Uma Análise Sobre o Uso de DLTs no Tratamento de Dados Pessoais: Aderência aos Princípios e Direitos elencados na LGPD

Anderson Boa Morte[1], Anália Meira[1], Rostand Costa[2], Dênio Mariz[1]

[1]Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB)
Caixa Postal 58.015-435 – João Pessoa – PB – Brasil

[2]Laboratório de Aplicações de Vídeo Digital (LAVID)
Centro de Informática – Universidade Federal da Paraíba (UFPB)
Caixa Postal 58.055-000 – Paraíba – PB – Brasil

{vieira.anderson, analia.meira}@academico.ifpb.edu.br,

rostand@lavid.ufpb.br, denio@ifpb.edu.br

Abstract*. Distributed Ledger Technology (DLT) can be very useful for the processing of personal data in accordance with the Brazilian Data Protection Law (LGPD), due to characteristics such as transparency and security. However, other characteristics such as immutability and distributed character can make this task difficult. Thus, this paper analyzes the challenges of reconciling DLT and data processing in accordance with the LGPD. As an object of analysis, the Datavalid project of SERPRO - Federal Data Processing Service, hired by Uber, was used in a hypothetical scenario in which data processing was performed using the Hyperledger Fabric.*

Resumo. *A tecnologia de registro distribuído (DLT – Distributed Ledger Technology) pode ser muito útil para o tratamento de dados pessoais em conformidade com a Lei Geral de Proteção de Dados Pessoais (LGPD), devido a características como transparência e segurança. No entanto, outras características como a imutabilidade e o caráter distribuído podem dificultar essa tarefa. Assim, este trabalho analisa os desafios da conciliação entre DLT e o tratamento de dados em conformidade com a LGPD. Como objeto de análise, utilizou-se o projeto Datavalid do SERPRO - Serviço Federal de Processamento de Dados, contratado pela Uber, em um cenário hipotético em que o tratamento de dados foi realizado utilizando-se o Hyperledger Fabric.*

1. Introdução

A Lei 13.709/2018 conhecida como Lei Geral de Proteção de Dados Pessoais (LGPD) surgiu como mecanismo para regular o tratamento de dados pessoais em território nacional, permitindo a garantia dos direitos fundamentais de liberdade e privacidade não só dos cidadãos brasileiros, mas de qualquer pessoa natural cujos dados pessoais sejam aqui tratados. Essa iniciativa adequa a legislação brasileira aos regulamentos já aplicados em outras partes do mundo, como é o caso do GDPR (*General Data Protection Regulation)* no contexto da União Europeia e no qual a LGPD foi baseada [Brasil 2018].

A LGPD define "tratamento de dados" como qualquer operação realizada com dados pessoais, tais como: utilização, avaliação, processamento, classificação, armazenamento e comunicação [Brasil 2018]. O GDPR possui definição semelhante

[PARLIAMENT 2016], mas o regulamento europeu é considerado mais específico e detalhado que a lei brasileira, de forma que se pode afirmar que um tratamento em conformidade com o GDPR também estará em conformidade com a LGPD, embora o contrário não seja necessariamente verdadeiro.

Ambos os regramentos definem perfis. O **Titular** do dado pessoal (*data subject* no GDPR) é o detentor do direito fundamental de privacidade e transparência no tratamento de dados. O **Controlador** (*Controller* no GDPR) é o responsável pela tomada de decisões sobre o tratamento. O **Operador** (*Processor* no GDPR) é aquele que realiza a operação de tratamento em nome do controlador. Há também a figura das Autoridades de Proteção de Dados (DPA - *Data Protection Authorities)* que é um órgão governamental que, dentre outras atribuições, possui o dever de fiscalizar e orientar, inclusive tecnicamente, sobre o tratamento de dados pessoais [ITSRIO 2019].

No contexto da orientação técnica, algumas instituições como o Parlamento Europeu [Finck 2019] e a Autoridade Francesa de Proteção de Dados [CNIL 2018] têm publicado documentação técnica abordando a utilização de tecnologias de livro-razão distribuído (DLT– *Distributed Ledger Technology*) - mais especificamente a *blockchain* (cadeia de blocos) - para o tratamento de dados pessoais, reforçando a aplicabilidade desta tecnologia para este fim. No entanto, o que se pode desprender destas documentações, bem como da análise de outros estudos como os realizados por [Onik et al. 2019] e [Politou et al. 2019] é que o uso de *blockchain* para esse propósito deve ser realizado de maneira bastante criteriosa. Deve-se analisar as características da *blockchain* que podem auxiliar ou dificultar o tratamento em conformidade com a legislação. Alguns dos direitos dos titulares dos dados elencados na LGPD tais como o direito ao esquecimento, ou seja, ao apagamento dos seus dados, e o direito à retificação (edição) dos dados, esbarram em características da *blockchain* como a imutabilidade. Para se alinhar com esses direitos previstos, o trabalho de [Onik et al. 2019] observa a importância de salvar os dados pessoais fora da cadeia (off-*chain)*. Esse processo é feito a partir da separação dos dados em duas categorias: **dados pessoais** (PII - *Personally Identifiable Information* e PPII - *Potential Personally Identifiable Information*) e **dados não pessoais** (NPII - *Non-personally Identifiable Information*), sendo o armazenamento de PII e PPII (dados pessoais) realizado fora da cadeia (*off-chain*), enquanto um *hash* desses dados pessoais juntamente com os NPIIs são armazenados na cadeia.

A Autoridade Francesa de Proteção de Dados (CNIL[1]) emitiu um relatório técnico abordando outros aspectos relacionados ao uso de *blockchain* para o tratamento de dados pessoais [CNIL 2018]. O documento aborda, entre outros pontos, questões como a da responsabilização, ressaltando que em um contexto naturalmente distribuído como o de *blockchain* não é uma tarefa trivial identificar quais entes da rede poderiam ser responsabilizados em caso de alguma infração à lei de proteção de dados. Em linhas gerais, sugere-se cautela na adoção de ***blockchains* públicas** para fins de tratamento de dados pessoais, tendo em vista as dificuldades de identificar responsáveis nesse contexto.

A CNIL ainda aponta a questão da localização dos nós *blockchain,* que podem estar em países que não possuem legislação quanto à proteção de dados pessoais, o que naturalmente seria considerado inadequado sob essa ótica.

---

[1] CNIL - *Commission Nationale de l'Informatique et des Libertés*

Considerando o exposto, é possível observar diversos pontos que merecem ser detalhados e submetidos à análise durante a fase de definição de requisitos funcionais e não-funcionais do projeto de um novo sistema de informação ou de ajustes em sistemas existentes visando a sua conformidade com a LGPD. Deste modo, o objetivo deste trabalho é analisar aspectos técnicos referentes à utilização da DLT para o tratamento de dados pessoais em conformidade com a LGPD, sobretudo na observância a alguns princípios e direitos dispostos na lei.

A metodologia utilizada se baseia na análise de documentos regulatórios relacionados com o GDPR, além de outras fontes na literatura científica, para compilar definições, destacar problemas e desafios inerentes ao processo de utilização da DLT visando estabelecer a conformidade com as normas.

Adotou-se como objeto de análise e discussão uma aplicação existente desenvolvida pelo SERPRO - Serviço Federal de Processamento de Dados chamada Datavalid [SERPRO 2020], no cenário onde essa foi contratada pela Uber para a validação dos dados pessoais dos seus motoristas. Aplicando-se a metodologia proposta, foram avaliados os possíveis impactos da utilização da DLT sobre o processo adotado no Datavalid no tratamento (compartilhamento) dos dados dos motoristas e sugeridos caminhos para conciliar a dinâmica da aplicação com alguns aspectos da LGPD, como por exemplo, os direitos ao esquecimento e à retificação de dados pessoais. O restante do documento está organizado como segue. Na Seção II são abordados os trabalhos relacionados a este artigo. Na Seção III são discutidos os desafios de conciliação entre o uso de DLT e LGPD. Na Seção IV são listadas as recomendações gerais decorrentes dessa discussão. Na Seção V é apresentado o projeto Datavalid em maiores detalhes. Na Seção VI realiza-se a discussão da aplicação dessas recomendações no projeto Datavalid. Por fim, a Seção VII apresenta a conclusão e possíveis trabalhos futuros.

2. Trabalhos relacionados

O trabalho apresentado por [Puccinelli 2019] discute os “pontos de tensão” referentes ao uso da DLT para o tratamento de dados pessoais em conformidade com o GDPR. Este trabalho baseia-se fortemente no relatório do CNIL [CNIL 2018] sobre o tema e no estudo do Parlamento Europeu [Finck 2019] que também aborda a temática.

O diferencial deste trabalho está em trazer esta discussão especificamente para o contexto da LGPD, além de avançar na análise trazendo para o arcabouço teórico contribuições como os trabalhos de [Onik 2019] e [Politou et al. 2019]. Adicionalmente, analisa como se daria o tratamento de dados considerando um projeto real, relevante no contexto brasileiro, o projeto Datavalid do SERPRO, considerando também uma DLT bastante utilizada na indústria de software: o Hyperledger Fabric. Assim, a análise avança na discussão teórica, e ademais se propõe também a analisar uma situação real de tratamento de dados em território brasileiro.

3. Desafios de Conciliação entre a DLT e a LGPD

Primeiramente, é importante salientar que, de acordo com os estudos de [CNIL 2018], [Finck 2019] e [Puccinelli 2019], não há uma DLT específica que seja compatível com o GDPR. O que se ressalta nesses estudos é que cada situação de tratamento de dados pessoais utilizando-se de *blockchain* deve ser analisada individualmente, caso a caso.

Baseando-se nos estudos de [Puccinelli 2019] referentes ao GDPR, pode-se afirmar que são quatro as áreas de tensão, aparentemente inconciliáveis, decorrentes da utilização da tecnologia *blockchain* para o tratamento de dados pessoais em conformidade com a legislação:

1. identificação clara dos responsáveis pelo tratamento;
2. anonimização para minimizar riscos à privacidade;
3. revisão de decisões automatizadas baseadas em *smart contracts;*
4. exercício de direitos e outros pontos do GDPR frente às características da cadeia.

3.1 Identificação dos responsáveis pelo tratamento

Todos os nós em uma rede *blockchain* possuem e mantêm uma réplica do registro de transações efetuadas na forma de uma livro-razão (ledger*)* distribuído. Essa definição reforça o caráter naturalmente distribuído de uma rede *blockchain.* [Greve 2018].

Em redes *blockchain* privadas (ou permissionadas), a exemplo do Hyperledger Fabric, a entrada de novos nós é realizada mediante permissão e acesso controlado. Assim, *blockchains* privadas são naturalmente mais adequadas para a identificação dos nós, o que viabiliza a eventual responsabilização em caso de alguma infração à lei de proteção de dados pessoais. Por outro lado, em *blockchains* públicas, a exemplo da Bitcoin e da Ethereum, não há controle da entrada de participantes, tornando a identificação de um responsável muito mais complexa, senão inviável [Greve et al. 2018], [Puccinelli 2019].

Ainda no âmbito da responsabilização é importante a apresentação do conceito de nós validadores, também chamados de mineradores. Esses nós são aqueles que competem entre si quanto à resolução de um protocolo de consenso, à base de desafios criptográficos, sobre a ordem em que as transações serão realizadas e armazenadas permanentemente na cadeia de blocos e depois replicadas em cada servidor (nó). O nó vencedor é recompensado por isso. No Bitcoin a recompensa se dá em criptomoedas [Greve et al.2018].

A CNIL [CNIL 2018] aborda sobre responsabilidade de nós validadores e de nós participantes em uma rede *blockchain,* reconhecendo três tipos de atores:

- nós clientes: que têm o direito de ler e manter uma cópia da cadeia;
- nós participantes: que têm o direito de fazer entradas (ou seja, fazer uma transação para a qual solicitam validação); e
- nós mineradores: que validam uma transação e criam blocos.

Embora a CNIL reforce que esta não é uma questão pacífica, ela sinaliza que os mineradores não seriam responsáveis, já que apenas validam transações e não determinam as finalidades nem os meios de tratamento, tarefa esta atribuída ao **Controlador,** no contexto do GDPR. Para a CNIL, seriam classificados como responsáveis pelo tratamento, todos aqueles que introduzem dados pessoais na *blockchain,* sejam pessoas jurídicas ou pessoas físicas cujo tratamento relaciona-se à sua atividade profissional ou comercial.

Neste contexto, a CNIL cita dois exemplos de responsáveis: a) os notários que registram títulos de seus clientes em uma *blockchain;* e b) as instituições bancárias que armazenam dados de seus correntistas em uma *blockchain* como parte do processo de administração de clientes.

3.2 Anonimização como minimização de riscos à privacidade

De acordo com o Art. 12 da LGPD, dados anonimizados não são considerados dados pessoais, de modo que não se aplicam a esses as regras dispostas nessa legislação. A lei, entretanto, preocupa-se com uma possível utilização de técnicas que permitam a reversão do dado anonimizado, vide artigo 5º da mesma, onde encontra-se a sua definição.

Preocupação semelhante pode ser observada no GDPR que, segundo [Polido 2018], adotou um conceito dito expansionista, definindo que dado pessoal pode referir-se a qualquer informação que permita a identificação do titular, ainda que o vínculo não seja estabelecido de imediato. É uma definição que parte da premissa de que dados anônimos podem ser passíveis de reversão. Aprofundando-se nesta temática, devem ser avaliados os seguintes riscos quando se adota alguma técnica de anonimização [Puccinelli 2019]:

- **risco de reversão**: quando o processo de criptografia aplicado à anonimização pode ser revertido, e assim reconstituir-se os dados originais (e.g. desencriptação por força bruta);
- **risco de vinculação:** quando se pode vincular os dados criptografados a um determinado indivíduo mediante a análise de padrões de uso ou contexto, ou por comparação com outras partes da informação. Seriam como os PPII – *Potential Personally Identifiable Information,* descritos em [Onik et al. 2019].

A CNIL esforça-se por sumarizar a questão apontando cinco pontos de orientação, são eles:

1. Ao gravar dados pessoais na *blockchain* deve-se antes aplicar os princípios de *privacy by design:* Devem ser priorizadas ações desde o início do projeto no intuito de preservar a privacidade do titular dos dados;
2. Caso a escolha seja de fato utilizar a *blockchain* para o tratamento de dados pessoais, deve-se adotar o armazenamento destes fora da cadeia (*off-chain*), tal como apresentado no modelo proposto por [Onik et al. 2019];
3. Se mesmo assim, a decisão for por armazenar dados na cadeia, estes devem ser armazenados sob a forma de um *commitment* criptográfico, ou de um *hash,* ou de um *hash* aplicado sobre o dado pessoal cifrado.

   A ideia em priorizar-se a técnica de *commitment* (compromisso) a um *hash* é – no caso do *hash* - evitar que se possa deduzir o texto original submetido à função. Se o universo de dados submetido ao *hash* for muito pequeno, se for binário, por exemplo, quente ou frio bastaria a um adversário submeter a função de *hash* a ambos os valores para poder deduzir o valor original. A fim de evitar esta dedução, uma técnica bastante utilizada é concatenar um valor aleatório ao valor que se vai aplicar ao *hash* de forma a dificultar a dedução do valor original. [Puccinelli 2019] chama estas técnicas de "salgar e apimentar" (do espanhol, salado y pimentado).

   Por sua vez, um *commitment* (compromisso), grosso modo, pode ser entendido como o estabelecimento de um contrato onde a mensagem original (texto plano) seria colocada em um envelope e lacrada. A ideia é o comprometimento digital com a mensagem original, que será ocultada, e quando necessário, poderá ser revelada. Uma vez firmado o compromisso (ou seja, o contrato) ele não pode ser alterado. Um *nonce* é utilizado para cada novo contrato,

lembrando que *nonce* é um número aleatório que pode ser utilizado somente uma vez (n-*number, once-uma vez)*. Posteriormente, o conteúdo do envelope pode ser revelado e verificado [Greve et al. 2018, p. 5-7].

4. Se nenhuma dessas técnicas for possível, a finalidade do tratamento está justificada e a avaliação de impacto diz que o risco residual é aceitável, podendo-se então subir à cadeia o dado pessoal em texto plano ou submetido à função de *hash.*
5. Quanto a medidas de segurança, a CNIL preocupa-se quanto ao potencial de falha do algoritmo de anonimização, de suas vulnerabilidades ou quanto à confidencialidade da cadeia como um todo.

3.3 Revisão de decisões automatizadas baseadas em *smart contracts*

O artigo 20 da LGPD versa sobre o direito do titular do dado pessoal de solicitar a revisão de decisões automatizadas, incluídas aí as decisões acerca da construção automatizada de perfis, por exemplo, perfil de crédito, perfil pessoal ou profissional.

Importante enfatizar o conceito de contratos inteligentes (*smart contracts*). Esses permitem a realização de transações de forma automatizada. Uma vez que os nós participantes da rede firmem entre si este contrato, e os termos sejam atendidos, a transação é realizada [Greve et al. 2018].

O relatório da CNIL [CNIL 2018] descreve que a decisão exclusivamente automatizada decorre da execução do contrato inteligente pelos entes da rede *blockchain.* Neste cenário, no que diz respeito às medidas adequadas para salvaguardar os direitos do titular dos dados, estes devem poder requerer intervenção humana nesse processo, de forma a contestar a decisão tomada após a execução do contrato inteligente. O controlador deve, portanto, fornecer a possibilidade de intervenção humana, permitindo que o titular conteste a decisão, mesmo que o contrato já tenha sido executado e independentemente do que está registrado na *blockchain.*

3.4 Exercício de direitos e outros pontos do GDPR frente às características da cadeia

A seguir uma discussão sobre direitos elencados no GDPR e sua compatibilidade com o uso de DLT.

3.4.1 Consentimento

A LGPD estabelece como uma condição para o tratamento de dados pessoais o consentimento expresso do titular. A lei é ainda mais rígida quando o tratamento envolve dados sensíveis (dados biométricos, étnicos, religiosos etc.), condicionando o seu tratamento a um consentimento específico e destacado.

Em *blockchains* privadas (permissionadas) é possível identificar-se claramente para quem o titular está dando consentimento para o tratamento. No entanto, em *blockchains* públicas, onde por padrão a entrada de novos entes não requer permissionamento, é difícil determinar para quem está se dando o consentimento, pois não fica claro quem é o controlador (responsável pelo tratamento) [Puccinelli 2019].

3.4.2 Direito à informação

[Puccinelli 2019] destaca a compatibilidade entre o uso da tecnologia *blockchain* para o tratamento de dados pessoais e o direito à informação, salientando que controlador terá que proporcionar informação concisa, clara e de fácil acesso ao titular antes de enviá-la para validação.

Esta compatibilidade pode ser explicada pela necessidade de consenso distribuído entre os participantes da rede. Imaginando-se que o titular seria um dos nós da rede, e como há uma cópia do *ledger* em cada nó, a informação sobre o tratamento de dados pessoais torna-se do conhecimento de todos os integrantes. Garante-se assim, no contexto da LGPD, a adequação ao princípio da transparência - vide artigo 6º, inciso VI - onde é garantido ao titular a visibilidade em torno do que venha a ser realizado com os seus dados pessoais.

3.4.3 Direito de acesso e portabilidade

O artigo 18 da LGPD, respectivamente em seus parágrafos segundo e quinto, garante que o controlador forneça ao titular acesso aos seus dados pessoais, garante também a portabilidade desses dados a um outro controlador. [Brasil 2018].

O relatório da [CNIL 2018] considera que estes não são direitos incompatíveis com a tecnologia *blockchain.* Há, entretanto, a reiterada ressalva, de que em *blockchains* públicas não é trivial a identificação do controlador.

3.4.4 Minimização de dados pessoais

Considerando-se a anonimização do dado pessoal através de uma criptografia de chave pública, deve-se armazenar na cadeia, além do próprio dado anonimizado, apenas e tão somente, a sua chave pública. Desta forma, apenas o titular do dado pessoal - detentor da chave privada de criptografia - teria acesso ao conteúdo, garantindo assim o seu direito à privacidade. Entretanto, vale ressaltar o cuidado que se deve ter na guarda da chave privada, pois a perda dessa corresponderia à perda irrecuperável do dado [Puccinelli 2019].

3.4.5 Direito ao esquecimento

Caso o dado pessoal seja armazenado em texto claro na *blockchain* seria, por definição,

devido à sua característica de imutabilidade, impossível a sua exclusão ou edição. Embora, recentemente alguns estudos tenham oferecido alternativas a esta imutabilidade. Estudos como os apresentados por [Politou et al. 2019], frente aos novos requisitos como o direito de ser esquecido do GDPR, RtbF, do inglês *Right to be Forgotten,* têm apresentado soluções alternativas e técnicas criptográficas que permitem uma *blockchain* mutável.

3.4.6 Direito de retificação

O vínculo entre os blocos em uma rede *blockchain* é realizado através do *hash* (resumo criptográfico) de todas as transações presentes em um determinado bloco. Caso se desejasse alterar o conteúdo de um determinado bloco, por exemplo, para retificar um determinado dado pessoal lá armazenado, o *hash* seria alterado e o vínculo entre os blocos seria quebrado.

Portanto, pode-se afirmar que a característica da imutabilidade é algo intrínseco à tecnologia *blockchain.* Embora, reitera-se, já existam iniciativas como as listadas em [Politou et al. 2019] que descrevem *blockchains* mutáveis. Assim, a impossibilidade de modificar os dados ingressos em um bloco, leva o controlador a inserir os dados atualizados ou retificados em um novo bloco, considerando-se que uma transação posterior cancelaria uma operação anterior, mesmo que esta siga presente na cadeia [Puccinelli 2019].

3.4.7 Transferência internacional de dados e territorialidade

A LGPD em seu artigo 3º, inciso I, restringe a aplicabilidade da lei à "operação de tratamento em território nacional". No entanto, em *blockchains* **públicas**, onde os membros da rede podem estar em qualquer lugar do globo, não há a garantia de que os entes que realizam tratamento de dados estejam em território nacional.

Uma outra questão referente à territorialidade está no artigo 33, inciso I. Este versa sobre a transferência internacional de dados. Grosso modo, é permitida a transferência para países ou organismos internacionais que tenham grau de proteção de dados pessoais equivalente ao previsto na LGPD. Assim, essa transferência pode ser muito problemática em *blockchains* públicas sem permissionamento ou mesmo em *blockchains* privadas cujas exigências quanto às transferências não estejam convenientemente implementadas [Puccinelli 2019].

4. Recomendações gerais quanto à utilização de DLTs para o tratamento de dados pessoais em conformidade com a LGPD

Discutidas as questões relativas ao uso de DLTs para o tratamento de dados pessoais em conformidade com o GDPR/LGPD, consolida-se no Quadro 1 as recomendações quanto ao uso dessa tecnologia nesse contexto. A ideia é com base nessas ter-se subsídios para analisar a sua utilização no contexto do compartilhamento de dados pessoais no projeto Datavalid.

Quadro 1: Recomendações gerais sobre o uso da DLT para tratamento de dados pessoais em conformidade com a LGPD

Questão ou direito da LGPD	Recomendação
identificação dos responsáveis pelo tratamento	utilizar *blockchains* privadas (permissionadas); mineradores não podem ser responsabilizados (CNIL, 2019).
anonimização	seguir recomendações da CNIL hierarquicamente, nesta ordem: 1. adotar *privacy by design;* 2. realizar armazenamento *off-chain;* 3. se não for possível, usar *commitment*; 4. se não for possível, usar *hash;* 5. se nenhum dos anteriores for possível, assumir risco residual
possibilidade de revisão de decisões automatizadas baseadas em *smart contracts*	implementar processo para possibilitar a intervenção humana, caso haja contestação da decisão automatizada realizada pela execução do *smart contract.*
consentimento	utilizar *blockchains* privadas (permissionadas) com *smart contracts;* implementar consentimento nos *smart contracts;* demais orientações: não utilizar *blockchains públicas* pois não se sabe para quem está se dando o consentimento;

	obs. não seria recomendável, por exemplo, utilizar o Ethereum (pública, embora possua *smart contracts*), nem o Bitcoin (pública e sem *smart contracts*).
direito à informação	utilizar *blockchains* privadas com *smart contracts,* implementar direito à informação nos *smart contracts.*
direito de acesso	idem direito à informação, implementar acesso no *smart contract.*
direito à portabilidade	sem maiores problemas quanto à sua implementação com *blockchain* (CNIL, 2019).
minimização de dados pessoais na cadeia	somente armazenar na cadeia o dado anonimizado e a sua chave pública, assim seria respeitada a privacidade do usuário, detentor da chave privada associada.
direito ao esquecimento	utilizar armazenamento *off-chain*
direito à retificação	utilizar armazenamento *off-chain;* o *commitment* dos dados retificados deve ser armazenado em um novo bloco, anulando logicamente o bloco com o dado original.
transferência internacional e tratamento em território nacional	utilizar *blockchains* privadas (permissionadas); não utilizar *blockchains* públicas pois não se pode garantir a localização dos nós. obs. estes podem estar em países que não têm legislação compatível com a LGPD; obs. os nós que tratam os dados pessoais devem estar em território nacional

5. Objeto de análise: Projeto Datavalid

Apresenta-se, nessa seção, o objeto de análise deste artigo: o projeto Datavalid.

5.1. Contextualização

Em linhas gerais, o projeto Datavalid consiste no provimento de uma interface que permite a validação de dados pessoais por meio de consultas a bases de dados oficiais do Estado brasileiro, como CPF, CNPJ e CNH. [SERPRO 2020]. A escolha deste projeto como objeto de análise deve-se a três fatores:

1. o projeto realiza operações de **tratamento de dados pessoais,** portanto dentro do contexto da LGPD;
2. o projeto é relevante, considerando-se que foi desenvolvido pela maior empresa pública de tecnologia da informação do Estado Brasileiro, o SERPRO; possui **abrangência nacional** e se utiliza de grandes bases de dados: CPF, CNH e CNPJ.
3. o projeto teve seus serviços contratados por uma **empresa de alcance global:** a Uber [SERPRO 2019].

Importante ressaltar que não se tem informações técnicas detalhadas acerca de como fora modelado e implementado o projeto Datavalid tal como o mesmo encontra-se atualmente. Assim sendo, o objetivo da análise é discutir como se daria uma hipotética utilização do Hyperledger Fabric (*blockchain* privada) aplicado ao tratamento (compartilhamento) de dados pessoais neste projeto, baseado nas recomendações quanto à conciliação entre DLT e o tratamento de dados pessoais em conformidade com a LGPD, reunidas no Quadro 1.

5.2. Escopo: Características e Serviços

O projeto Datavalid [SERPRO 2020] oferece dois tipos de validação. A validação da identidade do usuário a partir do fornecimento dos seus dados biométricos (impressão digital ou reconhecimento facial), e a validação cadastral, permitindo a validação de dados pessoais como CPF (Cadastro de Pessoas Físicas), nome, sexo, data de nascimento etc. Esses serviços são realizados através de consultas a bases de dados originais do Estado brasileiro para validar os dados fornecidos, retornando, dentre outras informações, um índice de similaridade que permite avaliar por exemplo se a foto da face disponibilizada está ou não associada aos demais dados pessoais informados.

6. Análise do tratamento (compartilhamento) de dados

Considerando os desafios de conciliação entre a DLT e o tratamento de dados em conformidade com a GDPR/LGPD e as recomendações apresentadas no Quadro 1, esta seção discute como se daria o compartilhamento de dados no projeto Datavalid assumindo-se a hipotética adoção do Hyperledger Fabric no escopo da contratação desta solução pela Uber.

6.1. A escolha do tipo de DLT

Uma análise das recomendações apresentadas no Quadro 1, permite afirmar que *blockchains* privadas (permissionadas) são mais apropriadas para o tratamento de dados pessoais em conformidade com o GDPR/LGPD. Sendo assim, optou-se em realizar a análise do tratamento de dados no Datavalid utilizando-se hipoteticamente o Hyperledger Fabric.

Vale destacar que o Hyperledger engloba vários projetos, cada um com a sua característica em particular, sendo o Hyperledger Fabric a iniciativa mais consolidada e que tem por característica o foco na modularidade. Esta arquitetura modular permite, por exemplo, o uso de diferentes protocolos de consenso e diferentes linguagens de programação para a criação de *smart contracts* [Rebello et al. 2019].

6.2. A identificação clara dos responsáveis pelo tratamento

O serviço de validação de dados disponibilizado pelo Serpro foi contratado pela Uber para a validação em tempo real da fotografia da face do motorista de aplicativo [SERPRO 2019]. Nesta situação, e considerando-se os papéis elencados na LGPD, **o motorista de aplicativo é o titular dos dados pessoais, a Uber o controlador e o SERPRO o operador**.

O SERPRO qualifica-se como operador por ser o provedor da solução de tratamento de dados pessoais. A Uber, contratante do Datavalid, seria qualificada como controladora pois o SERPRO realiza o tratamento de dados pessoais em seu nome, vide artigo 5º incisos V I e VII da LGPD[2].

---

[2]VI - controlador: pessoa natural ou jurídica, de direito público ou privado, a quem competem as decisões referentes ao tratamento de dados pessoais.

VII - operador: pessoa natural ou jurídica, de direito público ou privado, que realiza o tratamento de dados pessoais em nome do controlador.

Neste cenário, considerando a adoção de uma ***blockchain*** **privada (permissionada)** como é o caso do Hyperledger Fabric é possível a identificação do Controlador, responsável pelo tratamento, neste caso, a Uber.

Por analogia, conforme o relatório CNIL [CNIL 2018], da mesma forma que o notário que registra o título de seu cliente em uma *blockchain* seria o responsável pelo tratamento, a Uber seria caso inserisse os dados dos motoristas de aplicativo em uma *blockchain*.

Ademais, vale destacar, que o Fabric provê um serviço de filiação à rede chamado MSP - *Membership Service Provider* onde cada participante registra sua participação e obtém acesso ao sistema. A responsabilidade de cada nó é atribuída durante o processo de filiação. Alguns nós podem ser configurados para ter poder de administração, podendo modificar as permissões dos demais.

Portanto, a tecnologia permite a identificação dos responsáveis, no entanto tal como apontado por [Puccinelli 2019] e [CNIL 2018] a responsabilização deve ser tratada caso a caso, conforme for configurada a rede no caso concreto.

6.3. Anonimização como minimização de riscos à privacidade

Adotando-se os princípios de *privacy by design* mencionados [CNIL 2018] e [Puccinelli 2019], em se considerando uma rede *blockchain* formada pelo motorista de aplicativo (titular dos dados pessoais), a Uber e o SERPRO, é recomendável o armazenamento dos dados pessoais compartilhados pelo titular fora da cadeia (*off-chain*), tal como apresentado no modelo proposto por [Onik et al. 2019]. Portanto, sem recorrer-se a nenhuma técnica de anonimização dos dados compartilhados.

6.4. O exercício de direitos e outros pontos da LGPD frente às características da cadeia

Acerca dos direitos elencados na LGPD considerados incompatíveis com a característica de imutabilidade da cadeia: direito ao esquecimento e direito à retificação, analisando-os sob a perspectiva de compartilhamento de dados pessoais em uma rede formada pelo motorista (titular dos dados), a Uber (controladora) e o SERPRO (operador) pode-se afirmar que o modelo proposto por [Onik et al. 2019] permite um tratamento em conformidade com a LGPD.

Este modelo fundamenta-se na premissa de que **um novo bloco será criado e adicionado à cadeia** *blockchain* sempre que houver um compartilhamento de dados pessoais bem-sucedido envolvendo pelo menos dois dos três perfis descritos na LGPD: titular, controlador e operador. Isto é, um bloco será criado e adicionado à cadeia toda a vez que houver um compartilhamento de dados pessoais bem-sucedido entre o titular e o controlador, ou entre o controlador e o operador.

Importante salientar que o uso de uma ***blockchain*** **privada (permissionada),** como o Hyperledger Fabric e, por conseguinte, da necessidade de consenso distribuído entre os participantes da rede, permite que se dê ciência da operação de compartilhamento de dados pessoais a todos os integrantes da rede, inclusive ao motorista (titular dos dados).

Antes de descrever o bloco, ressalta-se que cada nó na rede *blockchain* representa um dos perfis descritos na LGPD: titular, controlador e operador. O bloco em questão é formado pelos seguintes componentes:

- **cabeçalho do bloco:** guarda informações gerais sobre um evento de compartilhamento de dados pessoais. Armazena: *hash* de dados do cabeçalho anterior; tempo da transação (*timestamp)*; e estilo de codificação (*charset*).
- **contador de transações:** Guarda o número de compartilhamentos de dados já realizados entre o titular e o controlador, ou entre o controlador e o operador. É incrementado em um (+1) sempre que um compartilhamento de dados pessoais bem-sucedido é realizado entre esses entes.
- **dados da transação**: importante ressaltar que essa seção do bloco não pode armazenar dados pessoais (PII e PPII) diretamente, pois devido à imutabilidade da *blockchain* esses não poderiam ser modificados ou excluídos, infringindo assim respectivamente os incisos III e VI do artigo 17 da LGPD, que versam sobre os direitos do titular dos dados pessoais.

  Portanto, esta seção do bloco armazena:

  - *commitment* de dados pessoais (PII e PPII), em cumprimento ao proposto no relatório CNIL [CNIL 2018];
  - termos de compartilhamento (*smart contract)*
  - NPII;

6.5. Fluxo de dados

Considerando-se a disponibilização dos dados pessoais (PII e PPII) e não pessoais (NPII) do motorista (titular) para o controlador (Uber).

Em um primeiro momento, nesta primeira comunicação entre usuário e controlador ocorre a execução do *smart contract*, contendo os termos de compartilhamento, ou seja, as regras que devem ser cumpridas por usuário e controlador para que este compartilhamento seja efetuado. Vale destacar que o *smart contract* também engloba o algoritmo de consenso. Nestes termos, deve estar presente também o consentimento do usuário, de modo a cumprir este fundamental requisito da LGPD.

Uma vez executado com sucesso o *smart contract*, firmado o consenso entre as partes, ocorre de fato o compartilhamento de dados entre titular e controlador. Em um segundo momento, no controlador, um separador é executado, ocorrendo assim a separação entre NPII e dados pessoais (PII e PPII).

Em um terceiro momento, são registrados no bloco:

- os termos de uso: (*smart contract*), englobando neste os termos de privacidade, regulamentação, usabilidade, regras de distribuição, processo de notificação de violação e consenso;
- NPII;
- um *commitment* de PII e PPII;
- a identificação do titular;
- a identificação do controlador;
- são registrados no banco de dados local: PII e PPII

7. Conclusão e trabalhos futuros

Este trabalho se propõe a analisar aspectos técnicos a respeito da utilização da DLT para o tratamento de dados em conformidade com a LGPD. Devido a características como segurança, transparência e descentralização, DLTs podem ser bastante úteis para o tratamento de dados pessoais em conformidade com LGPD. No entanto, devido a outras características como a imutabilidade e mesmo a descentralização, o uso requer bastante cautela.

Diversos aspectos foram analisados de forma a trazer recomendações de como esta tecnologia pode ser bem empregada no tratamento de dados pessoais em conformidade com a legislação. Pode-se afirmar que o uso de ***blockchains* privadas (permissionadas)**, como o Hyperledger Fabric, e o **armazenamento *off-chain*** mostram-se apropriados aos desafios de conciliação entre DLTs e tratamento de dados pessoais em conformidade com o GDPR/LGPD. A conciliação, no entanto, torna-se bastante problemática em *blockchains* públicas não permissionadas, notadamente quanto à dificuldade de responsabilização, e da garantia que o tratamento de dados seja realizado em território nacional.

Como trabalho futuro recomenda-se a propositura de um modelo seguindo-se as recomendações decorrentes das análises aqui efetuadas, e a implementação desse utilizando-se do Hyperledger Fabric, de modo a enriquecer a análise com os dados decorrentes dessa implementação.

Referências

Brasil (2018). Lei n o 13.709, de 30 de agosto de 2018. Disponível em: https://bit.ly/2VfiMWX. Acesso em: 20 de dezembro de 2019.

CNIL (2018). Blockchain. Solutions for a responsible use of the blockchain in the context of personal data. *CNIL Report* , page pp. 10.

Finck, M. (2019). Blockchain and the General Data Protection Regulation: can distributed ledgers be squared with European data protection law?: study. European Parliament, Brussels.

Greve, F., Sampaio, L., Abijaude, J., Coutinho, A., Valcy, Í., and Queiroz, S. (2018). Blockchain e a Revolução do Consenso sob Demanda. *Minicursos do Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos,* page 52.

ITSRIO (2019). Lei Geral De Proteção De Dados Pessoais (Lgpd) E Setor Público. page 40. Acesso em: 20 de dezembro de 2019.

Onik, M. M. H., Kim, C. S., Lee, N. Y., and Yang, J. (2019). Privacy-aware blockchain for personal data sharing and tracking. *Open Computer Science* , 9(1):80–91.

PARLIAMENT, E. (2016). REGULATION (EU) 2016/679 of 27 April 2016. Disponível em: https://bit.ly/3c6dDag. Acesso em: 20 de dezembro de 2019.

Polido, F. B. P. e. a. (2018). *GDPR e suas repercussões no direito brasileiro – Primeiras impressões de análise comparativa.* IRIS-BH.

Politou, E., Casino, F., Alepis, E., and Patsakis, C. (2019). Blockchain Mutability: Challenges and Proposed Solutions. *IEEE Transactions on Emerging Topics in Computing* .

Puccinelli, O. R. (2019). Blockchains y otras formas de contabilidad distribuida (DLT) y su impacto en la protección de los datos personales. Disponível em: https://bit.ly/3akRFjl. Acesso em: 16 de março de 2020.

Rebello, G., Camilo, G., Silva, L., Souza, L., Guimarães, L., Alchieri, E., Greve, F., and Duarte, O. (2019). Correntes de Blocos: Algoritmos de Consenso e Implementacão na Plataforma Hyperledger Fabric. *Jornada de Atualização em Informática 2019,* pages 93–148.

SERPRO (2019). Release: Uber fecha contrato com o SERPRO. Disponível em: https://bit.ly/2ya167b. Acesso em: 20 de fevereiro de 2020.

SERPRO (2020). Datavalid. Disponível em: https://bit.ly/2x9qx8S. Acesso em: 18 de fevereiro de 2020.