XEQUE-MATE

o tripé da
proteção de dados pessoais
no jogo de xadrez das
iniciativas legislativas
no Brasil

PRIVACIDADE E VIGILÂNCIA

GPoPAI/USP
grupo de pesquisa em políticas públicas
para o acesso à informação

são paulo
02.07.2015

GPOPAI

APRESENTAÇÃO

GPOPAI USP

Grupo de Pesquisa em Políticas Públicas Para o Acesso à Informação da Universidade de São Paulo

PROJETO
DE PESQUISA
PRIVACIDADE
E VIGILÂNCIA

O GPoPAI é um grupo de pesquisa sediado na Universidade de São Paulo (Campus Leste) que se dedica, em linhas gerais, à investigação dos impactos das novas tecnologias em termos de políticas públicas. Dentre os seus projetos de pesquisa, encontra-se, em andamento, o projeto intitulado "Estruturação de um campo de estudos sobre privacidade e vigilância no Brasil". Um dos eixos deste projeto de pesquisa destina-se ao acompanhamento e à execução de ações participativas no campo acadêmico e da sociedade civil para fins de aprimoramento legislativo da proteção de dados pessoais no Brasil.

FINANCIAMENTO

O projeto de pesquisa **Privacidade e Vigilância** é financiado pela Fundação Ford.

COORDENADORES

Jorge A. S. Machado
Pablo Ortellado
Márcio Moretto Ribeiro

PESQUISADOR AUTOR DO TEXTO

Bruno Ricardo Bioni

EQUIPE DE PESQUISA

Amanda Monteiro
Eduardo Sales
Marina Salles
L. Almeida
Caio Canic Silva
Leonardo Piccioni
Carybé Silva

PROJETO GRÁFICO

Goma Oficina

A

OBJETIVO E SUMÁRIO EXECUTIVO

O objetivo deste guia é organizar conceitos e referenciais teóricos básicos sobre três temas basilares da proteção dos dados pessoais, verificando-se como eles foram endereçados pelas iniciativas legislativas que pretendem regulamentar de forma abrangente a matéria no Brasil[1]. Não se trata de uma obra monográfica que irá esgotar cada um dos temas, mas que pretende ser informativa para quem queira pesquisar sobre o assunto, especialmente para fins de engajamento junto ao debate legislativo sobre proteção de dados pessoais.

Os resultados encontrados que contribuem para a discussão regulatória são:

1 A definição do conceito de dados pessoais pode seguir uma orientação expansionista ou reducionista. Há um vocabulário que, respectivamente, alarga (pessoa identificável) ou restringe (pessoa identificada) o escopo de uma lei de proteção de dados pessoais;

1.1 Ainda que divergentes, tais teorizações detêm o mesmo centro gravitacional. Ambas demandam uma análise contextual de onde está inserido um dado, a fim de aferir o seu grau de identificabilidade para compreender se uma determinada informação está relacionada a uma pessoa identificada ou identificável;

1.2 Essa análise pouco trivial foi ilustrada através do exemplo de uma base de dados relacionais. A simples inclusão ou exclusão de colunas ou linhas nessa base de dados, estruturada em formato de uma tabela, são fatores determinantes para um dado ser adjetivado como pessoal;

1.3 A rivalização desses referenciais teóricos aparece também nas chamadas base de dados não estruturadas (NoSQL). Diferentemente da base de dados relacionais, elas não têm uma linguagem organizada sendo dificilmente alcançadas pelo conceito reducionista de dados pessoais. Com isso, o conceito expansionista tem maior aderência à realidade, na medida em que tais bases de dados têm se tornando cada vez mais comuns em decorrência da viralização de tecnologia do *Big Data*;

1. Seria equivocado dizer que no Brasil não há normas de proteção dos dados pessoais, isto porque, existem diversas normas legais e infralegais sobre a matéria. Falta, contudo, uma lei que regulamente de forma abrangente a proteção dos dados pessoais dos cidadãos, a fim de extrapolar uma proteção meramente setorial. Essa é a pretensão das iniciativas legislativas que serão analisadas por esse guia.
Veja-se, a título de exemplo, as seguintes leis setoriais de proteção de dados pessoais vigentes no ordenamento jurídico brasileiro:
a) proteção dos dados pessoais dos consumidores: Código de Defesa do Consumidor (artigos 43 e 44); **b)** proteção dos dados pessoais para fins de análise de crédito: Lei do Cadastro Positivo (Lei nº 12.414/2011); **c)** proteção dos dados no âmbito da transparência da Administração Pública: Lei de Acesso à Informação Pública (artigo 31 da Lei nº 12.527/2011); **d)** proteção de dados pessoais dos usuários de internet: Marco Civil da Internet (artigo 3º, incisos II e III, 7º a 17 da Lei nº 12.965/2014).

2 A antítese dos dados pessoais seriam os dados anônimos. Contudo, eles encerram uma falsa promessa (semântica), na medida em que sempre haverá a possibilidade de reversão do seu processo e a conseguinte identificação de um indivíduo;

2.1 Por isso, propostas legislativas que adotem, ao mesmo tempo, o conceito expansionista de dados pessoais e dados anônimos correm o risco de serem tautológicas. Haveria uma redundância entre tais conceitos, na medida em que dados anônimos são, potencialmente e provavelmente, dados relacionados a uma pessoa identificável;

2.2 Para se evitar tal tipo incoerência, surge o conceito de razoabilidade como delimitador da elasticidade do conceito expansionista de dados pessoais. Não basta a mera possibilidade de que um dado seja atrelado a uma pessoa para ser enquadrado como identificável, essa vinculação deve ser objeto de um esforço razoável;

2.3 A problemática reside na significação desse conceito equívoco, já que o termo razoabilidade comporta várias interpretações: **I.** o próprio mercado poderia se autorregular, extraindo-se das suas práticas correntes o que seria um esforço que extrapolaria o razoável para fins de reversão do processo de anonimização; **II.** a legislação poderia franquear tal capacidade autorregulatória ao mercado, mas observado critérios pré-estabelecidos (*ex ante*) a orientar a significação do termo razoável; **III.** o órgão fiscalizador poderia ter tal ingerência, que poderia não estar vinculado a critérios pré-estabelecidos em lei, atuando, dessa forma, com maior discricionariedade (*ex post*);

2.4 Mas, por trás de todo esse debate teórico, deve-se observar que para direcionar conteúdo, publicidade e sujeitar uma pessoa a um processo de decisão automatizada não é necessário saber a sua identidade. Basta lhe atribuir um identificador eletrônico, sem qualquer correlação com a sua identidade real (e.g., nome, número de CPF, etc.). Reconhecendo-se o dispositivo conectado, já é possível moldar toda a sua navegação e formar um perfil completo de todas as suas atividades;

2.5 E, dessa forma, o cidadão, intermediado pela máquina a qual ele está conectado, pode ser afetado de alguma

forma. Consumidores já foram discriminados por sistemas de flutuação de preços, que lhe direcionaram ofertas mais caras, por serem usuários de Mac/Apple ao invés de Android. Essas inferências poderão se proliferar com a internet das coisas, onde cada dispositivo permitirá uma catalogação completa e precisa da pessoa não identificável, mas alcançável por tais práticas;

2.6 Se o objetivo de uma lei geral de proteção de dados pessoais é tutelar o cidadão exposto a tais práticas, deve-se, então, pensar de forma consequencialista e normatizar tais situações nas quais a mera atividade de tratamento de dados para fins de profiling tem repercussões sobre a esfera de uma pessoa;

3 O consentimento tem sido o pilar da maioria, senão de todas as leis de proteção de dados pessoais. O único dissenso reside em torno da sua adjetivação, que procura restringir ou expandir a participação do cidadão na dinâmica da proteção de seus dados. Notou-se haver uma escala progressiva da qualificação do consentimento que são variações a respeito de qual deve ser a sua carga participativa: **I.** informado: básico/pressuposto; **II.** livre: mínima; **III.** para finalidades determinadas: pré-intermediária; **IV.** inequívoco: intermediária; **V.** específico e expresso: máxima;

3.1 Como para toda regra existem exceções, não é diferente com o consentimento. Destaca-se a hipótese denominada de interesses legítimos que, em termos práticos, autoriza o tratamento de dados pessoais sem que haja ulterior consentimento do seu titular. O termo ulterior sinaliza que tal hipótese não se dá no vácuo, mas dentro da dinâmica da regra geral em que há uma relação pré-estabelecida na qual o titular consentiu previamente para um uso específico ou para uma finalidade determinada de seus dados;

3.2 Com isso, abre-se espaço para que os dados sejam reutilizados, não estando adstritos a uma finalidade anterior para a qual foi dado o consentimento do seu titular. Dessa forma, tecnologias como *Big Data* encontram ressonância na figura dos legítimos interesses, por ser uma tecnologia focalizada na reutilização de uma mesma base dados para diversos propósitos;

3.3 Deve-se observar, contudo, a existência de freios e contrapesos para que tal exceção não seja tão elástica a ponto de distorcer a regra geral do consentimento. A previsão de que deve haver mecanismos de transparência para que o cidadão possa se opor a tal tipo de tratamento de dados pessoais, a atuação fiscalizatória do órgão regulador para auditar tais práticas do mercado e a exigência de padrões de segurança que minimizem os riscos à privacidade dos cidadãos são fatores a serem levados em conta para desvendar se há um saldo positivo estabilizador da regra do consentimento;

3.4 A prometida capacidade do cidadão em controlar seus dados pessoais deve ser nutrida:

3.4.1 Prever que o cidadão não deve se submeter à lógica do tudo ou nada dos contratos de adesão, na qual ele deve consentir com o uso de seus dados pessoais sob pena de não ter acesso a um produto ou serviço. Deve-se dar granularidade a tal escolha, de modo que ele possa consentir sobre os diversos usos, por quanto tempo e frequência, sobre o compartilhamento com terceiros e etc. Trata-se de franquear um leque de escolhas ao cidadão, a invés de reduzí-la a uma escolha binário do tudo ou nada;

3.4.2 Tal como na telefonia móvel, deve-se permitir que o cidadão possa migrar de aplicações, levando consigo seus dados pessoais para serviços ou produtos mais *privacy-friendly*. Esse direito de portabilidade tende a elevar a privacidade como elemento de competitividade e, por conseguinte, maximizar as escolhas dos cidadãos frente a uma plêiade de atores que buscarão cativá-lo com ofertas mais pró-proteção de dados pessoais.
Pelos resultados de pesquisa supracitados, percebe-se que esse guia pode servir como ponto apoio para um estudo introdutório do tema da proteção de dados pessoais, não se limitando, apenas, a uma análise comparativa estática das três propostas legislativas analisadas. Por isso, ressalta-se que esse guia busca ser atemporal, contendo outra faceta, mais ampla e genérica, que é contribuir para o debate regulatório da proteção de dados pessoais no Brasil.

B

CONTEXTO: A AGENDA LEGISLATIVA DE PROTEÇÃO DE DADOS PESSOAIS NO BRASIL

Atualmente, existem 03 (três) iniciativas legislativas que pretendem regulamentar de forma geral o tema:

PROJETO DE LEI DE AUTORIA DO EXECUTIVO (PLPDP/EXE)

Em fevereiro de 2015, o Ministério da Justiça retomou a pauta de 05 (cinco) anos atrás de um anteprojeto de lei de proteção de dados pessoais. Por meio de uma nova plataforma, foi aberta para consulta pública uma nova versão do anteprojeto que recebeu contribuições de toda a sociedade até julho de 2015. Em outubro daquele ano, veio a lume a versão-pós consulta pública que foi encaminhada ao Congresso Nacional com poucas modificações em 12 de maio de 2016, transformando-se no Projeto de Lei nº 5.276/2016. Esse projeto lei tramita[2] em caráter de urgência constitucional, o que significa, em termos práticos, um trâmite legislativo mais célere, com ausência de algumas formalidades, e, o mais importante, há uma espécie de sanção ou mecanismo de pressão para que esse rito mais abreviado seja levado a cabo. Ao invés 10 (dez) sessões legislativas e uma tramitação progressiva nas Comissões designadas para a sua apreciação , o PLPDP/EXE deverá ser analisado em 05 sessões e de forma simultânea por todas elas. Essa economia procedimental deve resultar, necessariamente, na sua análise em até 45 (dias) a contar da data do seu recebimento na Câmara, sob pena do trancamento da pauta da casa. Se até o dia 27 de junho de 2016 tal proposição não for analisada, haverá obrigatoriamente a sua inclusão na pauta do Plenário e o sobrestamento de todas as matérias sujeitas à apreciação[3] da Câmara dos Deputados.

2. Levando em consideração o período em que este documento está sendo elaborado.

3. (Comissão de Constituição e Justiça e de Cidadania (CCJC) – Relator: Deputado Alessandro Molon; Comissão de Trabalho, de Administração e Serviço Público (CTASP) – Relator: Deputado Orlando Silva. Já houve o pedido de redistribuição do Projeto de Lei para Comissão de Ciência Tecnologia e Informática, que aguarda análise pela Presidência da Casa.

PROJETO DE LEI DO SENADO FEDERAL DE PROTEÇÃO DE DADOS PESSOAIS (PLPDP/SEN)

Em abril de 2015, os projetos de leis/PLs nºs 330/2013, 181/2014[4] e 131/014 passaram a tramitar sob a relatoria do Senador Aloysio Nunes. Em outubro daquele ano, após a realização de audiência pública, o senador apresentou substitutivo a tais projetos de leis. Com tal fusão, tais iniciativas legislativas ganharam o corpo de uma lei geral de proteção de dados pessoais (PLPDP/SEN). Aprovado na Comissão de Ciência e Tecnologia (CCT), em 13 de outubro de 2015,e, mais recentemente, no último dia 10 de maio de 2015 na Comissão e Meio Ambiente, Defesa do Consumidor e Fiscalização e Controle (CMA), a matéria segue agora para a Comissão de Assuntos Econômicos (CAE). Se aprovado na CAE, a sua última parada, antes de ir para votação no Plenário, será a Comissão de Constituição e Justiça (CCJ).

4. Esse projeto de lei em específico foi fruto dos trabalhos de Comissão de Inquérito Parlamentar que foi instalada em decorrência das revelações do ex-analista da Agência Nacional de Segurança dos Estados Unidos, Edward Snowden, de que o governo brasileiro teria sido alvo de espionagem do governo americano. No relatório final da chamada CPI da Espionagem é que foi gestada tal iniciativa legislativa.

PROJETO DE LEI DA CÂMARA DOS DEPUTADOS DE PROTEÇÃO DE DADOS PESSOAIS (PLPDP/CAM)

Desde 2012 tramita na Câmara dos Deputados o PL nº 4060/2012. Entre arquivamentos (janeiro) e desarquivamentos (fevereiro) no ano de 2015, houve a realização de audiências públicas para auxiliar os parlamentares na análise do projeto. Atualmente, tal iniciativa legislativa encontra-se na comissão de Ciência e Tecnologia, Comunicação e Informática. No último dia 04 de maio, o deputado-relator Sóstenes Cavalcante apresentou parecer de aprovação com algumas emendas substitutivas. Com isso, o PLPDP/CAM já foi pautado diversas vezes na CCTCI, mas ainda não foi apreciado por todos os seus deputados-integrantes.

São três, portanto, as iniciativas legislativas que visam regulamentar de maneira ampla a proteção de dados pessoais no Brasil. Todas elas se movimentaram e continuam a se movimentar concomitantemente, havendo, especialmen-

te, uma sobreposição de acontecimentos no último mês de maio de 2016. Essa forte movimentação recente, associada ao regime de urgência constitucional atribuída a uma delas, desenha uma conjuntura favorável para que o Brasil venha a ter finalmente uma lei geral de proteção de dados pessoais (Bioni & Monteiro, 2016).

Nesse contexto, esse guia visa ser informativo sobre como tais peças do xadrez do legislativo brasileiro convergem e divergem entre si. O recorte dessa análise centra-se em três aspectos centrais de qualquer lei de proteção de dados pessoais, o que chamamos de **tripé da proteção dos dados pessoais.**

C

O TRIPÉ DA PROTEÇÃO DOS DADOS PESSOAIS

Os três temas escolhidos são:

I conceito de dados pessoais;
II dados anônimos e;
III consentimento do titular dos dados pessoais.

A escolha dessa tríade não é acidental. O primeiro e o segundo são o que definem o escopo de qualquer normativa de proteção de dados pessoais, filtrando quais dados merecem proteção. O terceiro é o pilar normativo da grande maioria das leis ao redor do mundo, senão de todas elas.

O protagonismo desses três temas confirmou-se com a experiência do debate público do então Anteprojeto de Lei de Proteção de Dados Pessoais (PLPDP/EXE). Ao todo foram 1.856 (um mil e oitocentos e cinquenta e seis) comentários[5] , que se debruçaram sobre os mais de 13 (treze) eixos temáticos do debate público, sendo que os três temas acima destacados foram os mais debatidos.

O consentimento foi o **campeão** de todos eles com 336 (trezentos e trinta e seis) comentários. O conceito de dados pessoais, juntamente com dados anônimos, foram os **vice-campeões** com 295 (duzentos e noventa e cinco) comentários.

A importância desses três temas reflete-se, portanto, no engajamento da sociedade em debatê-los, o que legitima a escolha deles para compor o denominado tripé da proteção de dados pessoais.

5. Esse número não contabiliza as contribuições realizadas em formato PDF, mas, tão somente, os comentários realizados por meio de uma interface gráfica da plataforma online criada para o debate público.

D

ORGANIZAÇÃO E PREMISSAS METODOLÓGICAS DO GUIA

Primeiro, haverá um breve mapeamento dos referenciais teóricos com o objetivo de introduzir o tema ao leitor e emitir um diagnóstico do "estado atual da arte" da matéria, verificando-se como a elaboração de conceitos e estratégias regulatórias nas normativas de proteção de dados pessoais é orientada por tais fundações teóricas.

Em um segundo momento, confrontar-se-á tal diagnóstico com o texto das iniciativas legislativas para que o leitor possa visualizar as escolhas encampadas pelo legislador brasileiro, possibilitando-lhe uma percepção crítica frente ao leque de opções antes revelados pelo mapeamento teórico desse guia.

Ao longo dessas duas frentes, com objetivo de tornar mais concreto e prático esse guia, recorrer-se-á a exemplos práticos que serão acomodados em "caixas" para que o leitor possa distingui-los da primeira parte (**xeque ao rei**). E, ao se comparar os projetos de leis, utilizar-se-á grifos de colorações diferentes para destacar as semelhanças e diferenças entre eles.

Pretende-se, com isso, dar maior concreção às implicações que as diferentes escolhas em jogo são capazes de produzir, reduzindo-se, em última análise, a assimetria informacional para quem queira se engajar nesse debate legislativo e regulatório.

SUMÁRIO

1
DEFINIÇÃO DE DADOS PESSOAIS

1.1 QUAIS OS TAMANHOS DOS POSSÍVEIS GARGALOS PARA A PROTEÇÃO DOS DADOS PESSOAIS?

O conceito de dados pessoais é um elemento chave. Ele filtra o que deve estar dentro ou fora do escopo de uma lei de proteção de dados pessoais, demarcando o terreno a ser por ela ocupado. Diferenças sutis em torno da sua definição implicam em consequências drásticas para o alcance dessa proteção. Por isso, compreender se um dado pode ser adjetivado como pessoal é, antes de tudo, um exercício de interpretação detido sobre cada palavra utilizada para prescrever a sua conceituação.

De forma sistemática, o seu vocabulário é composto por palavras que restringem ou alargam o gargalo dessa proteção. Há uma bipartição do seu léxico que ora retrai (reducionista), ora expande (expansionista), a moldura normativa de uma lei de proteção de dados pessoais (Solove & Schwartz, 2011: 1872).

1.1.1 REDUCIONISTA VERSUS EXPANSIONISTA

A orientação reducionista baseia-se em uma lógica restritiva pela qual dado pessoal é uma informação que deve estar associada a uma pessoa específica. Ele deve ser um signo que permita estabelecer de forma imediata ou direta um vínculo com o seu titular, individualizando-o de forma precisa. Um dado para ser pessoal deve ser, portanto, a projeção de uma pessoa una e inequívoca (identificada) (Poullet & Dinant, 2004:29).

Enquanto que a expansionista aposta em uma lógica mais flexível, que desconsidera a associação exata entre uma informação e uma pessoa. Dado pessoal pode ser qualquer tipo de informação que permita a sua identificação, ainda que o vínculo entre o dado e um indivíduo não seja estabelecido de prontidão, mas de forma mediata ou indireta. Um dado para ser pessoal deve ser, portanto, a projeção de uma pessoa identificável.

Ainda que divergentes, tais teorizações detêm o mesmo *centro gravitacional*. Ambas demandam uma análise contextual donde está inserido um dado, aferindo-se o seu grau de *identificabilidade*[6] para, então, desencadear a compreensão se uma determinada informação está relacionada a uma pessoa identificada ou identificável (Solove & Schwartz, 2011: 1874).

Por isso, a intelecção do conceito de dado pessoal e, por conseguinte, das estratégias regulatórias possíveis para a sua definição é algo fluído que pode ser esclarecido a partir da dinâmica de conceitos básicos de sistemas de informação e de banco de dados. Somente, assim, o seu vocabulário ganhará uma análise mais concreta a demonstrar as diferenças e consequências práticas entre tais estratégias regulatórias distintas e como elas foram absorvidas em cada uma das propostas legislativas em trâmite no congresso nacional.

Os quadros ao lado sintetizam qual é esse vocabulário e como ele está presente nas proposições legislativas brasileiras para, em seguida, proceder-se a uma análise ilustrativa das suas repercussões práticas.

6. Os autores norte-americanos citados utilizaram o termo identifiability. Por falta de correspondência na língua portuguesa, nós optamos por tal neologismo.

QUADRO 1
VOCABULÁRIO ANALÍTICO PARA
A DEFINIÇÃO DO CONCEITO DE DADOS PESSOAIS

EXPANSIONISTA	REDUCIONISTA
Pessoa identificável	Pessoa identificada
Pessoa indeterminada	Pessoa específica/determinada
Vínculo mediato, indireto, impreciso ou inexato	Vínculo imediato, direto, preciso ou exato
Alargamento da qualificação do dado como pessoal	Retração da qualificação do dado como pessoal

QUADRO 2
TABELA ANALÍTICA DA DEFINIÇÃO DE DADOS PESSOAIS
DAS PROPOSTAS LEGISLATIVAS

PLPDP/EXE	PLPDP/SEN	PLPDP/CAM
Art. 5º, inciso I - dado pessoal: dado relacionado à pessoa natural identificada ou identificável, inclusive numeros identificativos, da dos locacionais ou identificadores eletrônicos quando estes estiverem relacionados a uma pessoa	**Art. 3º, inciso I** – dado pessoal: qualquer informa ção sobre pessoa natural identficável ou identificada;	**Art. 7º, inciso I** – dado pessoal: qualquer informação que permita a identificação exata e precisa de uma pessoa determinada;
Expansionista (com um rol exemplifcativo)	Expansionista (com um rol exemplifcativo)	Reducionista (sem rol exemplificativo)

1.1.2 ILUSTRAÇÕES DAS CONSEQUÊNCIAS DAS ESTRATÉGIAS REDUCIONISTA E EXPANSIONISTA

Para fins de ilustração, optou-se pelo exemplo de um banco de dados relacionais. Esse tipo de banco de dados é estruturado por tabelas (Mannino, 2008: 555), onde cada uma de suas colunas – que são chamadas de atributos (Rob, 2011: 33) – é a maneira pela qual os dados são organizados. É a correlação entre as colunas e as linhas dessa tabela que emprestam valor (significado) aos dados, permitindo que deles seja extraído algo inteligível (informações).

Há, assim, uma dinâmica composta pela entrada de um dado (linha), associado a um atributo (coluna), que empresta um significado a todo um amontoado de dados (ROB, 2011: 04). Esse é um exercício didático para entender como os referenciais reducionistas e expansionista distanciam-se completamente um do outro, apesar da proximidade léxica entre entres eles.

TABELA 1
BASE DE DADOS RELACIONAIS

NOME	CPF	CEP	IDADE	CLASSIFICAÇÃO | SEGMENTAÇÃO
Bruno dos Santos	123.456-77	04055-000	18	Jovem hipster
Bruno dos Santos	234.567-88	04055-111	17	Jovem poupador
Bruno dos Santos	345.678-99	04055-222	17	Jovem consumista
Bruno Souza	456.789-10	01201-000	65	Idoso com rentabilidade
Bruno Souza	567.891-01	04201-111	66	Idosa sem rentabilidade
Bruna Schonberg	222.333.44-55	04201-222	70	Idosa com rentabilidade
Maria Silva	157.890.88-66	09201-000	40	Adulto desempregado
Maria Silva	666.666.66-66	09201-111	38	Adulto perfil executivo
Maria Silva	987.354.22-99	09201-222	16	Jovem hipster

HIPÓTESE 1 (REDUCIONISTAS)

A presença de homônimos não permitiria que houvesse a individualização precisa de uma parcela das pessoas inseridas no banco de dados acima, caso não houvesse outros dados – identificadores (únicos) como, por exemplo, o CPF. Somente mediante tal associação identificou-se exatamente o(s) "Bruno(s)" e a(s) "Maria(s)", tornando-as pessoas identificadas, isto é, individualizados de maneira precisa, exata e inequívoca.

HIPÓTESE 2 (EXPANSIONISTAS)

Caso houvesse a eliminação da coluna "B", haveria incerteza a respeito de qual dos "Brunos" seria consumista, poupador ou *hipster* (coluna "E"), já que não haveria outro dado – identificador (único) – para diferenciar cada um dos homônimos. E, o mesmo, com relação a qual Maria seria a desempregada ou a executiva. Tais informações não estariam relacionadas a pessoas identificadas, ainda que elas pudessem vir a ser a partir da sua localização geográfica (coluna "B"), por exemplo. Nesse caso, a incerteza gerada por um grupo de pessoas com o mesmo nome tende a ser eliminada a partir da agregação dessa outra informação, tornando-as pessoas identificáveis. Há, em última análise, o potencial de individualizá-las.

Síntese: verificar se um dado pode ser adjetivado como pessoal é uma análise contextual que depende de qual tipo de informação poder ser extraída de uma base de dados. Essa análise circunstanciada pode ser mais dura ou flexível. Para os reducionistas somente na hipótese "A" haveria dados pessoais. Para os expansionistas as hipóteses "A" e "B" seriam abraçadas pelo conceito de dado pessoal.

XEQUE AO REI

BASE DE DADOS NÃO ESTRUTURADOS (NOSQL)

Uma outra diferença pragmática a ser apontada entre reducionistas e expansionistas seria o seu alcance com relação às chamadas base de dados não estruturadas (NoSQL, *not only structured query language*). Imagine-se que, ao contrário do exemplo anterior, todos os dados estivessem dispersos, não havendo uma correlação imediata estabelecida entre uma pessoa e uma informação.

Ou seja, eles não estariam em uma linguagem organizada (SQL – *structured query language*), caracterizada no nosso exemplo pela presença de colunas e linhas que permitissem estabelecer um vínculo direto entre os dados e os seus respectivos titulares. Logo, o conceito reducionista de dados pessoais seria dificilmente aplicável.

Dado o momento atual em que bases de dados não estruturadas têm ditado cada vez mais as práticas de mineração de dados, associadas à tecnologia do *Big Data* (vide: subcapítulo 3.3), a escolha por um conceito reducionista de dados pessoais tende a ter pouca penetração na realidade.

Ao passo que o conceito expansionista teria uma maior incidência. A desestruturação de uma base de dados não retira a possibilidade de que eles sejam associados a seus titulares. Pelo contrário, cada vez mais, novas tecnologias – e.g., algoritmos – são programadas para, justamente, estabelecer tal conexão que *a priori* mostrava-se remota e distante (vide: subcapítulo 2.2).

Nesse cenário, apesar da proximidade léxica entre expansionistas e reducionistas (identificável *versus* identificada), verifica-se um antagonismo significativo entre eles. O primeiro referencial coloca em xeque o segundo, por ter uma aderência normativa muito maior que é capaz de alcançar, inclusive, base de dados NoSQL.

2
DADOS ANÔNIMOS E A SUA PSEUDO-DICOTOMIA COM O CONCEITO DE DADOS PESSOAIS

2.1 ENTENDENDO A DINÂMICA DO PROCESSO DE ANONIMIZAÇÃO

A antítese do conceito de dado pessoal seria um dado anônimo, ou seja, aquele que é incapaz de revelar a identidade de uma pessoa[7] . Essa inaptidão pode ser fruto de um processo pelo qual é quebrado o vínculo entre o(s) dados e seu(s) respectivo(s) titular(es), o que é chamado de anonimização (Doneda, 2006: 157-58).

Esse processo pode se valer de diferentes técnicas que buscam eliminar tais elementos identificadores de uma base de dados (Council of Europe, 2014: 44-43), variando entre: a **I.** supressão, **II.** generalização (Ohm, 2010: 1713); **III.** randomização e **IV.** pseudoanonimização[8].

Foge do escopo desse guia abordar cada uma das citadas técnicas do processo de anonimização. Vale a pena, no entanto, abordar as duas primeiras para ilustrar a sua dinâmica e pavimentar a via de acesso para o objetivo central desse guia que é se debruçar sobre as implicações normativas de uma eventual dicotomia entre dados anônimos (anonimizados) e dados pessoais e como isso está presente nas propostas legislativas em curso no Brasil.

Retomando-se o nosso exemplo de base de dados relacionais estruturada, deve-se identificar quais elementos poderiam ser modificados – suprimidos ou generalizados – para que o seu grau de identificabilidade seja, em tese, eliminado.

I supressão do CPF: por ser um identificador que é capaz de diferenciar até mesmo pessoas homônimas, em razão de ser um identificador único. Logo, a sua disponibilização mesmo que parcial – e.g., 05 (cinco) primeiros dígitos – não seria prudente;

II generalização do nome completo: constaria apenas o prenome, desde que fosse observado que os nomes da base de dados não são comuns. Assim, o coletivo de pessoas que deles poderia se apossar tende a mitigar o risco de identificação, o que não ocorreria se os nomes fossem pouco usuais;

7. Diante do próprio significado linguístico, anônimo seria aquele que não tem nome diante da ausência de um signo identificador: HOUAISS, Antônio. VILLAR, Mauro de Salles. Op. Cit., p. 140-141.

8. Para muitos a pseudoanonimização não é considerada como uma técnica de anonimização. Isto porque, substitui-se, apenas, os identificadores diretos – e.g., nome, CPF e etc – por pseudônimos – e.g., números aleatórios, de modo que a pessoa permanece sendo identificável em razão de tais pseudônimos serem um retrato detalhado indireto delas (WP 29, 2014: 20).

III generalização da localização geográfica: ao invés de disponibilizar o número completo do CEP, poder-se-ia divulgar apenas os seus primeiros dígitos. Assim, haveria uma localização menos detalhada, a fim de quebrar o vínculo de identificação desta informação com um sujeito;

IV generalização da idade: ao invés de divulgar a idade exata, poder-se-ia apenas indicar a faixa etária para viabilizar a categorização dos indivíduos como jovens, adultos ou idosos (coluna "E") e, por outro lado, inviabilizar a sua individualização dada a exponencialidade de pessoas que se enquadram naquela mesma faixa etária.

Com maior ou menor grau de intensidade – e.g., supressão ou generalização – nota-se um método cujo mote é gerenciar circunstancialmente a *identificabilidade* de uma base de dados. As características de cada dado e a percepção deles estarem inseridos em um *pool* de informações devem orientar tal análise. Por isso, não há um único método ou uma combinação perfeita *ex ante* a parametrizar o processo de

TABELA 2
BASE DE DADOS RELACIONAIS ANONIMIZADA

NOME	CPF	CEP	IDADE	CLASSIFICAÇÃO | SEGMENTAÇÃO
Bruno [redacted]	[redacted]	04055-[redacted]	18 >	Jovem hipster
Bruno [redacted]	[redacted]	04055-[redacted]	18 >	Jovem poupador
Bruno [redacted]	[redacted]	04055-[redacted]	18 >	Jovem consumista
Bruno [redacted]	[redacted]	01201-[redacted]	60 <	Idoso com rentabilidade
Bruno [redacted]	[redacted]	04201-[redacted]	60 <	Idosa sem rentabilidade
Bruna [redacted]	[redacted]	04201-[redacted]	60 <	Idosa com rentabilidade
Maria [redacted]	[redacted]	09201-[redacted]	18 <	Adulto desempregado
Maria [redacted]	[redacted]	09201-[redacted]	18 <	Adulto perfil executivo
Maria [redacted]	[redacted]	09201-[redacted]	18 >	Jovem hipster

anonimização, devendo-se analisar contextualmente como ele deve ser empreendido para que os titulares dos dados anonimizados não sejam reidentificados, nem mesmo por quem procedeu a sua anonimização (Article 29 WP, 2014: 11).

2.2 O MITO DA IRREVERSIBILIDADE DOS DADOS ANONIMIZADOS E AS FRONTEIRAS ELUSIVAS JUNTO AO CONCEITO DE DADO PESSOAL (EXPANSIONISTA)

Amarrar o conceito teórico de dados anônimos a uma análise contextual com os olhos voltados para a irreversibilidade do processo de anonimização joga luz diretamente sobre o fator problemático dessa proposição: o seu caráter elusivo ou mesmo a sua impossibilidade teórica (Texeira, 2015).

Torna-se cada vez mais recorrente a publicação de estudos que demonstram ser o processo de anonimização algo falível. A representação simbólica de que os vínculos de identificação de uma base de dados poderiam ser completamente eliminados, garantindo-se, com 100% (cem por cento) de eficiência, o anonimato das pessoas (Narayana & Shmatikov, 2010: 24) é um mito.

Os pesquisadores Arvind Narayanan e Latanya Sweeney têm se destacado nessa área. Eles reidentificaram diversas bases de dados anonimizados, como a do censo americano, de hospitais e centros de pesquisa de saúde, bem como de provedores de aplicação de internet – o famoso caso Netflix Prize.

Nesse último caso, Arvind Narayanan, junto com outro cientista da computação (Vitaly Shamatikov), desenvolveram um algoritmo que calculava dentre outras coisas: **I.** quantos *bits* de informação seriam necessários para reverter um processo de anonimização; **II.** qual seria o melhor critério para a escolha de uma informação auxiliar – i.e., uma outra base de dados – a ser agregada para tanto; **III.** uma métrica sobre a probabilidade de acerto da re-indetificação, evitando-se "falsos positivos" – i.e., a vinculação errônea de indivíduos aos dados desanonimizados (Narayana & Shmatikov, 2008: 6).

À época a maior provedora de *streaming* de filmes do mundo criou um concurso, cujo desafio era melhorar o seu algoritmo de sugestão de filmes. Então, a Netflix disponibi-

lizou a sua base de dados com todas as avaliações dos filmes de seu catálogo do período de 1998 à 2005, suprimindo os nomes dos usuários avaliadores e deixando somente a data e a nota da avaliação.

A fim de tornar tal processo de de-identificação mais robusto, a Netflix se utilizou da técnica de randomização (vide: item 2.1). Ela alterou algumas datas e *ratings* das avaliações dos seus consumidores (vide: as regras do concurso), o que aumentaria o risco de "falsos positivos", a não ser pelo fato do algoritmo dos pesquisadores ter sido projetado para isso.

Os pesquisadores "rodaram" tal algoritmo na base de dados disponibilizada, descobrindo que seria necessário entre 03 (três) a 19 (dezenove) *bits* de informação para reverter o processo de anonimização e que esse *pool* de informações necessário estava publicamente disponível e acessível na Internet Movies Databases/IMDB (Narayana & Shmatikov, 2008: 12).

O IMDB é um *website* onde as pessoas compartilham suas impressões sobre filmes, utilizando-se, na maioria das vezes, dos seus nomes reais. Desta forma, os pesquisadores "cruzaram" essas informações com a base de dados da Netflix, o que gerou uma sinergia entre elas a respeito das datas das avaliações dos filmes e dos seus respectivos scorings. Assim, a peça-faltante do quebra-cabeça – a identidade dos usuários dos Netflix – foi desvendada com base nos nomes contidos das avaliações do IMDB (Narayana & Shmatikov, 2008: 11).

O exemplo em questão é simbólico, pois ele sublinha o "calcanhar de Aquiles" dos dados anônimos. Sempre existirá a possibilidade de uma base de dados anonimizada ser agregada a outra para a sua reidentificação. É o que se costuma chamar de entropia da informação (Ohm, 2010: 1749), que é o uso de uma informação auxiliar para a reversão do processo de anonimização (Narayana & Shmatikov, 2008: 4).

E o cenário atual é desafiador a esse respeito:

I nossas vidas têm sido, cada vez mais, *datificadas* (Mayer-Schoneberger & Cukier, 2013: 96), na medida em que:

I.a. cresce-se a cultura do *open data* e, com ela mais informações sobre os cidadãos estão dispersas e publicamente acessíveis na rede[9];

9. Veja-se, nesse sentido, a política de dados abertos do Poder Executivo Federal (Decreto Federal nº 8777/2016) que disponibilizará ao público e em formato aberto diversas bases de dados da administração pública da administração federal. A mais sensível delas é a do Sistema Nacional de Informações de Registro Civil – SIRC, que contém dados sobre nascimentos, casamentos, divórcios e óbitos de todos os cidadãos brasileiros.

I.b cresce-se nossa interação com o mundo *online*, havendo uma completa biografia digital das nossas vidas que é compartilhada com inúmeros atores dos nossos "relacionamentos *online*" (Solove, 2004: 46);

II. cada vez mais um seleto grupo de *players* consolidam a sua posição dominante e, com isso, há uma maior concentração dos nossos dados em poucas mãos;

III. a existência de uma indústria (*data brokers*) que tem como único propósito reunir o maior volume possível de dados pessoais para deles extrair a sua máxima rentabilidade (*Federal Trade Commission*, 2014).

Nesses dois últimos casos, vale ressaltar, há um poderio informacional que é uma ruína para qualquer base de dado anonimizada, na medida em que se trata de um gama de informações auxiliares com o potencial de viralizar a sua reidentificação (Ohm, 2010: 1748).

Em outras palavras, não faltam atores e um manancial de dados para desbancar qualquer processo de anonimização. Tem-se um contexto que é completamente antagônico à promessa semântica de dados anônimos, como aquele que não seria capaz, em hipótese alguma, de identificar um sujeito.

Por isso, a sua falibilidade tem se mostrado consensual, até mesmo entre aqueles que são seus defensores mais ferrenhos (El Emam & Malin, 2015). E, com isso, migra-se de um discurso que outrora sustentava a irreversibilidade do processo de anonimização para uma abordagem centrada na mitigação dos riscos dessa reversibilidade (UK *Information Commissioner´s Office*, 2012). Isso traz implicações normativas relevantes, sobretudo do ponto de vista de uma eventual dicotomia entre dados pessoais e dados anônimos.

XEQUE AO REI

(IN)COERÊNCIA NORMATIVA DE UMA DICOTOMIA ENTRE DADOS ANÔNIMOS E O CONCEITO DE DADOS PESSOAIS (EXPANSIONISTA VERSUS REDUCIONISTA)

Partindo do pressuposto que dados anônimos são sempre reversíveis, eles sempre terão o potencial de identificar alguém. Logo, atrai-se o conceito expansionista de dados pessoais que carrega consigo justamente essa *virtu* – uma informação relacionada a uma pessoa identificável (vide: subcapítulo 1.1). Propugnar o contrário, seria distorcer a compreensão semântica do sufixo em questão – identificável – que compõe tal definição mais alargada de dados pessoais.

Por isso, a princípio, uma eventual dicotomia entre dados pessoais e dados anônimos só guardaria coerência junto ao conceito reducionista de dados pessoais (vide: subcapítulo 1.1), na medida em que dados anônimos não são dados relacionados a uma pessoa identificada. Pelo contrário, demanda-se a reversão do processo de anonimização para se chegar aos respectivos titulares, sendo a sua identificabilidade remota (identificável) e não imediata (identificada).

Desta forma, propostas legislativas que adotam o conceito expansionista de dados pessoais e, ao mesmo tempo, estabelecem uma dicotomia deste com dados anônimos correm o risco de serem tautológicas. Isso porque, haveria uma redundância normativa entre eles, já que dados anônimos são, em última análise, potencialmente e provavelmente, dados relacionados a uma pessoa identificável.

Para não gerar tal incoerência, a única saída seria a adoção de um filtro que delimitasse a elasticidade do conceito expansionista – neste caso o sufixo identificável –, sob pena da fronteira entre dados pessoais e dados anônimos ser sempre transponível. Esse ponto é mais uma das diferenças do jogo de xadrez das iniciativas legislativas sob análise, a ser abordada em seguida.

2.3 RAZOABILIDADE DO IDENTIFICÁVEL: INFLUÊNCIA DA DIRETIVA DA UNIÃO EUROPEIA

Seja no texto da diretiva 95/46[10], seja em sua regulamentação[11], o direito comunitário europeu valeu-se do critério da razoabilidade para delimitar o espectro de incidência do conceito expansionista de dados pessoais. Não basta a mera possibilidade de que um dado seja atrelado a uma pessoa para atrair o sufixo identificável (Article 29 WP 29, 2007: 15). Essa vinculação deve ser objeto de um esforço razoável, sendo esse o perímetro de elasticidade do conceito de dado pessoal como aquele relacionado a uma pessoa identificável.

A *contrarium sensu*, de acordo com as regulações europeias, se para a correlação entre um dado e uma pessoa demanda-se um esforço fora do razoável, não há que se falar em dados pessoais. Nessa situação, o dado é considerado como anônimo, uma vez que o filtro da razoabilidade barra o seu enquadramento como aquele relacionado a uma pessoa identificável (Article 29 WP, 2007: 21).

Trata-se de uma estratégia normativa alinhada à premissa de que os dados anônimos são sempre passíveis de reversão. Daí porque, o critério da razoabilidade nada mais é do que uma diretriz acerca do que venha a ser um risco aceitável em torno da reversibilidade do processo de anonimização, a fim de que os dados anonimizados estejam fora do conceito de dados pessoais.

Com isso, há coerência em se estabelecer conceitos diferentes para tais espécies de dados, sobretudo sob o ponto de vista de uma dicotomia mutualmente excludente entre eles que é delimitada pelo fator da razoabilidade. Do contrário, repita-se, haveria uma redundância normativa, na medida em que dados anônimos – sem o critério da razoabilidade – seriam sempre enquadrados dentro do conceito de dado pessoal, como aquele relacionado a uma pessoa identificável.

Essa estratégia normativa foi incorporada nas duas iniciativas legislativas que adotaram o conceito expansionista de dados pessoais, notando-se uma forte influência do direito comunitário europeu:

10. Consideranda 26: Whereas the principles of protection must apply to any information concerning an identified or identifiable person; whereas, to determine whether a person is identifiable, account should be taken of all the **means likely reasonably** to be used either by the controller or by any other person to identify the said person; whereas the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable (...)

11. Consideranda 23: "(...)To determine whether a person is identifiable, account should be **taken of all the means reasonably likely** to be used, such as singling out, **either by the controller or by any other person to identify the individual directly or indirectly**. To ascertain whether means are **reasonable likely** to be used to identify the individual, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into considera-

QUADRO 3
TABELA ANALÍTICA DA DEFINIÇÃO
DE DADOS ANÔNIMOS E/OU ANONIMIZADOS

INICIATIVA LEGISLATIVA	PLPDP/EXE	PLPDP/SEN	PLPDP/CAM
Definição de dados anônmos	**Artigo 5º, inciso V**: dados anonimizados: dados relativos a um titular que não possa ser identificado;	**Artigo 2º, inciso IV** alínea "a": anonimizados e dissociados, desde que não seja possível identificar o titular;	Inexistente
Definição de anonimização	**Artigo 5º, inciso XII**: anonimização: qualquer procedimento por meio do qual um dado deixa de poder ser associado, direta ou indiretamente, a um indivíduo;	**Artigo 2º, inciso IV,§4º**: **§ 4º** Os dados desanomizados, assim compreendidos aqueles dados inicialmente anônimos que, por qualquer técnica, mecanismo ou procedimento, permitam, a qualquer momento, a identificação do titular, terão a mesma proteção dos dados pessoais, aplicando-se aos responsáveis por sua coleta, armazenamento e tratamento o disposto nesta Lei.	Inexistente
Filtro da razoabilidade	**Art. 13**. Os dados anonimizados serão considerados dados pessoais para os fins desta Lei quando o processo de anonimização ao qual foram submetidos for revertido ou quando, com esforços razoáveis, puder ser revertido (...) **§2º** O órgão competente poderá dispor sobre padrões e técnicas utilizadas em processos de anonimização e realizar verificações acerca de sua segurança	**Artigo 3º, inciso XIV**: dado anonimizado ou anônimo: dado relativo a um titular que não possa ser identificado, considerando a utilização dos meios técnicos razoáveis e disponíveis na ocasião de sua coleta ou tratamento.	Inexistente

2.4 O QUE É RAZOÁVEL?

tion both available technology at the time of the processing and technological development. The principles of data protection should therefore not apply to anonymous information, that is information which does not relate to an identified or identifiable natural person or to data rendered anonymous in such a way that the data subject is not or no longer identifiable. This Regulation does therefore not concern the processing of such anonymous information, including for statistical and research purposes.

12. A partir dessa dupla vertente do elemento custo (financeiro/tempo) para se reverter um processo de anonimização, criar-se-ia uma métrica para definir o que venha ser um esforço razoável para se reidentificar uma base de dados anonimizadas. Ter-se-ia, então, critérios pré-estabelecidos em lei para canalizar o sentido que deve ser empregado a tal conceito equívoco (Borgesius, 2016).

Razoabilidade é um conceito equívoco que comporta uma série de interpretações. Estabelecer critérios para demarcar uma interpretação objetiva do vocábulo e/ou indicar a quem compete o seu respectivo exercício de significação são possíveis estratégias para mitigar tal exponencialidade semântica. Trata-se, em última análise, de fornecer elementos adicionais para orientar qual sentido deve ser empregada à terminologia em questão.

Ao se estabelecer uma conceituação mutualmente excludente entre dados pessoais e dados anônimos/anonimizados delimitada pelo filtro da razoabilidade, deve-se arquitetar, portanto, como e/ou a quem compete aferir a suficiência (razoabilidade) do processo de anonimização – i.e., a mensuração do que venha a ser um risco aceitável (razoável) para um eventual ataque de reidentificação (vide subcapítulo supra 1.2.3).

Para tanto, três diferentes estratégias regulatórias emergem:

I silêncio (autorregulação "total"): a legislação apenas enunciaria a razoabilidade, nada dispondo sobre critérios para a sua interpretação e/ou a respeito de um órgão regulador para a sua aplicação. Diante desse silêncio, o próprio mercado se autorregularia. Desse modo, o significado de razoabilidade corresponderia às práticas correntes do mercado sem que houvesse qualquer tipo de ingerência externa para a sua significação;

II regulação ex ante (autorregulação "parcial" ou heterorregulação "parcial"): aprioristicamente (ex ante), a legislação poderia dispor quais seriam os critérios para a aferição do que venha ser razoável, como, por exemplo, o estado da arte da tecnologia e/ou a relação custo financeiro/tempo para se reverter um processo de anonimização[12]:

II.a como os atores deveriam se comportar para fins de uma eventual autorregulação. Diferentemente da hipótese I, o mercado não seria absolutamente livre para se autonormatizar ou;

II.b como o órgão regulador deveria atuar, reduzindo a sua discricionariedade na aplicação da norma;

III regulação ex post (heterorregulação "total"): delegar a definição de tais critérios ao órgão regulador, apostando-se em uma regulação *ex post* que:

III.a inviabilizaria a autorregulação do mercado, pelo menos até que sobreviesse a definição dos critérios e a sua respectiva fiscalização pelo órgão regulador e;

III.b conferiria maior discricionariedade ao órgão fiscalizador, já que ele não estaria vinculado a critérios pré-determinados em lei para a sua atuação – diferentemente da hipótese II.b.[13]

Esse leque de opções está presente nas inciativas legislativas sob análise, que adotaram o conceito expansionista de dados pessoais e de dados anonimizado.

13. Veja-se que o que é razoável será objeto uma regulação *ex post*. Caberia ao órgão regulador definir quais os padrões e as técnicas adequadas para que os processos de anonimização sejam operacionalizados, assegurando-se um risco aceitável (razoável). O "estado da arte da tecnologia" atrelado ao custo para a execução do processo de desanonimização poderiam ser os fatores a serem levados em conta para se estabelecer uma possível métrica do que venha a ser razoável, deflagrando, em última análise, a conceituação mutualmente excludente entre dados anônimos e dados pessoais.

QUADRO 4
TABELA ANALÍTICA
DO FILTRO DA RAZOABILIDADE E SUA REGULAÇÃO

INICIATIVA LEGISLATIVA	PLPDP/EXE	PLPDP/SEN	PLPDP/CAM
Filtro da razoabilidade	**Art. 13, caput**: Os dados anonimizados serão considerados dados pessoais para os fins desta Lei quando o processo de anonimização ao qual foram submetidos for revertido ou quando, com esforços razoáveis, puder ser revertido.	**Artigo 3º, inciso XIV**: dado anonimizado ou anônimo: dado relativo a um titular que não possa ser identificado, considerando a utilização dos meios técnicos razoáveis e disponíveis na ocasião de sua coleta ou tratamento.	Inexistente
Qual é a Estratégia regulatória para a significação da razoabilidade no processo de anonimização?	**Art. 13, e §2º**: O órgão competente poderá dispor sobre padrões e técnicas utilizadas em processos de anonimização e realizar verificações acerca de sua segurança (Regulação ex post a cargo do órgão regulador que definirá os critérios e a sua fiscalização)	Inexistente (Silêncio quanto aos critérios e a sua fiscalização, ficando, totalmente, a cargo do mercado se autorregular)	Inexistente
	Heterorregulação total [14]	Autorregulação total	Inexistente

2.5 O QUE ESTÁ EM JOGO EM MEIO A TODO ESSE DEBATE TEÓRICO CONCEITUAL?

Certa vez, um engenheiro do *Google* teria dito que eles não coletam informações associadas ao nome das pessoas, pois eles gerariam desinformação – "ruído" nas palavras dele. (New York Times, 2012). Em outra oportunidade, o então chefe de assuntos de privacidade do *Facebook*, Erin Egan, afirmou paradoxalmente que apesar da rede social fornecer publicidade com base na identidade dos seus usuários, isso não significaria que eles sejam pessoas identificáveis (The Wall Street Journal, 2012).

De fato, o modo pelo qual a internet funciona não se faz necessário saber a identidade do usuário para lhe direcionar um conteúdo, ou, mesmo, sujeitá-lo a um processo de decisão automatizada. Basta lhe atribuir um identificador eletrônico único que permita separá-lo dos milhões de usuários da rede, como por exemplo, com relação ao computador a qual ele está conectado, o que é feito através do número de conexão a ele atribuído (o chamado protocolo de endereço IP: NIC.Br, 2014)[15]. A partir desse identificador eletrônico, reconhece-se o dispositivo conectado, o que permite, dentre outras coisas, a memorização dos *logins* e/ou senhas para um acesso mais dinâmico às aplicações da *web*. É dessa maneira que se melhora a experiência do usuário – mantra tão repetido nos diais atuais – que nada mais é que a formação de um perfil comportamental da sua navegação.

É possível, portanto, compilar um perfil "*browsing*" (navegação) – vernáculo na língua inglesa que define o ato de "surfar" na internet –, ainda que não se tenha certeza a respeito do sujeito que pratica tal ação. Pode ser, por exemplo, uma família ou mesmo um conglomerado de funcionários de uma empresa que fazem uso da mesma conexão, o que acaba por tornar difícil o enquadramento de tais dados como aqueles relacionados a uma pessoa identificável.

Por isso, é tão sedutor o argumento de que tais corporações não coletam dados pessoais. Por outro lado, é inconteste que tais corporações se valem desse expediente para modular a interação das pessoas na internet.

14. Tais padrões seriam cogentes e funcionariam como standard mínimo para quem anonimizasse uma base de dados. Nada impede, contudo, que o seu responsável adote padrões de anonimização mais rígidos, como um dos desdobramentos de "boas práticas" a serem objeto validação pelo órgão competente (artigo 50, §2º, do PLPDP). Nessa situação, migrar-se-ia da heterorregulação para uma corregulação, na qual o estado, na sua capacidade fiscalizatória, e fiscalizados atuariam colaborativamente. Vide: (Zanatta, 2015).

15. O Marco Civil acabou por definir o conceito de protocolo IP em seu artigo 5º, inciso III: "III - endereço de protocolo de internet (endereço IP): o código atribuído a um terminal de uma rede para permitir sua identificação, definido segundo parâmetros internacionais;"

Se a premissa da causa regulatória da proteção de dados pessoais é tutelar o cidadão que é cada vez mais exposto a tais tipos de práticas que afetam a sua vida (Mayer-Schoneberger, 1997:219-221). Então, uma compartimentalização "dura" entre dados pessoais e dados anônimos deixaria de fazer sentido. Em especial quando está em questão a formação de perfis comportamentais que tem por objetivo precípuo influenciar de alguma forma a vida do ser humano, que está atrás de um dispositivo e pouco importa ser este identificável ou não (Poullet & Dinant, 2004: 33).

Abre-se espaço, assim, para uma escolha normativa consequencialista. Não se normatiza apenas pela lente da conceituação mutualmente excludente entre dados pessoais e dados anônimos, mas, também, através da relação de causa e efeito que a mera atividade de tratamento de dados pode exercer sobre um indivíduo (Nissembaum & Barocas, 2014:45)[16].

Em jogo está, portanto, a própria eficácia da uma lei de proteção de dados pessoais que pode ter a sua causa regulatória esvaziada – a proteção do cidadão – caso viesse a prevalecer uma teorização a ela estéril entre dados pessoais e dados anônimos. Essa é mais uma das diferenças entre as iniciativas legislativas sob análise, sendo que apenas uma delas procura "tapar" tal furo ao prever que dados anonimizados estarão no escopo da lei, quando se destinarem para a formação de perfis comportamentais:

16. Os autores advertem: ainda que os indivíduos não sejam "identificáveis", eles são alcançáveis por tais práticas de mineração de dados. Inferências e outras predições podem ser estabelecidas, de sorte que tais atividades podem "bater à porta" deles, de modo que a proteção de seus dados (privacidade) deve abarcar tais situações.

17. Por uma questão de coerência com a própria definição de dados pessoais (identificado e identificável), ambos os qualificadores deveriam ser invocados nesse dispositivo. Caso contrário, abre-se margem argumentativa que toda a normativa de proteção de dados pessoais não seria aplicada aos dados anonimizados, mesmo que para a formação de perfis comportamentais em razão deles estarem atrelados a uma pessoa identificável e não a uma identificada. Quando, na verdade, o espírito da norma seria tutelar a pessoa natural, sendo ela não identificada ou não identificável. O quadro abaixo, que trata da hipótese de discriminação de preço, sinaliza que essa prática é levada a cabo pela inferência que se faz dos dispositivos conectados, de modo que o indivíduo não é identificável nesse caso.

QUADRO 5
TABELA ANALÍTICA DE UMA NORMATIZAÇÃO CONSEQUENCIALISTA DE DADOS ANONIMIZADOS

PLPDP/EXE	PLPDP/SEN	PLPDP/CAM
Art. 13. Os dados anonimizados serão considerados dados pessoais para os fins desta Lei quando o processo de anonimização ao qual foram submetidos for revertido ou quando, com esforços razoáveis, puder ser revertdo. **§ 1º** Poderão ser igualmente considerados como dados pessoais para os fins desta Lei os dados utilizados para a formação do perfil comporta mental de uma determinada pessoa natural, ainda que não identificada. [17]	Inexistente	Inexistente

XEQUE AO REI

IDENTIFICADORES ELETRÔNICOS E INTERNET DAS COISAS

Já bate a nossa porta o cenário em que todos as coisas ao nosso redor estarão conectadas à internet. Da geladeira ao carro, existirão inúmeros sensores capazes de virtualizar por completo nossas vidas e, mais do que isso, automatizá-las como, respectivamente, as compras no supermercado e por quê não a reserva e o pagamento online do vallet (estacionamento) mediante a sincronização do GPS do veículo e o local apontado (restaurante) de destino. Esse é um prato cheio para que sejam estruturados perfis comportamentais cada vez mais precisos, já que todos os nossos hábitos serão passíveis de catalogação. Isso sem falar que cada sensor permitirá uma série de inferências sobre quem os utiliza (ainda que não se trate de uma pessoa identificada ou identificável).

Nesse cenário, uma escolha normativa, que não abra espaço para que tais profilings estejam fora do escopo da lei, parece ser o mais prudente. É bem crível que o discurso de que os dados coletados, processados e armazenados

identificam apenas um dispositivo e não uma pessoa seja transposto para o cenário da internet das coisas, escancarando-se uma porta para que as pessoas atrás desses devices sejam, de alguma forma, afetadas.

Basta pensar, por exemplo, em termos de price-discrimination. Da mesma forma que consumidores acabaram por receber ofertas com precificação diferentes por conta do identificador único do seu dispositivo conectado, mediante a inferência de que usuários de Mac/Apple teriam mais condições financeiras do que de Windows (Wall Street Journal, 2012), não é difícil imaginar que donos de jaguar terão a reserva online da vaga do estacionamento com precificação maior do que de carros populares, tal como o pedido de compra vindo de uma “Brastemp” em comparação com outras mais marcas de segunda linha. Daí porque, a importância em se pensar uma normatização consequencialista, que supera o entrave teórico entre dados pessoais e dados anônimos, para proteger o cidadão toda vez que ele possa ser afetado pela mera atividade de tratamento de dados.

3
CONSENTIMENTO

3.1 INTRODUÇÃO: O PAPEL DO CONSENTIMENTO NA PROTEÇÃO DE DADOS PESSOAIS

A não ser nas primeiras leis de proteção de dados pessoais em que se procurou domesticar a tecnologia (Mayer-Schoneberger, 1997:219-221), o consentimento figura e prevalece como protagonista ao longo de todo seu ciclo evolucional. Ele surgiu como uma espécie de **carta coringa regulatória** (Tene & Polonetsky, 2011: 47) em vista da complexidade em se estabelecer um sistema completo de autorizações e proibições para regular a atividade de tratamento de dados pessoais (Rodotà, 2008:76).

Essa é a relação de causa e efeito pela qual o consentimento adquiriu tal papel normativo de centralidade, a ponto da proteção de dados ser equiparada ao direito do cidadão autogerenciar as suas informações pessoais (Solove, 2013). Uma menor ou maior carga participativa do indivíduo em autodeterminar as suas informações pessoais (autodeterminação informacional) é o que está por trás da equívoca adjetivação, mais ou menos extensa, que deve ser atribuída ao consentimento (subcapítulo 3.2) e, por fim, a releitura que dele deve ser feita em um cenário no qual a inovação e os modelos de negócios da economia digital encontram nos dados pessoais a sua base de sustentação (subcapítulo 3.3).

3.2 A EQUÍVOCA ADJETIVAÇÃO DO CONSENTIMENTO

Consentimento é um típico elemento do direito contratual. É por meio dele que os indivíduos exprimem a sua vontade de contratar, dando ciência uma a outra da sua intenção negocial para que seja selado um compromisso entre elas (Gomes, 2008). Nada mais é do que a liberdade que todo o cidadão tem de reger a sua vida (Amaral, 2008: 383), criando, modificando e extinguindo as suas relações em meio à sociedade. Essa mesma autonomia é captada e transportada para a proteção dos dados pessoais, na medida em que é o próprio cidadão quem deve governar seus dados pessoais.

A problemática gira em torno de como esse controle deve ganhar vida. Tal como na seara contratual em que o consentimento pode se vestir de várias roupagens, não é diferente quanto à proteção de dados pessoais. Essa é a questão que está por trás da adjetivação que lhe é atribuída, fornecendo-se pistas sob qual deveria ser a carga participativa exigida do cidadão para a circulação de seus dados pessoais.

Para fins de sistematização, chegou-se à seguinte **abordagem escalável** do consentimento que busca estabelecer uma análise progressiva:

I informado (básica): para que um indivíduo possa controlar seus dados pessoais, ele precisar ter ciência quanto a sua coleta, uso e compartilhamento. Ou seja, o fluxo informacional precisa tomar forma (ser informado), sendo a etapa prévia e necessária para que o consentimento comece a ser gestado (Bioni, 2014). O adjetivo informado é, portanto, a porta de entrada (pressuposto) para que o cidadão ingresse (tenha participação) dentro da dinâmica de proteção dos dados pessoais, viabilizando-se, em última análise, a racionalização de um processo de tomada de decisão a seu respeito (Lima & Bioni, 2015);

II livre (mínima): uma vez informado, o cidadão pode, então, racionalizar a sua intenção em controlar seus dados pessoais. O adjetivo livre cinge-se, assim, à concepção de um ato volitivo que deve não ser fruto de coação (física ou moral), a fim de que a autotderminação informacional seja vetorizada de forma genuína. Pense-se, por exemplo, no tratamento de dados pessoais no contexto das relações laborais em que, no mais das vezes, a situação de subordinação do empregado tende a minar a voluntariedade do consentimento, especialmente quando a não autorização em questão pode resultar em demissão (Article 29 WP, 2011: 14). Por isso, faz-se necessária uma análise casuística para se chegar à conclusão se o consentimento pode ser adjetivado ou não como livre e, por via de consequência, se o controle exercido pelo titular dos dados pessoais corresponde a uma manifestação de vontade voluntária.

III finalidade determinada (pré-intermediária): essa manifestação de vontade deve ter um direcionamento. Não pode o cidadão consentir que seus dados pessoais sejam tratados com base em propósitos totalmente genéricos, emitindo-se uma espécie de verdadeiro "cheque em branco" que esvaziaria qualquer esfera de domínio sobre seus dados (Article 29 WP 29, 2011: 17). Ao mesmo tempo, a locução "finalidade determinada" abre margem para que o ato de consentir não desça a descrições tão restritas, mas a uma leque de situações que façam sentido dentro de um determinado contexto (vide: item infra "V"). Por exemplo, em serviços de *internet banking* seria razoável que os dados pessoais dos consumidores fossem tratados não só para operacionalizar o próprio serviço em si de transferências financeiras, como, também, para prevenção de fraudes. Tais finalidades são determinadas de acordo com o escopo da relação em questão.

IV inequívoco (intermediária): esse último exemplo é ilustrativo de que de que o controle dos dados pessoais não demanda necessariamente uma ação afirmativa por parte do seu titular, mas poder ser implicitamente extraída do contexto de uma relação. O adjetivo inequívoco trabalha, assim, com a hipótese de que o consentimento pode ser tácito (InternetLAB, 2016: 94), perfazendo-se um leque de autorizações subentendido pelo contexto da relação no qual fluem os dados pessoais (Bioni, 2016) – (vide: item infra "V").

V específico e expresso (máxima): diferentemente da locução finalidades determinadas, o consentimento adjetivado pelo termo específico deveria descer a um escopo de consequências claramente e precisamente detalhado (WP 29, 2011: 24). Isso implicaria, por via de consequência, que o titular dos dados pessoais indicasse assertivamente a sua vontade quanto ao trânsito de seus dados, desembocando em uma dupla adjetivação a ser completada pelo termo expresso[18]. Haveria, assim, a carga máxima de participação do cidadão dentro da dinâmica da proteção dos dados pessoais baseada na acepção de que ele deveria seguir seus dados em todos os seus movimentos (Rodotà, 2008:17). Essa adjetivação potencializa ao extremo a concepção da autodeterminação informacional, diferenciando-se, qualitativamente, do qualificador inequívoco e da locução para finalidade determinadas, na medida em que se afasta de qualquer tipo de autorização passiva, tácita ou implícita por parte do titular dos dados pessoais (InternetLAB, 2016: 95)

18. No ordenamento jurídico brasileiro, o adjetivo expresso já se faz presentes nas leis setoriais de proteção de dados pessoais no âmbito da internet (artigo 7º, inciso VII, do Marco Civil da Internet) e no fornecimento de crédito com base em informações de adimplemento (artigo 9º, caput, da Lei de Cadastro Positivo). Em ambos os casos, exige-se o consentimento ou autorização expressa para que os dados pessoais sejam fornecidos ou compartilhados com terceiros.

QUADRO 6
ESCALADA PROGRESSIVA DA ADJETIVAÇÃO DO CONSENTIMENTO E A CORRESPONDENTE CARGA PARTICIPATIVA DO TITULAR DOS DADOS PESSOAIS

ADJETIVOS DOS CONSENTIMENTO	CARGA DE PARTICIPAÇÃO
Expresso e Específico	Máxima
Inequívoco	Intermediária
Finalidade determinada	Pré-intermediária
Livre	Mínima
Informado	Básica

Desta vez, as iniciativas legislativas sob análise diferenciam-se substancialmente.

I PLPDP/CAM: sequer adota o consentimento como estratégia regulatória central, na medida em que só se cogita a vontade do titular dos dados pessoais após o início da atividade de tratamento de dados pessoais (*opt-out*). Ou, seletivamente, para um grupo restrito (menores de idade) e, ainda, sem que haja qualquer tipo de qualificação em torno do consentimento[19];

II PLPDP/EXE vs PLPDP/SEN: ambos adotam o consentimento como estratégia regulatória central, mas com grau de intensidades díspares em razão da variação da adjetivação a ele empregada.

19. *Opt-in* x *opt-out*: o sistema opt-in exigiria uma autorização prévia do cidadão para que os dados sejam coletados, armazenados e tratados. Ao passo que, o sistema opt-out prescinde de qualquer tipo de consentimento para tais práticas, porque se entende que o indivíduo deveria manifestar a posteriori sua discordância. Ao se adotar uma sistemática (*opt-out*) que transfere a capacidade do cidadão de controlar sobre seus dados pessoais, sempre para depois deles já estarem sendo tratados, acaba-se por marginalizar a sua carga participativa. Isto porque, como visto, para que seja extraído um consentimento válido – qualquer que seja a sua adjetivação, mais ou menos extensa –, demanda-se, antes tudo, que o sujeito seja informado e, em sua carga participativa intermediária, uma análise contextual para verificar se haveria um consentimento inequívoco (implícito). O discurso binário *opt-in* e *opt-out* simplifica toda essa análise, artificializando a autodeterminação informacional a ponto de reduzir a zero a carga participativa do cidadão quanto ao fluxo de suas informações pessoais.

QUADRO 7
TABELA ANALÍTICA COMPARATIVA DA ADJETIVAÇÃO DO CONSENTIMENTO

PLPDP/EXE	PLPDP/SEN	PLPDP/CAM
Art. 5º, inciso VII consentimento: manifestação livre, informada e inequívoca pela qual o titular concorda com o tratamento de seus dados pessoais para uma finalidade determinada (...) **Art. 7º**: O tratamento de dados pessoais somente poderá ser realizado nas seguintes hipóteses: **I** - mediante o fornecimento pelo titular de consentimento livre, informado e inequívoco	**Art. 6º, inciso IV** consentimento livre, específico, inequívoco e informado sobre coleta, armazenamento e tratamento de dados pessoais, que deverá sempre ocorrer de forma destacada; (...) **Art. 12**. O tratamento de dados pessoais somente pode ser realizado nas seguintes hipóteses: **I** – mediante consentimento livre, específico, inequívoco e informado concedido pelo titular dos dados;	(***opt-out***) **Art. 15**. O Titular tem direito a autodeterminação das informações e dados pessoais prestados ou coletados, por qualquer meio.Parágrafo Único. O tratamento de dados e o envio de comunicações comerciais ou sociais é permitido, salvo se o titular solicitar o bloqueio do tratamento dos seus dados ou tiver manifestado diretamente ao respon sável pelo envio a opção de não recebê-la. (***opt-in***) **somente para menores Art. 17**. O tratamento de dados pessoais de crianças somente será possível mediante o consentimento dos seus pais, responsáveis legais ou por imposição legal
Intermediária	Máxima [20]	Nula

3.3 *BIG DATA* E A HIPÓTESE CAMALEÃO DOS INTERESSES LEGÍTIMOS: CONSENTIMENTO COMO REGRA OU EXCEÇÃO?

Toda e qualquer atividade precisaria de um espaço não previamente definido para criação, de modo que exigir um escopo inventivo pré-definido seria inviabilizá-la. Esse discurso encontra seu ápice na tecnologia do *Big Data* (Bioni, 2016: 265) que, por se tratar de uma tecnologia que permite reutilizar uma mesma base de dados para propósitos diferentes, seria incompatível com a dinâmica normativa centrada no consentimento específico (InternetLAB, 2016: 95).

Diferentemente da técnica "tradicional" de mineração de dados, o *Big Data* é uma metodologia que descarta a etapa prévia de estruturação de dados. É desnecessário relacionar os dados em entidades e atributos para minerá-los (Sampaio, 2013) – (subcapítulo 1.1.2), o que restringiria o seu uso a uma única finalidade pré-determinada por essa estruturação prévia e necessária.

Com isso, o processamento de dados se dá em um volume, velocidade e principalmente variedades maiores (os seus três famosos "Vs"). É tal progresso qualitativo e quantitativo que torna exponencial os diversos usos que se pode fazer de uma mesma base de dados, que passa a ser mutável, por não se exigir sua prévia estruturação para minerá-la.

Nesse cenário, como delimitar um único uso para os dados pessoais, se a própria tecnologia visa alargá-los, tornando-os indeterminaveis *a priori*? (Nissembaum & Barocas, 2014:60).

Nessa conjuntura, uma abordagem normativa mais flexível seria necessária, o que foi endereçado, mesmo que não voluntariamente para tais desafios mais contemporâneos, por algumas legislações ao redor do mundo através de exceções à regra do consentimento.

Previsões legais para o tratamento adicional dos dados pessoais sem qualquer tipo de consentimento ulterior do titular, como do interesses legítimos na Diretiva da União Europeia, são um bom exemplo disso.

A terminologia legítimos interesses é equívoca, o que permitiria uma série de interpretações se estas fossem realizadas de forma assistemática das demais disposições da diretiva e, sobretudo, frente à regra geral do consentimento. Ou seja, tal hipótese não se dá no vácuo, mas dentro da dinâmica de regra geral em que há uma relação pré-estabelecida na qual o titular consentiu para um uso específico ou para uma finalidade determinada de seus dados (Bioni, 2016: 265).

Contudo, quem os coletou demanda utilizá-los novamente, de modo que tal hipótese alcança o tratamento dos dados pessoais para usos secundários sem que haja um consentimento ulterior do titular (Article 29 WP 29, 2013: 21). A legitimidade dessa demanda (interesse) é aferida de acordo com a noção de compatibilidade entre o uso adicional e aquele que originou a coleta dos dados pessoais (Kuner, 2007:100).

Como dito, tal exceção não se dá no vazio, mas na conjuntura de que houve um consentimento prévio que parametriza o novo uso (adicional) dos dados pessoais. O propósito secundário e o original da coleta dos dados pessoais devem ser próximos um do outro (Article 29 WP 29, 2013: 24), demandando-se uma análise contextual da relação (vide: item "IV" do subcapítulo 3.2) para verificar se esse uso secundário está de acordo com as legítimas expectativas do titular dos dados pessoais. (WP 29, 2013: 24).

Veja-se, portanto, que há parâmetros claros para que a exceção dos interesses legítimos tome um lugar restrito na dinâmica normativa da proteção de dados pessoais (Bioni, 2016: 267). Do contrário, a regra geral do consentimento tornar-se-ia exceção, tamanha a elasticidade e as diversas facetas que a hipótese camaleão dos interesses legítimos poderia alcançar. Um sistema de freios e contrapesos deve ser arquitetado para tanto:

20. Ainda que não se tenha observado o adjetivo "expresso", o qualificador específico acaba por ser preponderante na atribuição da "carga máxima" de participação do titular dos dados pessoais. Isto porque, como dito, ao se exigir que o consentimento desça às minúcias de ser especificado de acordo com todo o escopo de consequências a que o tratamento dos dados pessoais está sujeito, reclama-se, por coerência, ações afirmativas do indivíduo que perfaçam tal esfera de controle detalhista e minuciosa. O qualificador inequívoco é capaz de se transmudar no adjetivo expresso, quando está acompanhado do termo específico, sob pena de ser desvirtuada a significação desse último. Por isso, sob o ponto de vista da técnica legislativa e com base na escala progressiva acima sistematizada, seria mais coerente normativamente a utilização do termo específico vir sempre acompanhado do termo expresso, por serem duas faces da mesma moeda que simbolizam a potência máxima da autodeterminação informacional.

XEQUE AO REI

FREIOS E CONTRAPESOS ENTRE A HIPÓTESE DE INTERESSES LEGÍTIMOS E A REGRA GERAL DO CONSENTIMENTO

A hipótese de interesses legítimos deve ser desenhada de forma cuidadosa para que o consentimento permaneça sendo a regra geral da dinâmica da proteção dos dados pessoais, sob pena de tornar, de forma incoerente, artificial a mencionada adjetivação a ele empregada.

É, nesse sentido, que as iniciativas legislativas em questão se diferenciam. Elas variam desde a inserção da hipótese dos interesses legítimos sem qualquer tipo de freio e contrapesos, até estacas menos ou mais rígidas que resultam, respectivamente, em um saldo menos ou mais estabilizador da regra geral do consentimento frente à hipótese dos interesses legítimos.

Sob o ponto de vista de coerência normativa centrada na regra geral do consentimento e da autodeterminação informacional, um sistema de freios e contrapesos mais rígido acaba por não esvaziar a promessa de que o cidadão deve exercer controle sobre seus dados pessoais.

Por exemplo: **I.** mecanismos de transparência quanto ao tratamento de dados pessoais via a hipótese de interesse legítimos, em conjunto com meios pelos quais o titular dos dados possa a ela se opor; **II.** padrões de segurança que minimizem os riscos à privacidade, como a anonimização dos dados, e, ainda; **III.** resguardar ao órgão fiscalizador o poder de auditar tais práticas do mercado, mediante a exigência de relatórios de impacto à privacidade.

Todas essas hipóteses e outras mais podem compor um arranjo normativo harmônico ao objetivo regulatório central de uma lei geral de proteção de dados pessoais, que é garantir ao cidadão o controle de seus dados pessoais. Do contrário, a hipótese guarda-chuva de interesses legítimos poderá ser super-utilizada, desbancando para o (des)controle dos dados pessoais.

QUADRO 8
TABELA ANALÍTICA COMPARATIVA DA HIPÓTESE
DE INTERESSES LEGÍTIMOS: FREIOS E CONTRAPESOS

PLPDP/EXE	
Artigo 6º, inciso I: finalidade: pelo qual o tratamento deve ser realizado para finalidades legítimas, específicas, explícitas e informadas ao titular; titular, não podendo ser tratados posteriormente de forma incompatível com essas finalidades; (...) **Art. 10**. O legítimo interesse do responsável somente poderá fundamentar um tratamento de dados pessoais, respeitados os direitos e liberdades fundamentais do titular, devendo ser necessário e baseado em uma situação concreta. **§ 1º** O legítimo interesse deverá contemplar as legítimas expectativas do titular quanto ao tratamento de seus dados, de acordo com o disposto no art. 6º, II. **§ 2º** O responsável deverá	adotar medidas para garantir a transparência do tratamento de dados baseado no seu legítimo interesse, devendo fornecer aos titulares mecanismos eficazes para que possam manifestar sua oposição ao tratamento de dados pessoais. **§ 3º** Quando o tratamento for baseado no legítimo interesse do responsável, somente os dados pessoais estritamente necessários para a finalidade pretendida poderão ser tratados, devendo ser anonimizados sempre que compatível com a finalidade do tratamento. **§ 4º** O órgão competente poderá solicitar ao responsável relatório de impacto à privacidade quando o tratamento tiver como fundamento o seu interesse legítimo.
Freio e contrapesos mais rígidos	

PLPDP/SEN	PLPDP/CAM
Artigo 4º, inciso I: coleta, armazenamento e processamento de forma lícita, com observância do princípio da boa-fé e adstritos a finalidades determinadas, vedada a utilização posterior incompatível com essas finalidades (...) **Art. 12, inciso VIII** – quando necessário para atender aos interesses legítimos do responsável pelo tratamento ou do terceiro a quem os dados sejam comunicados, desde que não prevaleçam sobre os interesses ou os direitos e liberdades fundamentais do titular dos dados.	Dispositivo similar inexistente **Art. 9º, caput**: Os dados pessoais serão tratados com lealdade e boa fé, de modo a atender aos legítimos interesses dos seus titulares.[21]
Freio e contrapesos brando	Inexistência de freios e contrapesos

21. Veja-se que, erroneamente, sob o ponto de vista da experiência estrangeira, não se estabeleceu a figura do interesse legítimo em favor do responsável pelo tratamento de dados pessoais para abrir justamente a possibilidade do uso secundário e o correspondente tratamento adicional dos dados pessoais. Mesmo assim, a inserção de tal terminologia equívoca sem parâmetros interpretativos restritivos em favor do titular dos dados pessoais, tal como aqueles existentes entre o PLPDP/EXE e o PLPDP/SEN, em conjunto com a ausência em se estabelecer o consentimento como regra geral (quadro 6) torna essa proposta legislativa a mais dispare de todas e, em total dissonância com o padrão normativo da autodeterminação informacional (subcapítulo 3.1).

3.3.1 CONSENTIMENTOGRANULAR

O modelo de negócio prevalente na Internet hoje é o chamado *zero-price advertisement business.* O consumidor não paga uma quantia monetária para acessar e utilizar serviços e produtos. Isto porque, eles são rentabilizados pelo uso de seus dados pessoais minerados para a entrega de publicidade comportamental e/ou conteúdo direcionado.

Esse arranjo econômico é, portanto, conflitante com a "livre" escolha (consentimento) da utilização de dados pessoais (vide: subcapítulo 3.2), já que o cidadão deve assentir com tal prática, sob pena de não ter acesso ao serviço ou produto. É o famoso "pegar ou largar", cujo leque de opções do processo de tomada de decisão para o controle dos dados pessoais passar a ser sufocado pela lógica binária do *take-it* ou *leave-it*.

Nesse contexto, faz-se necessário, ao menos, que as normas gerais de proteção de dados pessoais contrabalanceiam tais interesses contrapostos. Caso contrário, pode haver uma distorção da própria proteção de dados pessoais ancorada no consentimento do seu titular.

Tais modelos de negócio deveriam ser, a escolha do cidadão, menos invasivos, minimizando-se, sempre que possível, a coleta e o tratamento de seus dados pessoais. Atualmente, por exemplo, eles detêm controle apenas sobre quais informações suas serão acessíveis por outros usuários, mas não em face da própria aplicação, via de regra.

Por isso, ainda que os dados pessoais sejam indispensáveis ou necessários para tais modelos de negócio, deve-se impor certos limites. Tal como abrir espaço para o chamado consentimento "granular" (Article 29 WP 29, 2014: 22), por meio do qual o titular dos teria a opção de escolher (Article 29 WP 29, 2012: 15):

I quais os tipos de dados pessoais seriam coletados (geolocacionais, contatos telefônicos, voz, identificador do dispositivo e etc.);

II quais os tipos de tratamento seus dados pessoais (para a entrega de conteúdo direcionado, para fins de publicidade comportamental, para ativar determinadas funcionalidades de um aplicativo *mobile* e etc.);

III por quanto tempo e frequência duraria o tratamento de seus dados;

IV se haveria ou não o compartilhamento dos dados com terceiros, os chamados "parceiros comerciais" tão presentes nos temos de uso e políticas de privacidade.

O consentimento granular estabelece, portanto, limites à microeconomia dos dados pessoais, na medida em que resguarda a opção do titular em emitir autorizações fragmentadas no tocante ao fluxo de seus dados pessoais.

Desta forma, uma aplicação pode oferecer inúmeras funcionalidades, cujo funcionamento demanda, indispensavelmente, uma gama de dados pessoais para a sua operacionalização. Com a ressalva do consentimento granular, o titular poderá fazer o uso de tal aplicação, determinando, de forma correlacionada, quais dados pessoais seus serão tratados de acordo com as funcionalidades que pretende fazer uso.

Essa é uma escolha normativa que busca dar um ponto de apoio a figura do consentimento adjetivado como livre e, mais do que isso, efetivá-lo por conferir um "poder de barganha" ao cidadão que dá vazão a uma escolha voluntária de sua parte.

A presença ou a ausência desse contrabalanceamento entre a necessidade e/ou indispensabilidade dos dados pessoais, como condição para o acesso a um produto ou serviço, frente à previsão expressa de que devem ser assegurados ao cidadão meios para o controle dados pessoais a abrir espaço para o consentimento granular (Monteiro & Bioni, 2016), é mais uma das diferenças entre os projetos de lei sob análise:

QUADRO 9
CONSENTIMENTO GRANULAR

PLPDP/EXE	PLPDP/SEN	PLPDP/CAM
§ 4º Quando o consentimento para o tratamento de dados pessoais for condição para o fornecimento de produto ou serviço ou para o exercício de direito, o titular será informado com destaque sobre tal fato e sobre os meios pelos quais poderá exercer controle sobre o tratamento de seus dados. **§ 5º**: O órgão competente poderá dispor sobre os meios referidos no parágrafo anterior. § 4º.	Dispositivo similar inexistente	Dispositivo similar inexistente
Consentimento granular	Inexistente	Inexistente

3.3.2 DIREITO DE PORTABILIDADE

O consentimento somente será significativo se o usuário tiver opções, dentre os fornecedores de produtos e serviços, que lhes ofereçam melhores práticas quanto à proteção de seus dados pessoais. A privacidade pode e deve ser um elemento de competição, a fim de que o próprio titular possa se valer de tal competitividade para fazer escolhas mais significativas (Bioni & Monteiro, 2015).

Por isso, dentre os direitos do titular dados pessoais, surge a portabilidade pela qual o titular possa "trocar" de prestador de serviço ou fornecedor de produto, levando consigo seus dados pessoais.

Veja-se o notório exemplo da telefonia móvel. Ao se criar regras que facilitaram a migração dos consumidores de uma companhia para outra, criou-se um ambiente mais competitivo no qual os consumidores passaram a angariar ofertas melhores nesse segmento.

O ponto focal da portabilidade foi a atribuição de todos os ônus procedimentais às operadoras e, simetricamente, o seu exercício facilitado por parte do consumidor mediante simples requisição.

Por isso, tal como na telefonia móvel, o ônus de estabelecer a transmissão dos dados deve ser imputado aos fornecedores de produtos ou serviços, cabendo a eles garantir a interoperabilidade entre suas bases de dados para funcionalizar tal migração e, ainda, sem qualquer tipo de custo ao cidadão.

Cria-se, assim, um arranjo normativo capaz de fortalecer o direito do cidadão em controlar seus dados pessoais (CTS/FGV, 2015: 52). Isto porque, além da mera previsão de um consentimento recheado de adjetivos, uma lei geral de proteção de dados pessoais deve, concomitantemente, pensar em obrigações e deveres correlatos que funcionalizem essa prometida esfera de controle. A portabilidade insere-se nesse contexto, sendo mais uma das diferenças entre as iniciativas legislativas sob análise.

QUADRO 10
DIREITO DE PORTABILIDADE

PLPDP/EXE	PLPDP/SEN	PLPDP/CAM
Artigo 8º Art. 18. O titular dos dados pessoais tem direito a obter, em relação aos seus dados: V - portabilidade, mediante requisição, de seus dados pessoais a outro fornecedor de serviço ou produto; **§ 2º** Os direitos previstos neste artigo serão exercidos mediante requerimento do titular a um dos agentes de tratamento, que adotará imediata providência para seu atendimento.(...) **§ 4º** A providência de que trata o § 2º será realizada sem custos para o titular.	Dispositivo similar inexistente	Dispositivo similar inexistente
Direito de portabilidade	Inexistente	Inexistente

4
CONCLUSÃO

TABULEIRO DO TRIPÉ DA PROTEÇÃO DE DADOS PESSOAIS E SEUS DESDOBRAMENTOS NO JOGO DE XADREZ DAS INICIATIVAS LEGISLATIVAS BRASILEIRAS

Ao longo desse guia teórico procurou-se reportar analiticamente três pontos centrais de qualquer lei geral de proteção de dados pessoais e seus respectivos desdobramentos. Procurou-se não só mapear os referenciais teóricos em torno desse tripé, mas, igualmente, identificar como as três iniciativas legislativas brasileiras endereçaram tais questões.

As similaridades e diferenças entre os projetos de leis são, por vezes, sutis, sobretudo do ponto de vista de uma análise sistemática de como algumas questões estão relacionadas entre si. Essas variações são tão exponenciais quanto as de um jogo de xadrez.

Por isso, a conclusão desse estudo é direcionada para reduzir tal assimetria informacional, facilitando-se uma análise comparativa do tripé da proteção de dados pessoais entre tais peças do jogo de xadrez do legislativo brasileiro. A seguir há um "tabuleiro" com todas as comparações elaboradas de forma sistematizada. Xeque-mate!

QUADRO 11
TABULEIRO DO TRIPÉ DA PROTEÇÃO
DE DADOS PESSOAIS NO LEGISLATIVO

	PEÇA (TEMA) COMPARADA	PLPDP/EXE	PLPDP/SEN	PLPDP/CAM
1.	Definição de Dado Pessoal	Expansionista	Expansionista	Reducionista
1.1	Rol Exemplificativo	Sim	Não	Não
1.2	Alcança base de dados NoSQL	Sim	Sim	Não
2.	Definição de Dados Anônimos(izados)	Sim	Sim	Não
2.1	Definição de Anonimização	Sim	Sim	Não
2.2	Filtro da Razoabilidade	Sim	Sim	Não
2.2.1	Estratégia regulatória para sua definição	*Ex post* via órgão fiscalizador	Autorregulação do mercado sem critérios estabelecidos de forma *ex ante*	–
2.3	Dentro do escopo legal para fins de *profiling*	Sim	Não	–
2.3.1	Alcança possíveis práticas discriminatórias no cenário da internet das coisas	Sim	Não	–
3.	Menção ao consentimento	Sim	Sim	Sim
3.1	Consentimento como estratégia central	Sim	Sim	Não
3.1.1	Adjetivo Informado	Sim	Sim	Não
3.1.2	Adjetivo Livre	Sim	Sim	Não
3.1.3	Adjetivo Inequívoco	Sim	Sim	Não
3.1.4	Locução para finalidades determinadas	Sim	Não	Não
3.1.5	Adjetivo específico e/ou expresso	Não	Sim	Não
3.1.6	Carga Participativa do Cidadão	Intermediária	Máxima	Nula
4.	Interesses Legítimos como hipótese de tratamento de dados pessoais	Sim	Sim	Sim
4.1	A quem ele é direcionado	Responsável pelo tratamento dos dados pessoais	Responsável pelo tratamento dos dados pessoais	Ao próprio titular dos dados pessoais
4.2	Favorecer e regula o uso secundário no *Big Data*	Sim	Sim	Não
4.3	Sistema de freio e contrapesos frente à regra geral do consentimento	Sim	Sim	Não
4.3.1	Qual o nível do sistema de freio e contrapesos	Rígido	Brando	–
4.4	Consentimento granular	Sim	Não	Não
4.5	Direito de portabilidade para fortalecer o consentimento	Sim	Não	Não

BIBLIOGRAFIA

AMARAL, Francisco. **Direito civil:** introdução. Rio de Janeiro: Renovar, 2008.

Article 29 Data Protection Working Party. Opinion 02/2013 on apps on smart devices. Disponível em: <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp223_en.pdf>.

_____. **Opinion 8/2014 on the on Recent Developments on the Internet of Things.** Disponível em: < http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp223_en.pdf>.

_____. **Opinion 3/2013 on Purpose Limitation.** Disponível em: <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf>. Acesso em 02 de dezembr de 2015.

_____. **Opinion 04/2007 on the concept of personal data.** Disponível em:<http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>.

_____. **Opinion 05/2014 on anonymisation techniques.** Disponível em:<http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>.

_____. **Opinion 15/2011 on the definition of consent anonymisation techniques**. Disponível em:< http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp187_en.pdf>.

BAROCAS, Solon; NISSENBAUM, Helen. Big Data's End Run Around Consent and Anonymity. *In* **Privacy, Big Data and the Public Good** (J. Lane, V. Stodden, S. Bender, H. Nissenbaum, editores). Cambridge: Cambridge University Press, 2014. Disponível em: <https://www.nyu.edu/projects/nissenbaum/papers/BigDatasEndRun.pdf>.

BIONI, Bruno Ricardo. **Autodeterminação informacional**: paradigmas inconclusos entre a tutela dos direitos da personalidade, a regulação dos bancos de dados eletrônicos e a arquitetura da internet. Dissertação (Mestrado). Faculdade de Direito da Universidade de São Paulo. São Paulo: USP-SP, 2016.

_____, Bruno Ricardo; MONTEIRO, Renato Leite. **Será que enfim o Brasil caminha na direção de uma Lei Geral de Proteção de Dados Pessoais?** Disponível em: <http://www.privacidade.net/?p=150).

_____, Bruno Ricardo. O dever de informar e a teoria do diálogo das fontes para a aplicação da autodeterminação informacional como sistematização para a proteção dos dados pessoais dos consumidores: convergências e divergências a partir da análise da ação coletiva promovida contra o Facebook e o aplicativo 'Lulu'. **Revista de Direito do Consumidor**, v. 94, p. 283-326, 2014.

_____, Bruno Ricardo; MONTEIRO, Renato Leite. **Você trocaria de serviço online caso não pudesse levar consigo todo a sua 'vida digital'?** Disponível em: < http://www.brasilpost.com.br/renato-leite-monteiro/voce-trocaria-de-servico-_b_7283542.html>.

BORGESIUS, Frederik J. Zuiderveen, Singling Out People Without Knowing Their Names – Behavioural Targeting, Pseudonymous Data, and the New Data Protection Regulation (February 16, 2016). Disponível em: http://ssrn.com/abstract=2733115.

Centro de Tecnologia e Sociedade da Fundação Getúlio Vargas Rio. Contribuição à Consulta Pública do Anteprojeto de Lei de Proteção de Dados Pessoais. Disponível em: <http://pensando.mj.gov.br/dadospessoais/wp-content/uploads/sites/3/2015/07/5c5fb198bc34294eb44cc88dab6a0706.pdf>.

Council of Europe. **Handbook on European Data Protection Law.** Publications Office of the Europe Union: Luxembourg, 2014. Disponível em: <http://www.echr.coe.int/Documents/Handbook_data_protection_ENG.pdf>.

DEVRIES, Jennifer Valentino. VINE, Jeremy Singer. **They Know What You're Shopping For.** Disponível em: <http://www.wsj.com/articles/SB10001424127887324784404578143144132736214>.

DONEDA, Danilo. **Da privacidade à proteção dos dados pessoais.** Rio de Janeiro: Renovar, 2006.

EMAM, Khaled El. MALIN, Bradley. Appendix B: concepts and methods for deidentifying clinical trial data. Institute Of Medicine (Iom), Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk 7 (2015)

Federal Trade Commission/FTC. **Data brokers**: a call for transparency. Maio de 2014. p. i. Disponível em: <https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>. Acesso em 11 de junho de 2015.

GOMES, Orlando. **Contratos**. 26. ed. 2ª tiragem. Rio de Janeiro: Forense, 2008.

Internetlab. **O que está em jogo no Anteprojeto de Lei de Proteção de Dados Pessoais?** Disponível em: < http://www.internetlab.org.br/pt/internetlab-reporta/o-que-esta-em-jogo-no-anteprojeto-de-lei-de-protecao-de-dados-pessoais/>. Acesso em 11 de junho de 2015.

LIMA, Cíntia Rosa Pereira de. BIONI, Bruno Ricardo. A proteção dos dados pessoais na fase de coleta: apontamentos sobre a adjetivação do consentimento implementada pelo artigo 7, incisos VIII e IX do Marco Civil da Internet a partir da *human computer interaction e da privacy by default*. *In* **Direito & Internet III:** Marco Civil de Internet – Tomo I (Newton de Lucca, Adalberto Simão Filho e Cíntia Rosa Pereira de Lima coordenadores). Quartier Latin, 2015. p. 263-290.

MANNINO, Michael V. **Projeto, desenvolvimento de aplicações e administração de banco de dados**. Tradução de Beth Honorato. São Paulo: McGraw-Hill, 2008.

MATTIOLI, Dana. **On Orbitz, Mac Users Steered to Pricier Hotels**. Disponível em: < http://www.wsj.com/articles/SB10001424052702304458604577488822667325882>.

MAYER-SCHONEBERGER, Viktor. CUKIER, Kenneth. **Big Data**: A revolution will transform how we live, work and think. New York: Houghton Mifflin Publishing, 2013.

MONTEIRO, Renato Leite; BIONI, Bruno Ricardo. **Principais inovações da nova versão do Anteprojeto de Lei de Proteção de Dados Pessoais/APLPDP**. Disponível em: <http://www.privacidade.net/?p=74>.
NARAYANAN, Arvind. SHMATIKOV, Vitaly. Myths and Fallacies of "Personally Identifiable Information". **Comunications of the ACM**. Vol. 53, Nº 06, Junho, 2010. Disponível em: <www.cs.utexas.edu/~shmat/shmat_cacm10.pdf >.

_____, Arvind. SHMATIKOV, Vitaly. **Robust De-anonymization of Large Sparse Datasets**. Disponível em: < https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf>.

Núcleo de Informação e Corordenação.Br. **Como funciona a Internet? Parte 1**: **O protocolo IP**. Disponível em: < https://www.youtube.com/watch?v=HNQD0qJ0TC4>.

OHM, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization (August 13, 2009). **UCLA Law Review**, Vol. 57, 2010; Colorado Law Legal Studies Research Paper No. 9-12. Disponível em http://ssrn.com/abstract=1450006.

Hardy Quentin. **Rethinking Privacy in an Era of Big Data**. Disponível em: <http://bits.blogs.nytimes.com/2012/06/04/rethinking-privacy-in-an-era-of-big-data/?_r=0>.

ROB, Peter. **Sistemas de bancos de dados: projeto e implementação.** Tradução de All Taska. São Paulo: Cengage Leraning, 2011.

SAMPAIO, Luciana. **NoSQL, SQL e Big data**. Disponível em:<http://lucianasampaio.wordpress.com/2013/10/03/nosql-sql-e-big-data/>. Acesso em 08 de março de 2013.

SCHWARTZ, Paul M. SOLOVE, Daniel J. **The PII Problem:** Privacy and a New Concept of Personally Identifiable Information. Disponível em: <http://scholarship.law.berkeley.edu/facpubs/1638>.

SOLOVE, Daniel. J. **The digital person: technology and privacy in the information age**. New York University Press: New York, 2004.

_____, Daniel. Introduction: Privacy self-management and the consent dilemma. **Harvard law review**, 2013. Vo. 126: p. 1884. Disponível em: http://www.harvardlawreview.org/media/pdf/vol126_solove.pdf. Accesso: em 05 de janeiro de 2014. p. 1880-1903.

TEIXEIRA, Lucas. **Teoricamente impossível: problemas com a anonimização de dados pessoais.** Disponível em: < https://antivigilancia.org/pt/2015/05/anonimizacao-dados-pessoais/>.
UK Information Commissioner Office. Anonymisation: managing data protection risk code of practice. Disponível em: <https://ico.org.uk/media/1061/anonymisation-code.pdf>.

ZANATTA, Rafael. A Proteção de Dados Pessoas entre Leis, Códigos e Programação: os limites do Marco Civil da Internet. *In* **Direito & Internet III**: Marco Civil de Internet – Tomo I (Newton de Lucca, Adalberto Simão Filho e Cíntia Rosa Pereira de Lima coordenadores). Quartier Latin, 2015. p. 447-470.