

O direito à explicação no Regulamento Geral sobre a Proteção de Dados

GABRIELA CALDAS*

Resumo: Os desafios da evolução tecnológica, incluindo novas formas de tratamento automatizado de dados pessoais, impõem a adoção de regulamentação apropriada, tanto do ponto de vista da proteção dos direitos fundamentais como da preservação da inovação. O Regulamento Geral sobre a Proteção de Dados situa-se nessa zona de compromisso marcada pela importância em dispor de regras flexíveis que, protegendo o essencial, mantêm validade numa perspetiva de futuro. O reforço do papel da interpretação associado a este modelo exigirá esforço acrescido por parte dos reguladores na consideração das diferentes questões técnicas e jurídicas evidenciadas no debate sobre o direito à explicação e na identificação das melhores formas de garantir a instrumentalização da inovação digital ao serviço dos indivíduos e da sociedade em geral.

Palavras-chave: *Regulamento Geral sobre a Proteção de Dados; decisões individuais automatizadas; direito à explicação; interpretação da lei.*

Abstract: The challenges of technological developments, including new forms of automated processing of personal data, require the adoption of appropriate regulations, from the point of view of protecting fundamental rights and preserving innovation. The General Regulation on Data Protection is located in this commitment area marked by the importance in having flexible rules that, while protecting the essential, maintain the validity in a perspective of future. The reinforcement of the role of interpretation associated with this model will require greater efforts by regulators to deal with the different issues addressed in the debate on the right to explanation identifying the best ways to ensure the exploitation of digital innovation is at the service of individuals and society in general.

Keywords: *General Regulation on Data Protection; automated individual decision-making; right to explanation; interpretation of the law.*

* Doutoranda em Direito e Segurança na Faculdade de Direito da Universidade Nova de Lisboa (FDUNL). *Licence Spéciale* em Direito Europeu no Instituto de Estudos Europeus da Universidade Livre de Bruxelas. Membro do Centro de Investigação & Desenvolvimento sobre Direito e Sociedade (CEDIS).

Introdução

A revolução digital a que assistimos desde meados do século passado acelerou de forma significativa nos últimos anos em resultado dos progressos tecnológicos verificados sobretudo na área das redes neurais e da *machine learning*¹. Muito embora o impacto das tecnologias digitais não tenha ido, até agora, além da criação de novos produtos e serviços que introduzem melhorias qualitativas na nossa vida diária, certo é que o seu desenvolvimento faz antever alterações de fundo na organização da sociedade e do nosso modo de vida em geral. O potencial das tecnologias de inteligência artificial² é imenso, sendo-lhe reconhecida capacidade suficiente para gerar dentro de pouco tempo sistemas dotados de larga autonomia, aptos a substituírem a intervenção humana em praticamente todos os estádios da formação da vontade e da tomada de decisões. Daí a absoluta necessidade de salvaguardar o conjunto complexo de direitos fundamentais e princípios democráticos que serve de imprescindível pano de fundo à manutenção dos equilíbrios da nossa sociedade.

Para além das perturbações relacionadas com o emprego e a organização da economia de que se começa a falar insistentemente, importa referir que as novas tecnologias facilitam substancialmente uma vasta tipologia de meios de controle político e social, tanto por parte de autoridades públicas como de entidades privadas, sobre largas franjas da população, a custos reduzidos e com elevada eficácia. Os instrumentos utilizados, embora impercetíveis nalguns casos, podem mesmo assim ter um impacto significativo no comportamento das pessoas, levando-as não só a agir de determinada maneira, mas também a perceber a realidade e a estruturar o próprio pensamento em função da informação que lhes é canalizada.

Acresce que as novas tecnologias da informação estimulam os indivíduos a disponibilizar publicamente cada vez mais dados pessoais e permitem que empresas privadas e autoridades públicas façam uso de tais dados

¹ A *machine learning* é o ramo da inteligência artificial que explora processos de fazer com que os sistemas autónomos melhorem o seu desempenho com base na experiência.

² O conceito de inteligência artificial é aplicado a sistemas que interagem com o seu ambiente e são capazes de atuar para atingir determinados objetivos.

no exercício das suas atividades numa escala sem precedentes através de procedimentos automatizados difíceis de controlar³.

Estes fatores, aliados às divergências entre as regulamentações nacionais e à inconsistência das regras dos nossos principais parceiros externos, estão na base da atualização e reforço das normas europeias em matéria de proteção de dados que culminaram em abril de 2016, após mais de cinco anos de intensas negociações, com a adoção do RGPD⁴.

Na exposição de motivos do RGPD é desde logo reconhecida a importância em dotar a União Europeia dum quadro jurídico de proteção dos dados pessoais mais consistente e rigoroso na sua aplicação, que permita à economia digital desenvolver-se no espaço interno e aos cidadãos controlar os seus próprios dados, num ambiente generalizado de segurança para os indivíduos, operadores privados e entidades públicas.

É com este propósito de aprofundamento e facilitação das regras de proteção de dados que, nomeadamente o art. 5.º do RGPD, acrescenta aos princípios anteriormente consagrados na Diretiva 95/46/CE, o princípio da transparência em relação ao titular dos dados, o da minimização de dados e o da responsabilidade por defeito do responsável pelo tratamento dos dados.

Das salvaguardas introduzidas pelo RGPD para reforçar o controle individual sobre os dados pessoais salientam-se as que se prendem com a tomada de decisões automatizadas. Entre elas inclui-se a obrigação de fornecer “informações significativas sobre a lógica envolvida”, expressamente referida nos art. 13.º a 15.º, bem como o direito de “obter intervenção humana” mencionado no art. 22.º, a que se convencionou chamar “direito à explicação” com base na expressão “...obter uma explicação sobre a decisão tomada ...” utilizada no considerando 71.

³ BENTLEY, Peter, BRUNDAGE, Miles, HAGGSTRÖM, Olle, METZINGER, Thomas. *Should we fear artificial intelligence?*, 2018, p. 15. Disponível em <http://www.ep.europa.eu/stoa/> (acedido a 13/5/2018).

⁴ Regulamento (UE) 2016/679 relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE, p. 1–88.

Uma parte dos autores⁵ vê o “direito à explicação” como um instrumento imprescindível de promoção da responsabilidade e transparência dos sistemas de *machine learning* que o RGPD não pode deixar de contemplar dada a generalização e impacto do uso de tais sistemas, enquanto outros⁶ se interrogam em que medida o “direito à explicação”, a existir, poderá efetivamente ser posto em prática tal a falta de clareza dos mecanismos de proteção consagrados no RGPD nesta matéria a exemplo, de resto, do que acontece relativamente a outros direitos como o direito de retificação e o direito ao apagamento dos dados.

O presente artigo tem por objetivo analisar os contornos desta discussão para, no final, tecer algumas considerações sobre o papel da regulação na promoção do desenvolvimento digital em conformidade com os valores e princípios que moldam as nossas sociedades.

1. Breve análise do debate em torno do direito à explicação no RGPD

O debate gerado em torno do direito à explicação que aborda, entre outras, questões complexas sobre a interpretação dos algoritmos enquanto responsáveis pelo tratamento de dados pessoais segundo uma lógica compreensível para os humanos, tem atraído a atenção tanto da comunidade científica como de um número crescente de juristas interessados em investigar a exequibilidade do direito, bem como o valor normativo do termo “explicação” dentro do contexto da inteligência artificial.

No campo do debate jurídico as contribuições mais proeminentes provêm de dois trabalhos de investigadores de Oxford que apresentam, cada um, conclusões distintas sobre como interpretar as disposições relevantes do RGPD nesta matéria. O primeiro, de Bryce Goodman e Seth

⁵ MALGIERI, Gianclaudio, COMANDÉ, Giovanni. ‘Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation’, in *International Data Privacy Law*, v. 7, Issue 4, 2017, p. 243–265. Disponível em <https://doi.org/10.1093/idpl/ix019> (acedido a 5/5/2018).

⁶ WACHTER, Sandra, MITTELSTADT, Brent and RUSSELL, Chris. “Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR”, in *Harvard Journal of Law & Technology*, 31, 2018, p. 842-888. Disponível em <http://dx.doi.org/10.2139/ssrn.3063289> (acedido a 18/1/2019).

Flaxman⁷, conclui que o RGDPD cria um direito à explicação, enquanto o segundo, de Sandra Wachter, Brent Mittelstadt e Luciano Floridi⁸, afirma que atualmente tal direito não existe.

Bryce Goodman, do Oxford Internet Institute, e Seth Flaxman, do Departamento de Estatística de Oxford, não hesitam em considerar que, embora o texto do RGDPD não inclua qualquer referência expressa a um direito à explicação, a sua existência não pode ser questionada face às disposições do RGDPD em matéria de decisões automatizadas. A parte do artigo que se refere ao direito à explicação (a outra diz respeito à não discriminação) é, no entanto, reduzida e os argumentos são sucintos. Numa primeira versão, os autores baseiam a existência deste novo direito no considerando 71 do RGDPD que, ao afirmar que o titular dos dados, no caso de tratamento automatizado, tem sempre “...o direito de obter uma explicação sobre a decisão tomada na sequência dessa avaliação...”, seria suficientemente claro. Numa versão posterior⁹, os autores referem que o direito resulta dos art. 13.º a 15.º, onde são especificadas as salvaguardas requeridas pelo art. 22.º no que toca à fundamentação das decisões individuais automatizadas.

Questionando-se sobre o significado da expressão “informações uteis relativas à lógica subjacente” empregue nas referidas disposições, Goodman e Flaxman relembram os três tipos de obstáculos à transparência algorítmica elencados por Burrell¹⁰: i) dissimulação intencional da informação por parte do responsável pelo tratamento; ii) informação que é ininteligível para as pessoas em geral; iii) discrepância entre os modelos matemáticos gerados por *machine learning* e as formas de raciocínio humano, sendo este último o principal desafio que se põe aos investiga-

⁷ GOODMAN, Bryce & FLAXMAN, Seth. “EU Regulations on Algorithmic Decision Making and ‘a Right to an Explanation’”, *2016 ICLM Workshop on Human Interpretability in ML*, 2016, p. 1.

⁸ WACHTER, Sandra et al.. “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation”, in *International Data Privacy Law*, 7, 2, 2017, p. 76.

⁹ GOODMAN Bryce & FLAXMAN, Seth. “EU Regulations on Algorithmic Decision Making and ‘a Right to an Explanation’”, in *AI Magazine*, vol. 38, n. 3, 2017, p. 56 e 57.

¹⁰ BURRELL, J. “How the machine “thinks”: Understanding opacity in machine learning algorithms”, in *Big Data & Society*, 2016, p. 1.

dores e programadores, em especial com o aparecimento das técnicas de *deep learning*¹¹.

Sem deixar de relevar os obstáculos técnicos que se levantam atualmente, bem como as imprecisões do texto legislativo, Goodman e Flaxman insistem na importância de fundamentar decisões individuais automatizadas, sustentando que “qualquer explicação adequada deve, no mínimo, descrever como os dados recolhidos se articulam com as previsões”,¹² o que pressupõe que o modelo algorítmico deva ser claro e compreensível para um ser humano desde a sua programação inicial, para o que é necessário que o trabalho dos especialistas na matéria seja encorajado e prossiga.

O segundo artigo, de Wachter, Mittelstadt e Floridi, investigadores do Oxford Internet Institute, é antes de tudo uma forte contestação não só à existência do direito à explicação no RGPD como à sua viabilidade. Os autores não deixam de reconhecer, no entanto, que a inexistência de um tal direito constitui uma lacuna importante na proteção dos dados pessoais no caso de decisões automatizadas, para o que recomendam a adoção de uma série de medidas, incluindo legislativas, destinadas a explicitar as ambiguidades e omissões de que o atual texto padece¹³.

Para melhor compreensão da análise, o artigo distingue à partida dois significados para o termo “explicações”: o da funcionalidade do sistema de processamento e o das decisões individuais específicas. A esta distinção acresce ainda a das explicações *ex ante* e *ex post* conforme sejam fornecidas antes ou depois da tomada de decisão automatizada¹⁴.

Depois de examinar as diferentes disposições suscetíveis de fundamentar um direito à explicação no RGPD, ou seja os art. 22.º, 13.º, 14.º e 15.º, os autores concluem que nenhuma delas fornece uma base legal suficientemente forte para o que se convencionou chamar um “direito à explicação”, determinando tão só um direito limitado à obtenção de informação por

¹¹ O *deep learning* é uma forma particular de *machine learning* que envolve o uso de redes neurais com capacidade de aprendizagem e de compreensão de grandes quantidades de dados de forma não supervisionada.

¹² GOODMAN, Bryce & FLAXMAN, Seth. “EU Regulations on Algorithmic Decision Making and ‘a Right to an Explanation’”, in *ICML Workshop on Human Interpretability in ML*, 2016, p. 6.

¹³ WACHTER, Sandra et al. “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation”, *op. cit.*, p. 96 e ss.

¹⁴ *Idem*, p. 81.

parte do titular dos dados sobre a lógica envolvida no processamento dos sistemas automatizados de tomada de decisão que apelidam de “direito de ser informado”.

Esperando que novos desenvolvimentos venham a introduzir o direito à explicação num futuro próximo, enumeram três cenários que lhes parecem plausíveis nesse sentido: i) a adoção de normas de direito nacional que vão além do RGPD e que criem um direito à explicação de decisões específicas (semelhantes às medidas legislativas alemãs tomadas ao abrigo da Diretiva de 1995); ii) uma iniciativa por parte dos responsáveis pelo tratamento de dados no sentido de incluir o direito à explicação de decisões específicas no conceito “salvaguarda adequada” tal como se encontra expresso no n.º 3 do art. 22.º; e iii) finalmente, decisões judiciais que vão no mesmo sentido, isto é, que interpretem amplamente as salvaguardas do art. 22.º ou então que estabeleçam que o direito de acesso inscrito no art. 15.º integra o direito à explicação no rol dos requisitos enunciados na alínea h) do seu n.º 1.

O texto de Wachter e outros é, por sua vez, fortemente criticado num artigo de Andrew Selbst e Julia Powles¹⁵, que o qualificam de “reação exagerada” ao artigo de Goodman e Flaxman e de contribuir para “distorcer o debate” em torno do direito à explicação no RGPD.

Através de uma análise do texto e do espírito da lei, que pretendem mais construtiva e direcionada para os seus reais objetivos, Selbst e Powles concluem que as “informações significativas sobre a lógica envolvida” exigidas no caso de decisões automatizadas, não podem deixar de conter “algo como” um direito à explicação para permitir que o titular dos dados exerça efetivamente os seus direitos relativamente a decisões que o afetem, direitos que lhe são conferidos não só pelo RGPD mas que resultam dos próprios direitos fundamentais¹⁶.

Na defesa da existência deste direito os autores referem, por um lado, as obrigações dos responsáveis pelo tratamento de dados de prestar informações adicionais, em que se incluem elementos úteis sobre a lógica em causa, o significado e as consequências previsíveis para o titular dos dados, resultantes da alínea f) do n.º 2 do art. 13.º e da alínea g) do n.º 2

¹⁵ SELBST, Andrew D. and POWLES, Julia. “Meaningful Information and the Right to Explanation”, in *International Data Privacy Law*, 7, 4, 2017, p. 234. Disponível em <https://ssrn.com/abstract=3039125> (acedido a 5/5/2018).

¹⁶ *Idem*, p. 242.

do art. 14.º e, por outro, o direito de acesso conferido ao titular dos dados pela alínea h) do n.º 1 do art. 15.º.

Este entendimento é, segundo eles¹⁷, suportado pelo art. 22.º que, depois de estabelecer no n.º 1 o princípio geral de que “o titular dos dados tem o direito de não ficar sujeito a nenhuma decisão tomada exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar”, introduz no n.º 3 uma lista não exaustiva de medidas de salvaguarda dos direitos, liberdades e legítimos interesses do titular dos dados, designadamente o direito de, pelo menos, obter intervenção humana, de exprimir o seu ponto de vista e de contestar as decisões, caso estas sejam tomadas no âmbito das exceções enumeradas nos n.º 2 a 4¹⁸.

Na definição do que são as garantias adequadas, o considerando 71 introduz uma clarificação importante ao indicar que o titular dos dados tem o direito de obter intervenção humana, de manifestar o seu ponto de vista, de obter uma explicação sobre a decisão tomada na sequência dessa avaliação e de contestar a decisão. O considerando acrescenta, ainda, que “A fim de assegurar um tratamento equitativo e transparente no que diz respeito ao titular dos dados, tendo em conta a especificidade das circunstâncias e do contexto em que os dados pessoais são tratados, o responsável pelo tratamento deverá utilizar procedimentos matemáticos e estatísticos adequados à definição de perfis, aplicar medidas técnicas e organizativas que garantam designadamente que os fatores que introduzem imprecisões nos dados pessoais são corrigidos e que o risco de erros é minimizado, e proteger os dados pessoais de modo a que sejam tidos em conta os potenciais riscos para os interesses e direitos do titular dos dados e de forma a prevenir, por exemplo, efeitos discriminatórios contra pessoas singulares em razão da sua origem racial ou étnica, opinião política, religião ou convicções, filiação sindical, estado genético ou de saúde ou orientação sexual, ou a impedir que as medidas venham a ter tais efeitos...”.

Embora os considerandos não tenham natureza vinculativa certo é que assumem um papel determinante na interpretação do clausulado, como reconhecem Selbst e Pawels, que concluem, assim, pela existência de um direito à explicação no RGPD, resultante da alínea f) do n.º 2 do art. 13.º,

¹⁷ *Idem*, p. 237.

¹⁸ *Idem*, p. 236.

da alínea g) do n.º 2 do artigo 14, da alínea h) do n.º 1 do art.15.º e do art. 22.º, mesmo que a expressão não conste de nenhum destes artigos. Acrescentam que esta conclusão tem suporte na lógica subjacente ao conjunto das disposições do RGPD e nos objetivos que ele expressa de reforço da proteção dos direitos e garantias individuais no que toca ao controle dos dados pessoais, base sobre a qual o texto legislativo tem de ser interpretado¹⁹.

Esta convicção de que o RGPD deve ser visto como um todo sustentável, em que a par de direitos específicos, como o direito de acesso e o direito de retificação e de apagamento, e de mecanismos de salvaguarda, tais como a avaliação de impacto e a certificação, estão presentes importantes objetivos transversais como o de explicar e tornar os algoritmos mais responsáveis e compreensíveis que têm de ser tidos em consideração, é partilhada por Lilian Edwards e Michael Veale²⁰.

Outra questão, segundo estes autores²¹, refletida no debate sobre o direito à explicação, é a da exequibilidade das disposições do RGPD que, embora sejam claras quanto à existência legal do direito e das garantias em causa, não deixam de levantar inúmeros problemas sobre a sua operacionalidade no estado atual da evolução tecnológica. A este propósito chamam a atenção para os art. 17.º e 20.º, sobre o direito de retificação e apagamento e o direito à portabilidade dos dados, que, apesar da precisão do seu enunciado, de igual modo colocam questões de ordem técnica quanto à sua implementação.

2. A interpretação do RGPD à luz da metodologia clássica e da autorregulação

Para além da tecnologia envolvida, a questão da natureza dos algoritmos usados na tomada de decisões é, sobretudo, uma questão legal na medida

¹⁹ SELBST, Andrew D. and POWLES, Julia. “Meaningful Information and the Right to Explanation”, *op. cit.*, p. 242.

²⁰ EDWARDS, Lilian and VEALE, Michael. “Slave to the algorithm? Why a ‘right to an explanation’ is probably not the remedy you are looking for”, in *Duke Law and Technology Review*, 2017, p. 81.

²¹ *Idem*, p. 67 e ss.

em que os cidadãos têm, num Estado de Direito, de dispor de meios eficazes para preservar os direitos, liberdades e garantias de que são titulares e, se for caso disso, para contestar quaisquer decisões que os infrinjam. Os sistemas automatizados permitem hoje criar perfis e categorizar os indivíduos em níveis de alto ou baixo risco em matéria de saúde, crédito, emprego, classificá-los como potenciais infratores ou mesmo como potenciais criminosos ou terroristas, sendo que tais decisões têm impacto significativo na passagem de fronteiras, no acesso à segurança social, na assistência médica, na entrada em estabelecimentos de ensino, etc.

No entanto, o tratamento dos dados pessoais baseado em algoritmos é ainda caracterizado pela opacidade dos mecanismos operacionais e pela ininteligibilidade da lógica envolvida na preparação e obtenção dos *outputs*. A rede inextricável de algoritmos que se sucedem e se encadeiam resulta num imbróglcio de decisões obscuras, que criam incerteza e desconfiança sobre como um ambiente controlado por máquinas irá interagir com o mundo real e interpretar e avaliar os nossos comportamentos²². Daí que tenhamos de criar novos meios de proteger a liberdade individual e os direitos de cidadania, num quadro jurídico de regras previsíveis e credíveis, capaz de assegurar o funcionamento dos mecanismos de impugnação das decisões individuais na base das quais se constrói a sua legitimidade. Nas palavras de M. Hildebrandt, “sendo certo que os remédios do passado já não são eficazes, é preciso repensar e reinventar o Estado de Direito, tornando o imbróglcio algorítmico transparente e as suas decisões compreensíveis e contestáveis”²³.

Embora as opiniões divirjam sobre o conteúdo e interpretação das diferentes disposições, a doutrina tende a concordar em atribuir ao RGPD o objetivo de acionar os mecanismos de salvaguarda do direito fundamental à proteção de dados e do direito à transparência do perfil em todas as situações em que haja decisões automatizadas que afetem significativamente os indivíduos.

Em concreto, como resulta do RGPD, as decisões individuais automatizadas são sempre geradoras de obrigações e de direitos com determinada

²² HILDEBRANDT, Mireille. “The New Imbroglcio – Living with Machine Algorithms”, in *The Art of Ethics in the Information Society*, 2016, p. 5. Disponível em https://works.bepress.com/mireille_hildebrandt/75/ (acedido a 7/8/2018).

²³ *Idem*, p. 2.

especificidade. Os art^o 13.^o e 14.^o exigem que o titular dos dados, quer tenha sido ele ou não a facultar os dados pessoais, receba informações sobre a existência de tais decisões, sobre a sua lógica e sobre a importância e consequências previstas que possam ter, enquanto o art. 15.^o atribui ao titular dos dados o direito de obter confirmação sobre o tratamento dos dados pessoais e de lhes aceder, bem como a informações específicas adicionais, designadamente sobre a lógica subjacente caso haja decisões automatizadas. Por sua vez, o art^o 22.^o proíbe decisões individuais tomadas exclusivamente com base no tratamento automatizado a não ser nos casos excecionais que enumera de forma taxativa, sendo que nestes casos as salvaguardas obrigatórias incluem, no mínimo, o direito de obter intervenção humana, de manifestar o seu ponto de vista e de contestar a decisão.

A asserção de direitos e obrigações no RGPD relativamente a decisões individuais automatizadas não é incompatível com o reconhecimento dos obstáculos científicos e técnicos que se põem quanto à efetividade do seu exercício. É antes a aceitação de que todas as decisões que afetam significativamente qualquer indivíduo são passíveis de contestação, pelo que se lhes exige, primeiro que tudo, que obedeçam aos princípios básicos da fundamentação e da transparência. Se isso não for possível, ou, desde que isso não seja possível, tais decisões devem ser tidas por incompatíveis com as normas vigentes.

Em grande parte, graças à investigação e publicação de trabalhos de relevo na matéria, tem-se notado nos últimos anos uma atitude menos complacente por parte da sociedade no que toca à “prepotência” algorítmica sem que tal atitude tenha, nem era isso o que se pretendia, impedido o desenvolvimento das técnicas de *machine learning*²⁴. Nos setores mais diversos – das redes sociais à administração da justiça – é cada vez maior a presença de sistemas de *machine learning* devido à sua capacidade de melhorar exponencialmente a gestão dos dados a um custo e escala que seriam impraticáveis com recursos humanos. Esta é uma tendência que se irá acentuar, sendo unânime entre os especialistas o sentimento de que os

²⁴ CASEY, Bryan, FARHANGI, Ashkon and VOGL, Roland. “Rethinking Explainable Machines: The GDPR’s ‘Right to Explanation’ Debate and the Rise of Algorithmic Audits in Enterprise”, in *Berkeley Technology Law Journal*, 2018, p. 5 e ss. Disponível em <https://ssrn.com/abstract=3143325> (acedido a 3/6/2018).

algoritmos de aprendizagem vão desempenhar um papel cada vez maior na organização da nossa vida. Como os deveremos então regular?

A responsabilidade algorítmica assenta na transparência. Questão que se tem posto é a de saber se essa transparência deve implicar a partilha do funcionamento dos algoritmos com os indivíduos cujos dados pessoais foram objeto de tratamento, ou se, por a informação se afigurar demasiado complexa, uma explicação sobre esse funcionamento será bastante. Embora reconhecendo as dificuldades práticas da primeira opção, autoridades nacionais como o ICO têm-na defendido sempre que os algoritmos em questão afetem os direitos e liberdades dos indivíduos²⁵.

A resposta a esta questão não resulta clara do RGPD. De resto, uma das características que alguns autores lhe atribuem é a de estabelecer um quadro normativo proativo que repousa, e bem, sobre o contributo da governação digital para agilizar políticas e procedimentos capazes a cada momento de melhor assegurar os equilíbrios entre a defesa dos direitos individuais e a preservação da inovação²⁶. Parece, com efeito, razoável apontar o RGPD como um exemplo da solução Goldilocks²⁷, ou seja, de um compromisso entre a necessidade de resolver problemas específicos e a de deixar espaço para a investigação e criação.

A adoção de disposições tendencialmente flexíveis que deixem margem para acolher novos desenvolvimentos tecnológicos acarreta, no entanto, novos desafios para a interpretação e poderes acrescidos para quem as vai aplicar. Se, na verdade, a linguagem do RGPD deixa pairar alguma ambiguidade sobre conceitos como o do direito à explicação das decisões individuais automatizadas, afigura-se por outro lado importante atentar na significativa mudança introduzida pela regulamentação no sentido de ampliar e reforçar os poderes de execução conferidos às autoridades de proteção de dados, especialmente nos capítulos VI e VIII.

²⁵ *Guide to the General Data Protection Regulation (GDPR)*. Disponível em <https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/> (acedido a 9/8/2018).

²⁶ FLORIDI, Luciano. “Soft Ethics and the Governance of the Digital”, in *Philosophy & Technology*, 31, 2018, p. 3.

²⁷ Princípio segundo o qual a solução ideal se situa no meio termo. “High-Level Hearing: a European Union Strategy for Artificial Intelligence”, 27 março 2018, p. 24. Disponível em https://ec.europa.eu/epsc/events/high-level-hearing-european-union-strategy-artificial-intelligence_en (acedido em 10/5/2018).

A CE tem também incentivado a atividade de grupos de especialistas na elaboração e disponibilização de linhas de orientação em matéria de interpretação do RGPD para benefício das autoridades nacionais e do público em geral. É o caso do GT29 que, em outubro de 2017, publicou um conjunto de diretrizes sobre a tomada de decisões individuais automatizadas e a definição de perfis no contexto do RGPD²⁸.

O GT29, criado pela diretiva 95/46 CE e agora substituído pelo Comité Europeu para a Proteção de Dados, foi desde a sua constituição o principal órgão consultivo da CE em matéria de proteção de dados e de questões de segurança da informação. Embora as suas recomendações não sejam vinculativas, elas constituem um referencial para as diferentes autoridades dos Estados-Membros e são, portanto, cruciais para compreender como essas autoridades vão interpretar o RGPD²⁹.

As referidas diretrizes do GT29 relativas a decisões individuais automatizadas constituem, pois, um utilíssimo instrumento de clarificação na compreensão de alguns temas polémicos gerados em torno do direito à explicação, como sejam os da determinação do conteúdo do direito “a ser informado”, de “obter intervenção humana” e de “aceder a informações úteis relativas à lógica subjacente”.

Depois de identificar os princípios da transparência e da qualidade da informação enunciados nos art. 5.º e 12.º, como princípios essenciais para a compreensão do RGPD, o documento do GT29 sublinha, quanto ao “direito a ser informado”, que a informação sobre o tratamento dos dados e objetivos pretendidos tem de ser clara e inteligível para o respetivo titular, ou seja, entenda-se, adequada às características do interessado³⁰.

²⁸ “*Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/67*”, adotadas em 3 de outubro de 2017 e revistas em 6 de fevereiro de 2018. Disponível em http://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826 (acedido a 9/8/2018).

²⁹ Cabe agora ao Comité Europeu para a Proteção de Dados, a funcionar desde maio de 2018, desempenhar essa importante tarefa de emitir orientações sobre a interpretação dos principais conceitos do RGPD e assim contribuir para uma aplicação uniforme e coerente da legislação europeia em matéria de proteção de dados de acordo com a Comunicação da Comissão ao Parlamento Europeu e ao Conselho “Orientações da Comissão relativas à aplicação direta do Regulamento Geral sobre a Proteção de Dados a partir de 25 de maio de 2018”, COM (2018) 43 de 24/1/2018, p. 11.

³⁰ GT29, *Guidelines*, *op. cit.*, p. 16 e ss.

Relativamente ao significado de “intervenção humana” nos termos nomeadamente do art. 22.^o, refere que o mero envolvimento simbólico não é suficiente, sendo exigível que tenha um papel decisivo na tomada de decisão, para o que esta deve estar a cargo de alguém com autoridade e competência para a alterar, depois de revistos todos os fatos pertinentes³¹.

Numa ótica de dar primazia ao princípio do controle dos dados pessoais pelo titular, o GT29 não deixa, no entanto, de equacionar as dificuldades decorrentes do desenvolvimento e complexidade das técnicas de *machine learning*, nem os legítimos interesses do responsável pelo tratamento dos dados³².

Daí que o RGPD obrigue o responsável pelo tratamento a fornecer ao titular dos dados as informações sobre a lógica envolvida na tomada de decisão que sejam necessárias para que este exerça os seus direitos, mas que não o obrigue, no entanto, a uma explicação complexa dos algoritmos usados ou à sua divulgação completa. O direito à explicação pode pois não exigir a total abertura das “caixas pretas”³³, desde que o interesse individual do titular seja respeitado em todos as fases de tratamento dos dados, através de informação suficiente, atempada e adequada à sua efetiva proteção.

No sentido de facilitar o cumprimento dos requisitos legais do art.^o 22, o GT29 faz ainda algumas sugestões de boas práticas³⁴ que, no fundo, correspondem a medidas de autorregulação por parte dos responsáveis pelo tratamento dos dados a ter em consideração em todas as fases do processo, desde o *design* do modelo até à sua efetiva utilização.

Finalmente, convém sublinhar que o dispositivo do art. 22.^o não constitui para o GT29 um mero instrumento de contestação dependente da iniciativa individual, mas sim a afirmação de que são proibidas decisões individuais automatizadas que não se enquadrem em nenhuma das exceções taxativamente enumeradas³⁵.

³¹ *Idem*, p. 21.

³² *Idem*, p. 25.

³³ Conceito decorrente da ciência da computação que designa sistemas de que se conhecem apenas os dados de entrada e de saída sem acesso ao seu funcionamento interno.

³⁴ Por ex., verificações regulares da qualidade dos modelos e dos algoritmos usados e realização de auditorias independentes para garantir que os indivíduos são tratados de forma justa, correta e não discriminatória. GT29, *Guidelines*, *op. cit.*, p. 31 e ss.

³⁵ GT29, *Guidelines*, *op. cit.*, p. 19.

Esta interpretação corresponde à vontade do legislador de deslocar da esfera individual para o campo da autorregulação a efetiva proteção dos direitos conferidos ao titular dos dados pelo RGPD, com vista a minorar as inevitáveis dificuldades práticas do exercício de tais direitos se considerarmos a assimetria de recursos entre uns e outros que o RGPD não deixa de evocar.

Considerações finais

A previsão da irrupção de novas formas de inteligência artificial que, a curto prazo, vão não só melhorar a qualidade do nosso quotidiano, como transformar a realidade político-social em que vivemos, exige uma atenção especial no que toca à proteção das liberdades e garantias fundamentais do indivíduo em que se inclui o direito à proteção dos dados pessoais. O RGPD reflete a preocupação de adequar a proteção dos dados pessoais aos novos desafios da tecnologia, por forma a proporcionar ao titular dos dados os instrumentos necessários para supervisionar e controlar a utilização que deles é feita.

Entre os direitos que o RGPD atribui ao titular dos dados está o de contestar quaisquer decisões que o afetem significativamente desde que tomadas unicamente com base no tratamento automatizado dos seus dados pessoais, bem como o de obter informações úteis sobre tais decisões, a que se tem chamado direito à explicação e que resultaria, segundo alguns autores³⁶, em especial dos art. 13.º a 15.º e 22.º, como descrito anteriormente.

A polémica em torno da questão de saber se tal direito se encontra ou não inscrito no RGPD não se afigura verdadeiramente relevante.

O conhecimento dos fatos e da lei que nos rege, a possibilidade de exprimir outros pontos de vista e de contestar quaisquer decisões que nos afetem, o direito a um processo equitativo, são tudo parte integrante dos valores essenciais que o Estado de Direito tem de garantir e que nenhuma regulamentação pode desrespeitar e de que é inseparável o princípio geral da motivação das decisões, mormente das que limitem os direitos e liberdades individuais.

³⁶ GOODMAN, Bryce e FLAXMAN, Seth. “EU Regulations on Algorithmic Decision Making and ‘a Right to an Explanation’”, *op. cit.*, p.1.

A questão relevante do debate sobre o direito à explicação não é assim a de pretender que tal princípio, para ser atendível no quadro da proteção de dados, tenha de ser especificamente enunciado no RGPD.

O ponto central é que o legítimo exercício do direito à explicação ganha particular significado quando na base da decisão litigiosa se encontram tratamentos automatizados, incluindo a definição de perfis, obtidos por meio de algoritmos indecifráveis para o titular dos dados, colocando-se então a verdadeira questão de saber até que ponto a opacidade algorítmica e o receio de entravar a inovação poderão impedir o funcionamento do princípio da motivação das decisões.

Em abstrato ninguém duvida da necessidade de um direito à explicação de decisões automatizadas, a dúvida está em como o exercer enquanto os algoritmos forem inexplicáveis.

Face à imprecisão da legislação caberá à jurisprudência e aos reguladores manter um olhar atento sobre os caminhos da inovação por forma a assegurar que o desenvolvimento das tecnologias digitais não se faz em detrimento dos direitos e liberdades dos cidadãos, mas sim ao seu serviço.

No seu Parecer de 26 de julho de 2017 sobre o projeto de acordo entre o Canadá e a UE em matéria de transferência de dados PNR³⁷, o Tribunal de Justiça refere expressamente a necessidade de “prever que os modelos e os critérios utilizados no âmbito do tratamento automatizado dos dados dos registos de identificação dos passageiros serão específicos, fiáveis e não discriminatórios”, sem o que tal tratamento não será compatível com as disposições da CDFUE.

Para ser legítimo, o tratamento de dados pessoais deve ser rodeado de medidas claras e precisas que permitam que os titulares dos dados disponham de garantias suficientes para proteger eficazmente os seus dados pessoais contra os riscos de abuso, sendo que “a necessidade de dispor de tais garantias é ainda mais importante quando os dados pessoais são sujeitos a tratamento automatizado”³⁸.

Tratando-se de uma decisão judicial que se prende com a proteção dos dados pessoais num contexto de cooperação policial por maioria de razão será de supor que o Tribunal, chamado a pronunciar-se sobre a matéria no

³⁷ Parecer 1/15 do TJ de 26 de julho de 2017, para. 172. Disponível em <http://curia.europa.eu/juris/liste.jsf?pro=AVIS&num=C-1/15> (acedido a 24/10/2017).

³⁸ *Idem*, para. 141.

âmbito do RGPD, não tenha dúvidas em declarar a nulidade de decisões individuais automatizadas a que falte fundamentação consistente.

Importantes avanços no sentido de explicar a inteligência artificial estão em curso³⁹, sendo a própria CE a eleger a inteligência artificial explicável, na sua Comunicação de 25 de abril 2018 “Inteligência artificial para a Europa”⁴⁰, como uma das prioridades em matéria de investigação e inovação nesta área. Esse é, no entanto, um processo em desenvolvimento que não pode sobrepor-se à absoluta exigência de respeitar e proteger os direitos fundamentais no tempo presente, independentemente de eventuais avanços tecnológicos, como foi reconhecido por investigadores de diversas áreas na Declaração de Toronto de 16 de maio de 2018⁴¹.

Até lá, se a explicação tem ou não de passar por abrir “caixas pretas” é algo que só deverá ser avaliado em função dos legítimos interesses dos cidadãos e nunca em função dos meios técnicos disponíveis.

³⁹ O programa “Explainable Artificial Intelligence” da *Defense Advanced Research Projects Agency* tem como objetivo criar um conjunto de técnicas de *machine learning* que produza modelos mais inteligíveis, mantendo um alto nível de aprendizagem (precisão de previsão) e que permita que os utentes humanos compreendam, confiem e controlem com eficácia a geração de novos parceiros artificialmente inteligentes.

⁴⁰ COM (2018) 237 de 25/4/2018, p. 19.

⁴¹ Disponível em <https://www.accessnow.org/cms/assets/uploads/2018/05/Toronto-Declaration-D0V2.pdf> (acedido a 11/1/2019).