a partir de um “Recurso (Resource)”, que podem ser dados privados. Por exemplo, em um sistema de saúde, dados privados de pacientes geram “Vulnerabilidades” para o sistema (relacionamento criado entre Vulnerability e Resource e obtido de IEEE1).

De acordo o modelo conceitual original [3], uma preocupação com a privacidade (aqui indicada como "Mecanismo de Privacidade") pode ser considerada como um objetivo. Portanto, o "Mecanismo de Privacidade" pode ser especializado em "Autenticação (Authentication)", "Autorização (Authorization)", "Anonimato (Anonymity)", "Pseudonimato (Pseudonymity)", "Desligabilidade (Unlinkability)", "Integridade (Integrity)", "Detectabilidade (Detectability)", "Interventividade (Intervenability)", "Confidencialidade (Confidentiality)", "Autonomia (Autonomy)", "Controle de Acesso (Access Control)", "Salvaguardas ou Proteções (Safeguards)", "Consciência (Awareness)", "Abertura (Openness)", "Precisão (Accuracy)", "Acordo (Agreement)", "Obrigação (Obligation)", "Socialização (Socialization)", "Intencionalidade (Intentionality)", "Não Repúdio (Non Repudiation)", "Disponibilidade (Availability)", "Garantia (Assurance)", "Medida (Measure)", "Auditabilidade (Auditability)", "Conflito (Conflict)", "Oportunidade (Opportunity)", "Fraqueza (Weakness)", "Força (Strength)", "Permissão/Consentimento (Permission/Consent)" e "Responsabilização (Accountability)". Segundo ARXIV1 (descrita na Tabela 1), também entram nessa categoria “Transparência (Transparency)”, “Aviso (Notice)”, “Inobservância (Unobservability)” (destacados em verde na Fig. 2). Além desses novos conceitos, “Equidade (Fairness)” (em azul na Fig. 2) também foi obtido de SPRINGER1 (também detalhado na Tabela 1).

5 Limitações da Pesquisa e Ameaças à Validade

Segundo Mendes et al. [12], em atualizações de RSLs sempre há o risco de uma atualização prematura, que não gera conteúdo relevante para o tema. Portanto seguimos a metodologia traçada por Mendes et al. [12] para uma validação de atualização e chegamos a um resultado positivo.

Inerente ao processo de RSL, uma limitação da pesquisa é encontrar todos os artigos relevantes existentes. Buscando contornar esse problema, foram seguidas as recomendações de Wohlin et al. [13] para um levantamento eficaz de artigos relevantes para uma atualização no tema. Ainda assim, há um número considerável de bases não agregadas pelo Google Scholar, ferramenta recomendada pela metodologia adotada, assim como artigos não acessíveis.

Ainda segundo Wohlin et al. [13], deve-se ter cuidado com o viés introduzido pelos pesquisadores na seleção dos para atualização. Para reduzir essa possibilidade, uma das recomendações para a seleção de novos estudos foi realizar a seleção com pelo menos uma dupla de pesquisadores, porém o trabalho corrente teve que se limitar a apenas um pesquisador. Esse risco se torna menos relevante quando se observa o número de artigos novos e a riqueza de novos conceitos/relações achados em um período relativamente curto entre as duas pesquisas.

6 Conclusões e Trabalhos Futuros

Este trabalho buscou destacar quais os trabalhos existentes na área de Linguagens de Modelagem que capturam Privacidade lançados após Peixoto et al. [3], de forma a se

construir um guia em forma de modelo conceitual para auxiliar o desenvolvimento de aplicações que lidam com o tópico. Foi utilizado então a metodologia em Mendes et al. [12] para confirmar, a partir de um fluxograma, que uma atualização de Peixoto et al. [3] traria conteúdo válido e interessante para à comunidade acadêmica.

A fim de obter a resposta para a questão de pesquisa elaborada, foi realizada uma atualização da RSL original, onde 811 estudos foram obtidos a partir de forward snowballing sobre a seleção final de artigos originais e então filtrados buscando aplicar os mesmos critérios de inclusão e exclusão, chegando por fim a 26 estudos considerados relevantes.

Dos 26 artigos finais foram analisadas informações diversas sobre as pesquisas e linguagens de modelagem usadas. Destes, 10 artigos contribuiram diretamente com a extração de novos conceitos para atualizar o modelo conceitual de privacidade original de Peixoto et al. [3]. Foi possível incrementar este modelo conceitual com 21 novos conceitos e suas relações, assim como suas devidas descrições e bases teóricas. Esse aumento não veio apenas na forma de expansão nas dimensões já existentes no modelo conceitual original, mas também possibilitou a criação de novas dimensões de conceitos de privacidade. Por exemplo, foi criada uma nova dimensão que especifica a construção de "Políticas de Privacidade". Isso faz alusão à vivacidade da pesquisa em Conceitos de Privacidade, já que muitas das pesquisas novas achadas tinham modelos com dimensões não presentes na seleção original de Peixoto et al. [3].

De fato, há vários trabalhos focados em privacidade que poderiam ter sido considerados, mas neste estudo nós só consideramos trabalhos que envolvam linguagens de modelagem, para seguir o protocolo da RSL original [3], conforme as diretrizes de Wohlin et al. [13] para atualizar RSLs. Um primeiro trabalho futuro seria realizar uma nova atualização considerando trabalhos publicados a partir do 2o semestre de 2020 e também projetar um sistema de atualização automática do modelo conceitual, dada a importância do tema.

Pesquisas futuras podem ampliar o escopo da pesquisa apresentada, realizando uma RSL com foco em trabalhos que abordem privacidade e leis de proteção de dados, sem obrigatoriamente estar usando uma linguagem de modelagem de requisitos. Uma possível análise desses trabalhos poderia promover o entendimento de como os conceitos do modelo conceitual de privacidade se relacionam com leis de proteção de dados específicas e são implementados na prática. Além disso, relatos técnicos poderiam ser analisados para enriquecer o modelo utilizando grey literature. Por fim, um estudo também poderia ser feito para analisar qual a melhor abordagem para modelar conceitos de um domínio.

Agradecimentos. Este trabalho foi parcialmente apoiado pelo CNPq, CAPES e FACEPE.

Referências

1. Gharib, M., Salnitri, M., Paja, E., Giorgini, P., Mouratidis, H., Pavlidis, M., ... & Della Siria, A. (2016). Privacy requirements: findings and lessons learned in developing a privacy platform. In 2016 IEEE 24th International Requirements Engineering Conf. (RE) (pp. 256-265). IEEE.

2. Hong, J. I., Ng, J. D., Lederer, S., & Landay, J. A. (2004, August). Privacy risk models for designing privacy-sensitive ubiquitous computing systems. In Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques (pp. 91-100).
3. Peixoto, M. M., Silva, C., Maia, H., & Araújo, J. (2020). Towards a Catalog of Privacy Related Concepts. In REFSQ Workshops.
4. Kalloniatis, C., Kavakli, E., & Gritzalis, S. (2008). Addressing privacy requirements in system design: the PriS method. Requirements Engineering, 13, 241-255.
5. Mouratidis, H., & Giorgini, P. (2007). Secure Tropos: a security-oriented extension of the Tropos methodology. International Journal of Software Engineering and Knowledge Engineering, 17(02), 285-309.
6. Giorgini, P., Massacci, F., Mylopoulos, J., & Zannone, N. (2006). Requirements engineering for trust management: model, methodology, and reasoning. International Journal of Information Security, 5, 257-274.
7. Gharib, M., Giorgini, P., & Mylopoulos, J. (2017). Towards an ontology for privacy requirements via a systematic literature review. In Conceptual Modeling: 36th International Conference, ER 2017, Valencia, Spain, November 6–9, 2017, Proceedings 36 (pp. 193-208). Springer International Publishing.
8. Beckers, K. (2012, August). Comparing privacy requirements engineering approaches. In 2012 Seventh International Conference on Availability, Reliability and Security (pp. 574-581). IEEE.
9. Deng, M., Wuyts, K., Scandariato, R., Preneel, B., & Joosen, W. (2011). A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. Requirements Engineering, 16(1), 3-32.
10. Hadar, I., Hasson, T., Ayalon, O., Toch, E., Birnhack, M., Sherman, S., & Balissa, A. (2018). Privacy by designers: software developers' privacy mindset. Empirical Software Engineering, 23, 259-289.
11. Omoronyia, I., Cavallaro, L., Salehie, M., Pasquale, L., & Nuseibeh, B. (2013, May). Engineering adaptive privacy: on the role of privacy awareness requirements. In 2013 35th International Conference on Software Engineering (ICSE) (pp. 632-641). IEEE.
12. Mendes, E., Wohlin, C., Felizardo, K., & Kalinowski, M. (2020). When to update systematic literature reviews in software engineering. Journal of Systems and Software, 167, 110607.
13. Wohlin, C., Mendes, E., Felizardo, K. R., & Kalinowski, M. (2020). Guidelines for the search strategy to update systematic literature reviews in software engineering. Information and software technology, 127, 106366.
14. Agostinelli, S., Maggi, F. M., Marrella, A., & Sapio, F. (2019). Achieving GDPR compliance of BPMN process models. In Information Systems Engineering in Responsible Information Systems: CAiSE Forum 2019, Rome, Italy, June 3–7, 2019, Proceedings 31 (pp. 10-22). Springer International Publishing.
15. Labda, W., Mehandjiev, N., & Sampaio, P. (2014). Modeling of privacy-aware business processes in BPMN to protect personal data. In Proceedings of the 29th Annual ACM Symposium on Applied Computing (pp. 1399-1405).
16. SOLOVE, D. J. A taxonomy of privacy. U. Pa. L. Rev., v. 154, p. 477, 2005.
17. Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering version 2.3. Engineering, 45(4ve), 1051.

18. Cruzes, D. S., & Dyba, T. (2011, September). Recommended steps for thematic synthesis in software engineering. In 2011 International Symposium on Empirical Software Engineering and Measurement (pp. 275-284). IEEE.
19. Netto, D., Peixoto, M. M., and Silva, C. (2019). Privacy and Security in Requirements Engineering: Results from a Systematic Literature Mapping. In Anais do WER19 - Workshop em Engenharia de Requisitos, Recife, PE, Brasil, Editora PUC-Rio.
20. Calazans, A. T. S., & Jefferson, A. (2020) "Empathy and Criativity in Privacy Requirements Elicitation: Systematic Literature". Anais do WER20 - Workshop em Engenharia de Requisitos, São José dos Campos, SP, Brasil, Editora PUC-Rio.
21. Valença, G., Sarinho, M., Polito, V., & Lins, F. (2022). Do platforms care about your child's data a proposal of legal requirements for children's privacy and protection. In Anais do WER22 - Workshop em Engenharia de Requisitos, Natal, Brasil, Editora PUC-Rio.
22. Ferrão, S. É. R., & Canedo, E. D. (2022). Uma taxonomia para requisitos de privacidade e sua aplicação no Open Banking Brasil. In Anais do WER22 - Workshop em Engenharia de Requisitos, Natal, Brasil, Editora PUC-Rio.
23. Santana, E., Vilela, J., & Peixoto, M. M. (2022). Diretrizes para apresentação de políticas de privacidade voltadas à experiência do usuário. In Anais do WER22 - Workshop em Engenharia de Requisitos, Natal, Brasil, Editora PUC-Rio.
24. Terra, A., Vilela, J., & Peixoto, M. (2022). A catalog of quality criteria to guide the assessment of applications' privacy policies. In Anais do WER22 - Workshop em Engenharia de Requisitos, Natal, Brasil, Editora PUC-Rio.
25. Araújo, E., Vilela, J., Silva, C., & Alves, C. (2021). Are my business process models compliant with LGPD? The LGPD4BP method to evaluate and to model LGPD aware business processes. In XVII Brazilian Symposium on Information Systems (pp. 1-9), SBC.
26. Peixoto, M., Ferreira, D., Cavalcanti, M., Silva, C., Vilela, J., Araújo, J., & Gorschek, T. (2023). The perspective of Brazilian software developers on data privacy. Journal of Systems and Software, 195, 111523.