Ensino da Adequação à LGPD no Desenvolvimento de Software através da Aprendizagem Ativa e Centrada no Discente

Juliana Saraiva, Juliana Araújo, Sérgio Soares
{julianajags, julianaaraujo}@dcx.ufpb.br, scbs@cin.ufpe.br
Departamento de Ciências Exatas, Universidade Federal da Paraíba (UFPB)
Centro de Informática, Universidade Federal de Pernambuco (UFPE)

RESUMO

As abordagens de ensino-aprendizagem nos cursos de Engenharia de Software precisam aproximar a teoria da prática e assim, metodologias ativas e centradas no discente vêm sendo propostas e avaliadas. Ademais, demandas legislativas impõem, multidisciplinarmente, teorias, métodos e técnicas que abordem o desenvolvimento de software seguro, conforme prevê a LGPD que exige que o desenvolvimento de produtos e soluções de software considerem a privacidade de dados pessoais desde a concepção e incorpore esta prática durante toda a vida do software: Princípios de *Privacy by Design* e *Privacy by Default*. Este trabalho realizou um experimento com 142 discentes de graduação que produziram 947 artefatos ágeis – Estórias de Usuário e Cenários BDD - a partir do Inventário de Dados LGPD. Foi possível concluir que a Aprendizagem Significativa, o Pensamento Computacional e a Aprendizagem baseada em Problemas demonstraram-se apropriadas no ensino de requisitos legais para a Engenharia de Software.

PALAVRAS-CHAVE

LGPD, Engenharia de Software Seguro, Metodologia Ativa, Requisitos Ágeis

1 INTRODUÇÃO

A distância entre o que é exposto em aulas teóricas e a aplicabilidade do conhecimento técnico adquirido pelos alunos nos cursos de graduação não é nenhuma novidade quando se trata de Ensino Superior no Brasil e no mundo. Essa realidade é observada em todas as áreas de conhecimento, não sendo diferente na Ciência da Computação, e consequentemente, na Engenharia de Software.

Por mais que o tema já tenha sido amplamente debatido na academia e na indústria, reportado em diversas produções científicas, o problema ainda parece persistir. Estudos apontam que este cenário, muitas vezes, tem como causas

(i) a insistência no ensino tradicional moldado em aulas unicamente expositivas, (ii) componentes curriculares contendo uma vasta carga de conteúdo teórico, que acaba por não ser fixada pelos alunos e (iii) por ausências de iniciativas mais práticas que desenvolvam nos alunos habilidades de resolução de problemas reais [1]. Com frequência, o resultado dessas abordagens tradicionais no ensino da Engenharia de Software é insuficiente no dia a dia da indústria.

A dinâmica frenética e as transformações sociais e computacionais cotidianas, demandam da(o) engenheira(o) de software, habilidades técnicas e não técnicas, algumas vezes não aprimoradas nos modelos tradicionais de ensino. Complementarmente, o acontecimento da pandemia de COVID-19 acabou impulsionando a adoção de abordagens como Metodologias Ativas, Sala de Aula Invertida e Aprendizagem Significativa, a fim prover o desenvolvimento de *hard* e *soft skills* num cenário de ensino híbrido e remoto [2] [3] [4] [5] [6].

Algumas vezes, as grades curriculares dos cursos de Ciência da Computação e áreas afins, condensam o amplo conteúdo de Engenharia de Software em uma ou poucas disciplinas. Como resultado, força professores a ensinarem o conteúdo em tempo recorde, implementando recursos educacionais obsoletos, faltando inovação tecnológica e de pesquisa nessa abordagem [7]. Assim, habilidades demandadas pelo mercado de trabalho volátil, proativo, empreendedor e mutável, acabam sendo insuficientes para a tal exigência.

Neste contexto, num mundo embasado em tecnologia, coleta e processamento massivo de dados (incluindo os pessoais), com uso de inteligência artificial, a Cibersegurança é um fator crítico [8]. Levando em consideração que esse tipo de dados faz parte da identidade e autodeterminação dos sujeitos, regulamentos sobre Privacidade e Proteção de Dados Pessoais vêm sendo propostos ao redor do mundo, como é o caso da GDPR (*General Data Protection Regulation*) na Europa [9] e a LGPD (Lei Geral de Proteção de Dados) no Brasil [10].

Essas duas legislações impõem que o desenvolvimento de produtos e soluções de software considerem a privacidade de dados pessoais desde a concepção e incorpore esta prática durante toda a vida do software: Princípios de *Privacy by Design* e *Privacy by Default* [11]. Esta obrigação legal impacta diretamente no processo de desenvolvimento de software, que precisa se atentar à privacidade e proteção de dados pessoais

EduComp24, Abril 22-27, 2024, São Paulo, São Paulo, Brasil (On-line)

durante todo desenvolvimento software, trazendo um novo desafio para a Engenharia de Software [12] [13].

Ao longo dos anos, estudos apontam que muitos incidentes de Cibersegurança cresceu, causando desde penalidades monetárias até perdas de vidas humanas [14]. E muitos desses incidentes tiveram como causa principal a baixa qualidade de código e consequente vulnerabilidades. Portanto, o desenvolvimento de sistemas seguros são cada vez mais requeridos, demandando atenção e cuidados adicionais no processo de desenvolvimento de software.

Não só no meio educacional, mas também no âmbito industrial, a transferência de conhecimento referente à Cibersegurança vem impondo esforços às(aos) engenheiras(os) de software. Estudos relatam técnicas de aplicação de jogos sérios para compartilhamento de recomendações e boas práticas de codificação segura como uma das alternativas [15]. Entretanto, resultados indicam a falta de capacitação e experiência prática desses profissionais, exigindo novos paradigmas de ensino da Engenharia de Software. Além disso, a LGPD, o Código Penal e o Código Civil dispões claramente sobre a possibilidade de responsabilidade da(o) engenheira(o) de software caso haja um incidente de segurança ou se os dados pessoais foram utilizados indevidamente, nos campos administrativos, penais e civis, respectivamente.

Num cenário onde o ensino da Engenharia de Software é desafiado a se adaptar às demandas impostas por requisitos legais, como é o caso da LGPD, já em vigor no Brasil, o **problema de pesquisa** deste trabalho é a implantação iminente do tema "segurança de dados pessoais", conforme prevê a LGPD, como conteúdo obrigatório em componentes curriculares relacionados à Engenharia de Software. Portanto, o objetivo deste trabalho é analisar se as técnicas de ensino-aprendizagem de Metodologias Ativas, Aprendizagem Significativas e Pensamento Computacional podem ser adotadas para introduzir o conteúdo de Proteção e Privacidade de Dados Pessoais LGPD no contexto do ensino da Engenharia de Software.

Assim, as seguintes Questões de Pesquisa (QP) foram formuladas: **QP01 -** A Aprendizagem Significativa auxilia o ensino da Engenharia de Software, quando há introdução de novos temas acerca de desenvolvimento de software seguro, conforme prevê a LGPD? **QP02 -** A Metodologia Ativa auxilia o ensino da Engenharia de Software, quando há introdução de novos temas acerca de desenvolvimento de software seguro, conforme prevê a LGPD? **QP03 -** O Pensamento Computacional auxilia o ensino da Engenharia de Software, quando há introdução de novos temas acerca de desenvolvimento de software seguro, conforme prevê a LGPD?

Este trabalho está organizado em 4 seções incluindo esta. A Seção 2 apresenta o detalhamento sobre quais artefatos ágeis estão sendo analisados neste estudo. Os trabalhos relacionados encontram-se na Seção 3, enquanto a Seção 4 descreve a metodologia adotada. Já a Seção 5 apresenta e discute os resultados. Por fim, as considerações finais e possibilidades de trabalhos futuros estão evidenciados na Seção 6.

2 ARTEFATOS DA ENGENHARIA DE SOFTWARE ÁGIL

O Processo de Engenharia de Requisitos pode ser compreendido como um conjunto de atividades que lida com a identificação, análise, documentação e gerenciamento dos requisitos de um sistema ou software [16]. Neste contexto, diferentes modelos de processo adotam procedimentos distintos para sua realização, gerando, consequentemente, diferentes artefatos.

Neste cenário, as Metodologias Ágeis, há muito tempo, destacam-se no mercado de software como sendo algumas das mais proeminentemente adotadas [17], por causa da resposta rápida às mudanças, maior envolvimento do cliente, melhoria contínua, redução de riscos e foco na qualidade dos produtos gerados. Assim, este trabalho foca no estudo de dois dos artefatos gerados no processo de Engenharia de Requisitos abordado por algumas metodologias ágeis: Estórias de Usuário e Cenários BDD (Behavior Driven-Development).

A criação das Estórias de Usuário é uma técnica utilizada na Engenharia de Software que descreve requisitos funcionais vistos sob a ótica do usuário. Esta técnica é uma maneira concisa e simples de comunicar as necessidades e expectativas dos usuários em relação a um sistema ou software. Sendo elas escritas em linguagem natural, fornecem uma descrição breve e centrada na funcionalidade desejada pelo usuário, normalmente definidas seguindo uma sintaxe básica, através de 3 (três) cláusulas:

1. Eu como <ator/usuário/papel/função>
2. Quero <atividade/serviço - objetivo/meta>
3. Para <resultado/finalidade do processamento>

Já a construção dos Cenários BDD é uma estratégia usada no desenvolvimento de software para descrever o comportamento de um software ou funcionalidade, em termos do que se é observável [18]. Essa abordagem se concentra na colaboração entre desenvolvedores, analistas de negócios e partes interessadas para garantir um entendimento comum dos requisitos e comportamentos esperados. A construção de Cenários BDD auxiliam os engenheiros de software a validar se o sistema está funcionando corretamente e se atende aos requisitos esperados. Assim como as estórias, são escritos em linguagem natural apresentados de acordo com a estrutura a semelhante a seguinte:

1. Dado <pré-condições/contexto>
2. Quando <um (ou mais) evento ocorrer>
3. Então <saída obtida/comportamento esperado>

Para a construção das Estórias de Usuário e Cenários BDD várias informações precisam ser coletadas pelos engenheiros de software. Sob essa perspectiva, os Requisitos Funcionais

(RF) normalmente são descritos através das Estórias, enquanto os Cenários BDD, focam no comportamento das funcionalidades, descrevendo Requisitos Não-Funcionais (RNF) e de Domínio. É importante destacar que os RNF normalmente são mais complexos de compreender, elicitar, documentar e validar, uma vez que possuem uma natureza mais abstrata, são mais conflituosos entre si para atender e de difícil medição [19] [20]. Visando auxiliar o processo de Engenharia de Requisitos, este estudo busca mapear elementos do IDP com as cláusulas-padrão que determinam as estruturas de construção das Estórias de Usuário e Cenários BDD, mostrados anteriormente.

3 TRABALHOS RELACIONADOS

A junção de estudos nas Áreas de Engenharia de Software e Segurança da Informação não sendo cada vez mais demandada nos últimos anos, uma vez que leis estão impondo a implantação de medidas de segurança às empresas de tecnologia para proteger a privacidade de dados pessoais. A PCM Tool: *Privacy Requirements Specification in Agile Software Developme*nt, por exemplo, foi proposta como ferramenta para dar suporte à especificação de software [21]. Os autores pontuaram que os desenvolvedores de software possuem uma dificuldade em especificar requisitos de privacidade. Assim, eles propuseram essa ferramenta visando dar suporte a um guia, também proposto por eles, de especificação de requisitos de privacidade no desenvolvimento de software ágil.

Dando continuidade a este trabalho, [22] investigaram o nível de compreensão dos desenvolvedores de software com relação à privacidade. Foram analisados fatores pessoais, comportamentais e do ambiente que poderiam influenciar no processo de tomada de decisão quando os requisitos de privacidade devem ser modelados. Através da replicação de entrevistas semiestruturadas com 30 engenheiros, 9 fatores pessoais, 5 comportamentais e 7 externos foram considerados relevantes para delimitar a compreensão e especificação dos requisitos de privacidade.

Já um estudo prático sobre a implantação da LGPD no processo de desenvolvimento de software é o de Nardelli e colaboradores (2021) [23]. Nesta pesquisa, o foco foi expor e analisar práticas implementadas por uma empresa de desenvolvimento de software, especificamente no âmbito da Segurança da Informação de acordo com o que recomenda a Lei Geral de Proteção de Dados (LGPD). Teve como objetivo compreender as abordagens empregadas durante a concepção e manutenção de aplicações seguras. 42 Práticas de segurança foram identificadas sob três eixos principais: Técnico, Cultural/Pessoal e Jurídico.

Já em 2022, o mesmo autor realizou um estudo em uma empresa que desenvolve Sistemas de Gestão de Governo Eletrônico [24]. O resultado do trabalho levou à elaboração de um modelo temático que se fundamenta em eixos principais: Técnico, Cultural/Pessoal e Jurídico. No escopo deste artigo, foram discutidas as primeiras 20 práticas analisadas, visando fornecer uma contribuição para outras instituições, tanto públicas quanto privadas, que buscam adotar práticas de segurança eficazes no ambiente de desenvolvimento de software.

Visando auxiliar o processo de compreensão da LGPD pelos desenvolvedores, Bertan et al. [25 também propôs um GitBook com os principais conceitos trazidos pela LGPD. Com este trabalho eles objetivaram capacitar os profissionais de tecnologia da informação a aderir eficazmente às diretrizes da LGPD ao longo do ciclo de desenvolvimento de software. Como resultado um Objeto de Aprendizagem foi concebido, proporcionando uma ferramenta eficaz e prática para o alinhamento do desenvolvimento de software às exigências legais da LGPD.

Melo e colaboradores [26] realizaram uma pesquisa bibliográfica sobre elementos do Scrum que podem ser utilizados para auxiliar as organizações na implantação da LGPD. Os autores alegam que adoção de Scrum proporciona maior agilidade e adaptabilidade às exigências legais. Além da LGPD, o investigou a importância de outras leis brasileiras, como o Marco Civil da Internet. Como resultado, os autores propuseram uma Metodologia Scrum como uma abordagem eficiente e versátil para a implantação e promoção da conscientização sobre a LGPD nas organizações.

Na mesma direção, o trabalho [27] desenvolveram um trabalho objetivando apresentar um guia prático para que as empresas brasileiras consigam se adequar a LGPD. A metodologia SGPD – Sistema de Gestão de Proteção de Dados foi apresentada para gerenciamento de processos, implementação e governança de privacidade e proteção de dados. Como conclusão, foi observado que o framework Scrum demonstrou robustez em ferramentas e técnicas que facilitam o gerenciamento e implantação de cada uma das cinco fases do SGPD serve como guia de boas práticas para atingir o objetivo do projeto de implementação da LGPD.

Também no mesmo ano, [28] propuseram o G-Priv: Um Guia para Apoiar a Especificação de Requisitos de Privacidade em Conformidade com a LGPD. Eles apontaram as dificuldades para extrair e operacionalizar requisitos de privacidade de dados pessoais. Após a realização de um survey com 18 profissionais, como resultado, propuseram um catálogo de padrões de privacidade e um guia G-Priv, para auxiliar a especificação de requisitos de privacidade em conformidade com a LGPD.

É possível observar como soluções, técnicas e ferramentas adotadas no processo de desenvolvimento de software podem facilitar a garantia da proteção e privacidade dos dados pessoais. Entretanto, estudos que analisem se os artefatos gerados pelos analistas em Segurança da Informação também podem ser adotados no processo de desenvolvimento do software ainda são demandados. Assim, o trabalho aqui proposto visa analisar se é possível utilizar o Inventário de Dados Pessoais (IDP) como ferramenta para dar suporte ao

processo de Engenharia de Requisitos, especificamente focando na criação de Estórias de Usuário e Cenários BDD.

É importante pontuar, que apesar dos trabalhos mencionados tratarem especificamente da LGPD, que é o regulamento brasileiro que lida com privacidade de dados pessoais, a tendência mundial de proteção dos dados apresenta-se de forma semelhante. Inclusive, a própria LGPD é uma inspiração da GDPR, vigente na Europa. Ressalta-se que a relevância desses trabalhos, assim como nosso, reforça a necessidade de comunicação de trabalho multidisciplinar entre as áreas de Engenharia de Software e Segurança da Informação, na garantia da proteção e privacidade de dados pessoais, conforme impões as leis e regulamentos ao redor do mundo.

4 METODOLOGIA

4.1 Descrição do Estudo e Aplicação de Técnicas de Ensino-Aprendizagem

Um experimento controlado foi realizado com alunos de graduação em universidades públicas federais no Brasil. O único pré-requisito na seleção dos sujeitos foi que eles estivessem cursando ou já cursado a disciplina de Engenharia de Software e/ou Engenharia de Requisitos na graduação. O experimento teve como objetivo avaliar se o Inventário de Dados Pessoais (IDP), documento obrigatório exigido pela LGPD, a ser construído pelas instituições que fornecem bens ou serviços no Brasil, poderia ser utilizado no processo de desenvolvimento de software, especificamente na Engenharia de Requisitos.

O IDP é um artefato que toda a instituição, pública ou privada, precisa construir e manter, caso realize tratamento de dados pessoais. Ele deve conter o registro das operações de tratamento de dados pessoais, conforme prevê o art. 37 da LGPD. Segundo esta lei, "*tratamento*" é toda operação realizada com dados pessoais, a citar: coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração.

Uma vez que a LGPD não explicita quais informações precisam estar registradas no inventário, a Autoridade Nacional de Proteção de Dados (ANPD) publicou um Guia de Elaboração de Inventário de Dados Pessoais com o objetivo de auxiliar esta construção. Ademais, ressalta-se que a criação do IDP, objetiva-se atender não só a LGPD, mas também outras normas vigentes sobre o tema de privacidade e segurança da informação. Normalmente, o IDP é construído por especialistas LGPD em conjunto com analistas de segurança da informação, e mantido por encarregados(as) de dados pessoais, sendo esta a pessoa indicada pela instituição para atuar como canal de comunicação entre o controlador e os titulares dos dados, garantir o Programa de Privacidade de dados e responder junto à ANPD.

Não limitadas a essas, o IDP, obrigatoriamente, precisa conter: Processo/Tratamento de dados pessoais; Atores envolvidos (agentes de tratamento); Titular dos dados pessoais; Finalidade do tratamento; Hipótese legal que embasa o tratamento (arts. 7º e 11 da LGPD); Dados pessoais tratados;

Categoria dos dados pessoais; Tempo de retenção dos dados pessoais; Com quais instituições os dados pessoais são compartilhados; Como é feita a transferência internacional de dados; e Medidas de segurança da informação são adotadas.

Para realização deste estudo, os passos metodológicos ilustrados na Figura 1 foram seguidos. Primeiramente foi realizada uma aula expositiva abordando a LGPD e revisão conteúdo de técnicas de criação de Estórias de Usuário e Cenários BDD (*Behavior-Driven Development*). Esta etapa teve o propósito de introduzir novos conceitos para aqueles que estavam vendo pela primeira vez estes conteúdos, ou aprimorar, para o caso daqueles que já haviam estudado esses temas. Em seguida, em cada turma, 3 (três) grupos foram divididos aleatoriamente, onde o GRUPO 01: teve acesso à Descrição do Sistema, o GRUPO 02: teve acesso ao Inventário de Dados Pessoais (IDP) e o GRUPO 03: teve acesso, simultaneamente, à Descrição do Sistema e ao IDP. Todos os artefatos utilizados no estudo estão disponíveis [1] para eventuais replicações do estudo.

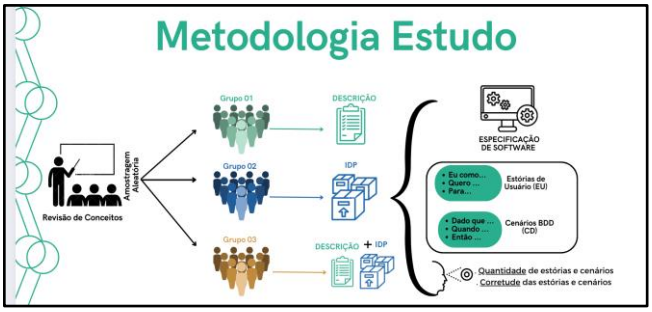

Figura 1: Passos Metodológicos do Estudo

O estudo consistiu na criação de Estórias de Usuário e Cenários BDD pelos alunos de graduação de Engenharia de Software, dos cursos de Ciência da Computação, Ciência de Dados e Inteligência Artificial, e Tecnologia em Análise e Desenvolvimento de Sistemas. É importante ressaltar que cada aluno construiu individualmente seus artefatos. Ao final, eles enviaram esses artefatos, de forma anonimizada para os autores deste trabalho. Além disso, responderam a um questionário de avaliação da atividade.

Neste contexto, a "Aprendizagem Significativa" pode ser compreendida quando o aluno é capaz de relacionar o novo conhecimento com conceitos relevantes já existentes e integrar o novo conteúdo ao seu cognitivo, tornando-o mais compreensível e retido por um período mais longo [29].

[1] https://www.julianasaraiva.net/pesquisas

Portanto, a introdução do conteúdo de desenvolvimento seguro, contemplando a privacidade e proteção de dados pessoais, foi introduzida aos alunos através do estabelecimento da relação os novos conceitos trazidos pela LGPD e o conhecimento prévio que tinha sobre atributos de qualidade de software, construção de Estórias de Usuário e Cenários BDD.

A "Metodologia Ativa" foi abordada através do PBL (*Problem-Based Learning* - Aprendizagem Baseada em Problemas), onde os discentes foram apresentados a situações-problema, relevantes para a área de estudo em questão, e ao invés de receberem informações prontas, são encorajados a investigar, analisar, discutir e resolver o problema [30]. Para isto, durante a execução do experimento, eles receberam descrições de um software real, o IDP e instruções de execução da atividade. A partir desses documentos construíram Estórias de Usuário e Cenários BDD, contemplando, quando possível, requisitos de segurança.

Adotou-se um cenário fictício de um Software de Controle Acadêmico (SCA), focando nas funcionalidades de "Cadastro de Aluno" e "Emissão de Histórico Escolar". Este cenário foi escolhido porque os discentes de graduação possuem, inevitavelmente, conhecimento e experiência de uso neste ambiente. Desta forma, mesmo que esses alunos utilizassem diferentes SCA, há uma familiaridade dessas funcionalidades, tornando mais fácil a compreensão sobre um possível cenário real de aplicação.

Levando em consideração que a maioria dos alunos estavam tendo o primeiro contato com os conceitos da LGPD e com o próprio IDP, para facilitar o processo de aprendizagem e execução das atividades do experimento, um terceiro paradigma de ensino foi adotado: o "Pensamento Computacional". Ele pode ser compreendido como uma abordagem mental e uma habilidade cognitiva que envolve a resolução de problemas, o raciocínio lógico e a capacidade de decompor um problema complexo em partes menores, buscando soluções eficientes e sistêmicas [31]. Para isso, é necessário desenvolver quatro habilidades principais: (i) Decomposição de Problemas, (ii) Reconhecimento de padrões, tendências e regularidades, (iii) Abstração e (iv) Projeto de algoritmos, construindo uma sequência de passos a serem realizados.

Assim, foi desenvolvido pelos autores deste trabalho, um mapeamento entre os elementos do IDP que podem ser utilizados no processo de construção das Estórias de Usuário e Cenários BDD - Figura 2. Este mapeamento contém a relação entre os artefatos de Segurança da Informação e Engenharia de Software [32] Desta forma, os alunos puderam abstrair as informações desnecessárias para a construção dos artefatos de requisitos contidas no IDP, e seguindo um conjunto de passos, construir de forma mais eficiente os artefatos de Engenharia de Software exigidos no experimento.

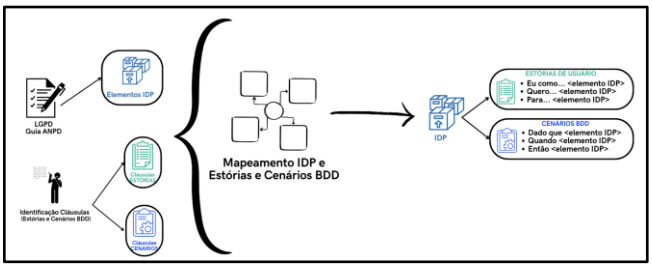

Figura 2: Passos Metodológicos do Estudo

Uma vez que a LGPD ainda é uma lei bastante principiológica, há a necessidade de complementos normativos para expressar de forma clara como ela deve ser aplicada. A ANPD vem publicando, há alguns anos, vários documentos como o Guia de Elaboração de Inventário de Dados Pessoais contendo instruções práticas acerca da implementação da LGPD. Assim, neste trabalho, ele foi a maior fonte de extração de dados para esta atividade.

Apesar da ANPD não exigir um formato para sua construção, ele é elaborado geralmente em planilhas eletrônicas. Por mais simples ou pouca a quantidade de atividades de tratamento de dados pessoais numa instituição de desenvolvimento de software, essa planilha contém várias linhas (centenas ou milhares) - representando cada um dos dados pessoais utilizados em cada atividade – e várias colunas (dezenas ou centenas), contendo as informações (jurídicas e tecnológicas) a serem registradas sobre o dado pessoal, de acordo com a exigência da LGPD – os elementos do IDP.

Apesar do IDP contém informações que subsidiem a construção das Estórias e Cenário, sua utilização pode ser dificultada pelo seu tamanho, complexidade e mistura de informações técnicas jurídicas e de segurança da informação. Consequentemente, esquematizar onde os atores, descrição da funcionalidade, pré-condições e finalidade de execução podem ser encontradas no inventário, torna essa tarefas mais eficiente e rápida. Portanto, o mapeamento visa dar suporte à resolução do problema inspirado na forma como os computadores resolveriam - Pensamento Computacional (i) decompondo os problemas, (ii) abstraindo informações do IDP não úteis para a construção dos artefatos em Engenharia de Software, (iii) gerando pensamento algorítmico, e (iv) reconhecendo padrões a serem seguidos ao adotar o IDP.

O IDP também descreve os processos do negócio que realizam o tratamento de dados pessoais, além do ciclo de vida desses dados quando tratados. A Figura 3 ilustra como os dados contidos no IDP podem ser utilizados na elaboração das Estórias de Usuário.

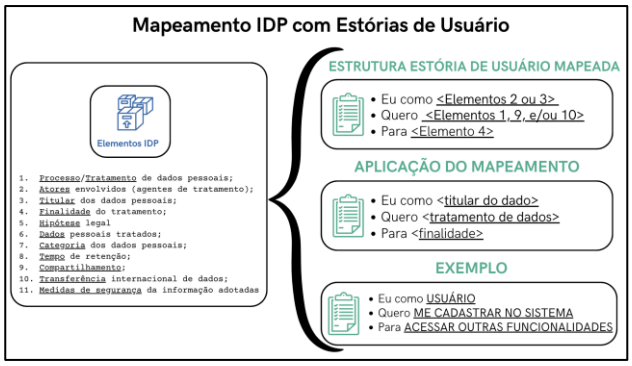

Figure 3: Mapeamento entre IDP e Estórias de Usuário

A funcionalidade que precisa ser descrita na Estória pode ser encontrada no IDP através da descrição do Processo de Negócio ou da descrição Tratamento de Dados (Elemento 1), da descrição do Compartilhamento – Interno, Externo ou Internacional (Elementos 9 e 10). Por fim, a finalidade da execução da funcionalidade descrita na Estória de Usuário pode ser obtida através da finalidade de tratamento descrita no IDP (Elemento 4).

Analogamente à construção das Estórias de Usuário, para construir os Cenários BDD, os engenheiros de software podem utilizar várias informações contidas no IDP. A Figura 4 ilustra qual o resultado do mapeamento entre os artefatos de Segurança da Informação (IDP) e de Engenharia de Software (Cenários BDD).

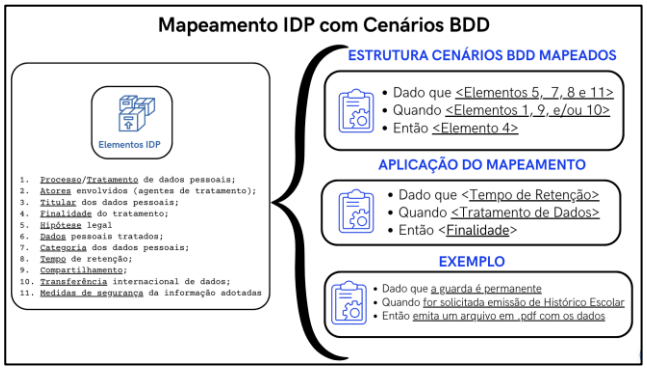

Figure 4: Mapeamento entre IDP e Cenários BDD

Através da Figura 4 é possível compreender o contexto em que a funcionalidade será executada, ou seja, as pré-condições de operação, pode ser facilmente encontrado através das informações contidas no IDP sobre a Hipótese Legal, que autoriza o controlador a realizar o tratamento (Elemento 5), a Categoria do dado pessoal – Comum ou Sensível (Elemento 7), o Tempo em que aquele dado pode ser retido, e consequentemente tratado (Elemento 9) e as Medidas de Segurança que precisam ser implantadas para o tratamento (Elemento 11). É importante deixar claro não há obrigatoriedade da presença de todos esses elementos na descrição do contexto, dependendo do caso concreto de funcionalidade a ser disponibilizada pelo software.

Para a cláusula do Cenário BDD que determina a execução da funcionalidade (<quando>), as informações sobre os diferentes tipos de tratamento e compartilhamento de dados pessoais podem ser encontradas através dos Elementos 1, 9 ou 10 do IDP. Novamente ressalta-se que todos esses elementos não estarão obrigatoriamente presentes nos Cenários BDD, ficando dependente do contexto em que a funcionalidade será executada. Por fim, a finalidade do tratamento (Elemento 4) pode ser utilizada para descrever o que acontecerá com o software ou com processo de negócio após a execução da funcionalidade (fluxo principal e fluxos alternativos).

Este mapeamento visa facilitar o processo de Engenharia de Requisitos, fazendo com que requisitos não-funcionais, mais especificamente, de segurança da informação, ou de domínio, relacionado às hipóteses legais que permitem ou proíbam as empresas de software tratarem dados pessoais, sejam elicitados, documentados, validados e testados.

Ademais, medidas de segurança à nível de camada aplicação, onde geralmente encontra-se o desenvolvimento dos software numa estrutura computacional de distribuição de serviço tecnológico, serão levadas em consideração de forma mais sistemática. Isso pode acontecer devido à exigência da implantação dessas medidas conforme prevê o IDP, devendo ser abordadas explicitamente no processo de Engenharia de Requisitos – Princípio de *Privacy by default*.

4.2 Coleta de Dados

A coleta se deu através do envio dos artefatos de Engenharia de Software gerados para o e-mail de um dos autores deste trabalho. É importante ressaltar que os dados foram anonimizados, sem ser possível a identificação exata do discente que participou do estudo. Complementarmente, os alunos responderam ao questionário de avaliação do experimento online através do Google Forms, também sem coleta de dados pessoais dos discentes. As respostas do questionário foram persistidas em planilhas eletrônicas para futuras análises.

O Quadro 1 sumariza as métricas e indicadores apresentados através da abordagem GQM [33] escolhidos com o intento de responder às QPs elencadas. A primeira coluna indica qual métrica foi coletada, enquanto a segunda coluna apresenta o objetivo da coleta da métrica - descrição da métrica. A pergunta guia para gerar possíveis valores de referência para as métricas (última coluna), está apresentada na terceira coluna.

Para as métricas PCEU e PCCB foi observado se todos os alunos escreveram ao menos 1 EU e 1 CB para cada funcionalidade proposta. Levando em consideração que as Estórias de Usuário (EU) visam expressar as necessidades do usuário, elas precisam abordar o principal serviço a ser entregue pela funcionalidade do software. A Qualidade das EU

e dos CB se caracterizou por expressar a **corretude** da principal tarefa a ser realizada pela funcionalidade.

Quadro 1: Métricas e Indicadores do Estudo - GQM

MÉTRICA	Abordagem GQM		
	Objetivo (descrição)	*Questão*	*Valor*
PCEU	Possibilidade de Criação de Estória de Usuário	É possível criar ES?	**SIM/ NÃO**
PCCB	Possibilidade de Criação de Cenários BDD	É possível criar CB?	**SIM/ NÃO**
#EU	Número de Estórias de Usuário criadas	Quantas EU foram criadas?	**Quantidade**
#CB	Número de Cenários BDD criados	Quantos CB foram criados?	**Quantidade**
CtEU	Corretude das Estória de Usuário	Qual % erros encontrados nas EU?	**%Erros**
CtCB	Corretude dos Cenários BDD	Qual % erros encontrados nos CB?	**%Erros**

Ela foi estipulada através da porcentagem de erros na construção das EU e CB. As EU e os CB foram considerados errados quando não seguiam as cláusulas estruturais informadas no início do estudo ou quando não foram escritas de acordo com o especificado na Descrição do Sistema ou no detalhamento dos processos contidos no IDP. Quando isso aconteceu, as EU e os CB foram descartados como artefato de Engenharia de Software válido para análise estatística. É importante ressaltar que a análise da corretude desses artefatos foi realizada pelos pesquisadores que lecionam, há muitos anos, disciplinas de Engenharia de Software e correlatas.

4.3 Extração e Análise de Dados

Após o envio do e-mail, os dados foram extraídos e persistidos em planilhas Google Sheets. A fim de coletar as métricas adotadas neste trabalho (apresentadas na Tabela 1), foram mantidas as seguintes informações na planilha: (i) as EU criadas, (ii) os CB criados, (iii) a quantidade de EU criadas, (iv) a quantidade de EU criadas com erros, (v) a quantidade de CB criadas e, (vi) a quantidade de CB criadas com erros.

Para auxiliar a avaliação da quantidade e qualidade de artefatos de Engenharia de Software gerados pelos três grupos, uma análise estatística foi realizada utilizando oSoftware RStudio [2] . Primeiramente, observou-se a normalidade dos dados coletados através do teste de Saphiro-Wilk [Royston 1992]. Para esse tipo de análise, se o p-valor é menor que 0.05, então há indícios para descartar a normalidade dos dados. Isto aconteceu para todos os grupos de dados deste estudo.

Como consequência matemática, o Teste de Kruskal-Wallis foi escolhido como teste estatístico para comparar a semelhança entre 03 grupos de discentes [Kruskal and Wallis 1952]. É importante lembrar que para analisar a qualidade - através da Corretude - representada pelo percentual de erros encontrados nos artefatos gerados (%errors), os valores das métricas CtEU e CtCB de cada grupo foram aferidos.

5 DISCUSSÃO DE RESULTADOS

5.1 Visão Geral do Estudo

Demograficamente, o estudo foi realizado em 7 universidades públicas federais, abarcando todas as regiões do país, com exceção do sul, conforme é possível observar na Figura 5. Participaram da pesquisa 142 discentes, produzindo 947 artefatos de Engenharia de Software, sendo 307 Estórias de Usuário (EU) e 640 Cenários BDD (CB). Diferentes cursos da área de Tecnologia e Engenharia colaboraram com a pesquisa, e a maioria dos discentes estava matriculada na disciplina de Engenharia de Software, quando o estudo aconteceu.

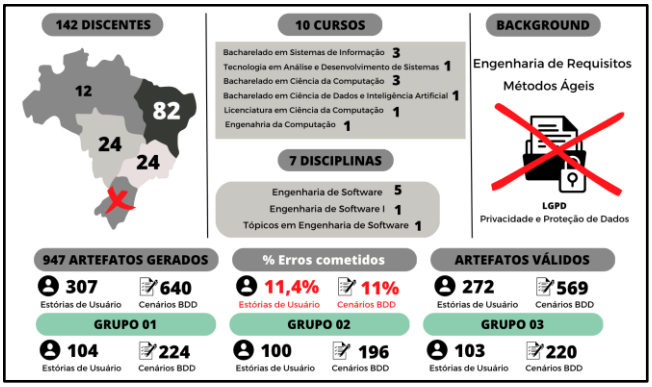

Figura 5: Visão Geral – Resultados do Estudo

Outro ponto a ser destacado é que, como *background* acadêmico, todos já tinham estudado Engenharia de Requisitos e Métodos Ágeis, alguns deles inclusive, já haviam realizado atividades práticas de construção de Estórias de Usuário e Cenários BDD. Entretanto, para minimizar possibilidades de vieses entre as universidades, entre os cursos e o background dos alunos, como parte do estudo, um treinamento foi realizado com todos os participantes antes da execução da atividade do experimento abordando conceitos sobre a LGPD e uma revisão sobre como construir Estórias de Usuário e Cenários BDD.

Por outro lado, é importante frisar que nenhum dos cursos havia abordado o conteúdo da LGPD no processo de desenvolvimento de software, apesar da lei estar em vigor desde 2018. No que se refere à quantidade e qualidade dos artefatos gerados, os discentes tiveram autonomia de decidir quantos e como as Estórias de Usuário e Cenários BDD deveriam ser criados. Foram considerados incorretos em torno de 11% do total produzido, de acordo com a definição de corretude apresentada anteriormente. Levando em consideração que os alunos tiveram o primeiro contato com o conteúdo da LGPD, este número pode ser considerado baixo.

[2] RStudio. RStudio: Integrated Development Environment for R. Disponível em https://www.rstudio.com/. Acesso em: Julho 10, 2023.

5.2 Metodologias Ativas e Centradas no Discente para o Ensino da Engenharia de Software

Considerando que a Aprendizagem Significativa visa construir conexões entre o novo conhecimento (LGPD) e o conhecimento prévio (Engenharia de Requisitos), a adoção do mapeamento entre os elementos do IDP e o elementos que compõem as Estórias e os Cenários, fez com que os discente conseguissem, com uma certa facilidade, desenvolver as atividades de criação dos artefatos. O mapeamento deu suporte à resolução do problema inspirado na forma como os computadores resolveriam – Pensamento Computacional. Assim adotaram (i) **decomposição** em problemas menores, (ii) **abstração** das informações do IDP não úteis para a construção dos artefatos em Engenharia de Software, (iii) **pensamento algorítmico**, uma vez que eles precisam seguir uma série de passos para executar as atividades do experimento, e (iv) **reconhecimento de padrões** a serem adotados para as duas funcionalidades as quais eles precisaram analisar.

Apesar da maioria dos discentes terem realizado todas as atividades, alguns relataram que, no início, não foi fácil compreender a lógica entre os elementos que poderiam ser utilizados no IDP para construir Estórias e Cenários:

SUJEITO 23: "IDP apenas fiquei um pouco perdido com muita informação."
SUJEITO 71: "Muitas colunas no IDP, o que complicou o entendimento. Mas com o mapeamento ficou mais fácil depois."
SUJEITO 110: "Em IDP, a tabela com as informações era extensa o que torna a compreensão complicada com este formato."

Ao realizar o experimento, os discentes foram forçados a compreender o assunto através de PBL (*Problem Based Learning*) - Aprendizagem Baseada em Problemas, uma vez que foi apresentado um cenário que requereu investigação e resolução levando em consideração a LGPD. Como observado na execução do estudo, os alunos acabam se engajando mais no processo de aprendizagem, uma vez que sentem estarem se preparando para desafios do mundo real. Isso os prepara para lidar com problemas e soluções complexas, adaptar-se a novas tecnologias e enfrentar situações reais com confiança, como é possível observar em seus próprios comentários:

SUJEITO 12: "Interessante a forma como isso pode afetar em um desenvolvimento de software".
SUJEITO 75: "Bastante útil e interessante, pois foram mostradas outras áreas que não mexem diretamente com programação, mas que são de Engenharia de Software."
SUJEITO 69: "Apenas elogiar que foi uma ótima dinâmica!!!"

Na aplicação da Engenharia de Software, espera-se que práticas como essa impulsionem o desenvolvimento de profissionais competentes, criativos e capazes de enfrentar desafios complexos na área da computação. Complementarmente, essa abordagem de ensino acaba estimulando a inovação, pois incentiva a criatividade e a busca por soluções inovadoras, encorajando os alunos a pensarem "fora da caixa" e a explorar abordagens diferentes para resolver problemas, fomentando a capacidade de pensamento criativo.

Consequentemente, ao utilizarem organizadores prévios (o mapeamento) e aprendizagem baseada em problemas, os discentes conseguiram, com quase 90% de corretude geral, apreender um novo conceito - adoção do IDP da LGPD na Engenharia de Requisitos. Outro apontamento na mesma direção é a autoavaliação feita pelos sujeitos do estudo, que indicam que antes do experimento, apenas 10,7% deles já tinham tido contato com IDP LGPD.

Ressalta-se que mesmo aqueles que já haviam tido visto a LGPD, não sabiam da possibilidade de utilizá-lo como instrumento no processo de Engenharia de Requisitos. Neste sentido, com a realização deste estudo, espera-se que os alunos sejam capazes de integrar essas novas informações de maneira mais efetiva e duradoura para o processo de desenvolvimento de software. É imprescindível pontuar que o percentual de erros, ao construir os artefatos de Engenharia de Software, entre os grupos são estatisticamente **semelhantes**.

Como consequência, há indícios de que o conhecimento recém adquirido - LGPD na Engenharia de Software - foi apreendido de forma eficaz. Esta inferência foi feita porque o GRUPO 01, que utilizou descrição do sistema - maneira mais tradicional de elicitar e documentar requisitos na Engenharia de Software, possuíram desempenho semelhante aos GRUPOS 02 e 03, que adotaram o IDP como base para construir os artefatos de Engenharia de Software.

Por fim, apesar do elevado número de artefatos gerados (**947**) e um baixo índice de erros (~11% - 841 corretos) no que foi construído, alguns discentes relataram tempo insuficiente para realização da atividade. Isso pode ter ocorrido porque alguns deles, apesar de já terem estudado Metodologias Ágeis e Engenharia de Requisitos, nunca tinham realizado atividades práticas como a criação de Estórias de Usuários e Cenários BDD, simulando um problema real. Desta forma, não só o conteúdo da LGPD era um assunto novo, mas também conteúdos concernentes à própria Engenharia de Software.

6 CONSIDERAÇÕES FINAIS

Este trabalho avaliou a possibilidade de inserir novos conceitos nos cursos de Engenharia de Software, especificamente "segurança de dados pessoais", conforme prevê a LGPD através de metodologias ativas e centradas no discente. Neste sentido, foi possível compreender que Metodologias Ativas, Aprendizagem Significativas e Pensamento Computacional podem ser adotadas para introduzir o conteúdo de Proteção e Privacidade de Dados Pessoais LGPD, no contexto do ensino da Engenharia de Software.

Para isto, um experimento foi realizado com 10 cursos de ensino superior que possuem como disciplina obrigatória ou optativa, o componente de Engenharia de Software (ou correlatas). O estudo envolveu 142 alunos, resultando na

criação de 947 artefatos de Engenharia de Requisitos Ágeis - Estórias de Usuário e Cenários BDD.

Através da avaliação quantitativa (análise estatística) e qualitativa (corretude dos artefatos e autoavaliação dos sujeitos), foi possível observar que as técnicas de Metodologias Ativas e Centradas no discente mostram-se apropriadas para desenvolver nos sujeitos a capacidade de adquirirem um novo conhecimento técnico interdisciplinar de maneira rápida e engajada. Adicionalmente, o conhecimento teórico conseguiu ser aplicado na realização de atividades práticas através da criação de artefatos de Engenharia de Software a partir da avaliação de problemas reais.

E mais, suscitou-se um novo paradigma de compreender que existe uma obrigação do processo de desenvolvimento de software seguro como padrão e não mais como opção, uma vez que a LGPD impõe legalmente.

Apesar da colaboração apresentada, não é pretensão neste estudo generalizar as contribuições, sendo necessária a replicação deste experimento em mais universidades, inclusive contemplando a região Sul do Brasil, para termos mais diversificação dos sujeitos. Adicionalmente, a replicação da metodologia adotada, abordando outras naturezas de conhecimento (técnicos ou não-técnicos) precisa ser realizada para saber se o conteúdo novo abordado no estudo - Engenharia de Requisitos adotando o IDP da LGPD - foi um fator que tendenciou os resultados encontrados.

AGRADECIMENTOS

Este trabalho é parcialmente financiado pelo INES 2.0 (www.ines.org.br), bolsa CNPq 465614/2014-0, bolsa FACEPE APQ-0399-1.03/17 e APQ/0388-1.03/14, bolsa CAPES 88887.136410/2017- 00. Sérgio Soares é parcialmente apoiado pela bolsa CNPq 306000/2022-9.

REFERÊNCIAS

[1] Oguz, D., and Oguz, K. (2019). Perspectives on the gap between the software industry and software engineering education. IEEE Access, 7, 117527-117543. DOI: 10.1109/access.2019.2936660.

[2] M. Cinque (2016). Lost in translation". Soft skills development in European countries. Tuning Journal for Higher Education, 3(2), 389–427. https://doi.org/10.18543/tjhe-3(2)-2016pp389-427.

[3] Dempsey, M., and Brennan, A. (2017). Turbocharging the journey into the liminal space and beyond. development, 27, 28. https://pdfs.semanticscholar.org/0320/768107faed2a722df15ac557b978f8efee97.pdf.

[4] Tang, J., Zhang, S. X., and Lin, S. (2021). To reopen or not to reopen? How entrepreneurial alertness influences small business reopening after the COVID-19 lockdown. Journal of Business Venturing Insights, 16, e00275. https://doi.org/10.1016/j.jbvi.2021.

[5] Tseng, H., Yi, X., and Yeh, H.-T. (2019). Learning-related soft skills among online business students in higher education: Grade level and managerial role differences in self-regulation, motivation, and social skill. Computers in Human Behavior, 95, 179–186. https://doi.org/10.1016/j.chb.2018.11.035.

[6] Wei-Chang Kong. 2001.E-commerce and cultural values. IGI Publishing, Brennan, A., Dempsey, M., McAvoy, J., O'Dea, M., O'Leary, S., & Prendergast, M. (2023). How COVID-19 impacted soft skills development: The views of software engineering students. Cogent Education, 10(1). doi:10.1080/2331186X.2023.2171621.

[7] García-Morales, V. J., Garrido-Moreno, A., and Martín-Rojas, R. (2021). The transformation of higher education after the COVID disruption: Emerging challenges in an online learning scenario. Frontiers in Psychology, 12, 196. https://doi.org/10.3389/fpsyg.2021.616059.

[8] Gasiba, T. E., Iosif, A.-C., Suppan, S., Lechner, U., & Pinto-Albuquerque, M. (2023). Reflections on Training Next-Gen Industry Workforce on Secure Software Development. In Proceedings of the 5th European Conference on Software Engineering Education (ECSEE '23) (pp. 1–10). Association for Computing Machinery. doi:10.1145/3593663.3593665.

[9] General Data Protection Regulation (GDPR), Regulation (EU) 2016/679, 27 April 2016.

[10] Lei Geral de Proteção de Dados (LGPD), Lei nº 13.709, de 14 de agosto de 2018.

[11] Alkubaisy, D., Piras, L., Al-Obeidallah, M.G., Cox, K., & Mouratidis, H. (2022). A framework for privacy and security requirements analysis and conflict resolution for supporting GDPR compliance through privacy-by-design. In R. Ali, H. Kaindl, & L.A. Maciaszek (Eds.), Anais do XXXI Workshop sobre Educação em Computação (pp. 67-87). Cham: Springer. doi:10.1007/978-3-030-96648-5_4.

[12] Senarath, A.R., & Arachchilage, N.A.G. (2018). Understanding user privacy expectations: A software developer's perspective. Telematics and Informatics, 35, 1845–1862. doi:10.1016/j.tele.2018.05.012.

[13] Senarath, A., Grobler, M., & Arachchilage, N.A.G. (2019). Will they use it or not? Investigating software developers' intention to follow privacy engineering methodologies. ACM Transactions on Privacy and Security, 22, 1–30. doi:10.1145/3364224.

[14] Gasiba, T. E., Iosif, A.-C., Suppan, S., Lechner, U., and Pinto-Albuquerque, M. (2023). Reflections on Training Next-Gen Industry Workforce on Secure Software Development. In Proceedings of the 5th European Conference on Software Engineering Education (ECSEE '23) (pp. 1–10). Association for Computing Machinery. doi:10.1145/3593663.3593665.

[15] K. Pohl, C. Rupp. Requirements Engineering: Fundamentals, Principles, and Techniques. Springer; 2015.

[16] A Mihelič A, Vrhovec S, Hovelja T. Agile development of secure software for small and medium-sized enterprises. Sustainability. 2023;15(1):801.

[17] R Rose, S.; Wynne, M.; Helles√∏y, A. The Cucumber for Java Book: Behaviour-Driven Development for Testers and Developers. Birmingham: Pragmatic Bookshelf, 2015.

[18] Georges T, et al. Guiding feature models synthesis from user-stories: an exploratory approach. Synthesis. 2023. 30:31.

[19] Parsa S. Acceptance testing and behavior driven development (BDD). In: Software Testing Automation: Testability Evaluation, Refactoring, Test Data Generation and Fault Localization. Cham: Springer International Publishing; 2023. p. 79-158.

[20] Peixoto, M., Silva, C., Lima, R., Araújo, J., Gorschek, T., & Silva, J. (2019). PCM Tool: Privacy Requirements Specification in Agile Software Development. In Anais Estendidos do X Congresso Brasileiro de Software: Teoria e Prática, (pp. 108-113). Porto Alegre: SBC. doi:10.5753/cbsoft_estendido.2019.7666.

[21] Peixoto, Mariana, et al. ""The perspective of Brazilian software developers on data privacy."" Journal of Systems and Software 195 (2023): 111523.

[22] Cleber Nardelli. Segurança da Informação e LGPD Aplicado no Desenvolvimento de Software. In: ESCOLA REGIONAL DE ENGENHARIA DE SOFTWARE (ERES), 5. , 2021, Evento Online. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2021. p. 169-178. DOI: https://doi.org/10.5753/eres.2021.18462.

[23] F. Alexandrini and C. Nardelli, C. (2021). PRIMEIRA FASE PRIMEIRA FASE DA SEGURANÇA DA INFORMAÇÃO E LGPD APLICADO NO DESENVOLVIMENTO DE SOFTWARE GOVERNO ELETRÔNICO. REVISTA DE EXTENSÃO E INICIAÇÃO CIENTÍFICA DA UNISOCIESC, 9(1). Recuperado de https://reis.unisociesc.com.br/index.php/reis/article/view/332.

[24] Bertan, B. C., Portilho, M. D. O., Gurfinkiel, M. V., & Borges, N. N.(2022). Abordagem da LGPD no desenvolvimento de software.

[25] D. R. de Melo Filho, et al. Metodologia Scrum: Uma aliada na implementação da LGPD. Research, Society and Development. 2023;12(4):e22712441189-e22712441189.

[26] D. L. Cardoso and T. Cardoso. Adequação da LGPD via "Projetos Ágeis Scrum". Boletim do Gerenciamento. 2023. 35(35):28-41.

[27] Camílio MN, Alves CF. G-Priv: Um Guia para Apoiar a Especificação de Requisitos de Privacidade em Conformidade com a LGPD. iSys-Brazilian Journal of Information Systems. 2023. 16(1):2-1.

[28] C. P. Santiago,J. W. Menezes and Aquino, F. J. A. (2003). Proposta e Avaliação de uma Metodologia de Aprendizagem Baseada em Projetos em Disciplinas de Engenharia de Software através de uma Sequência Didática. Revista Brasileira de Informática na Educação, 31, 31-59. DOI: 10.5753/rbie.2023.2817.

[29] Cintra, C., and Bittencourt, R. (2023). As Experiências de Estudantes em um Curso de Engenharia de Computação Baseado em PBL. In Anais do XXXI Workshop sobre Educação em Computação, (pp. 327-338). Porto Alegre: SBC. doi:10.5753/wei.2023.229276.

[30] Beleti Junior, C. R. e de Faria Sforni, M. S. (2023) "Pesquisas experimentais no desenvolvimento do pensamento computacional: um mapeamento sistemático de literatura no ensino de conceitos de computação", Educação em Foco, 26(49). doi: 10.36704/eef.v26i49.6623.

[31] Juliana Saraiva and Sergio Soares. 2023. Adoption of the LGPD Inventory in the User Stories and BDD Scenarios Creation. In Proceedings of the XXXVII Brazilian Symposium on Software Engineering (SBES '23). Association for Computing Machinery, New York, NY, USA, 416–421. https://doi.org/10.1145/3613372.3613375.

[32] V. Basili, G. Caldiera, F., F. McGarry and H. D. Rombach (2007). GQM strategies--aligning business strategies with software measurement. In: First International Symposium on Empirical Software Engineering and Measurement (ESEM 2007).