Implementando Privacy by Design para Reimaginar a Privacidade no Instagram

Lucas Guimarães Campregher[1], **Fábio Leandro Rodrigues Cordeiro (Orientador)**[1]

[1]Instituto de Ciências Exatas e Informática
Pontifícia Universidade Católica de Minas Gerais
Ciência da Computação (PUC-MG)
Caixa Postal 30535-901 - Belo Horizonte - MG, Brasil
lucas@campregher.com, fabio@pucminas.br

Resumo. *Os desafios relacionados à privacidade de usuários nos sistemas modernos está intrinsecamente relacionado ao processo de engenharia de software em que são construídos os sistemas. Com decisões que impactam milhões de pessoas, e geralmente sem considerar uma visão determinista da tecnologia, funcionalidades implementadas em redes sociais acabam - por muitas vezes - impactando negativamente os usuários, em contrassenso às principais definições de privacidade existentes. Este artigo visa - a partir dos conceitos de Privacy by Design - identificar nas redes sociais os principais problemas relacionados à privacidade dos usuários, e a partir dessa problemática, utilizar-se de proposições de novas metodologias de engenharia de software existentes para reconstruir funcionalidades do Instagram. Nesse trabalho, foi validada a efetividade do uso de Privacy by Design na construção de redes sociais que respeitem ao máximo a privacidade dos seus usuários.*

1. Introdução

Ao longo dos últimos anos, têm-se observado impactos negativos massivos decorrentes da interação humano-tecnológica [Kugler 2020]. É notável que a evolução da tecnologia permitiu que, hoje, algoritmos de grandes empresas - com a quantidade de dados suficiente - consigam nos conhecer tão bem quanto nós mesmos [Papazoglou 2019]. Esta realidade trouxe consigo importantes reflexões acerca da privacidade dos usuários que utilizam desses sistemas, e como ela é protegida atualmente no processo de engenharia de *software* moderno.

O Determinismo Tecnológico é uma das teorias sociológicas mais populares que discute a relação entre tecnologia e sociedade, evidenciando que o avanço e o desenvolvimento de novas tecnologias influenciam os valores, a cultura, o comportamento e a estrutura social dos povos. Uma visão determinista da tecnologia, nos permite criticar a forma como algumas decisões são tomadas hoje na construção dos sistemas, bem como seu impacto social - por muitas vezes negativo.

Dentro desse contexto, a privacidade de dados têm sido cada vez mais protagonista de discussões importantes para a engenharia de *software moderna*. O conceito de *Privacy by Design* (PbD) [Cavoukian 2010], desenvolvido pela doutora Ann Cavoukian é a principal referência para novos trabalhos que buscam a reformulação da privacidade de dados como é apresentada hoje. Definições presentes na ISO/IEC 29100 [ISO/IEC 29100:2011 2011] também endossam princípios a serem seguidos para

a adequação dos sistemas à privacidade real. Avaliando os *softwares* atuais a partir desses parâmetros, fica claro a desadequação geral a diversos quesitos relacionados à privacidade dos usuários, e consequentemente são expostas diversas deficiências no processo de engenharia de *software*, abrindo espaço para proposição de melhorias.

"*Privacy by design* reivindica que a privacidade seja levada em consideração durante todo o processo de engenharia"[Cavoukian 2010].

"A preocupação com a privacidade dos dados deve, portanto, fazer parte de qualquer desenvolvimento de *software*, independentemente do setor a que se destina"[Morales-Trujillo et al. 2019].

As redes sociais estão no foco deste trabalho. Nos últimos anos, duras críticas foram feitas acerca do modelo de negócios desses sistemas, que preza por estratégias para captar a atenção das pessoas, e maximar o tempo de uso para lucrar mais [Rubinstein and Good]. Nesse momento, não somente a ética é deixada de lado na construção de um modelo de engenharia de *software* que não preza em momento algum pela privacidade do usuário, como também são feridos princípios fundamentais em defesa dos usuários, e termos legais como a lei europeia de regulação de dados GDPR (*General Data Protection Regulation*), e sua homônima brasileira LGPD (Lei Geral de Proteção de dados). A responsabilização de engenheiros de *software* sobre a construção de funcionalidades com impactos negativos em bilhões de pessoas deve ser levada em consideração, e chamam a atenção para a proposição de soluções que possam ser aplicadas nesses casos.

O grande interesse a respeito do tema surgiu dado que hoje, a despreocupação acerca da privacidade na construção dos produtos deixa de ser um problema meramente ético. Muito além da moralidade, os impactos da interação humana com os grandes aplicativos - principalmente redes sociais - têm levado à alteração dos nossos hábitos, pensamentos, e até a forma como nos enxergamos [Papazoglou 2019]. Além disso, num contexto mais amplo, é possível perceber a existência de uma rede de informações quebrada [Alkawaz et al. 2021], que gera impactos ainda mais urgentes, ocasionando a formação de cenários políticos enfermos, manipulação em massa da população, e o monopólio da informação. Tudo isso, porque os usuários estão expostos - nas redes sociais - à condições de uso as quais nunca foram informadas ou consentidas corretamente, além de que as consequências e riscos aos usuários são omitidas o tempo todo pelas grandes empresas.

Este trabalho tem o objetivo de reimaginar a privacidade da redes sociais por meio da proposição de soluções alternativas para algumas das funcionalidades presentes hoje no Instagram. Essa nova perspectiva não tem a intenção de subverter o viés produtivo das grandes empresas, de modo a questioná-las do ponto de vista moral. O propósito aqui é o de atestar a eficácia de um modelo de engenharia baseado em *Privacy by Design* [Cavoukian 2010] aplicado na construção de redes sociais, além de contribuir e provocar a discussão acerca da privacidade na comunidade de engenharia de software.

O trabalho não tem o objetivo de viabilizar ou justificar uma mudança drástica de padronização nos modelos de engenharia de *software* existentes hoje nas grandes empresas de redes sociais atuais. Uma mudança estrutural como essa necessita da mobilização de mudanças, não só do modelo de engenharia de *software*, mas em diversos outros pontos críticos para as tomadas de decisão dentro dessas empresas, como adequações legislativas, regulatórias, e de impacto financeiro direto nesses casos. Ainda assim, o valor

da elaboração do trabalho está no despertar dessa preocupação por parte dos principais responsáveis pela construção de sistemas problemáticos nesse sentido atualmente: os desenvolvedores. Além disso, a validação do uso de um novo modelo de engenharia de *software*, e a exemplificação de funcionalidades construídas na realidade das redes sociais é de grande valor para a colaboração de conhecimento na comunidade acerca do assunto.

As decisões de implementação das grandes empresas de redes sociais são, e continuarão sendo, orientadas pelo lucro [Rubinstein and Good]. Tendo isso em vista, este trabalho continua relevante para a abordagem dessas empresas, visto que a expectativa é de comprovar que é possível construir um sistema de rede social que respeite a privacidade dos usuários e aumente sua satisfação geral com o sistema, sem comprometer a experiência do usuário - sendo essa uma possível vantagem competitiva de mercado. Além disso, a perspectiva de evolução da legislação em torno de regulações mais pesadas acerca do acúmulo indiscriminado de dados, faz com que seja vislumbrada uma realidade próxima em que uma rede social que limita melhor a coleta de dados dos usuários por meio de políticas de privacidade seja também uma rede social que apresenta uma abordagem mais vantajosa economicamente.

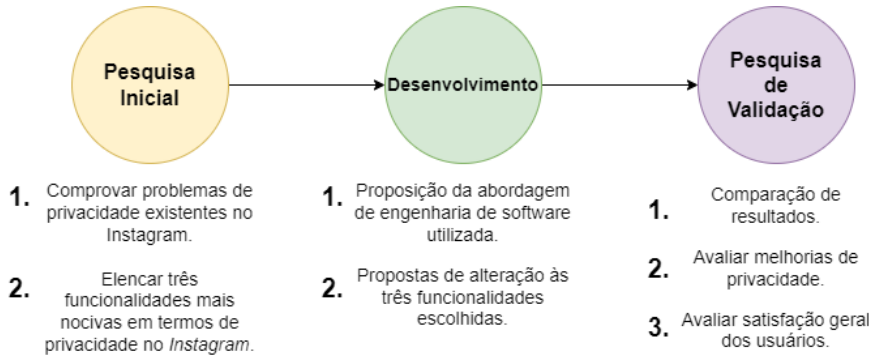

Figura 1. Metodologia

2. Fundamentação Teórica

2.1. Determinismo Tecnológico e Instrumentalismo Tecnológico

Em seu artigo de 2020, Newport defende uma visão mais determinista da tecnologia, visto que consequências complexas surgem a partir da interseção entre tecnologia e cultura. É ultrapassada a visão instrumentalista de que as novas tecnologias e sistemas são apenas ferramentas, e as consequências que podem vir a surgir do seu uso são de responsabilidade do usuário. Isso não é verdade, pois toda decisão de design e engenharia tomada na construção de um *software* interage e impacta diretamente na maneira como cada um vai interagir com o sistema, inclusive moldando comportamentos e pensamentos para atender interesses específicos de alguma empresa [Newport 2020]. O determinismo tecnológico defende que a tecnologia influencia diretamente nos fenômenos históricos e culturais, e portanto, devemos enquanto engenheiros de *software* nos preocupar e nos responsabilizar pelo impacto daquilo que construímos.

É preciso observar a maneira como os usuários interagem com nossos sistemas, e as consequências que podem surgir a partir daí, visto que por muitas vezes nossos sis-

temas geram efeitos colaterais complexos, que devem ser observados e tratados rapidamente. Efeitos colaterais complexos não são bem tratados pela atual ênfase acadêmica no instrumentalismo tecnológico. Os efeitos colaterais complexos devem ser incluídos em todo processo de engenharia iterativo.

2.2. Privacidade

O conceito de privacidade utilizado é o defendido pela doutora Ann Cavoukian. Privacidade é uma decisão individual. Não diz respeito a capacidade de manter informações em sigilo, mas corresponde ao controle pessoal e à liberdade de escolha de cada um enquanto cidadão de tomar decisões acerca de quais dados pessoais deseja divulgar, e para quem. Cada indivíduo é dono das informações que produz, e deveria portanto ter controle sobre essas informações.

2.3. Privacy by Design

A pressão emergente acerca da privacidade de dados é decorrente de um ecossistema tecnológico em que redes sociais, canais de mídia, governos, e empresas que prestam serviços altamente individualizados dependem de informações geradas pelos usuários. Em 2009, a doutora Ann Cavoukian, juntamente a equipe de comissários de informação de privacidade da província de Ontário, no Canadá, formalizaram o conceito de *Privacy by Design* [Cavoukian 2010]. Essa é uma abordagem que defende a ideia de que a privacidade deve ser considerada a partir da fase inicial de um projeto, e durante todo o ciclo de vida de um *software*. Para que isso seja aplicado na engenharia de software, são definidos sete princípios fundamentais:

1. Pró-ativo não reativo; preventivo não corretivo, de modo a evitar incidentes de violação à privacidade;

2. Privacidade como configuração padrão: as configurações padrão de determinado sistema devem ser ajustadas desde o início para preservar a privacidade do usuário;

3. Privacidade incorporada ao design, incluindo a arquitetura e modelos de negócio;

4. Funcionalidade total - soma positiva, não soma zero;

5. Segurança de ponta a ponta: proteção completa incorporada ao ciclo de vida da informação;

6. Visibilidade e transparência - mantê-lo aberto;

7. Respeito pela privacidade do usuário: mantê-lo centrado nos interesses do usuário.

Cada um dos princípios defende ideais extremamente valiosas a serem considerados na construção e evolução de um sistema. Entretanto, essa abordagem foi muito criticada por ser pouco específica acerca de sua aplicação prática na engenharia de software [Gurses et al. 2011]. Desde sua concepção inicial, muito foi discutido na comunidade científica acerca da aplicação do PbD na engenharia de *software* [Morales-Trujillo et al. 2019].

Empresas com modelo de negócios construídos com base na vigilância e manipulação dos dados dos usuários, como é o caso das redes sociais, representam

um desafio ainda maior para a implantação do PbD em seus processos de engenharia [Rubinstein and Good]. Apesar de soar inicialmente improvável, a utilização de princípios e práticas de PbD se provou eficiente para evitar incidentes de privacidade de redes sociais em cenários contrafactuais [Rubinstein and Good]. Foi constatado que uma engenharia de *software* baseada em privacidade, e o design de uma experiência de usuário com usabilidade otimizada em torno desse conceito, são altamente relevantes para avaliar e superar uma série de problemas de privacidade, incluindo questões emergentes que afetam os serviços de redes sociais [Rubinstein and Good].

As instituições regulatórias de privacidade pelo mundo têm apostado em PbD como uma abordagem viável para adotar reformas nos regimes de privacidade existentes [Rubinstein and Good]. A Comissão Federal do Comércio dos Estados Unidos, buscando promover a proteção dos consumidores, propõe recomendações que - baseados em PbD - buscam solucionar problemas de privacidade, e regular a implementação das funcionalidades por meio de novos formatos de engenharia de *software* [Staff 2012] . Nesse caso, há uma proposição mais objetiva em torno do que deve ser seguido para garantir a privacidade dos usuários em um sistema.

As definições presentes na ISO/IEC 29100 também colaboram na definição de princípios a serem seguidos em prol da privacidade dos usuários [ISO/IEC 29100:2011 2011]. Estudos apontam as relações existentes hoje entre os princípios de PbD e as definições da ISO/IEC 29100 [Morales-Trujillo et al. 2019]. Essa padronização é extremamente relevante para delimitar todos os aspectos em que a privacidade deve ser considerada na construção de um *software*.

3. Trabalhos Relacionados

3.1. Os impactos negativos na relação humano-tecnológica

A motivação do desenvolvimento deste trabalho está muito relacionada com a criticidade dos impactos negativos que decorrem da interação entre sociedade e tecnologia. Atualmente - sobretudo em sistemas de redes sociais - é possível observar consequências danosas não previstas decorrentes de decisões tomadas durante o processo de desenvolvimento.

O artigo "*When technology goes awry*"aborda criticamente os impactos negativos da abordagem do instrumentalismo tecnológico predominante hoje na construção de novos sistemas. Em defesa do determinismo tecnológico, o autor convida os engenheiros desenvolvedores a participar da importante discussão acerca dos efeitos da relação entre novas tecnologias e a sociedade. É importante ressaltar como mudanças - por menores que pareçam ser - podem acarretar em efeitos colaterais complexos que precisam ser combatidos, principalmente durante o processo iterativo de engenharia de *software* [Newport 2020].

Em outro artigo, é possível observar pontos de vista em relação à considerar, ou não, o uso excessivo de videogames ou *smartphones* um vício. Apesar de o vício em video games ter sido incluído na classificação internacional de doenças da OMS, nada ainda foi considerado quando se trata do uso compulsivo de *smartphones*. O ponto é que, independente dessa relação com as tecnologias serem classificadas como vícios patológicos, ou apenas um maus hábitos, muitas pessoas estão sofrendo gravemente com as

consequências negativas que isso pode gerar [Kugler 2020]. O vício na utilização de aparelhos celulares está intimamente relacionado com o uso de aplicativos de redes sociais, que consomem grande parte do tempo desses usuários. Desse artigo surgem dois questionamentos interessantes que corroboram com a discussão proposta: Estamos viciados em tecnologia? Estamos desenvolvendo tecnologias propensas a causar um vício?

A solução para a relação das pessoas com as redes sociais não pode passar apenas por uma reformulação na transparência das grandes empresas em esclarecer os motivos por trás da escolha de publicações e anúncios pelos algoritmos. Pelo contrario, a implementação de uma *feature* como essa poderia afetar mais ainda a maneira como nos vemos, e influenciar na nossa personalidade individual. Isso porque, a nossa crença de que os algoritmos são extremamente precisos é tao grande que, caso tenhamos a oportunidade de ver exatamente como eles nos descrevem, podemos acabar absorvendo essas informações como uma verdade, e corroborando mais ainda com o domínio gigante que essas companhias têm sobre nossas mentes [Papazoglou 2019]. Essa visão é importante para perceber que os processos modernos de engenharia de *software* devem se adequar a critérios coerentes de privacidade.

Outros problemas latentes presentes nas redes sociais atualmente são as *fake news* e construção de redes de desinformação. Pesquisas sugerem que, a principal fonte de consumo de noticias são as redes sociais, e concomitantemente, a maioria das pessoas já foram vitimas de noticias falsas por esses meios. Existem poucos avanços no que diz respeito ao combate desse problema atualmente, visto que a maioria das pessoas não tem ciência da existência de *sites* para validar as informações [Alkawaz et al. 2021]. Percebe-se que há muito espaço para medidas que busquem atacar este problema de maneira mais eficaz do que acontece hoje.

É possível concluir, a partir de uma visão crítica acerca dos impactos negativos causados pela relação humano-tecnológica nas redes sociais, que esses problemas estão relacionados aos processos de engenharia que envolvem decisões acerca das funcionalidades presentes nas redes sociais. Aprimorar esses processos em termos de privacidade é um dos caminhos fundamentais para tornar essa relação mais saudável.

3.2. *Privacy by Design* na engenharia de software

O artigo "A Systematic Mapping Study on Privacy by Design in Software Engineering"realiza um estudo sistemático de mapeamento que vai servir de base para as ideias desenvolvidas neste trabalho. Os autores consolidaram os diversos estudos existentes acerca do tema de *Privacy by Design*, apresentando conceitos fundamentais, propondo discussões acerca dessas definições, expondo estatísticas interessantes relacionadas às contribuições já realizadas, e identificando as principais pendências associadas à privacidade dos usuários unida aos processos de engenharia de *software*. O estudo [Morales-Trujillo et al. 2019], publicado pelo *CLEI Eletronic Journal (2019)*, destacou a falta de:

1. Metodologias e atividades de engenharia que abordam questões de privacidade.
2. Apoio à tradução dos seus princípios em atividades de engenharia.
3. Detalhes em termos de como pode ser implementado.
4. Orientações claras e detalhadas para abordar as questões de privacidade.

5. Ferramentas concretas para ajudar os desenvolvedores de *software* a projetar e implementar sistemas amigáveis à privacidade.

6. Especificidade em sua definição, sua imprecisão e seu alto nível de abstração.

O estudo também considera que "PbD está em fase inicial; seus fundamentos e princípios estão em processo de serem estabelecidos e um antigo conjunto de práticas, cuja intenção é seguir os princípios, foi proposto recentemente. O próximo passo para PbD é criar mais práticas e comprovar sua utilidade e aplicabilidade no desenvolvimento de *software*. Nesse cenário, acreditamos firmemente que é importante integrar as melhores práticas de PbD nos processos de desenvolvimento de *software*. O objetivo dessa integração é fortalecer os sistemas que são e serão desenvolvidos pelas organizações. Além disso, unificará as melhores práticas que orientam o desenvolvimento de *software* com PbD, o que, por sua vez, protegerá a privacidade dos dados sensíveis nos sistemas em constante crescimento". Essa ideia influencia fortemente na motivação para realização deste estudo e abre margem para outros trabalhos neste sentido.

4. Metodologia

O estudo sistemático de mapeamento de *Privacy by Design* na engenharia de software [Morales-Trujillo et al. 2019], em *CLEI Eletronic Journal (2019)*, analisou diversos trabalhos publicados na área e destacou a falta de, entre outros itens:

1. Metodologias e atividades de engenharia que abordam questões de privacidade.

2. Apoio à tradução dos seus princípios em atividades de engenharia.

3. Especificidade em sua definição, sua imprecisão e seu alto nível de abstração.

Dadas essas deficiências e as preocupações latentes acerca da privacidade de dados na construção de redes sociais, este trabalho possui o objetivo de identificar as principais funcionalidades existentes no Instagram que ferem à privacidade dos usuários, descrevê-las explicitando a problemática envolvida, e reimaginá-las a partir de uma perspectiva de Pbd. Conforme pode ser observado na Figura 1, esse processo foi dividido em três partes:

4.1. Pesquisa Inicial

Foi realizada uma pesquisa com usuários do Instagram para avaliação da privacidade do sistema, com perguntas baseadas no ciclo de vida da informação [Arass et al. 2017]. Também foram incluídas perguntas para avaliar a usabilidade do sistema e confiança dos usuários na empresa. Além disso, os respondentes elencaram as funcionalidades que mais ferem sua privacidade.

4.2. Desenvolvimento

Os problemas identificados foram traduzidos em atividades práticas de engenharia de *software*. As funcionalidades elencadas como as três principais dores dos usuários em termos de privacidade foram reimaginadas a partir de uma perspectiva de PbD. Para isso, foi feita uma análise atual da situação de cada uma delas, buscando documentar e propor soluções que busquem se adequar à realidade da privacidade dos usuários. Essa proposição foi construída com base nos principais estudos de caso, e propostas de *frameworks* relacionados à privacidade de dados.

4.3. Pesquisa de Validação

Finalmente, uma pesquisa foi aplicada com as mesmas perguntas de avaliação de privacidade feitas no primeiro momento, porém desta vez em relação as novas propostas de solução geradas pelo trabalho. O resultado pretende avaliar se as novas soluções geradas atendem ou não os ideais de privacidade, e permite a avaliação de como essas abordagens impactam nas percepções de usabilidade e confiança dos usuários. A análise desses resultados vai enriquecer o entendimento da viabilidade da aplicação prática de PbD na engenharia de sistemas modernos de redes sociais. Além disso, as soluções propostas geradas para cada funcionalidade constituem estudos de caso inovativos, que podem ser aplicados futuramente em novas abordagens de redes sociais que respeitem genuinamente a privacidade dos seus usuários.

5. Resultados Esperados

Espera-se que, com a realização das atividades descritas no artigo, sejam construídas propostas mais adequadas para o desenvolvimento das principais funcionalidades problemáticas para a privacidade dos usuários de redes sociais. A expectativa é de que seja comprovada a eficiência de um modelo de engenharia de *software* baseado em *Privacy by Design* quanto a capacidade de desenvolver sistemas mais conscientes aplicados na realidade atual dos usuários.

Na prática, espera-se que na pesquisa inicial seja comprovada a existência de dores latentes de privacidade de dados pelos usuários; e que na pesquisa final, a reconstrução das funcionalidades identificadas seja validada pelos próprios respondentes como soluções mais adequadas ao respeito à privacidade dos usuários - sem descumprimento dos seus objetivos primários.

Esses resultados seriam de grande valor para voltar a preocupação dos engenheiros de *software* acerca da responsabilidade na construção de novos sistemas e funcionalidades. Além disso, seria de extrema importância para aplicar conceitos recentes de engenharia de software na realidade das redes sociais, que possui poucos trabalhos com modelos propositivos a seu respeito.

6. Desenvolvimento

6.1. Pesquisa Inicial

A pesquisa inicial foi respondida majoritariamente por jovens. O número de respostas nos fornece garantia estatística de que os resultados são confiáveis. Em respeito a privacidade dos usuários, a pesquisa foi anônima.

Como pode ser visto na análise das idades dos respondentes (figura 2), foram 159 respostas com 54% de respondentes entre 18 e 24 anos, e 37% de 25 a 35 anos. A faixa etária é coerente com a dos usuários do aplicativo.

A primeira pergunta da pesquisa (figura 3) confirma a hipótese prévia de que os usuários não têm ciência clara de quais dados pessoais são coletados pelo Instagram, e a forma como o *app* coleta esses dados. (Ex.: Dados informados; dados de utilização do aplicativo; uso de câmera, localização, e outros recursos do meu *smartphone*; etc...)

Figura 2. Idade dos respondentes da pesquisa inicial

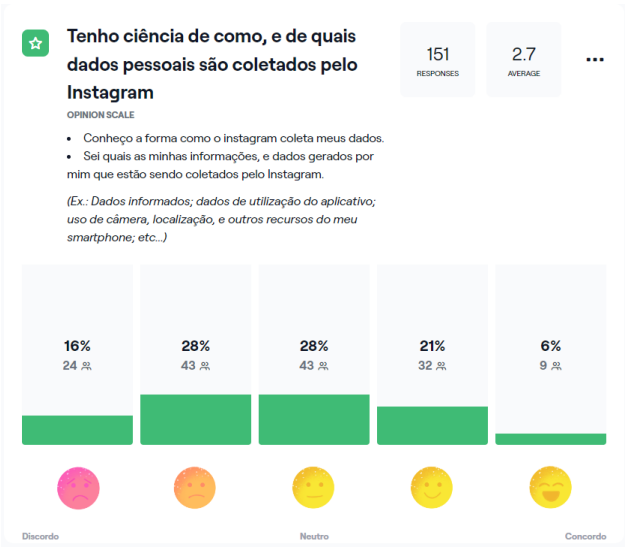

Figura 3. Ciência da coleta de dados

A finalidade de uso dos dados pessoais coletados pelo Instagram - incluindo compartilhamento de dados à terceiros - é ainda mais incerta aos usuários, como ilustrado na figura 4.

Estudos realizados em 2022 constataram que, no *Facebook* e no *Twitter*, a política relacionada a grande parte dos dados coletados pelos aplicativos representa uma ameaça à privacidade dos usuários por serem apresentados de forma vaga ou não pedirem a permissão de forma alguma [Miller et al. 2022], mesmo se tratando muitas vezes de dados sensíveis. Constatações como essa reforçam a ideia de que hoje, a forma como as redes sociais informam aos seus usuários acerca da coleta e uso de dados pessoais é ineficaz em termos de privacidade. *Privacy by Design* defende que todas as informações acerca do ciclo de vida das informações coletadas [Arass et al. 2017] devem ser informadas ao usuário no momento em que os dados são coletados, de maneira clara e amigável. Atualmente isso não acontece para a maioria dos casos. Os termos de privacidade do Instagram são apresentados de forma discreta no momento da criação de uma nova conta, e nele

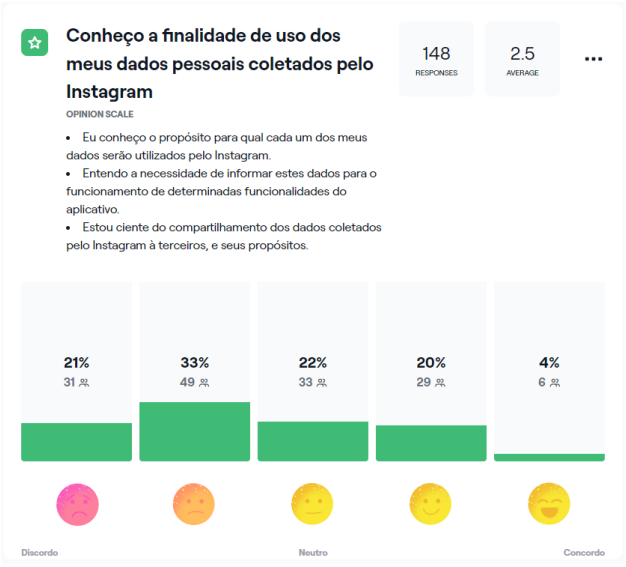

Figura 4. Ciência da finalidade dos dados

estão presentes - em um texto denso e vago - todas as políticas específicas de coleta e uso dos dados em todo o *app*.

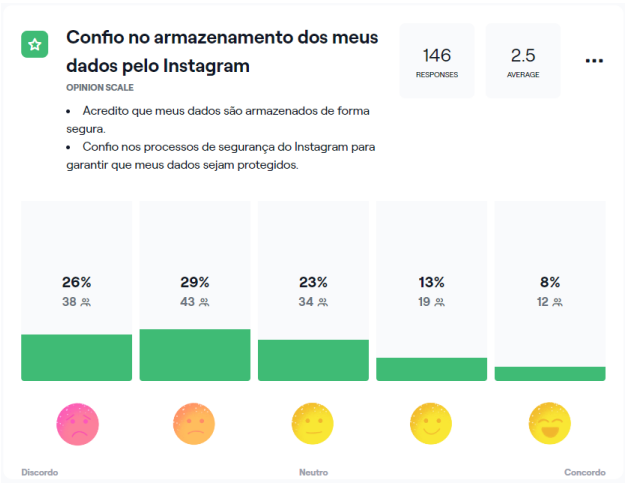

Figura 5. Confiança no armazenamento dos dados

As respostas da pergunta representada na figura 5 apontam que a maioria dos usuários do Instagram não confiam no armazenamento dos dados pela empresa. Isso significa que eles não acreditam que os dados são armazenados de maneira segura ou que o Instagram possui processos de segurança eficientes o suficiente para garantir que os dados sejam protegidos.

Como observado na figura 6, a maioria dos usuários não considera que manifestou consentimento livre e explícito sobre todo o uso dos seus dados pessoais no Instagram. Porém, curiosamente, um total de 30% dos respondentes afirmaram que concordam, ou concordam completamente com a afirmação de que manifestaram esse consentimento. Foi observado que esse alto índice está relacionado ao desconhecimento dos usuários acerca do significado de consentimento. Com o passar dos anos, e com a consolidação de

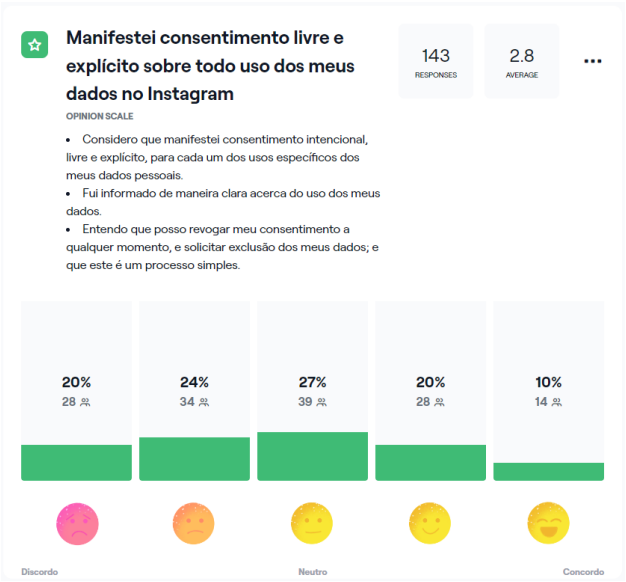

Figura 6. Consentimento livre e explícito

um padrão deficiente de apresentação dos termos de privacidade, muitos usuários acabam por considerar que, por ter clicado que "concordam com os termos"manifestaram consentimento com o uso de seus dados. Essa é uma visão deturpada do termo. Em privacidade de dados, consentimento deve ser dado de forma intencional, livre, explícita, e reversível, para cada um dos usos específicos dos dados pessoais.

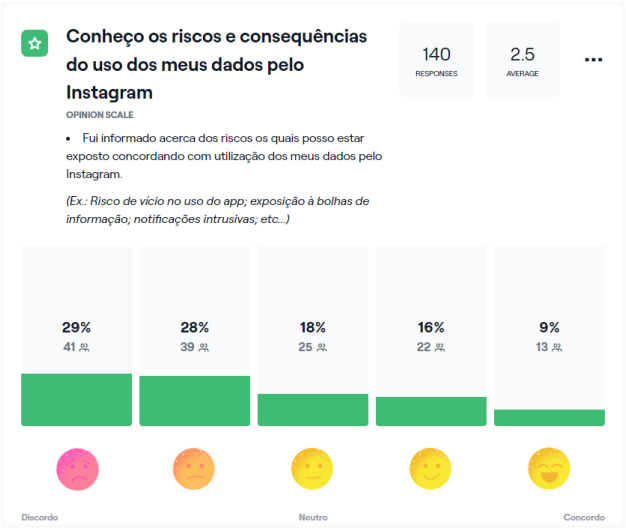

Figura 7. Riscos e consequências

Os diversos riscos relacionados à utilização de redes sociais são comprovados, porém, a grande maioria dos respondentes não consideram que foram informados sobre os riscos os quais podem estar expostos ao concordar com a utilização de seus dados pelo Instagram (figura 7). Evidentemente, não parece vantajoso para as empresas de redes sociais expor os danos que podem causar aos seus usuários, visto que seu modelo de negócio é baseado em obter o máximo de atenção possível das pessoas [Rubinstein and Good], e isso poderia afastá-los.

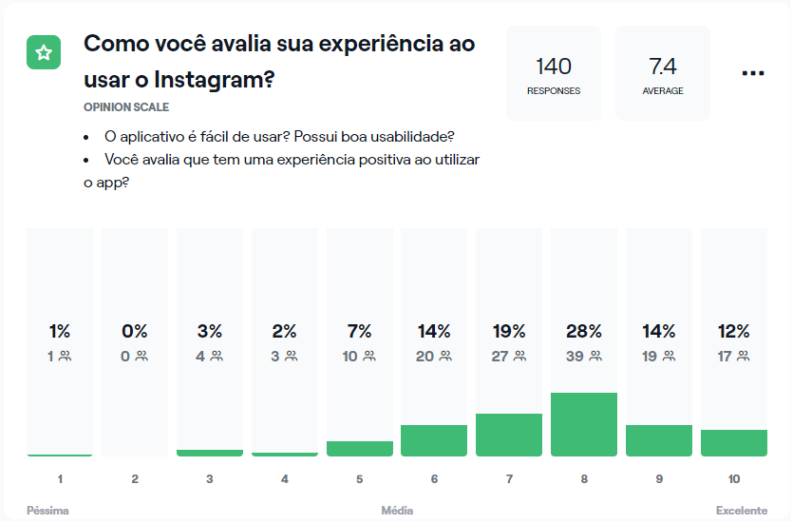

Figura 8. Avaliação de usabilidade

A figura 8 evidencia que, em relação à avaliação da experiência e usabilidade do Instagram, os respondentes classificaram a rede em uma nota média de 7.4 em 10, com a moda sendo a nota 8. Já era esperada uma nota alta neste ponto. O objetivo dessa pergunta é o de avaliar se as novas propostas de soluções podem impactar negativamente nesse ponto.

Figura 9. Confiança no Instagram

A figura 9 revela que, os respondentes da pesquisa avaliaram sua confiança no Instagram em uma nota média de 4.6 em 10. A pergunta considera a transparência e visibilidade em relação às práticas e objetivos da empresa, e a confiança no uso honesto e seguro de dados pessoais. Essa pergunta é importante para comparar essa nota na pesquisa final, e avaliar como as novas soluções propostas podem influenciar no nível de confiança dos usuários na empresa.

6.2. Reimaginando Funcionalidades do Instagram

A pesquisa inicial assumiu alguns critérios importantes para a avaliação de uma funcionalidade como mais ou menos nociva em termos de privacidade. Para classificar os itens, os usuários se basearam nas seguintes afirmações:

- Considero essa funcionalidade intrusiva.
- Considero que essa funcionalidade pode ser prejudicial, ou trazer riscos a minha saúde.(Ex.: Vício; distúrbios emocionais; etc...)
- Acredito que não tenho controle o suficiente dessa funcionalidade como gostaria.
- Considero que essa funcionalidade foi configurada sem o meu consentimento.
- Acredito que essa funcionalidade tira o meu livre arbítrio de utilização do sistema.

Os resultados da pesquisa inicial apontaram a seguinte classificação para as funcionalidades apresentadas (ordenadas pela mais nociva primeiro):

1. Propaganda direcionada.
2. Algoritmo de recomendação.
3. Geolocalização.
4. Controle de privacidade do perfil. (Controle de quem pode interagir comigo na rede)
5. *Scroll* Infinito. (*Posts* são carregados sempre ao rolar a página)
6. Notificações.
7. *Likes*, comentários e reações
8. Filtros em fotos e vídeos.

Sendo assim, foram submetidas as três primeiras a um processo de engenharia de *software* baseado em *Privacy by Design* para proposição de novas soluções; **Propaganda direcionada, algoritmo de recomendação, e Geolocalização**.

6.2.1. Abordagem de Engenharia de *Software* Utilizada

A implementação de *Privacy by Design* na engenharia de *software* é defendida pela doutora Ann Cavoukian em diversos trabalhos, como em seu artigo "Understanding How to Implement Privacy by Design, One Step at a Time", atestando o sucesso da aplicação desse conceito no mercado em grandes empresas como: TELUS, Intel, GE, e IBM [Cavoukian 2020].

Entretanto, as abordagens de Cavoukian são criticadas por serem vagas, ou pouco específicas em relação ao modo como os princípios de privacidade devem ser aplicados na prática. Essa definições vagas são sintomas da desconexão entre políticos e desenvolvedores ao tentar chegar em um acordo do que é preciso para cumprir tecnicamente com a proteção de dados. Esse parecer culminou no surgimento de novas proposições de metodologias de engenharia de *software* que destacam os pontos mais relevantes a serem aplicados durante o processo de desenvolvimento [Gurses et al. 2011].

Somente mecanismos de controle e transparência não são suficientes para mitigar os riscos relacionados à coleta massiva de dados das grandes empresas de redes sociais. Ademais, os dados sensíveis armazenados por essas empresas, e as informações que podem ser tiradas deles aumentam a atratividade desses dados por agentes mau intencionados. Sendo assim, riscos de grande magnitude podem ser associados a essas bases de dados, evidenciando a necessidade de uma preocupação especial acerca de quais dados devem ser armazenados, e sobre a segurança envolvida no processo [Gurses et al. 2011].

Mesmo que as organizações sejam honestas, e não tenham interesses em explorar os dados além do devido - o que muitas vezes não é verdade - os riscos da coleta massiva de dados ainda estão presentes. Por isso, a minimização de dados é apontada como um passo fundamental da implementação do PbD [Gurses et al. 2011].

A Comissão Federal de Comércio dos Estados Unidos (FTC) também propôs em 2012 um relatório de recomendações que - dentre outras coisas - propõe um *framework* para a implementação de um modelo de engenharia (baseado em Pbd) que reúna as melhores práticas para a proteção da privacidade dos consumidores. Esse relatório contempla diversas medidas legislativas relacionadas à privacidade de dados, que são fundamentais para que as empresas passem a ter uma abordagem de engenharia de software cada vez mais consciente. A adequação legislativa em relação à privacidade de dados não consegue acompanhar o ritmo da evolução da tecnologia, por isso, as propostas da comissão buscam acelerar as medidas de autorregulação por parte da indústria de *software* [Staff 2012].

A abordagem de engenharia de *software* utilizada neste trabalho reúne portanto as principais recomendações de estudos relacionados levando em consideração o que mais faz sentido para sua aplicação em um contexto de redes sociais. Para isso, a proposta é de que, durante a construção de cada funcionalidade de um sistema, seja seguido um fluxo de trabalho como o ilustrado na figura 9.

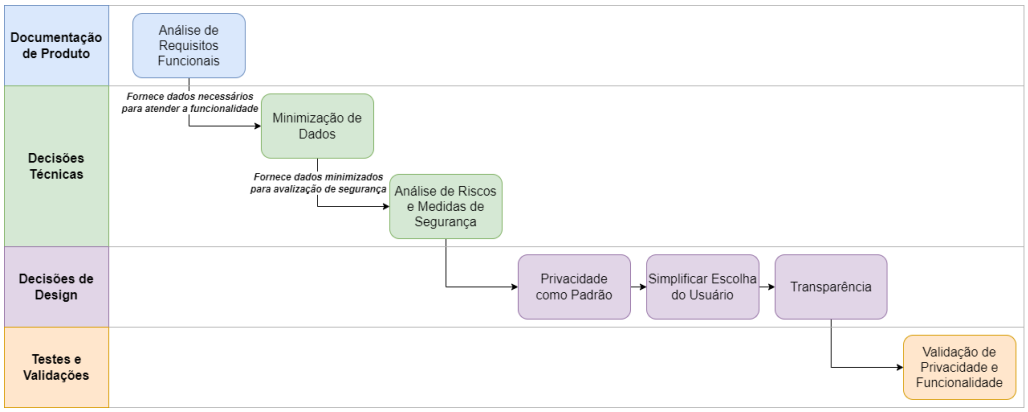

Figura 10. Fluxograma: Abordagem de Engenharia de *Software* Utilizada

6.2.1.1. Análise de Requisitos Funcionais

O primeiro passo para o desenvolvimento de um sistema com privacidade incorporada em todo o projeto é descrever claramente a sua funcionalidade. O objetivo do sistema deve ser bem definido e viável. Descrições vagas ou implausíveis têm um alto

risco de forçar os engenheiros a projetar um sistema que coletaria mais dados do que o necessário para atender os requisitos funcionais [Gurses et al. 2011].

Essa interpretação pode ser estendida a uma análise de todos os momentos em que houver a coleta de dados em um sistema, avaliando o seu propósito específico, e levantando preocupações acerca do tratamento do dado em todo o ciclo de vida dessa informação [Arass et al. 2017].

6.2.1.2. Minimização de Dados

A Lei Geral de Proteção de Dados diz o seguinte em relação à coleta e uso de dados: "... limitação do tratamento ao mínimo necessário para a realização de suas finalidades, com abrangência dos dados pertinentes, proporcionais e não excessivos em relação às finalidades do tratamento de dados;"[Brasil 2018].

Devem ser coletados e tratados somente os dados estritamente necessários para o cumprimento de uma funcionalidade proposta [Gurses et al. 2011]. As companhias deveriam limitar sua coleta de dados em prol de maior privacidade em seus produtos [Staff 2012].

Uma vez definidos os requisitos funcionais, é o momento de avaliar quais os dados são fundamentais para atingir esses objetivos. A proposição de técnicas de minimização de dados e possíveis soluções criptográficas de anonimização também devem ser levadas em consideração. Dados que extrapolam os requisitos definidos não devem ser coletados, ou devem ter seu uso explicitamente consentido por cada usuário [Staff 2012].

6.2.1.3. Análise de Riscos e Medidas de Segurança

É essencial que para a implantação do *Privacy by Design* exista segurança de ponta a ponta durante todo o ciclo de vida dos dados [Cavoukian 2010]. As bases de dados massivas presentes nas redes sociais são altamente visadas, portanto, um passo importante para atingir a privacidade de dados dos usuários é definir ações técnicas de segurança que protejam ao máximo o sigilo dessas informações [Gurses et al. 2011].

Certamente as grandes empresas de redes sociais já possuem um grande número de medidas protetivas em relação à segurança de seus dados. A ideia desse passo não é a de criticar as práticas de cibersegurança atuais, mas de identificar possíveis riscos específicos relacionados ao tipo de dados que é coletado em determinado momento, e chamar atenção para necessidade de implementação de tratativas de segurança mais assertivas.

6.2.1.4. Privacidade como Padrão

O pensamento de privacidade como padrão é considerado um divisor de águas para a implantação de PbD. Devemos assegurar que os dados pessoais são automaticamente protegidos em qualquer sistema. Se um indivíduo não fizer nada, a sua privacidade deve permanecer intacta. Não deveria ser necessária qualquer ação por parte do usuário para

proteger a sua privacidade - está integrado no sistema, por padrão [Cavoukian 2010].

Esse conceito deve ser considerado no momento do desenvolvimento de novas funcionalidades. É preciso identificar qual o comportamento padrão do sistema que preserva a privacidade do usuário, e alterá-lo somente mediante à manifestação clara e intencional do usuário.

6.2.1.5. Simplificar a Escolha do Usuário

Atualmente é comum que o consentimento do usuário em relação ao tratamento dos seus dados esteja consolidado em um longo documento de termos de privacidade, que é apresentado ao usuário logo em seu primeiro contato com o sistema. As escolhas acerca das práticas das empresas devem ser simplificadas e oferecidas num momento e num contexto em que o consumidor está prestes a tomar uma decisão sobre os seus dados [Staff 2012].

No contexto das redes sociais a escolha do tratamento dos dados pessoais em cada funcionalidade deveria ser apresentada no primeiro momento em que haveria a coleta da informação. Caso não haja o consentimento explícito do usuário em relação ao uso dos seus dados naquela funcionalidade, eles não devem ser coletados. Para funcionalidades que não fazem parte fundamental dos processos essenciais de funcionamento do sistema, deve haver um comportamento alternativo que preserve a privacidade dos dados.

A privacidade está intimamente ligada à capacidade de escolha dos usuários, e simplificar esse processo é fundamental para assegurá-la. Revogar um consentimento deve ser possível a qualquer momento e de maneira simples.

6.2.1.6. Transparência

As empresas devem aumentar sua transparência em relação às práticas relacionadas aos dados dos usuários. Os termos de consentimento devem ser mais claros, mais curtos, e mais padronizados para permitir melhor compreensão e comparação das práticas de privacidade [Staff 2012]. Esses termos devem contemplar de maneira clara e objetiva acerca de todas as especificidades do ciclo de vida daquela informação [Arass et al. 2017]. Além disso, as empresas devem ser claras e transparentes em relação aos objetivos de negócio e tecnologias envolvidos, além do compartilhamento desses dados com terceiros, se houver [Cavoukian 2010].

As empresas devem proporcionar acesso razoável dos usuários aos dados mantidos, e a extensão desse acesso deve ser proporcional à sensibilidade dos dados e à natureza da sua utilização [Staff 2012].

Além disso, devem ser expandidos os esforços para informar e educar os usuários em relação ao comércio de dados pessoais. Caso haja a comercialização de dados por parte da empresa, isso deve ser explicitamente informado e consentido pelo usuário [Staff 2012].

Os possíveis riscos de saúde relacionados à funcionalidades das redes sociais

também devem ser transparentes aos usuários, de forma que haja o consentimento do uso dos dados mesmo suscetível aos danos apresentados. Analogamente, podemos atribuir a transparência desse processo assim como é feita atualmente na comercialização de cigarros, em que os malefícios de saúde ficam estampados na traseira da embalagem, e o consumidor que consome esse tipo de produto está sendo explicitamente informado acerca dos danos de maneira clara, e ativamente consentindo com o risco.

6.2.1.7. Validação de Privacidade e Funcionalidade

Finalmente, é preciso validar que as soluções propostas estão respeitando as definições de privacidade consideradas, e que continuam atendendo os requisitos funcionais do produto. Práticas de desenvolvimento que prezam pela privacidade, ou segurança, mas deixam de atingir os objetivos de negócio, são inválidas [Gurses et al. 2011].

6.2.2. Propaganda direcionada

A propaganda direcionada é uma funcionalidade que, para fins de otimizar a recomendação de anúncios publicitários, realiza uma segmentação que leva em consideração interesses previamente demonstrados pelo público-alvo. Após a pesquisa inicial, notou-se um grande incômodo dos usuários em receber esse tipo de publicidade, apontando essa funcionalidade como a mais nociva em termos de privacidade.

6.2.2.1. Análise de Requisitos Funcionais

O objetivo da funcionalidade de propaganda direcionada é o de **otimizar a ferramenta de anúncios no aplicativo, aumentando a acurácia da recomendação de publicidade por público-alvo.**

A propaganda direcionada é a base do modelo de negócios do Instagram, pois a partir do perfil dos usuários o aplicativo é capaz de vender para seus clientes (anunciantes) a certeza de que os anúncios serão mostrados para as pessoas certas.

Encontrar na política de privacidade do Instagram exatamente quais dados são necessariamente utilizados nesse processo é uma tarefa complicada. As abas de configurações do aplicativo não nos fornecem essa informação, sendo necessário consultar diretamente a política de privacidade da empresa. Os termos explicitam o uso de dados provenientes de atividades dos usuários e informações fornecidas; seguidores e conexões; informações de aplicativos, navegadores e dispositivos; e informações de parceiros, fornecedores, e terceiros. Não há uma diferenciação clara sobre quais desses dados são utilizados especificamente para a propaganda direcionada, e muito menos um controle sobre a permissão de cada uma dessas fontes, porém é evidente que para atingir o requisito funcional definido não são necessários o uso de todos esses dados.

6.2.2.2. Minimização de Dados

Por se tratar de algoritmos de recomendação construídos por inteligência artificial, certamente há uma acúmulo enorme de dados para que sejam moldados os perfis dos usuários. Quanto mais informações obtidas, mais insumos o algoritmo têm para realizar as recomendações, e portanto, a funcionalidade se torna mais assertiva.

Nesse caso, a minimização de dados está relacionada ao armazenamento e tratamento somente daqueles dados os quais foram consentidos de serem compartilhados pelos usuários.

6.2.2.3. Análise de Riscos e Medidas de Segurança

Os tipos de dados envolvidos nessa funcionalidade são de grande atratividade, por se tratarem de informações sensíveis de construção de perfil de indivíduos. Esse tipo de informação deve ser submetido à medidas severas de segurança.

A política de privacidade do Instagram aponta que não compartilha nenhum dos dados coletados com terceiros.

6.2.2.4. Privacidade como Padrão

Percebe-se uma grande desadequação da funcionalidade de propaganda direcionada quanto à privacidade como padrão. Ao aceitar os termos de privacidade iniciais do Instagram, essa funcionalidade é dada como ativa, e os dados do usuário já começam a ser coletados e utilizados para construir o seu perfil.

Esse comportamento fere o conceito de privacidade como padrão defendido no PbD. Nesse caso, o correto seria que, por padrão, a funcionalidade de propaganda direcionada estivesse desligada desde o inicio do uso do aplicativo. Caso o usuário opte por ativá-la, isso deveria poder ser feito posteriormente com o seu consentimento.

6.2.2.5. Simplificar a Escolha do Usuário

Atualmente existe um controle de privacidade aos usuários nas configurações do Instagram que permite personalizar minimamente os dados compartilhados com a propaganda direcionada. No menu de configurações de anúncios é possível desativar o uso de dados pessoais provenientes de parceiros do Instagram. Ao desativar esse item, a empresa deixa de consultar informações pessoais de seus usuários a partir de fontes que não sejam suas atividades no aplicativo. Esse fluxo pode ser visualizado na figura 11.

Essa não é uma seção de fácil acesso aos usuários no aplicativo. Não é possível limitar quais dados de atividades no *app* são monitorados nessas funcionalidade, ou nenhuma outra configuração relacionada à limitação do uso dos dados.

A Comissão Federal de Comércio dos Estados Unidos defende que “As empresas devem dar aos consumidores uma escolha antes de coletar dados pessoais para Marketing próprio”[Staff 2012]. Essa é uma recomendação clara de que, as empresas podem até

coletar esses dados, mas o consentimento dos usuários deve ser dados de maneira mais simples e explícita.

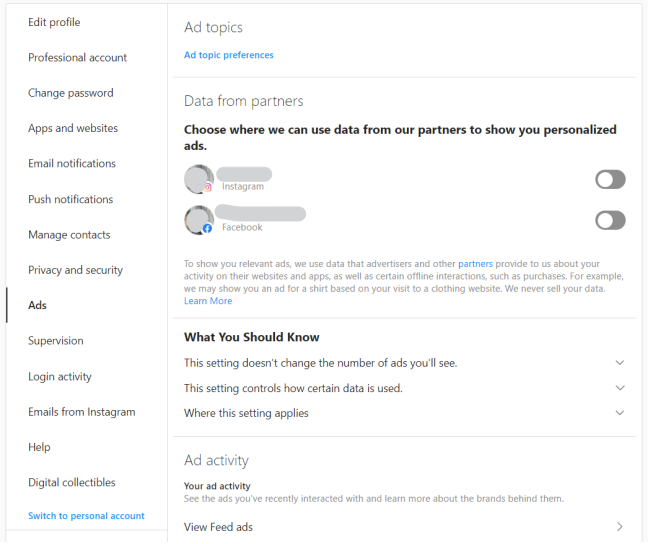

Figura 11. Controle atual de privacidade de propaganda direcionada

O momento da escolha acerca da privacidade de dados envolvida na funcionalidade deveria ocorrer no primeiro contato de um usuário com uma propaganda no aplicativo. Nesse momento, o usuário deveria ser informado de maneira clara e objetiva sobre a intenção de se utilizar dados pessoais para otimizar a ferramenta de recomendação de anúncios do Instagram, e exatamente quais são esses dados. O usuário então, poderia escolher entre compartilhar ou não os seus dados para este fim, sendo esta uma decisão que pode ser revogada a qualquer momento. Um modelo de alerta foi construído para ilustrar esse fluxo e pode ser visto pela figura 12.

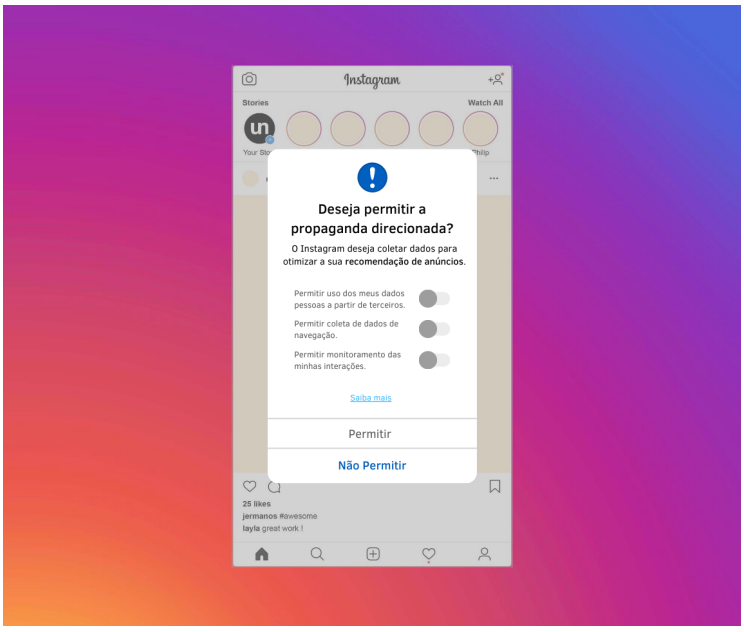

Figura 12. Modelo de alerta de consentimento para propaganda direcionada

6.2.2.6. Transparência

Os controles de privacidade existentes hoje no Instagram relacionados a propaganda direcionada são longos e vagos. A seção de configurações do aplicativo não deixa claro quais são exatamente quais são os dados que estão sendo coletados.

A política de privacidade da empresa, por sua vez, é bem clara em relação aos dados utilizados para o propósito de publicidade direcionada. Ainda assim, não é possível consultar de forma prática quais são os meus dados pessoais armazenados pelo Instagram, tão pouco gerenciá-los de alguma forma.

A transparência está presente na documentação do Instagram, mas deveria ser apresentada de forma mais simples. As configurações internas do aplicativo poderiam ser mais objetivas em relação aos dados coletados e seus objetivos.

6.2.2.7. Validação de Privacidade e Funcionalidade

Após as alterações sugeridas, em termos de privacidade, denota-se um ganho relacionado à capacidade dos usuários de decidirem acerca do uso de seus dados para propaganda direcionada no aplicativo. Desabilitar essa funcionalidade por padrão prioriza a privacidade dos usuários. Um alerta customizado permite com que os usuários possam ter um consentimento mais livre e explícito, além de habilitar um maior controle de suas permissões.

É necessário prover a escolha do usuário para essa funcionalidade visto que não se trata de um recurso fundamental para o funcionamento do sistema [Staff 2012]. As propagandas ainda podem ser apresentadas aos usuários sem que para isso haja o uso de dados pessoais específicos.

Após as sugestões realizadas a avaliação é de que, em termos funcionais, o Instagram tende a ser menos preciso em seu objetivo de negócio de otimizar a ferramenta de recomendação de anúncios para aqueles usuários que não permitirem o uso de seus dados. Ainda assim, o aplicativo não deixa de ser funcional, e pode ganhar uma confiança maior de seus usuários por oferecer melhores possibilidades de controle de privacidade.

6.2.3. Algoritmo de Recomendação

Os algoritmos de recomendação do Instagram se baseiam em dados dos usuários para decidir quais publicações serão exibidas em sua navegação no aplicativo. Se tratam de softwares de inteligência artificial que são capazes de traçar perfis altamente precisos para seus usuários.

6.2.3.1. Análise de Requisitos Funcionais

O objetivo dessa funcionalidade é o de **melhorar a experiência dos usuários no aplicativo por meio da recomendação de publicações coerentes com cada perfil**. Os

algoritmos de recomendação são incrivelmente eficientes, e são capazes de fornecer uma experiência personalizada e individualizada para cada usuário no Instagram. Aprimorar a recomendação de publicações significa fazer com que os usuários passem mais tempo navegando pelo aplicativo, e quanto mais tempo investido pelos usuários, mais rentável é o negócio [Rubinstein and Good].

Os dados envolvidos nessa funcionalidade são diversos. A política de privacidade do Instagram evidencia que são coletadas informações a partir de conteúdos criados; interações; câmera e voz; trocas de mensagens; metadados; interações com os conteúdos e o modo como isso é feito (pouco específico); uso de outros aplicativos por logins do Facebook; informações de transações efetuadas (incluindo informações de cartão de crédito); uso de *hashtags* e horários, frequência e duração das atividades.

Evidentemente, é possível criar recomendações de publicações personalizadas sem que haja necessariamente a coleta de todos esses dados. É possível também recomendar somente as publicações que estejam relacionadas as redes de conexões dos usuários. Sendo assim, a coleta de nenhuma dessas informações deveria ser obrigatória.

6.2.3.2. Minimização de Dados

A minimização de dados nessa funcionalidade - assim como aplicado para as propagandas direcionadas - está relacionada ao armazenamento e tratamento somente daqueles dados os quais foram consentidos de serem compartilhados pelos usuários.

6.2.3.3. Análise de Riscos e Medidas de Segurança

Vários dos dados coletados pelo Instagram que servem de insumo para os algoritmos de recomendação são de grande intrusividade, como o monitoramento de câmera, voz, trocas de mensagens, e informações de cartão de crédito. Esses dados constituem um conjunto de informações extremamente pessoais e específicas, as quais, se coletadas, devem ser submetidas à medidas de segurança especiais.

A política de privacidade do Instagram aponta - de maneira vaga - que já existem proteções especiais para dados de muito sensíveis, e que não compartilha nenhum dos dados coletados com terceiros.

6.2.3.4. Privacidade como Padrão

Há uma grande oportunidade de aplicação do conceito de privacidade como padrão para essa funcionalidade. O comportamento padrão que preza pela privacidade seria o de não utilizar ou coletar nenhum dos dados envolvidos nos algoritmos de recomendação para a construção do *feed* dos usuários. Nesse caso, as publicações apresentadas seriam apenas aquelas geradas por pessoas as quais o usuário segue na rede.

O usuário deve ser consultado para consentir acerca de cada tipo de dado o qual deseja permitir que seja coletado e utilizado para finalidade de recomendação de

publicações.

6.2.3.5. Simplificar a Escolha do Usuário

Atualmente não existe nenhum tipo de controle que permita personalizar desativar o algoritmo de recomendação, nem ao menos controla quais dados podem ou não ser utilizados nesse processo.

Para que o Instagram se adeque às recomendações de privacidade atuais é necessário que provenha a escolha do usuário quanto ao uso dos dados para este fim. Atualmente não nenhum tipo de escolha específico sobre essa funcionalidade.

O ideal seria prover a possibilidade de escolha no primeiro momento em que o usuário tem contato com o *feed* do Instagram. Nesse momento, o usuário deve ser informado de maneira clara e objetiva sobre a intenção de se coletar dados específicos para personalizar a experiência do usuário no Instagram. O usuário então, poderia escolher entre compartilhar ou não os seus dados para este fim, sendo esta uma decisão que pode ser revogada a qualquer momento. As figuras 13 e 14 ilustram sugestões aplicáveis para esse fluxo.

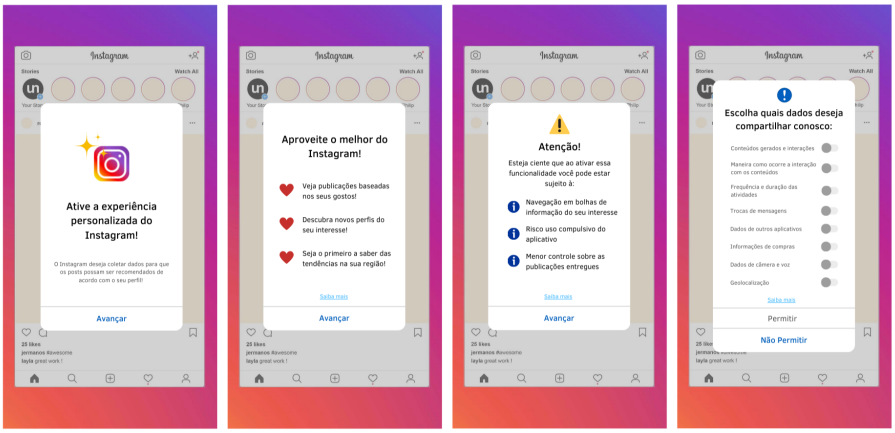

Figura 13. Modelo de alerta de consentimento para algoritmos de recomendação

6.2.3.6. Transparência

A política de privacidade do Instagram é a única fonte de transparência sobre o uso dos dados para algoritmos de recomendação. Não há informações de fácil acesso sobre a privacidade envolvida nessa funcionalidade em nenhuma configuração no aplicativo. Ainda assim, a política de privacidade se mostra vaga ao explicitar quais são os dados coletados em alguns casos, como o monitoramento da interação com os conteúdos; troca de mensagens; e uso de recursos de câmera e voz; por exemplo. As informações apresentadas também não esclarecem os motivos pelos quais alguns desses dados são coletados, e como eles podem ser úteis nos algoritmos de recomendação.

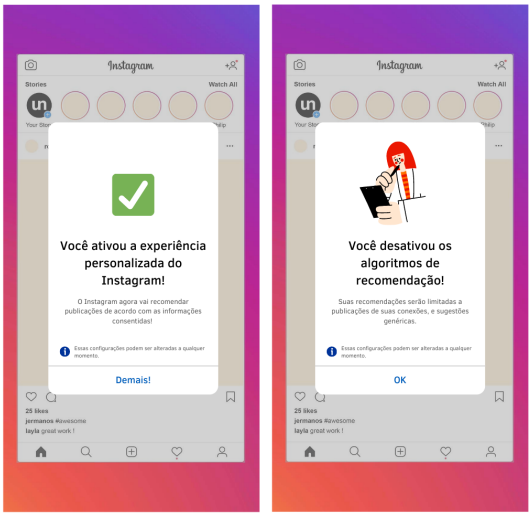

Figura 14. Modelo de confirmação das escolhas para algoritmos de recomendação

Além de adequações nos termos de privacidade, também é cabível a apresentação dessas informações - de forma mais simples e objetiva - nas configurações do aplicativo.

Outro problema identificado é a falta de transparência da empresa em relação aos riscos envolvidos na utilização dessa funcionalidade. Comprovadamente, o uso compulsivo dos aplicativos, e a formação de bolhas de informação têm se tornado grandes problemas relacionados ao uso das redes sociais [Papazoglou 2019] [Alkawaz et al. 2021]. A Comissão Federal do Comércio dos Estados Unidos ressalta a necessidade de todos os envolvidos ampliarem os esforços para educar os consumidores sobre dados comerciais e práticas de privacidade [Staff 2012]. Essas questões são importantes e devem ser apresentadas aos usuários antes de consentirem com o uso de seus dados e estarem sujeitos à essa funcionalidade.

6.2.3.7. Validação de Privacidade e Funcionalidade

Analisando o aplicativo em termos de privacidade, após a aplicação das sugestões sugeridas, nota-se que há uma grande diferença nas possibilidades de escolha dos usuários sobre o uso de seus dados em algoritmos de recomendação. O aumento do controle dos dados denota também num aumento da privacidade do sistema. Os termos de consentimento, que atualmente são apresentados de maneira difícil logo no cadastro, quando substituídos por termos claros e simplificados no momento do uso da funcionalidade permitem que o consentimento dos usuários seja mais consciente, livre, e explícito.

O Instagram é capaz de exercer seu pleno funcionamento sem que haja a recomendação de publicações a partir de dados dos usuários, portanto, prover a escolha do usuário para essa funcionalidade.

Mesmo após as sugestões de alterações propostas, o Instagram continua plenamente capaz de atender os requisitos funcionais definidos inicialmente. Por um lado, a não obrigatoriedade dos algoritmos de recomendação pode soar como um ponto negativo

ao negócio por fazer com que os usuários passem menos tempo navegando na rede social; porém, aumentar a privacidade, controle, e confiança dos usuários pode representar também uma vantagem competitiva para a empresa [Cavoukian 2020].

6.2.4. Geolocalização

Atualmente, grande parte dos computadores, celulares, relógios e *tablets* funcionam também como dispositivos GPS. O Instagram se utiliza dessa e de outras formas de localizar seus usuários para otimizar alguns de seus serviços.

6.2.4.1. Análise de Requisitos Funcionais

Existem dois objetivos principais para o uso da geolocalização no Instagram, são eles a **otimização de recomendações de eventos, anúncios e notícias**, e o uso dessas informações para **detectar atividades suspeitas** e ajudar a manter as contas seguras.

A geolocalização por si só não é uma funcionalidade do Instagram, mas sim um recurso que fornece informações relevantes para os objetivos mencionados.

Os dados de geolocalização presentes no Instagram podem ser obtidos via localização GPS dos dispositivos, e outros sinais como conexões via *bluetooth*. Mesmo que esse recurso esteja desativado nos aparelhos, outros tipos de dados de localização ainda são coletados, como endereços de IP, atividades no aplicativo como *check-ins* e informações fornecidas diretamente como as inseridas nos perfis de cada usuário.

É importante evidenciar que, os dados de geolocalização utilizados pelo Instagram não são fundamentais para o contexto de um produto de rede social. Essas informações são de grande relevância para aprimorar a acurácia dos serviços da empresa, mas nenhuma delas é obrigatória para o funcionamento do sistema ou de suas funcionalidades principais. Um aplicativo de rede social - enquanto um ambiente de comunicação, e compartilhamento de publicações - não precisa saber nada sobre a localização de seus usuários.

6.2.4.2. Minimização de Dados

As oportunidades de minimização de dados de geolocalização no Instagram estão relacionadas, primeiramente, a coletar somente aqueles dados cujo o uso foi consentido pelos usuários. Tendo isso em vista, para os dados que ainda forem coletados pela aplicação deve haver uma atenção especial quanto ao armazenamento mínimo dessa informações. Por exemplo, pode ser que para atender as demandas do Instagram não seja necessário armazenar todo o deslocamento dos usuários, mas somente os pontos estáticos por quais ele esteve. Outro ponto importante é realizar o descarte desses dados imediatamente após terem cumprido os objetivos envolvidos.

As propostas práticas de minimização dependem do conhecimento específico de quais dados são coletados e gerados sobre a localização dos usuários, além de como, e por

quanto tempo são armazenados; informações as quais não são especificadas na política de privacidade da empresa.

6.2.4.3. Análise de Riscos e Medidas de Segurança

Dados relacionados à geolocalização dos usuários são extremamente sensíveis, visto que podem mapear hábitos reais de deslocamento. Essas informações podem ser utilizadas por maus agentes para causar danos à essas pessoas no mundo real.

Considerando os riscos envolvidos no uso desses dados por pessoas erradas é de extrema importância a definição de políticas de segurança severas para o armazenamento desses dados. Ainda assim, a principal forma de mitigação desse risco seria a de minimizar os dados o máximo possível, afinal, se os dados não são armazenados não podem ser vazados. É importante entender a partir de quanto tempo após sua coleta esses dados perdem o valor funcional a ponto de serem desprezados pelo Instagram.

6.2.4.4. Privacidade como Padrão

O Instagram depende de configurações dos dispositivos para obter dados de localização GPS de seus usuários. O modo como isso é feito depende do sistema operacional utilizado, mas sempre existe a opção de ativar esse recurso o tempo todo, apenas durante o uso do aplicativo, ou nunca. A coleta desses dados deveria ser desativada por padrão no aplicativo, independentemente do sistema utilizado. Para esses dados em específico, o Instagram transfere sua responsabilidade, abrindo mão de uma oportunidade de corroborar com a privacidade de seus usuários.

Mesmo que o monitoramento via GPS seja desativado, outros dados relacionados à localização ainda são coletados, como endereços IP, informações fornecidas em cada perfil, e interações com o aplicativo como postagens e *check-ins*. Esses dados, por sua vez, são coletados e tratados por padrão, e não possuem nenhum tipo de controle relacionado à limitação de seu uso. Seguindo os conceitos de privacidade como padrão, o recomendado seria definir a coleta dessas como informações desativada. Posteriormente, caso desejável, o Instagram poderia prover a escolha ao usuário por fornecer ou não esses dados. Essa escolha deve ser feita separadamente no momentos específicos de contato com cada finalidade de uso.

6.2.4.5. Simplificar a Escolha do Usuário

É no mínimo chocante a forma como o Instagram limita as escolhas de seus usuários sobre o uso de dados de geolocalização. Para o caso da coleta de informações de localização do dispositivo os usuários ainda possuem alguma escolha, que está, porém, relacionada ao sistema operacional utilizado, e não ao aplicativo em si. Isso demonstra uma desadequação da empresa, ao transferir a responsabilidade, e não prover essa escolha aos usuários.

Para os outros tipos de dados coletados não há sequer escolha envolvida no processo, a não ser a de aceitar os termos e condições apresentados no momento do cadastro (que caso recusados inviabilizam o uso do aplicativo).

O processo das escolhas de privacidade dos usuários está constantemente relacionado às finalidades específicas para as quais os dados são utilizados. Nesse caso o uso dos dados de geolocalização possui dois objetivos, o de otimizar os algoritmos de recomendação, e o de proteger as contas detectando atividades suspeitas; portanto, é necessário prover a escolha ao usuário por compartilhar esses dados para cada funcionalidade em dois momentos distintos.

A sugestão proposta seria de incluir no alerta de configuração de privacidade dos algoritmos de recomendação um item relacionado ao consentimento para a coleta de dados de geolocalização, ccomo ilustrado na Figura 12.

Para prover a escolha do usuário quanto ao uso dessas informação com a finalidade de proteção da conta, pode ser incluído um botão do tipo *switch* - que deve estar desativado por padrão - no momento de um novo cadastro de usuário.

Ainda assim, configurações específicas acerca dos tipos de dados permitidos para cada finalidade deveriam estar disponíveis no aplicativos para serem alteradas a qualquer momento de maneira simples.

6.2.4.6. Transparência

É de suma importância que o Instagram seja transparente em relação ao uso dos dados de geolocalização - pelo menos de forma objetiva - também nas configurações do aplicativo, para que haja fácil acesso dos usuários a essas informações. Mesmo na política de privacidade da empresa existem problemas de transparência, como o fato de que não é possível visualizar ou gerenciar os dados individuais de geolocalização.

Para que haja mais transparência é preciso também prover a escolha dos usuários nos momentos relacionados a finalidade do uso dos dados, deixando claro assim o objetivo da coleta e tratamento dessas informações.

6.2.4.7. Validação de Privacidade e Funcionalidade

Há uma deficiência grande em termos de privacidade no Instagram quando se trata do gerenciamento dos dados de geolocalização. Após as alterações sugeridas percebe-se o ganho de privacidade quanto à inclusão de um controle de quais dados podem ser utilizados, e para quais fins.

O cumprimento dos requisitos funcionais definidos se mantém, mesmo após as alterações sugeridas. Ainda é possível que os algoritmos de recomendação - caso habilitados - exerçam seu papel de maneira eficiente com base em outras informações fornecidas pelos usuários. Permitir com que os usuários optem por não fornecer informações de geolocalização para a proteção das contas é um direito dos usuários. Impedir a coleta dessas informações nesse caso pode desabilitar essa funcionalidade, mas de maneira al-

guma se torna um grande problema de segurança, visto que ainda existem diversas outras maneiras - inclusive obrigatórias - de proteger o acesso indevido à essas contas.

Sendo assim, percebe-se que com a aplicação das propostas há um ganho notável de privacidade, e o aplicativo continua plenamente capaz de atender os requisitos funcionais delimitados por este contexto.

6.3. Análise de Resultados

Para avaliar como algumas das propostas desse trabalho podem impactar na visão dos usuários acerca de sua privacidade no Instagram foi realizada uma pesquisa final de validação com as mesmas perguntas presentes na pesquisa inicial. Desse modo, conseguimos obter métricas comparativas para validar a eficiência do uso da abordagem de engenharia de *software* proposta, no desenvolvimento de sistemas que forneçam maior privacidade aos seus usuários.

Nessa pesquisa foram apresentadas brevemente as alterações propostas relacionadas a **"Privacidade como Padrão"**, e a **"Simplificar a Escolha do Usuário"**. Antes de atender a pesquisa, os respondentes foram contextualizados sobre as alterações propostas nos fluxos de configuração de privacidade para **propaganda direcionada, algoritmos de recomendação, e geolocalização**. Em seguida, foi orientado para cada pergunta que as respostas devem ser relativas somente as alterações propostas, e não à aplicação do Instagram como um todo.

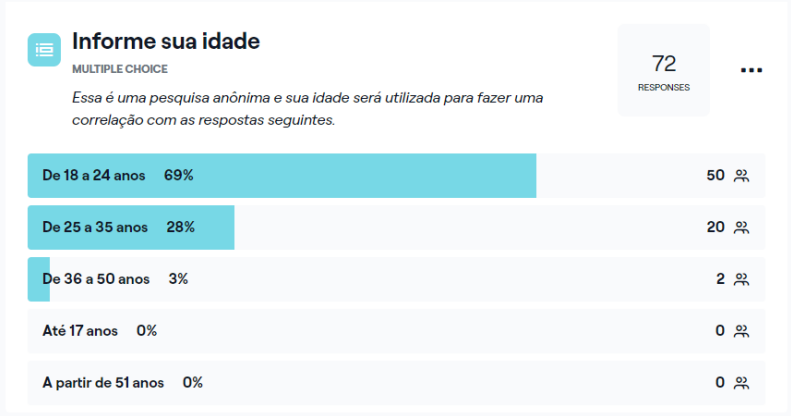

Figura 15. Idade dos respondentes da pesquisa de validação

Para essa última pesquisa, conforme visto na figura 15, houveram menos respostas, mas a média de idade dos respondentes permanece baixa, com absoluta maioria abaixo dos 35 anos.

Quanto a avaliação da ciência acerca de como, e quais dados são coletados pelo Instagram, houve um aumento da média das respostas de 2,7 para 3,2, em 5. Esses dados estão presentes na figura 16. É interessante observar que houve aumento expressivo na quantidade de pessoas que concordam completamente com a afirmação, representando apenas 6% na pesquisa inicial, e 24% na pesquisa de validação das propostas. Sendo assim, podemos afirmar que as propostas do trabalho influenciaram positivamente na visão dos usuários nesse ponto.

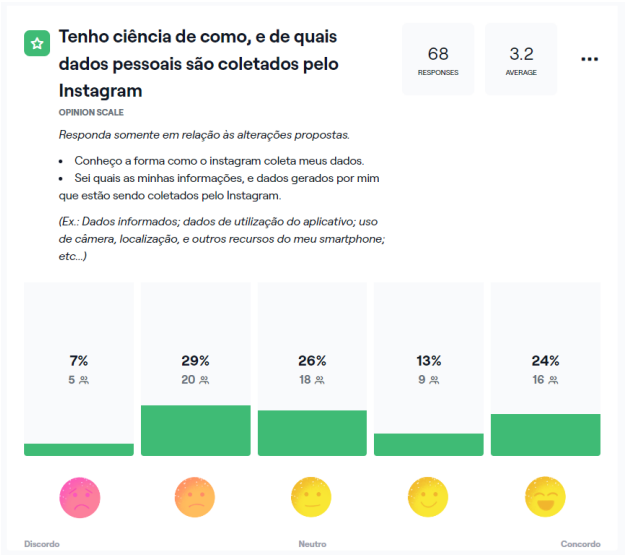

Figura 16. Ciência da coleta de dados - Validação

Nota-se que a percepção dos usuários em relação ao conhecimento da finalidade de uso dos seus dados melhorou para os novos fluxos propostos. A figura 17 traz os dados relacionados, e mostra que o percentual de pessoas que alegava ter total desconhecimento da finalidade do uso de seus dados no Instagram era de 21%, e na pesquisa de validação esse valor caiu para apenas 10%. Os respondentes que alegaram nota máxima em relação ao conhecimento da finalidade de uso dos dados saiu de 4% para 12%. A nota média subiu de 2,5 para 3.

Evidentemente, dadas as alterações propostas se tornou mais clara a finalidade do uso dos dados nesses casos. Ainda assim, há uma concentração de respostas em notas médias, o que pode ter ocorrido pois somente as alterações de fluxo apresentadas não são suficientes para garantir a privacidade completa dos usuários.

Apesar de haver um ligeiro aumento na nota dos respondentes em relação à confiança no armazenamento dos dados pelo Instagram (figura 18), percebe-se que as alterações apresentadas tiveram pouco impacto nesse quesito. Mesmo com os fluxos alternativos propostos, há uma desconfiança dos usuários do Instagram em relação ao armazenamento de seus dados pessoais, e essa é uma visão mais difícil de ser revertida. As demais alterações propostas no trabalho - que não somente de *design* - podem causar um impacto maior nesse aspecto.

Houve também uma melhoria da visão dos usuários acerca da manifestação de consentimento livre e explícito, como visto na figura 19. Essa melhoria, no entanto, está relacionada à uma maior concentração de notas médias nesse quesito.

A clareza dos usuários em relação aos riscos envolvidos nas funcionalidades apresentadas foi impactada positivamente, como visto na figura 20, porém a diferença em relação à pesquisa inicial é pouca, e as notas continuam concentradas em valores muito baixos para essa questão. Ainda há uma falta de esclarecimento dos usuários acerca das consequências envolvidas no uso de redes sociais, e para que essa visão seja alcançada são necessárias mais ações do que somente com alterações nos fluxos de configuração de

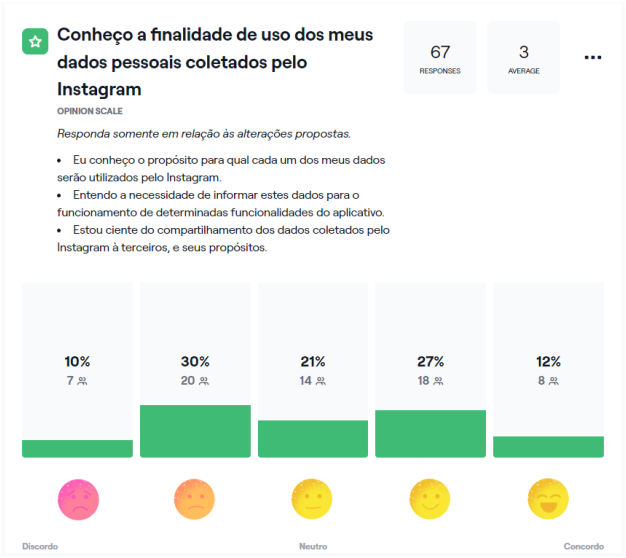

Figura 17. Ciência da finalidade dos dados - Validação

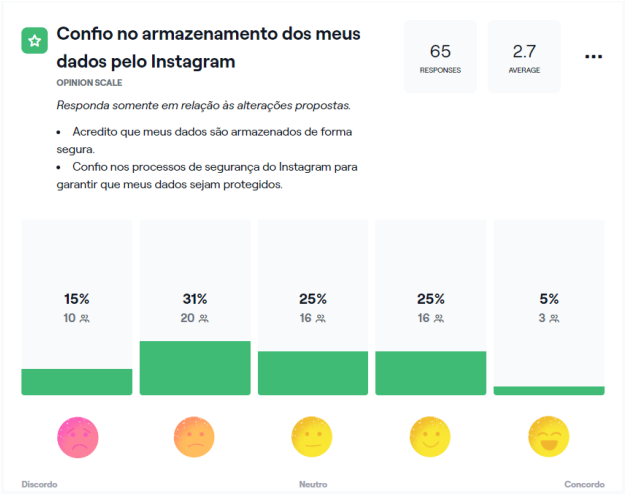

Figura 18. Confiança no armazenamento dos dados - Validação

privacidade, e alertas informativos breves.

A avaliação de usabilidade em relação aos fluxos propostos mantém uma avaliação muito próxima da obtida na pesquisa inicial (figura 21). Havia uma preocupação inicial relacionada a uma hipótese de que novos fluxos de configuração de privacidade poderiam impactar negativamente a visão dos usuários sobre a usabilidade do sistema, mas essa ideia não foi confirmada, visto que a média se manteve parecida em ambas pesquisas.

A despeito da média de avaliação baixa, a confiança dos usuários no Instagram teve um aumento considerável em sua nota, sendo 4,6 na pesquisa inicial e 5,4 na pesquisa de validação (figura 22). Esse aumento confirma uma ideia importante defendida em PbD de que a implementação de medidas corretas de privacidade não provoca necessariamente impactos negativos no produto, tendendo inclusive a se tornar uma vantagem competitiva de mercado para as empresas, visto que pode melhorar a percepção de confiança de seus usuários; como foi confirmado nesse caso [Cavoukian 2020].

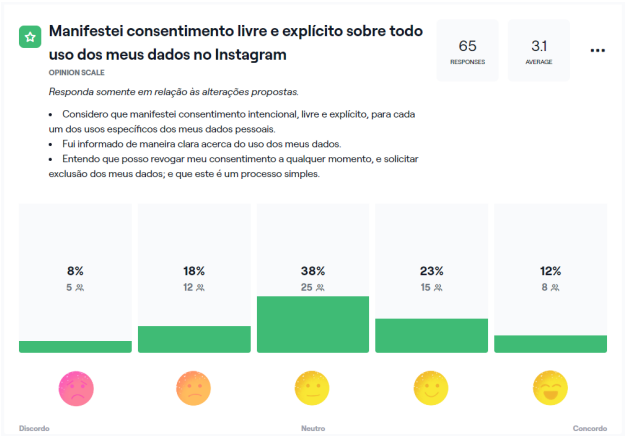

Figura 19. Consentimento livre e explícito - Validação

Figura 20. Riscos e consequências - Validação

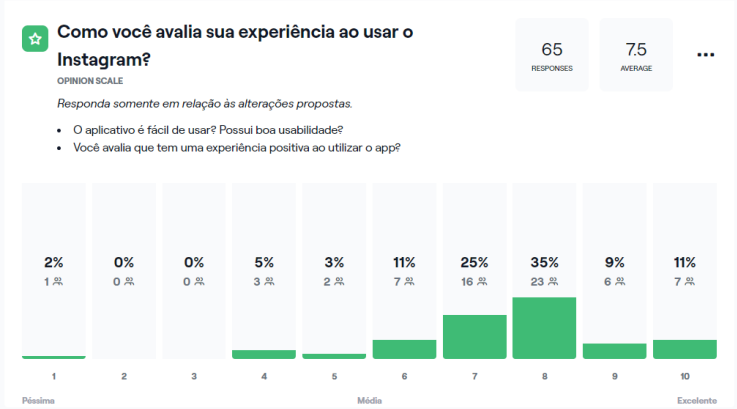

Figura 21. Avaliação de usabilidade - Validação

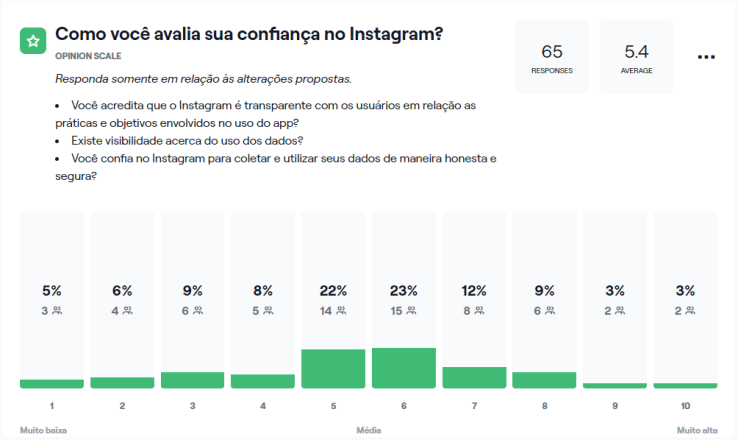

Figura 22. Confiança no Instagram - Validação

6.4. Conclusão

As deficiências de privacidade presentes no Instagram são diversas. Os usuários dessa rede social não possuem ciência clara sobre uso de seus dados pessoais, e há uma grande desconfiança em relação ao armazenamento e tratamento dos dados de maneira honesta e segura.

Esse trabalho foi capaz de identificar lacunas de privacidade no Instagram, principalmente quando se trata de controle do uso dos dados pelos usuários, e transparência em relação às práticas da empresa. Os passos presentes na abordagem de engenharia de *software* apresentada geraram propostas que impactaram positivamente nas percepções dos usuários em relação à sua privacidade no uso do sistema. Entre os principais resultados obtidos, nota-se que a confiança dos usuários no Instagram aumentou em 17% após as alterações propostas. As pesquisas realizadas apontam, principalmente, que após as proposições realizadas os usuários possuem maior ciência de como e quais dados pessoais são coletados pelo Instagram, e de qual a sua finalidade de uso.

Os principais resultados práticos gerados nesse trabalho envolveram proposições de novas abordagens de *design* para o Instagram. Evidentemente que, dadas essas propostas, houveram melhorias em todos os quesitos de privacidade avaliados, porém, para alguns casos como o nível de confiança dos usuários em relação ao armazenamento de

dados pelo Instagram, notou-se uma diferença de avaliação muito modesta. A principal justificativa observada é a de que algumas percepções dos usuários são mais complexas de serem subvertidas, e precisam de um conjunto de diferentes ações para gerar impactos reais em suas percepções.

Foi comprovada a expectativa inicial de que uma abordagem de engenharia baseada em *Privacy by Design* é eficiente no desenvolvimento de sistemas mais adequados à privacidade dos usuários de redes sociais. Os resultados apontam que é viável incluir privacidade nos processos de engenharia de *software* para o desenvolvimento de redes sociais, sem que o sistema deixe de atender os seus requisitos funcionais e sem que haja perda na qualidade da usabilidade dos sistemas. Pela comparação entre a pesquisa inicial e final nota-se que, após as alterações propostas, a avaliação dos usuários em relação à usabilidade do sistema se manteve muito próxima (7,4 na pesquisa inicial; e 7,5 na pesquisa de validação). Esses dados corroboram com o princípio de "Soma positiva"presente em PbD.

Os conceitos deste trabalho podem ser aplicados em outras funcionalidades do Instagram, ou ainda em novos estudos acerca de outras redes sociais. Para trabalhos futuros seria muito interessante o desenvolvimento completo de um sistema de rede social respaldado por uma abordagem de engenharia de software baseada em *Privacy by Design*.

Referências

Alkawaz, M. H., Khan, S. A., and Abdullah, M. I. (2021). Plight of social media users: The problem of fake news on social media. pages 289–293, Penang, Malaysia. IEEE.

Arass, M. E., Tikito, I., and Souissi, N. (2017). Data lifecycles analysis: Towards intelligent cycle. pages 1–8, Fez, Morocco. IEEE.

Brasil (2018). Lei nº 13.709, de 14 de agosto de 2018. *Diário Oficial [da] República Federativa do Brasil*.

Cavoukian, A. (2010). Privacy by Design: The 7 Foundational Principles. Revised: Oktober 2010.

Cavoukian, A. (2020). Understanding how to implement privacy by design, one step at a time. *IEEE Consumer Electronics Magazine*, 9:78–82.

Gurses, S., Troncoso, C., and Diaz, C. (2011). Engineering privacy by design. *Computers, Privacy Data Protection*, page 25.

ISO/IEC 29100:2011 (2011). Information technology — Security techniques — Privacy framework. Standard, International Organization for Standardization, Geneva, CH.

Kugler, L. (2020). Are we addicted to technology? page 15–16.

Miller, E., Rahman, M. R., Hossain, M., and Ali-Gombe, A. (2022). I don't know why you need my data: A case study of popular social media privacy policies. In *Proceedings of the Twelfth ACM Conference on Data and Application Security and Privacy*, CODASPY '22, page 340–342, New York, NY, USA. Association for Computing Machinery.

Morales-Trujillo, M. E., García-Mireles, G. A., Matla-Cruz, E. O., and Piattini, M. (2019). A Systematic Mapping Study on Privacy by Design in Software Engineering. *CLEI Electronic Journal*, 22(1).

Newport, C. (2020). When technology goes awry. page 49–52.

Papazoglou, A. (2019). Understanding facebook's algorithm could change how you see yourself.

Rubinstein, I. and Good, N. Privacy by design: A counterfactual analysis of google and facebook privacy incidents.

Staff, F. T. C. (2012). Protecting consumer privacy in an era of rapid change–a proposed framework for businesses and policymakers.