

Um olhar para as restrições à reutilização dos dados publicamente acessíveis: proposta de categorização

A Look at the Restrictions on the Reuse of Publicly Accessible Data: a Categorization Proposal

Maurício Antônio Tamer¹

Recebido em: 12.08.2024
Aprovado em: 30.06.2025

RESUMO

O artigo aborda a relevância dos dados na contemporaneidade e seu papel central nos negócios, com ênfase no fenômeno da datificação e na utilização de dados publicamente acessíveis como insumo estratégico para diversas atividades. Entretanto, embora tais dados sejam amplamente explorados, as restrições legais à sua reutilização constituem um aspecto igualmente relevante, mas que nem sempre recebe o devido destaque na literatura e na prática. O problema científico que orienta esta investigação é o seguinte: diante da ausência de maior visibilidade das restrições aplicáveis à reutilização dos dados publicamente acessíveis, seria possível evidenciá-las por meio de uma proposta de categorização ou classificação? Se a resposta for afirmativa, quais seriam essas categorias? Para responder a essas questões, este trabalho adota uma metodologia fundamentada em dados práticos e em experiências concretas de aplicação normativa, de modo a alinhar a análise teórica com a realidade enfrentada por organizações e pesquisadores. Com base nesse percurso metodológico, propõe-se uma categorização inicial que busca organizar e dar maior clareza às limitações jurídicas existentes, servindo como ponto de partida para debates e aperfeiçoamentos futuros.

Palavras-chave: dados; dados publicamente acessíveis; restrições; propriedade intelectual; proteção de dados pessoais.

ABSTRACT

The article addresses the relevance of data in contemporary society and its central role in business, with emphasis on the phenomenon of datafication and on the use of publicly accessible data as a strategic input for various activities. However, although such data are widely exploited, the legal restrictions on their reuse constitute an equally important aspect, which does not always receive the proper attention in academic literature or in

¹ Maurício Antonio Tamer. Pós-Doutorado pela Universidade do Estado do Rio de Janeiro, com pesquisa focada na relação entre Inteligência Artificial e as resoluções de conflitos. Doutor em Direito Político e Econômico pela Universidade Presbiteriana Mackenzie (2021). Mestre em Direito Processual Civil pela Pontifícia Universidade Católica de São Paulo (2016). Graduado em Direito pela Universidade Presbiteriana Mackenzie (2012). E-mail: mauriciotamer@gmail.com. Lattes: <https://lattes.cnpq.br/1292641448156094>.



practice. The scientific problem guiding this research is as follows: given the lack of greater visibility of the restrictions applicable to the reuse of publicly accessible data, would it be possible to highlight them through a proposal for categorization or classification? If so, what would these categories be? To answer these questions, this study adopts a methodology grounded in practical data and concrete experiences of normative application, thus aligning theoretical analysis with the reality faced by organizations and researchers. Based on this methodological approach, an initial categorization is proposed, aiming to organize and bring greater clarity to existing legal limitations, serving as a starting point for further debate and refinement.

Keywords: Data; Publicly available data; Restrictions; Intellectual property; Personal data protection.

1 INTRODUÇÃO

Reconhecidamente, a sociedade contemporânea tem em si uma característica muito pertinente: o fenômeno da datificação. Nas palavras de Viktor Mayer-Schönberger e Kenneth Cukier, um fenômeno social-tecnológico em que tudo pode ser descrito na forma de dados de modo que possa ser registrado, analisado e reorganizado (Mayer-Schönberger; Cukier, 2014, p. 77).

Fala-se mais do que “apenas” o perfil de sociedade da informação que, conforme lição de Yoneji Masuda, em que a sociedade está centrada na tecnologia de telecomunicações e informática (Masuda, 1982, p. 67). Disse o autor:

A espécie humana vive um processo de transformação silenciosa da sociedade: o surgimento de uma época da informação, centrada na tecnologia de telecomunicações e informática. Uma ‘Época da Informação’ é o período de tempo em que ocorre uma inovação na tecnologia da informação, em que se torna latente o poder de transformação da sociedade, capaz de acarretar uma expansão da quantidade e da qualidade da informação, e um aumento em larga escala no armazenamento da informação).

É um dos exemplos ou motores, também, do que tem se conhecido como Quarta Revolução Industrial (Schwab, 2016).

Em certa medida, todos os negócios e atividades empresariais, compreendidos estes como aqueles associados a uma atividade econômica organizada para a produção ou a circulação de bens ou de serviços (Conforme definição jurídica de empresa extraída do

Artigo 966 do Código Civil), têm ou almejam ter parte significativa de suas operações calcadas em dados. Sem desconsiderar as atividades empresariais em que a datificação é o coração de suas atividades e serviços, a exemplo dos *data brokers*, entidades cuja atividade empresarial principal é justamente identificar fontes diversas de dados, compilá-los, higienizá-los, classificá-los e comercializá-los a terceiros.

Fato é que os dados são, hoje, insumo significativo das atividades empresariais informando as decisões envolvidas em análises, resultados, produtos e serviços. O valor agregado de tais atividades, assim, tem relação direta com a quantidade e a qualidade dos dados envolvidos nos processos. Seja na coleta destes (diretamente de fontes privadas – usuários e outras empresas como os referidos *data brokers* – ou fontes publicamente acessíveis, como sites governamentais), no seu armazenamento, governança e até seletividade na eliminação.

Porém, há restrições jurídicas que recaem sobre os dados publicamente acessíveis. O valor agregado de produtos e serviços também tem na legalidade dos dados utilizados um componente importante de avaliação. Inclusive, todas estas situações convivem com outro fenômeno extremamente relevante na contemporaneidade: a abertura de dados. Ideia que, como se verá, em linhas gerais, propõe a regra de possibilidade de reutilização de dados públicos, se livres de qualquer restrição jurídica.

As facilidades de acesso propiciadas por tal dinâmica também geram uma falsa sensação de permissibilidade irrestrita de uso dos dados, o que não é verdade já que tais fontes também contém restrições.

Porém, nem sempre tais restrições são destacadas ou analisadas com o cuidado devido, seja na perspectiva acadêmica, seja na perspectiva prática ou negocial. Isto pode ocorrer, entre outros fatores, por uma rara ou talvez inexistente categorização de tais restrições.

Seria o típico caso em que uma classificação minimamente consistente se apresente de forma útil para realçar as restrições e melhor explicar o próprio objeto realçado. Não existem classificações certas ou erradas, mas úteis ou inúteis. A utilidade estaria, justamente, na potencialização de entendimento sobre o objeto classificado.

Feitas estas considerações, o problema científico que se apresenta é: diante da falta de maior destaque das restrições à reutilização dos dados publicamente acessíveis, seria possível realçá-las mediante uma proposta de categorização ou classificação? Se sim, quais seriam estas categorias?

Estes problemas ou fragilidade de real destaque parece se apresentar, por exemplo, em operações de fusões e aquisições que tenham empresas com negócios sustentados em dados, e mesmo no sucesso de mercado de produtos e serviços baseados em dados, cuja robustez e legalidade é cada vez mais avaliada no detalhe pelos potenciais clientes – estes também pressionados pelas expectativas de boas práticas da regulação.

Neste último caso, é o exemplo da boa prática esperada na diligência junto a terceiros conforme previsão do Artigo 50 da LGPD:

Art. 50. Os controladores e operadores, no âmbito de suas competências, pelo tratamento de dados pessoais, individualmente ou por meio de associações, poderão formular regras de boas práticas e de governança que estabeleçam as condições de organização, o regime de funcionamento, os procedimentos, incluindo reclamações e petições de titulares, as normas de segurança, os padrões técnicos, as obrigações específicas para os diversos envolvidos no tratamento, as ações educativas, os mecanismos internos de supervisão e de mitigação de riscos e outros aspectos relacionados ao tratamento de dados pessoais.

Com a metodologia hipotético-dedutiva, este artigo propõe uma das formas úteis ao enfrentamento do tema. Assim, começo estudando o tema importante e que potencializa a datificação dos negócios: os dados publicamente acessíveis. Após, será tratada análise das principais restrições aplicáveis, separando-as em categorias.

2 DADOS PUBLICAMENTE ACESSÍVEIS

Os dados publicamente acessíveis parecem-me que podem ser divididos em dois grandes grupos: (i) os dados abertos oriundos das atividades da Administração Pública; e (ii) os dados publicamente acessíveis em plataformas privadas (mídias sociais, marketplaces etc.).

2.1 Dados abertos

O movimento em prol do *Open Data* ou Dados Abertos ganhou consistência nos últimos trinta anos e se encontra em franca e elogiável expansão. Diferentemente de outros temas que contam com marco científico ou legislativo bem definido, o movimento não tem, por assim dizer, uma data de nascimento precisa. Seu desenvolvimento é difuso e amplo, o que, neste caso, tem se mostrado algo muito proveitoso.

Porém, é possível identificar algumas referências ou iniciativas de maior destaque (entre elas a iniciativa de *Open Government Data*. Em 1995, por exemplo, a expressão *Open Data* (Dados Abertos) teria sido utilizada pela primeira vez com autonomia e significado próprio em relatório de agência científica norte-americana em defesa da abertura dos dados entre os países como fator importante para a compreensão holística de fenômenos globais (Chignard, 2013).

Mesmo antes do próprio desenvolvimento da internet, já foi defendida a ideia de compartilhamento de conhecimento e informações em benefício da ciência, a exemplo dos reconhecidos trabalhos de Robert K. Merton em texto de 1942. Disse o autor:

The institutional goal of science is the extension of certified knowledge. The technical methods employed toward this end provide the relevant definition of knowledge: empirically confirmed and logically consistent statements of regularities (which are, in effect, predictions). The institutional imperatives (mores) derive from the goal and the methods. The entire structure of technical and moral norms implements the final objective. The technical norm of empirical evidence, adequate and reliable, is a prerequisite for sustained true prediction; the technical norm of logical consistency, a prerequisite for systematic and valid prediction. The mores of science possess a methodologic rationale but they are binding, not only because they are procedurally efficient, but because they are believed right and good. They are moral as well as technical- prescriptions. Four sets of institutional imperatives-universalism, communism, disinterestedness, organized skepticism-are taken to comprise the ethos of modern science. [...] The substantive findings of science are a product of social collaboration and are assigned to the community. They constitute a common heritage in which the equity of the individual producer is severely limited. An eponymous law or theory does not enter into the exclusive possession of the discoverer and his heirs, nor do the mores bestow upon them special rights of use and disposition. Property rights in science are whittled down to a bare minimum by the rationale of the scientific ethic. The scientist's claim to "his" intellectual "property" is limited to that of recognition and esteem which, if the institution functions with a modicum of efficiency, is

roughly commensurate with the significance of the increments brought to the common fund of knowledge. The institutional conception of science as part of the public domain is linked with the imperative for communication of findings. Secrecy is the antithesis of this norm; full and open communication its enactment.¹³ The pressure for diffusion of results is reenforced by the institutional goal of advancing the boundaries of knowledge and by the incentive of recognition which is, of course, contingent upon publication. A scientist who does not communicate his important discoveries to the scientific fraternity-thus, a Henry Cavendish-becomes the target for ambivalent responses. He is esteemed for his talent and, perhaps, for his modesty. But, institutionally considered, his modesty is seriously misplaced, in view of the moral compulsion for sharing the wealth of science. (Merton, 1973).

Em 2007, um grupo se dedicou ao tema de forma mais clara em reunião em Sebastopol, ao norte de São Francisco, Estado Unidos da América. Um dos objetivos, por exemplo seria definir um conceito de dados públicos abertos a ser adotado pelos candidatos à eleição presidencial daquele país. Um grupo formado por 30 pensadores e ativistas do tema, incluindo Tim O'Reilly e Lawrence Lessig, ambos amplamente conhecidos. O primeiro, especialmente, por sua contribuição notável em temas técnicos, tendo cunhado expressões como open source e Web 2.0. Lawrence Lessig, por sua vez, especialmente pela contribuição da organização Creative Commons² e pelas preciosidades literárias *The Future of Ideas* (2001), *Code and Other Laws of Cyberspace* (2000) e *Code: Version 2.0* (2006). Plataformas abertas e colaborativas como Wikipedia, por exemplo, também tem raiz histórica neste movimento.

Na oportunidade, o grupo estabeleceu oito princípios que estabeleceram a lógica básica da ideia de Open Data³, trazendo ainda mais base teórico-prática para a ideia de dados abertos. Embora não sejam definições exclusivas e restritivas, pode se dizer que são as mais aceitas contemporaneamente. São oito princípios de funcionamento e aproveitamento dos dados que devem ser constantemente buscados para que a proposta se desenvolva e consolide. A ideia do *Open Government Data*, de forma interligada, é de que dados públicos são de propriedade comum da sociedade, de forma muito similar ao

² Creative Commons é uma organização internacional sem fins lucrativos baseada nas concepções do Copyleft e voltada à divisão de conhecimento e à expansão de obras criativas e intelectuais, inclusive com a criação de licenças que permitem aos autores e desenvolvedores dos conteúdos com menor restrições. Mais informações disponíveis em: <https://creativecommons.org> Acesso em 06.03.2024.

³ A Organização para a Cooperação e Desenvolvimento Econômico também se dedica especialmente ao tema.

que se preconiza para as descobertas científicas. Está sustentada na premissa de compartilhamento e uso dos dados públicos com bem comum. Na mesma linha, para *Open Definition* definida e promovida pela *Open Knowledge Foundation*, *Open Data* ou dados abertos são os dados que podem ser livremente utilizados, reutilizados e distribuídos para qualquer pessoa (*Open Knowledge Foundation*).

Após os oito princípios iniciais, foram incorporados mais sete pela organização resultante de tal grupo e que devem ser igualmente observados. Todos vocacionados, principalmente, a meu ver, para potencializar a utilidade, a segurança e a licitude da prática.

Assim, os quinze princípios que devem ser observados na proposta do *Open Government Data* são: completude (todos os dados públicos devem estar disponíveis), primariedade (dados coletados diretamente de sua fonte e em maior granularidade possível), atualidade (os dados devem ficar disponíveis da forma mais rápida possível), acessibilidade (os dados devem ficar disponíveis ao maior número de usuários possíveis), possibilidade de processamento por sistema (os dados devem ser passíveis de tratamento automatizado), não discriminação (os dados devem ficar disponíveis a qualquer pessoa e sem necessidade de identificação), ausência de propriedade sobre os formatos (os dados devem ser disponibilizados em formatos que ninguém tenha propriedade), liberdade de licenças (os dados sejam livres de propriedades intelectuais), disponibilidade on-line e gratuita (os dados estejam disponíveis on-line e de forma gratuita), permanência (os dados devem estar disponíveis de forma estável e mais previsível possível), confiabilidade (a disponibilização deve assegurar a autenticidade e a integridade dos dados), presunção de abertura (os dados públicos são presumidamente acessíveis e abertos ao reuso, salvo restrições legítimas), registrabilidade (os dados devem estar acompanhados de informações e metadados que permitam aos usuários avaliar se os dados são precisos e atuais), ausência de mecanismos executáveis (os dados devem estar livres de qualquer software autoexecutável em preocupação clara associada à segurança da informação) e a disponibilidade contando com a participação dos entes públicos (participação direta do Poder Público).

No Brasil, a legislação brasileira acompanha o cenário estrangeiro sobre a promoção dos dados abertos, destacando-se no País um conjunto normativo formado, especialmente, por quatro atos normativos principais: Lei de Acesso à Informação, Decreto Regulamentador da Lei de Acesso à Informação, Lei do Governo Digital e a Política de Dados Abertos do Poder Executivo Federal. Todos que, em conjunto, estabelecem a abertura dos dados, observadas as restrições pertinentes. Paralelamente, muitos órgãos da Administração Pública se destacam na disponibilização e regulação de dados abertos afetos a cada um dos seus contextos próprios, porém, valendo-se das premissas fundantes deste regime normativo.

2.2 Dados publicamente acessíveis em plataformas privadas

Paralelamente aos dados abertos decorrentes de plataformas públicas, também há dados publicamente acessíveis em plataformas privadas. São os dados disponibilizados por pessoas nas aplicações privadas e que, como consequência dos serviços correlatos, ficam acessíveis ao público geral.

Alguns exemplos podem ser citados: (i) dados presentes em contas abertas de mídias sociais, assim disponibilizadas pelos usuários de tais serviços; (ii) dados presentes em serviços de marketplaces como os anúncios feitos pelos usuários; e (iii) dados presentes nas demais páginas de internet que, embora tenham objetivos privados (sítios eletrônicos de pessoas jurídicas por exemplo), resultam na disponibilização de uma série de dados.

Diferentemente da proposta dos dados abertos, inexistente uma regulação geral aplicável dispendo sobre o que deve ser disponibilizado e em qual qualidade. Há apenas leis específicas que determinam a publicidade de alguns dados como diretrizes oriundas do Código de Defesa do Consumidor, do Decreto Federal do SAC nº 11.034/2022, da própria Lei Geral de Proteção de Dados Pessoais (LGPD) e a recente Lei Federal nº 14.611/2023 que determina a publicação de relatórios de transparência salarial. Há, ainda, regulações setoriais específicas, a exemplo das regulações do Banco Central do Brasil e da Comissão de Valores Mobiliários.

A disponibilização dos dados e as condições de reutilização, observadas tais disposições específicas, fica sob critério dos entes privados, no escopo de sua livre iniciativa (Artigo 170, CF), e de aplicação de leis pertinentes.

3 AS RESTRIÇÕES JURÍDICAS APLICÁVEIS E SUAS RAZÕES DE SER

Como introduzido quando da exposição do problema científico, a datificação da sociedade e dos negócios resulta na experiência de que a maioria das atividades empresariais têm nos dados suporte relevante de práticas operacionais. Assim, a legalidade dos dados utilizados é fundamental para as empresas cujos produtos e serviços estão sustentados nos dados em si (data brokers por exemplo) e para as empresas que possuem ou acessam uma quantidade importante de dados para o desempenho de suas atividades. E, ao falar em legalidade dos dados utilizados, é falar também do valor decorrente das atividades que deles se valem, justamente em razão dos riscos jurídicos, de imagem e reputacionais da desconformidade.

As operações de fusões e aquisições associadas a todas estas atividades, produtos e serviços têm este valor como um fator chave. A prática de *due diligence* sobre os dados utilizados pelas empresas é cada vez mais recorrente e importante, olhando com lupa eventuais restrições, riscos e perdas de valor ou necessidades de investimentos associados a tal insumo.

Assim como o alcance e desempenho comercial e negocial de produtos e serviços calcados em dados publicamente acessíveis dependem cada vez mais da prestação de contas da legalidade dos dados utilizados em razão de uma governança cada vez mais criteriosa dos clientes sobre estes aspectos, eis que também pressionados pela regulação, especialmente de Proteção de Dados Pessoais.

Neste sentido, a ilegalidade parece estar, a meu ver, se desrespeitadas restrições de ordem jurídica à reutilização dos dados publicamente acessíveis. Em outras palavras, restrições que impedem ou limitam a reutilização de tais dados, podendo *contaminar* todas as atividades sequenciais que deles se valem. Porém, como adiantado, estas

restrições não parecem ter o destaque merecido e que, em última instância, justificaria a própria razão de existir das restrições em si.

Restringir algo juridicamente, a meu ver, em uma das interpretações possíveis, significa impedir ou restringir algum comportamento fático, um fazer, com respaldo legal. O sistema normativo jurídico identifica os bens jurídicos passíveis de proteção e, a partir desta identificação, desencadeia limitações ao comportamento humano que, pelo mesmo sistema, foram tidas por necessárias para proteção de tais bens.

Estas *limitações ao comportamento humano* podem ser chamadas de *restrições jurídicas*. Neste objeto de estudo, o comportamento humano seria justamente a reutilização de dados publicamente acessíveis. As limitações, porém, seriam aquelas decorrentes da identificação de bens jurídicos tutelados associados a estes dados. Restringe-se ou limita-se a reutilização, assim, para proteger tais direitos.

4 UMA PROPOSTA DE CATEGORIZAÇÃO DAS RESTRIÇÕES À REUTILIZAÇÃO

Feitas as considerações e a partir dos levantamentos de pesquisa feitos, as principais restrições podem ser segregadas em cinco categorias conforme bens jurídicos tutelados associados aos dados publicamente acessíveis:

- (i) restrições associadas à propriedade intelectual (propriedades intelectuais);
- (ii) restrições associadas à Proteção de Dados Pessoais (direitos de liberdade, privacidade e livre desenvolvimento da personalidade);
- (iii) restrições associadas à forma pela qual a coleta dos dados é possível (segurança da informação);
- (iv) outras restrições definidas pelas plataformas privadas (tutela da livre iniciativa constitucional); e
- (v) restrições associadas ao custeio sobre a coleta e reutilização dos dados (solidariedade dos custos do serviço).

A seguir, tenta-se detalhar cada uma delas.

4.1 Propriedade Intelectual

A forma como os dados estão dispostos e estruturados em repositório acessível publicamente pode constituir propriedade intelectual protegida e, nesta condição, passível de restrições legais conforme seu titular.

Fala-se aqui do conceito de base de dados, caracterizada como obra intelectual protegida nos termos da legislação de proteção de direito autoral. Isto, conforme previsão expressa do Artigo 7º, XIII, da Lei de Direitos Autorais ou Lei Federal nº 9.610/1998:

Art. 7º São obras intelectuais protegidas as criações do espírito, expressas por qualquer meio ou fixadas em qualquer suporte, tangível ou intangível, conhecido ou que se invente no futuro, tais como: [...] XIII - as coletâneas ou compilações, antologias, enciclopédias, dicionários, bases de dados e outras obras, que, por sua seleção, organização ou disposição de seu conteúdo, constituam uma criação intelectual.

Nesta condição, basta a existência pública da base de dados para que esta seja digna de direito e proteção, independentemente de registro, conforme Artigo 18 também da Lei de Direitos Autorais: “Art. 18. A proteção aos direitos de que trata esta Lei independe de registro.”. Inclusive, o que se protege não são os dados em si, mas a obra de construção e organização destes em base de dados de formato específico, conforme previsão limitativa do §2º do Artigo 7 da Lei: “§ 2º A proteção concedida no inciso XIII não abarca os dados ou materiais em si mesmos e se entende sem prejuízo de quaisquer direitos autorais que subsistam a respeito dos dados ou materiais contidos nas obras.”.

Esta proteção jurídica confere aos autores das bases de dados (da organização, da estruturação etc.), identificados no Artigo 11 da Lei, os direitos morais e materiais, conforme também previsões da Lei de Direitos Autorais.

Destaca-se, no contexto da problemática aqui estabelecida, o direito patrimonial exclusivo de utilizar os dados na forma como estão estruturados, ou seja, direito exclusivo sobre a base em si e não sobre os dados (Artigo 28). Além disto, a lei é clara ao estabelecer que qualquer reprodução (parcial ou integral), edição, adaptação etc. depende de autorização prévia e expressa do autor (Artigo 29).

Um exemplo sempre elucidativo e encontrado em pesquisa são os sites de anúncios de imóveis que compilam os dados dos imóveis e contatos em si de forma diferente. A maneira como os dados estão estruturados e são apresentados pode caracterizar a obra protegida na condição de base de dados. Os dados dos imóveis em si não são protegidos. Assim, terceiros sem autorização estão impedidos de replicar a base de dados como está disposta pelo site de anúncios.

Significa, portanto, que o terceiro que tiver acesso à base de dados fica impedido de replicá-la ou utilizá-la no formato em que está ou adaptá-la. E os dados nela presentes? A legalidade do uso destes fica afeta às demais restrições pertinentes, a exemplo daquelas a seguir tratadas de Proteção de Dados Pessoais.

4.2 Proteção de Dados Pessoais

Sendo os dados disponíveis publicamente dados pessoais, a legislação pertinente à proteção destes dados também atrai restrições à reutilização dos dados publicamente acessíveis desta condição.

No Brasil, trata-se especialmente da incidência da Lei Geral de Proteção de Dados Pessoais (Lei Federal nº 13.709/2018), já referendada como LGPD. Ela define o que vem a ser dado pessoal como “informação relacionada a pessoa natural identificada ou identificável”.

Tal definição reconhece que dado e informação, a meu ver, são conceitos diferentes. A informação é o resultado racional extraído do dado, conforme contexto em que o dado e seu uso estão. O dado será considerado um dado pessoal assim se, no contexto de seu uso, for possível a extração de uma informação pessoal, algo que identifique o indivíduo ou algo sobre ele. Sendo este o contexto sobre os dados publicamente acessíveis, a Proteção de Dados Pessoais e seus requisitos incidirão.

Assim, as restrições aplicáveis parecem se apresentar em duas ordens: (i) restrições gerais; e (ii) restrições específicas aos dados publicamente acessíveis.

Em relação às restrições gerais, todo o regime de Proteção de Dados Pessoais se aplica, com incidência plena da LGPD. Assim, por exemplo, a utilização dos dados

peçoais deve se dar em respeito aos princípios previstos na lei (Artigo 6º), em observância plena ao regime das bases legais de tratamento (Artigos 7º e 11, especialmente) e adoção das medidas técnicas e organizacionais de proteção prática de tais dados por quem os coleta em fontes publicamente acessíveis.

Porém, a LGPD, reconhecendo a ampla presença de dados pessoais publicamente acessíveis, estabelece restrições específicas ao reuso destes em seu Artigo 7º, §3º: “§ 3º O tratamento de dados pessoais cujo acesso é público deve considerar a finalidade, a boa-fé e o interesse público que justificaram sua disponibilização.”.

Assim, ao contrário do que muitas vezes se propaga de forma desavisada, o fato puro e simples do dado pessoal estar publicamente acessível não autoriza seu reuso descuidado ou irrestrito, mas deve respeito ao que a LGPD prevê.

4.3 Forma de coleta dos dados

Também são possíveis restrições relacionadas à forma de coleta dos dados publicamente acessíveis. A restrição, assim, não está relacionada à forma como os dados estão dispostos (*i.e.* restrições associadas à propriedade intelectual) ou a uma especial qualidade dos dados em si (*i.e.* Proteção de Dados Pessoais). Mas, sim, à forma pela qual a coleta se dá. A legalidade ou não do uso dos dados publicamente está diretamente associada ao como.

Comumente esta restrição está associada ao uso de ferramentas automatizadas de coleta. É o caso, por exemplo, dos Termos de Uso e Política de Privacidade da Receita Federal do Brasil. O documento estabelece que os usuários não usarão

robôs, sistemas de varredura e armazenamento de dados (como “spiders” ou “scrapers”), links escondidos ou qualquer outro recurso escuso, ferramenta, programa, algoritmo ou método coletor/extrator de dados automático para acessar, adquirir, copiar ou monitorar o serviço, sem permissão expressa por escrito do órgão. (Brasil, 2022).

4.4 Outras restrições estabelecidas pelas plataformas privadas

As plataformas privadas podem, dentro da sua livre iniciativa constitucional, estabelecer restrições legítimas ao reuso dos dados que estão acessíveis em seu ambiente. São providências, inclusive, plenamente razoáveis para proteção do negócio. Restrições desta natureza, normalmente, nos Termos de Serviço ou Uso também devem ser consideradas.

Para ilustrar, veja-se, por exemplo, as restrições dispostas pela empresa OLX. Os Termos e Condições Gerais de Uso da empresa dispõem que: “É proibida a utilização de textos, imagens, anúncios e qualquer outro elemento incluído ou disponível no Site para sua posterior inclusão em quaisquer veículos alheios ao Site sem a autorização prévia e por escrito da OLX.”.

4.5 Custeio sobre a coleta e reutilização dos dados

Há, ainda, uma outra restrição possível associada a custeios sobre a coleta e reutilização dos dados publicamente acessíveis. Tem sido crescente a discussão sobre a legitimidade de cobrança para a coleta de dados de plataformas públicas, em especial, pelos impactos à infraestrutura das entidades da Administração Pública. Existindo a previsão do custeio ou remuneração do ente público via taxa, caso esta não seja devidamente adimplida, a legalidade do uso dos dados pode ser questionada.

A exemplo, vale citar a Recomendação nº 74/2020 do Conselho Nacional de Justiça que, após Grupo de Trabalho específico, estabelece em seu Artigo 3º que os tribunais

poderão avaliar a conveniência e oportunidade de cobrança pelo acesso massificado a dados. §1º O valor da cobrança destina-se a suportar os custos de implantação e manutenção do sistema, devendo sua fixação ser efetuada na proporção do volume de dados utilizados. (CNJ, 2020).

5 CONSIDERAÇÕES FINAIS

Seguindo a abordagem hipotético-dedutiva, apoiada por evidências práticas e experiências de aplicação normativa, o estudo confirma a hipótese central: é possível — e útil — realçar as restrições à reutilização de dados publicamente acessíveis por meio de uma categorização.

A proposta inicial organiza tais limites em cinco eixos: (i) propriedade intelectual (tutela da seleção, organização e disposição da base de dados); (ii) proteção de dados pessoais (incidência integral do regime de privacidade e deveres correlatos); (iii) forma de coleta (licitude do “como”, inclusive quanto a automações e barreiras técnicas); (iv) regras contratuais e de plataforma (livre iniciativa e termos de uso aplicáveis); e (v) custeio/ acesso massificado (modelos de remuneração e sustentabilidade da infraestrutura). Essa estrutura evidencia que “acessível” não é sinônimo de “livre” e oferece um mapa para leitura jurídica consistente do reuso.

Do ponto de vista aplicado, a categorização serve como roteiro operacional para governança de dados e *due diligence*: começa-se pela verificação de titularidade e originalidade da base (propriedade intelectual), passa-se à identificação de dados pessoais e bases legais/limitações finalísticas (privacidade), valida-se o método de coleta (inclusive restrições a *scraping*), revisam-se contratos e termos de serviço (plataformas privadas) e, por fim, aferem-se condições de custeio e limites de acesso massivo. Ao alinhar teoria e prática, o modelo contribui para decisões de produto, contratos com terceiros, avaliação de riscos em M&A e prestação de contas a clientes e reguladores.

A análise também revela a natureza cumulativa e, por vezes, sobreposta dessas camadas: uma mesma iniciativa pode ser lícita sob o prisma autoral, mas vedada pela forma de coleta; ou compatível com a finalidade de disponibilização pública, porém incompatível com termos contratuais da plataforma. Por isso, a utilidade da categorização não está em compartimentos estanques, e sim na disciplina do processo decisório: triagem sequencial, registro de bases fáticas e justificativas, e revisão contínua diante de mudanças regulatórias e técnicas.

Como agenda de evolução, recomenda-se testar a matriz em setores específicos (financeiro, saúde, varejo digital), desenvolver checklists e árvores de decisão por caso de uso (p. ex., enriquecimento de leads, monitoramento de preços, treinamento de modelos), e consolidar um repertório de precedentes e entendimentos administrativos/judiciais.

A proposta aqui apresentada é ponto de partida: oferece linguagem comum e critérios verificáveis para reduzir assimetrias de avaliação, aumentar a segurança jurídica de iniciativas baseadas em dados e promover um equilíbrio mais maduro entre abertura informacional, inovação e tutela de direitos.

REFERÊNCIAS

BRASIL. **Constituição Federal do Brasil de 1988**. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 04 mar. 2024.

BRASIL. **Lei Federal nº 9.279/1996, de 14 de maio de 1996**. Regula direitos e obrigações relativos à propriedade industrial. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/19279.htm. Acesso em: 04 mar. 2024.

BRASIL. **Lei Federal nº 9.610/1998, de 19 de fevereiro de 1998**. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/19610.htm. Acesso em: 04 mar. 2024.

BRASIL. **Lei Federal nº 10.406/2002 de 10 de janeiro de 2002**. Código Civil. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/2002/110406compilada.htm. Acesso em 04 mar. 2024.

BRASIL. **Lei Federal nº 13.709/2018, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 04 mar. 2024.

BRASIL. Receita Federal Do Brasil. **Termos de Uso e Política de Privacidade da Receita Federal do Brasil**. Disponível em: <https://www.gov.br/receitafederal/pt-br/aceso-a-informacao/lgpd/termo-de-uso>. Acesso em: 26 mar. 2024.

CHIGNARD, Simon. **A brief history of Open Data In Paris Tech Review**, mar.2013. Disponível em: <http://www.paristechreview.com/2013/03/29/brief-history-open-data/> Acesso em: 06 mar. 2024.

CONSELHO NACIONAL DE JUSTIÇA. **Recomendação nº 74/2020**. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3487#:~:text=RESOLVE%3A-,%2C%20Art.,exce%C3%A7%C3%A3o%20do%20Supremo%20Tribunal%20Federal>. Acesso em: 26 mar. 2024.

MASUDA, Yoneji. **A sociedade da informação como sociedade pós-industrial** (trad. Kival Charles Weber e Angela Melim), Rio de Janeiro: Ed. Rio, 1982.

MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. **Big data: a revolution that will transform how we live, work, and think**, Nova Iorque: Houghton Mifflin Harcourt, versão eletrônica, 2014.

MERTON, Robert K. The normative structure of science *In* **The sociology of science: theoretical and empirical investigations**, editado por Norman Storer, Chicago e Londres: The University of Chicago Press, 1973.

OLX. **Termos e Condições Gerais de Uso**. Disponível em: <https://ajuda.olx.com.br/public/T%26C+OLX+v2+06-Jun-2022-172210.pdf>. Acesso em: 26 mar. 2024.

OPEN KNOWLEDGE FOUNDATION. **Open definition**. Disponível em: <https://opendefinition.org/od/2.1/en/> Acesso em: 06 mar. 2024.

SCHWAB, Klaus. **A quarta revolução industrial** (trad. Daniel Moreira Miranda), São Paulo: Edipro, 2016.