DOI: 10.1590/1413-81232020256.1.11792020

Preservação da privacidade no enfrentamento da COVID-19: dados pessoais e a pandemia global

Personal data usage and privacy considerations in the COVID-19 global pandemic

Bethania de Araujo Almeida (https://orcid.org/0000-0001-8918-2661) [1]
Danilo Doneda (https://orcid.org/0000-0001-9535-3586) [2]
Maria Yury Ichihara (https://orcid.org/0000-0001-8590-6212) [1]
Manoel Barral-Netto (https://orcid.org/0000-0002-5823-7903) [1]
Gustavo Correa Matta (https://orcid.org/0000-0002-5422-2798) [3]
Elaine Teixeira Rabello (https://orcid.org/0000-0002-8324-1453) [4]
Fabio Castro Gouveia (https://orcid.org/0000-0002-0082-2392) [5]
Mauricio Barreto (https://orcid.org/0000-0002-0215-4930) [1]

Abstract *Data has become increasingly important and valuable for both scientists and health authorities searching for answers to the COVID-19 crisis. Due to difficulties in diagnosing this infection in populations around the world, initiatives supported by digital technologies are being developed by governments and private companies to enable the tracking of the public's symptoms, contacts and movements. Considering the current scenario, initiatives designed to support infection surveillance and monitoring are essential and necessary. Nonetheless, ethical, legal and technical questions abound regarding the amount and types of personal data being collected, processed, shared and used in the name of public health, as well as the concomitant or posterior use of this data. These challenges demonstrate the need for new models of responsible and transparent data and technology governance in efforts to control SARS-COV2, as well as in future public health emergencies.*

Key words *Personal data, COVID-19, Data governance, Technology governance, Public health emergency*

Resumo *Dados ganham cada vez mais importância e valor na busca de respostas para enfrentar a COVID-19 tanto para a ciência quanto para as autoridades sanitárias. Em virtude da dificuldade de realizar diagnóstico da infecção na população em geral, iniciativas apoiadas em tecnologias digitais vêm sendo desenvolvidas por governos ou empresas privadas para possibilitar rastreamentos de sintomas, contatos e deslocamentos de modo a apoiar estratégias de acompanhamento e avaliação na vigilância de contágios. A despeito da importância e necessidade dessas iniciativas, questionamentos acerca da quantidade e tipos de dados pessoais coletados, processados, compartilhados e utilizados em nome da saúde pública, bem como os concomitantes ou posteriores usos desses dados, suscitam questionamentos éticos, legais e técnicos. Desafios que apontam para a necessidade de novos modelos de governança de dados e de tecnologias, responsáveis e transparentes, para controlar o Sars-Cov2 e as futuras emergências de saúde pública.*

Palavras-chave *Dados pessoais, COVID-19, Governança de dados, Governança de tecnologias, Emergências de saúde pública*

[1] Centro de Integração de Dados e Conhecimentos para Saúde, Fiocruz Bahia. R. Mundo, Trobogy. 41745-715 Salvador BA Brasil. baraujo2010@gmail.com
[2] Instituto Brasiliense de Direito Público. Brasília DF Brasil.
[3] Escola Nacional de Saúde Pública Sérgio Arouca, Fiocruz. Rio de Janeiro RJ Brasil.
[4] Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro. Rio de Janeiro RJ Brasil.
[5] Casa Oswaldo Cruz, Fiocruz. Rio de Janeiro RJ Brasil.

A crescente produção e uso de dados apoiados em tecnologias digitais cada vez mais potentes e especializadas possibilitou a emergência de novas formas de produção de conhecimento apoiadas em algoritmos e modelagens computacionais sofisticadas. Nesse novo contexto, dados ganham valor e importância através de mobilizações diversas que envolvem negociações, interesses sociais, políticos e econômicos[1].

Em tempos de pandemia pela COVID-19, pela necessidade urgente de responder de forma rápida aos desafios colocados pela introdução de um novo agente etiológico e pela peculiaridade da doença trazendo riscos à vida e à saúde das pessoas, a utilização de dados pessoais de diferentes fontes vem sendo requerida para explorar questões científicas a partir de características da população, de dados laboratoriais, hospitalares, entre outros, desde que orientada pelo embasamento ético e legal.

Um esforço mundial vem sendo desenvolvido para que as lacunas do conhecimento sobre a pandemia sejam respondidas rapidamente pela ciência e pela organização e prática nos serviços de saúde a fim de que medidas ágeis, oportunas e eficientes possam ser adotadas pelas autoridades sanitárias de cada país relacionadas a capacidade de diagnóstico, manejo clínico e reabilitação dos casos de COVID-19 e estratégias de prevenção. Ações que necessitam articular medidas governamentais e de diferentes segmentos da sociedade para maximizar os esforços no controle da doença.

A OMS orienta que a partir da avaliação de seus riscos, cada país se prepare para responder aos possíveis cenários visando executar as medidas necessárias rapidamente para a redução da transmissão do vírus e de seus impactos econômicos e sociais[2]. Sendo, portanto, necessários dados de boa qualidade para a descrição dos padrões epidemiológicos básicos. Contudo, as incertezas em torno da COVID-19 também se estendem à qualidade dos dados disponíveis para pesquisadores, não só para o entendimento do padrão epidemiológico da doença como para a construção de modelos matemáticos para gerar evidências que subsidiem as decisões de gestores nos seus diversos níveis.

Considerando a dificuldade de realizar diagnóstico da infecção na população em geral, iniciativas tecnológicas vêm sendo desenvolvidas para que seja possível rastreamentos de sintomas, contatos e deslocamentos considerados componentes importantes para subsidiar estratégias de monitoramento e vigilância de contágios pelos governos. Grandes apostas têm sido feitas no desenvolvimento de aplicativos que coletam dados pessoais, de geolocalização e circulação das pessoas[3]. Práticas que suscitam questionamentos sobre tipos e quantidade de dados necessários, e dos desafios éticos, legais e técnicos que permeiam a coleta, o acesso, o compartilhamento e a utilização desses dados[4,5].

A Apple e a Google recentemente firmaram uma parceria, visando garantir a interoperabilidade entre os sistemas iOS e Android, para a criação de uma ferramenta de rastreamento para a COVID-19. Segundo as empresas, as pessoas terão a opção de participar, mas não mencionam a opção de retirada de consentimento a qualquer tempo. O sistema, de acordo com especificações divulgadas[6], apresenta semelhanças com soluções que vêm sendo referidas como 'contact tracing' e inspiram-se, grosso modo, em implementação já operacional em Singapura e propostas em desenvolvimento na Europa como a DP-3T (*Decentralized Privacy-Preserving Proximity Tracing*)[7] ou o projeto PEPP-PT (*Pan-European Privacy-Preserving Proximity Tracing*)[8], esta proposta, assim como a *Safe Paths Platform* do MIT, buscam maximizar a privacidade[9].

Em termos gerais, estas soluções, que entram na classificação de sistemas de *Contact Tracing*, funcionam com a troca de identificadores anônimos entre telefones próximos via Bluetooth, após a instalação de um aplicativo disponibilizado pela autoridade de saúde nacional ou eventualmente pelo próprio sistema operacional, a depender de como opera a solução. Quando uma pessoa tiver resultado positivo para o coronavírus, ela irá fazer este registro no aplicativo, que o transmitirá para autoridades de saúde no seu respectivo país. Em seguida, as pessoas com as quais teve contato nos 14 dias anteriores serão alertadas que estiveram em contato com alguém que apresentou diagnóstico positivo para a doença[10]. Como se trata de tecnologias que ainda estão em fase de desenvolvimento e amadurecimento, há diferenças entre implementações que, com o passar do tempo, podem demonstrar ser muito significativas, como, por exemplo, já parece ser o enfoque centralizado do PEPP-PT em contraste ao descentralizado do DP-3T.

O panorama aponta para que, nas próximas fases em que a sociedade estará se adaptando a conviver com o vírus, caberá ao uso de dados pessoais e aplicativos ou dispositivos um papel de destaque não somente na medição do contato, porém para finalidades como verificar o cumprimento do isolamento, de quarentena, de verifica-

Ciência & Saúde Coletiva, 25(Supl.1):2487-2492, 2020

ção probabilística de contágio, do gerenciamento de permissões para a pessoa sair em público, entre muitas outras.

Lembramos que a coleta de dados através de aplicativos e smartphones requer acesso a essas tecnologias e familiaridade na sua utilização, apontando que os dados serão representativos de um determinado segmento populacional. Portanto, as medidas precisam considerar as desigualdades em saúde e os diferentes impactos do problema em diferentes segmentos da população.

Além de rastreamento de localização, incentivo à autodeclaração de sintomas por parte do usuário e envio automático de alertas sobre possível contato com paciente infectado, outras formas de utilização de dados pessoais têm sido empregadas como a utilização de dados de saúde. No Reino Unido, por exemplo, governo e empresas de tecnologia passaram a utilizar dados de pacientes para criar um repositório sobre a COVID-19. As empresas foram contratadas pelo Sistema de Saúde Britânico (*National Health Service* - NHS) para criar um repositório de dados e auxiliar na elaboração de modelos preditivos utilizando inteligência artificial. A justificativa para a iniciativa foi a de ter informações sobre as demandas nos serviços de saúde em tempo real a partir de dados de hospitalizações, disponibilidade de leitos para cuidados intensivos, necessidades de equipamentos e suprimentos.

A despeito dos dados serem confidenciais, anônimos e armazenados em um banco de dados do governo, e do NHS informar que permanecerão sob seu controle e sujeitos a severas restrições aderentes à legislação de proteção de dados, a iniciativa tem gerado desconfiança em torno do respeito a aspectos éticos, de privacidade e de proteção de dados dos cidadãos[11].

São colocadas questões e desafios relacionados à confiança nas instituições responsáveis pelo processamento de dados contendo informações pessoais, sejam governamentais ou privadas. Desconfianças e indagações que não visam impedir o uso dos dados para responder à pandemia, mas auxiliar no estabelecimento de salvaguardas para que haja equilíbrio entre os interesses individuais e os coletivos, além de aumentar a confiança da Sociedade e das instituições no tratamento de dados para fins sanitários, facilitando com que esta atividade seja realizada da forma mais eficaz possível[12,13].

No Brasil, a Lei Geral de Proteção de Dados (LGPD) foi aprovada e sancionada em 2018, e sua entrada em vigor, em agosto de 2020, pode ainda ser alterada, a depender do resultado de votações no Congresso Nacional de Projetos de Lei que procuram mudar a data de sua vigência para 2021. A LGPD é um marco na regulamentação de dados pessoais no país ao dispor sobre todas as operações de tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural[14].

A autodeterminação informativa é, indubitavelmente, um aspecto fundamental a ser levado em consideração para o uso de dados pessoais, conjuntamente com as garantias de transparência, segurança e minimização no uso de dados. Entretanto, existem situações em casos de emergência e de interesse público, como a saúde pública, em que o uso de dados pessoais é permitido, mesmo sem o consentimento do seu titular, desde que haja salvaguardas, proporcionalidade no uso dos dados para alcance das finalidades e especificidades relacionadas às credenciais dos órgãos autorizados a processar esses dados, conforme estabelecido na Lei Geral de Proteção de Dados brasileira e no Regulamento Geral de Proteção de Dados da União Europeia[15]. Assim, a LGPD possui diversos elementos capazes de facilitar a utilização de dados pessoais em políticas e sistemas idealizados para enfrentar a COVID-19, que poderão ser utilizados logo que a lei entre em vigor[16].

A anonimização, que consiste na aplicação de medidas técnicas para impossibilitar a associação direta ou indireta dos dados ao indivíduo, e a pseudoanonimização que geralmente remove identificadores e os substitui por um código chave único são estratégias de proteção de dados previstas em algumas leis. Em geral, dados anonimizados não são considerados dados pessoais ou o são com algumas ressalvas, enquanto dados pseudoanomizados são tidos como dados pessoais pelo potencial de reidentificação dos indivíduos através da utilização do código chave, ainda que disponham potencialmente de um nível maior de segurança. Em virtude da possibilidade de identificação dos dados, mesmo anonimizados[17], são necessárias combinações de vários procedimentos para preservar a privacidade dos indivíduos, particularmente quando ocorre integração entre bases de dados[18].

A conformidade com as leis gerais de proteção de dados, portanto, requer tecnologia, infraestrutura e pessoal especializado para que os dados sejam tratados de forma lícita, justa e

responsável em relação aos titulares dos mesmos, além de prever o princípio da responsabilização através de acompanhamento das atividades de processamento pelas autoridades designadas, que poderão aplicar sanções quando houver descumprimento da lei. Inclusive, alguns países possuem centros de dados criados por meio de parcerias entre governo, universidades e institutos de pesquisa para processar e prover acesso a dados anonimizados de forma segura e controlada para pesquisas de interesse público[19].

Dados anonimizados ou agregados não são considerados dados pessoais por leis de proteção de dados porque protegem a identificação dos indivíduos. Contudo, mesmo sem fazer referência a qualquer indivíduo, podem prejudicar grupos em virtude de informações sobre locais, etnicidade, situações de saúde e condições socioeconômicas, por exemplo requerendo escrutinamento ético sobre os potenciais benefícios gerados por tais evidências.

Linnet Taylor chama atenção que não existe proteção contra tecnologias irresponsáveis, pois as leis de proteção de dados voltam-se exclusivamente à proteção de dados pessoais não abarcando as liberdades e os direitos políticos de grupos. Sendo necessária a representação de diferentes grupos da sociedade civil na governança de tecnologias para que haja controle social. Defende que as empresas de tecnologia precisam ser transparentes e prestar contas à sociedade, ao menos as empresas que em virtude da pandemia passam a fazer parte de governos e da governança de dados dos cidadãos, para que tenham legitimidade para agir em nome do governo e da população[20].

Ao considerar que dados podem ser utilizados e compartilhados por diferentes pessoas e organizações simultaneamente, as questões principais a serem harmonizadas giram em torno da governança responsável dos dados baseada na transparência e empoderamento dos cidadãos para que haja confiança e estabelecimento de relacionamentos equilibrados e justos entre indivíduos e organizações[21].

A legitimidade de coleta, processamento, compartilhamento e uso de dados pessoais não advém do acesso aos dados, mas da confiança em quem os detém, tratando-os com transparência e dentro dos parâmetros legais. Nesta perspectiva, a utilização de dados pessoais para o enfrentamento da COVID-19 e das futuras emergências de saúde pública precisa ser pautada na transparência, verificação e responsabilização a partir dos propósitos da coleta, proporcionalidade das operações de tratamento dos dados em relação à finalidade de uso, por quem e por quanto tempo[22].

Dados coletados, compartilhados e utilizados em nome da saúde pública, especialmente por empresas privadas ou através de parcerias público-privadas, precisam ter termos e condições claros e transparentes sobre os propósitos de acesso, compartilhamento, usos e responsabilizações. As seguintes e outras questões precisam ser colocadas e respondidas de forma explicita. Por quem e como os dados serão acessados, processados e utilizados? Serão armazenados, reutilizados, descartados após o alcance da finalidade? Como serão protegidos? Em caso de abuso ou negligencia, quem será responsabilizado?

Outro aspecto regulatório que requer atenção volta-se aos direitos de propriedade intelectual, pois a seleção, a organização ou a disposição de dados em um banco de dados configura-se em direito autoral[23]. Bancos de dados que poderão ser integrados com dados de outras fontes para subsidiar o desenvolvimento de novas tecnologias, incluindo tecnologias de tratamento e prevenção para a COVID-19.

As parcerias entre governos, empresas de tecnologia e universidades são necessárias para viabilizar a extração de conhecimento confiável de grande volume de dados. Os acordos precisam ser claros quanto aos papéis dos envolvidos, resultados pretendidos e alcançados. O estabelecimento de protocolos com princípios orientadores voltados à aplicação ágil e prática de processamento de dados em casos de interesse coletivo, com regras e supervisão nacional e internacional, podem ser uma alternativa, como na atual situação de emergência em saúde.

A governança responsável de dados inclui também a descrição das metodologias de processamento e análise dos dados, pois dados têm valor de prova, de evidência, na tomada de decisão tanto para políticas públicas quanto para ciência[24]. Adicionalmente, todo algoritmo baseado em aprendizagem de máquina expressa a visão do padrão ou regularidade do que deve ser considerado para mensurar. Algoritmos são recursos poderosos e importantes que não podem prescindir de explicações causais em virtude de riscos de decisões baseadas apenas em resultados e predições automatizados. O método científico tem, portanto, papel preponderante para validar, aumentar a confiabilidade e a utilidade dos resultados. Inclusive no questionamento de suposições, valores e vieses que distinguem opiniões de evidências.

Somente regulamentações são capazes de estabelecer limites para o processamento de da-

Ciência & Saúde Coletiva, 25(Supl.1):2487-2492, 2020

dos pessoais por governos e empresas privadas, mesmo em crises sanitárias, para evitar impactos negativos decorrentes de flexibilizações momentâneas, que poderão se tornar permanentes como aconteceu nos Estados Unidos após o evento de 11 de setembro de 2001. Estratégias de vigilância da população foram introduzidas utilizando tecnologias existentes e emergentes com a justificativa de monitorar pessoas suspeitas e evitar ataques terroristas, ocasionando mudanças na lei decorrentes do medo incutido na sociedade[25].

Modelos de governança de dados mais justos, responsáveis e sustentáveis, que protejam e defendam princípios éticos e regulatórios, ampliam a confiança dos indivíduos e da sociedade na utilização de seus dados para responder a situações de legítimo interesse público. Aspectos relacionados ao direito à privacidade, direito à proteção de dados pessoais e direitos de grupos não inviabilizam o uso de dados pessoais e a possibilidade de seu uso para responder à pandemia. A emergência de saúde pública ocasionada pelo Sars-CoV-2 aponta para a premente necessidade de novas formas de governança de dados pessoais que incluam a sociedade civil para a promoção de benefícios equânimes para toda a sociedade.

Colaboradores

BA Almeida concebeu e liderou a escrita da primeira versão do artigo, e contribuiu igualmente na edição e revisão do texto final. M Barreto apoiou a concepção e escrita do artigo original e contribuiu igualmente na edição e revisão do texto submetido. Os demais autores, D Doneda, MY Ichihara, M Barral-Netto, GC Matta, ET Rabello e FC Gouveia apoiaram a escrita do artigo original e foram igualmente responsáveis pela edição e revisão do texto.

Referências

1. Leonelli S. Data – from objects to assets. *Nature* 2019; 574:317-320.
2. World Health Organization (WHO). *Critical preparedness, readiness and response actions for COVID-19*. Geneva: WHO; 2020.
3. The Economist. *Covid-19. App-based contact tracing may help countries get out of lockdown but only as part of a bigger system*. [acessado 2020 Abr 16]. Disponível em: https://www.economist.com/science-and-technology/2020/04/16/app-based-contact-tracing-may-help-countries-get-out-of-lockdown
4. Kim MJ, Denyer S. A 'travel log' of the times in South Korea: Mapping the movements of coronavirus carriers. *The Washington Post*. [acessado 2020 Abr 16]. https://www.washingtonpost.com/world/asia_pacific/coronavirus-south-korea-tracking-apps/2020/03/13/2bed568e-5fac-11ea-ac50-18701e14e06d_story.html
5. Gilbert D. Iran Launched an App That Claimed to Diagnose Coronavirus. Instead, It Collected Location Data on Millions of People. *Vice News*. [acessado 2020 Mar 14]. Disponível em: https://www.vice.com/en_us/article/epgkmz/iran-launched-an-app-that-claimed-to-diagnose-coronavirus-instead-it-collected-location-data-on-millions-of-people
6. Newsroom. *Apple e Google formam parceria para tecnologia de rastreamento de contato com COVID-19*. [acessado 2020 Abr 10]. Disponível em: https://www.apple.com/br/newsroom/2020/04/apple-and-google-partner-on-covid-19-contact-tracing-technology/
7. Payer M, Barman L. *Descentralized Privacy-Preserving Proximity Tracing*. [acessado 2020 Abr 10]. Disponível em: https://github.com/DP-3T
8. *Pan-European Privacy-Preserving Proximity Tracing*. Disponível em: [acessado 2020 Abr 10]. https://www.pepp-pt.org/
9. Project Safe Paths. *Massachusetts Institute of Technology*. [acessado 2020 Abr 10]. Disponível em: https://www.media.mit.edu/projects/safepaths/overview/
10. Panzzarino, M. Apple and Google are launching a joint COVID-19 tracing tool for iOS and Android. *Tech Crunch*. [acessado 2020 Abr 10]. Disponível em: https://techcrunch.com/2020/04/10/apple-and-google-are-launching-a-joint-covid-19-tracing-tool/
11. Lewis P, Conn D, Pegg D. UK government using confidential patient data in coronavirus response. *The Guardian*. [acessado 2020 Abr 12]. Disponível em: https://amp.theguardian.com/world/2020/apr/12/uk-government-using-confidential-patient-data-in-coronavirus-response?CMP=share_btn_tw&__twitter_impression=true
12. European Digital Rights. *EDRI calls for fundamental rights – based responses to COVID-19*. [acessado 2020 Mar 20]. Disponível em: https://edri.org/covid19-edri-coronavirus-fundamentalrights/
13. Mcknight G. Coronavirus surveillance concerns ramp up pressure to pass federal privacy law. *Internet Governance Hub*. [acessado 2020 Abr 10]. Disponível em: https://www.internetgovernancehub.blog/2020/04/10/coronavirus-surveillance-concerns-ramp-up-pressure-to-pass-federal-privacy-law/
14. Brasil. Lei n. 13.079, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). *Diário Oficial da União* 2018; 15 ago.
15. European Union (EU). *Regulamento Geral de Proteção de Dados da União Europeia – EU 2016/679 (GDPR)*. [acessado 2020 Mar 20]. Disponível em: https://gdpr-info.eu/
16. Doneda D. *Opinião e Análise. A proteção de dados em tempos de coronavírus*. [acessado 2020 Mar 20]. Disponível em: https://www.jota.info/opiniao-e-analise/artigos/a-protecao-de-dados-em-tempos-de-coronavirus-25032020
17. Rocher L, Hendrickx JM, De Montjoye YA. Estimating the success of re-identifications in incomplete datasets using generative models. *Nat Commun* 2019; 10:3069.
18. Harron K, Dibben D, Boyd J, Hjern A, Azimaee M, Barreto M, Goldstein H. Challenges in administrative data linkage for research. *Big Data&Society* 2017; 4(2):11-12.
19. Doneda D, Almeida BA, Barreto ML. Uso e proteção de dados pessoais na pesquisa científica. *Revista Direito Público* 2019; 16(90):179-194.
20. Nuffield Council on Bioethics and Ada Lovelace Institute. *Webinar - Beyound the exit strategy: ethical uses of data-driven technology in the fight against COVID-19*. [acessado 2020 Abr 20]. Disponível em: https://www.nuffieldbioethics.org/publications/covid-19/webinar-beyond-the-exit-strategy-ethical-uses-of-data-driven-technology-in-the-fight-against-covid-19
21. My Data Global Blog. *My Data vs COVID-19*. [acessado 2020 Abr 20]. Disponível em: https://mydata.org/2020/04/06/an-approach-for-fighting-covid-19-and-beyond-mydata/
22. Patel R. Removing the pump handle - stewarding data at times of public health emergency. [acessado 2020 Abr 8]. Disponível em: https://www.adalovelaceinstitute.org/removing-the-pump-handle-stewarding-data-at-times-of-public-health-emergency/
23. Guanaes P, Souza AR, Doneda D, Nascimento FJT. *Marcos legais nacionais em face da abertura de dados para pesquisa em saúde: Dados pessoais, sensíveis ou sigilosos e propriedade intelectual*. Rio de Janeiro: Fiocruz; 2018.
24. Leonelli S. Data Governance is Key to Interpretation: Reconceptualizing Data in Data Science. *Harvard Data Science Review* 2019. [acessado 2020 Abr 8]. Disponível em: https://hdsr.mitpress.mit.edu/pub/4ovhpe3v
25. Levi M, Wall DS. Technologies, Security, and Privacy in the Post-9/11 European Information Soci*ety. Journal of Law and Society* 2004; 31(2):194-220.

Artigo apresentado em 27/04/2020
Aprovado em 29/04/2020
Versão final apresentada em 01/05/2020