BIG DATA

NA INVESTIGAÇÃO CRIMINAL

DESAFIOS E EXPECTATIVAS
NA UNIÃO EUROPEIA

Laura Neiva

BIG DATA BIG

Laura Neiva é licenciada em Criminologia, pela Faculdade de Direito da Universidade do Porto (2017) e mestre em Crime, Diferença e Desigualdade, pelo Instituto de Ciências Sociais da Universidade do Minho (2019). É atualmente investigadora no Projeto "Exchange" – Forensic Geneticists and the Transnational Exchange of DNA data in the EU: Engaging Science with Social Control, Citizenship and Democracy, financiado pelo Conselho Europeu de Investigação (Contrato N.º [648608]), liderado por Helena Machado e sediado no Centro de Estudos de Comunicação e Sociedade (CECS) da Universidade do Minho.

A sua área de investigação centra-se no fenómeno do *Big Data* enquanto potencial técnica de investigação criminal, abrangendo o estudo das suas questões éticas, sociais, políticas e de direitos humanos. Procura deste modo compreender como é que as expectativas em torno do *Big Data* podem ter impacto na aplicação da lei, na justiça criminal e nas agências policiais.

BIG DATA

NA INVESTIGAÇÃO CRIMINAL

DESAFIOS E EXPECTATIVAS
NA UNIÃO EUROPEIA

Laura Neiva

húmus

***BIG DATA* NA INVESTIGAÇÃO CRIMINAL:**
DESAFIOS E EXPECTATIVAS NA UNIÃO EUROPEIA

Autora: Laura Neiva

Capa: António Pedro

Revisão e paginação: Margarida Baldaia

Apartado 7081
4764-908 Ribeirão – V.N. Famalicão
Telef. 926 375 305
humus@humus.com.pt

Impressão: Papelmunde, SMG, Lda. – V. N. Famalicão
1.ª edição: julho 2020
Depósito Legal: 471775/20
ISBN: 978-989-755-523-7

Este livro foi realizado com o apoio financeiro e de acesso a dados decorrentes de investigação realizada no âmbito do projeto "Exchange" – Forensic Geneticists and the Transnational Exchange of DNA data in the EU: Engaging Science with Social Control, Citizenship and Democracy, financiado pelo Conselho Europeu de Investigação (European Research Council) sob o programa de pesquisa e inovação da União Europeia Horizonte 2020 (Contrato N.º [648608]), liderado por Helena Machado no âmbito de uma Consolidation Grant e sediado no Centro de Estudos de Comunicação e Sociedade, Instituto de Ciências Sociais da Universidade do Minho (Portugal).

ÍNDICE

Índice de figuras **7**

Prefácio **9**
Helena Machado

Agradecimentos **11**

Introdução **13**

Contextos teóricos orientadores 15

Organização dos capítulos 17

PARTE I. O *BIG DATA*: UM MUNDO HETEROGÉNEO **21**

Capítulo 1. Mapeamento de conceitos, práticas e teorias **23**

1.1. Desenho conceptual: o que é o *Big Data*? 23

1.2. Desenho operacional: como se materializa o *Big Data*? 24

Capítulo 2. O rosto das revelações: Edward Snowden **27**

2.1. Desafios éticos e de direitos humanos 30

Capítulo 3. O *Big Data*: uma (nova) forma de exercer a vigilância? **39**

3.1. As novas tecnologias e a nova vigilância 41

3.2. Reconfigurações sociais: o que mudou? 43

PARTE II. *BIG DATA* E INVESTIGAÇÃO CRIMINAL **47**

Capítulo 1. Notas histórico-espaciais do *Big Data* na investigação criminal: uma realidade antiga? **49**

1.1. Projeto *Total "Terrorism" Information Awareness* 51

1.2. Sistema Echelon 51

1.3. Eurodac 52

Capítulo 2. Estudos empíricos **53**

2.1. Los Angeles 53

2.2. Austrália 55

2.3. França 56

2.4. Europol 57

PARTE III. TRAÇAR EXPECTATIVAS: AS NARRATIVAS SOBRE O *BIG DATA* **61**

Capítulo 1. Considerações metodológicas **63**

1.1. Perfil dos participantes 66

Capítulo 2. Lente conceptual analítica: Sociologia das Expectativas **69**

Capítulo 3. Traçar expectativas: resultados empíricos **75**

3.1. (Des)conhecimento e descrição técnico-profissional 75

3.2. Desenvolvimento, expansão e antecipação de aplicação do *Big Data* na investigação criminal 94

3.3. Perceção dos riscos e perigos do *Big Data* 105

3.4. Pareceres éticos e de direitos humanos 114

CONCLUSÃO **125**

Bibliografia **137**

ÍNDICE DE FIGURAS

Figura 1. Definição conceptual e operacional do *Big Data* 28

Figura 2. Transformações que operam na vigilância com o surgimento do *Big Data* 48

Figura 3: Relação entre *Big Data* e criminalidade 61

PREFÁCIO

Helena Machado

O presente livro é um retrato rico e desafiador da atual sociedade digital e da informação. Explorando o tema dos "Grandes Dados" (*Big Data*) e das expectativas sobre a sua aplicação na prevenção e governabilidade da criminalidade, convida o leitor a refletir sobre o poder da tecnologia na segurança e no controlo social.

O que é o *Big Data*? Quais as suas significações culturais e simbólicas? Que riscos e benefícios são associados à recolha massiva de dados sobre os cidadãos? Em que esferas da vida em sociedade são aplicadas as técnicas do *Big Data*? Que desafios e ameaças traz à privacidade dos cidadãos? Estas e outras questões imprescindíveis à compreensão geral do fenómeno dos "Grandes Dados" são mapeadas e respondidas de modo sistemático e organizado na primeira parte desta obra. A autora socorre-se de literatura académica das ciências sociais para delinear os principais conceitos e problemáticas associados a este fenómeno, apresentando exemplos concretos que facilitam a compreensão por parte do leitor e instigam a curiosidade e imaginação.

Numa segunda parte, este livro apresenta testemunhos fascinantes relacionados com expectativas em torno das potencialidades das técnicas do *Big Data* no controlo da criminalidade. Com base em 124 entrevistas realizadas em 25 países europeus, Laura Neiva desenvolveu uma aprofundada análise de expectativas em torno de cenários em que as técnicas do *Big Data* apoiam a atividade das autoridades policiais e agentes securitários, seja para prever o crime, seja para apoiar a investigação criminal.

As narrativas analisadas permitem responder a várias interrogações: O que é mais marcante em termos de expectativas sobre o *Big Data*? Que futuro se projeta relativamente ao seu uso na promoção da segurança

pública? Como é que o *Big Data* pode vir a mudar as práticas policiais e a investigação criminal?

A temática tratada neste livro envolve, de modo inquestionável, importantes questões éticas, designadamente as seguintes: Que tipo de informação se torna mais problemático recolher e inserir em grandes bases de dados? Que direitos humanos são protegidos ou ameaçados? De que forma os "Grandes Dados" podem ajudar na proteção de cidadãos tidos como "cumpridores da lei" e, simultaneamente, fragilizar os direitos cívicos de determinados grupos sociais, étnicos e raciais? Será que a utilização do *Big Data* vai perpetuar a tendência de criminalização de comunidades mais vulneráveis às malhas de suspeição e estigmatização do sistema de justiça criminal?

Este livro interessa a especialistas em investigação criminal, a cientistas sociais e juristas, a estudantes de diversas áreas, e a qualquer cidadão que tenha interesse no modo como a sociedade digital, na sua avidez pela recolha e circulação veloz de dados, pode gerir a criminalidade através de meios tecnológicos. Trata-se de uma obra pioneira no contexto português, que sem dúvida irá marcar um campo de estudos e de investigação científica.

Tive o privilégio de orientar a dissertação de mestrado que deu origem a este livro. Fui testemunha direta do empenho e talento de Laura Neiva. A sua maturidade e compromisso ético com a investigação faz prever um futuro muito promissor como académica profissional.

Este livro representa uma primeira incursão da autora na investigação científica, ao mesmo tempo que deixa a promessa de que muitos outros trabalhos marcantes e originais se lhe seguirão. Assim, a presente obra simboliza os motivos pelos quais devemos depositar esperança e otimismo em relação à mais nova geração de cientistas sociais em Portugal.

Braga, setembro de 2019

AGRADECIMENTOS

Este livro será, para sempre, meu e de todos aqueles que, junto a mim, o sonharam comigo. A eles, o meu sincero obrigado:

Ao Projeto "Exchange" – Forensic Geneticists and the Transnational Exchange of DNA Data in the EU: Engaging Science with Social Control, Citizenship and Democracy, financiado pelo Conselho Europeu de Investigação (European Research Council) sob o programa de pesquisa e inovação da União Europeia Horizonte 2020 (Contrato N.º [648608]), liderado por Helena Machado no âmbito de uma *Consolidation Grant* e sediado no Centro de Estudos de Comunicação e Sociedade, Instituto de Ciências Sociais da Universidade do Minho (Portugal), pelo apoio financeiro e pelo acesso aos dados que permitiram explorar a temática.

À Professora Doutora Helena Machado, por ter criado condições para que este livro fosse hoje uma realidade palpável, antes de ter sido um sonho muito desejado; por ter acompanhado de perto, vigilante, atenta e confiante, este percurso de crescimento pessoal, académico e profissional; pelos conselhos científicos, inspirações teóricas e recomendações imprescindíveis; pela confiança e motivação. Muito obrigada.

A todas as (minhas) colegas de equipa: Rafaela, Sheila, Nina, Marta, Sara, Filipa e Alícia, pela força que são e pelos pilares essenciais que foram na construção deste livro. A ajuda, o sorriso, a prontidão, a humildade, a aprendizagem, a vontade, os ensinamentos, a experiência e tudo aquilo que nem todas as palavras que existem poderiam dizê-lo. A todas, em geral, e a cada uma, em particular: muito obrigada, por tudo.

Às pessoas mais importantes da minha vida: Mãe, Pai, Flávia, Mahé e Bruno. As palavras nunca reproduzem com exatidão o que o coração sente. Mas eu sinto-vos e vocês sentem-me. A vossa existência embeleza todos os caminhos que percorro, sempre de mão dada a todos. Vocês serão sempre o destino dos meus sonhos. Obrigada por acolherem mais um. Muito, muito obrigada.

A todos os meus amigos, que são como família: Mónica, Tiago, Neto, Clara e todas as outras amizades que criei e solidifiquei ao longo deste caminho, onde me foram ouvindo, ajudando e voando comigo, sempre à procura de mais e melhor. Obrigada.

E a todas as outras pessoas com quem tive a sorte de me cruzar e o prazer de estar, obrigada. Por me darem tempo e espaço para ser. Fica o toque de todos os que me tocaram com histórias, conselhos, mensagens de ânimo, sorrisos e lágrimas. A vida é uma viagem, este foi um dos melhores destinos onde estive. Obrigada!

INTRODUÇÃO

Este livro pretende contribuir para o debate contemporâneo que se tem aprofundado em torno dos grandes dados – o *Big Data* – e dos seus potenciais benefícios e riscos para fins de investigação criminal. Atualmente, circulam diferentes visões sobre esta temática, maioritariamente difundidas pelos meios de comunicação social – importantes instrumentos de construção de sentidos e significações em volta dos mais diversos assuntos. Consequentemente, urge um interesse e uma necessidade académicos de criar condições para que se possa impulsionar e abrir um discurso – onde não existe um discurso efetivo agora – sobre as diferentes perspetivas, temporalidades, especialidades e materialidades que representam as bases de dados, os seus objetivos e as suas finalidades. Assim, esta obra afirma-se com o objetivo de projetar a máxima flexibilidade e permitir o possível para uma polifonia emergente.

A obra assenta numa investigação sobre o *Big Data* enquanto mecanismo promissor na vigilância e previsão de risco no combate à criminalidade. O aparato tecnológico desta técnica engloba a vigilância e observação indireta de indivíduos e a recolha de um elevado número de informações potencialmente convertidas em algoritmos. Com vista à produção de classificações numéricas e categorizações, tem como objetivo conceber ações de prevenção e repressão da criminalidade. Este livro apresenta o estudo deste fenómeno aplicado à segurança pública e policiamento transnacionais, na União Europeia.

A metodologia é qualitativa e ancora-se na análise de entrevistas realizadas a profissionais inseridos, laboralmente, nos sectores de segurança pública e policiamento transnacionais. Não obstante, inclui também a análise qualitativa das narrativas de profissionais da genética forense, investigadores académicos, membros de Organizações Não Governamentais e demais profissionais com

perfil relevante no que concerne à temática. Analisando minuciosamente as expectativas sociais em torno destas técnicas por parte dos pontos de contacto nacionais em rede transnacional de cooperação policial e judiciária, bem como geneticistas forenses e *stakeholders* de diferentes áreas (ética e regulação, investigação criminal, pesquisa universitária, empresas privadas e organizações não governamentais), visou-se explorar a forma como esta técnica é perspetivada no contexto da partilha transnacional de dados. Os resultados permitiram explorar as facetas sociais e culturais do *Big Data*, bem como identificar e contribuir para o debate contemporâneo acerca das questões ligadas à privacidade e proteção de dados.

Desta forma, pretende-se estimular a reflexão crítica sobre este tema e agitar a consciencialização pública, promovendo o debate em torno dos novos desafios que a União Europeia enfrenta, num contexto digital de partilha e cedência de dados permanente, embutido numa realidade que defende, cada vez mais, a privacidade e a proteção de dados. O objetivo é (re)pensar as formas como as expectativas sociais constroem novos caminhos nas rotas pré-definidas da investigação criminal europeia, dando-lhe novos sentidos; e a forma como o *modus operandi* do *Big Data* se matura na sociedade contemporânea. Urge debater este fenómeno: compreendê-lo, defini-lo, desmistificá-lo e desconstruí-lo. O *Big Data* não é uma realidade nova, é um paradigma sofisticado que encontra, cada vez mais, novos lugares para ser e se tornar técnica. Será a investigação criminal um meio propício? Antes de permeabilizar os contextos à inclusão do *Big Data*, não se deveriam estudar os meios promissores à penetração das redes de dados digitais? Está a sociedade preparada para assistir a um crescente desenvolvimento destes dados?

Esta obra pretende compreender, antes de concluir, o modo como os profissionais de segurança pública, genéticos forenses, professores universitários e demais pessoas alocadas a contextos relevantes perspetivam o *Big Data* no seu seio laboral. Como entendemos estas mudanças diárias nos grandes dados? Serão eles úteis à investigação criminal europeia? Como operariam na prática? Os estudos indicam que nos Estados Unidos da América esta é uma técnica promissora para o combate à criminalidade. E na Europa? O objetivo é conhecer a extensão do tema e questionar, operar numa reflexão crítica sociológica enriquecedora, que possa responder a questões existentes, mas que seja capaz de equacionar novas perguntas. O importante é (re)pensar.

A inexistência de outro trabalho de investigação equiparável a este converte esta obra numa contribuição única nacional e europeia para uma (melhor) compreensão do fenómeno *Big Data,* enquanto realidade social contemporânea

e enquanto potencial técnica de investigação, prevenção e repressão da criminalidade organizada e transfronteiriça. Privilegia-se aqui a ótica dos profissionais que lidam, laboralmente, com a partilha transnacional de dados na União Europeia, geneticistas forenses, juristas, professores/investigadores, investigadores criminais, profissionais que trabalham em empresas especializadas em proteção de dados e privacidade e demais pesquisadores com perfis diferenciados com conhecimentos relevantes no que concerne ao *Big Data* e à partilha transnacional de dados de ADN. Trata-se de um estudo pioneiro e único na área, que procura contribuir para o debate académico atual em torno dos grandes dados e das suas implicações éticas e de direitos humanos numa sociedade democrática.

CONTEXTOS TEÓRICOS ORIENTADORES

A literatura tem documentado que na última década dois grandes desenvolvimentos estruturais ocorreram em simultâneo: a vigilância quotidiana e os *grandes dados* (Brayne, 2017). Tudo começa com um objeto de atenção descrito como *mundo digital* ou *mundo dos dispositivos de geração de dados online* – o *Big Data*. Esse mundo é descrito como fluido, móvel e em permanente expansão e desenvolvimento. Consequentemente, está hoje cada vez mais omnipresente, instalando-se como um processo de rotina. E isso verifica-se, por exemplo, na proliferação de dados sociais transacionais que agora são armazenados e analisados diariamente (Frade, 2016; Youtie, Porter & Huang, 2016). Vivemos, desta forma, numa sociedade dinâmica que se move contínua e rapidamente numa única direção: para um crescimento e expansão cada vez mais rápidos, rumo ao futuro, mas um futuro sem promessa senão a continuação acelerada do presente. O tempo é assim definido: o tempo do digital, de comunicações em rede, com a promessa de ser cada vez mais eficaz e eficiente (Frade, 2016). Tudo isto acontece paralelamente a um processo de globalização que requer um alto nível de controlo sobre o espaço físico e o tempo, resultando na aceleração do ritmo de vida e das temporalidades organizacionais – e isso tem consequências. Assim, é importante debruçarmo-nos sobre a tecnologia, pois esta é considerada um grande recurso na compreensão do tempo e dos espaços – duas variáveis determinantes na compreensão dos fenómenos sociais e humanos (Araújo, 2008).

A datificação – a conversão de toda a informação em dados analíticos – tornou-se um paradigma aceite para compreender o comportamento social e a sociedade. Com o advento dos *sites*, das redes sociais e com o aperfeiçoamento

de todo o arsenal tecnológico, vários aspetos da vida social começaram a ser quantificados como nunca antes tinham sido. Assim, a transformação da sociedade num meio digital criou uma indústria de dados que os converteu em algo precioso, capaz de se inserir em variadas áreas (Van Dijk, 2014), sustentada na ideia de que a tecnologia permite a economização do tempo, contribuindo para ganhos no tempo-espaço e reforçando a sua utilização e expansão (Araújo, 2008). As Ciências Sociais foram uma das áreas que, nos últimos anos, se tornaram maleáveis a uma mudança na vanguarda da análise científica social (Haldford & Savage, 2017). Consequentemente, o Sistema de Justiça Criminal (SJC) também se foi expandindo e crescendo, à medida que a digitalização em massa de informações se foi tornando uma possibilidade crescente (Brayne, 2017). Por isso, importa explorar o contexto onde as realidades e existências emergem – atualmente entrelaçadas com o Capitalismo –, porque as capacidades tecnológicas nunca operam sozinhas. Concretamente, o *Big Data* não se *autogerou*, mas surgiu como fruto do c*apitalismo do conhecimento*, como resposta ao desenvolvimento sem precedentes do mundo tecnológico (Frade, 2016), penetrado pelas tecnologias da informação e comunicação, pela capacidade cada vez maior de fintar distâncias e eliminar tempos, através da velocidade de circulação e da híper-imposição do espaço virtual (Araújo, Cogo & Pinto, 2015). Todo este aparato móvel, digital, fluido e anónimo leva a que se debata a imersão do conhecimento científico neste palco interativo tecnológico e o surgimento de conhecimento científico a partir desta realidade informática. Neste contexto, uma das inúmeras facetas desta revolução digital é, particularmente, o surgimento do conceito de *dados* (Drewer & Miladinova, 2017). A forma como a sociedade adota, rejeita, usa e modifica estes grandes dados e as novas tecnologias pode provocar alterações nas relações de poder, na conceção de novas identidades e na produção de desigualdades (Selin, 2008).

O contexto europeu também sofreu modificações, algumas paralelas às descritas anteriormente, outras a elas contrárias. No entanto, todas estas mudanças se influenciam mutuamente e convergem para produzir alterações sociais. Atualmente, a Europa está sob ameaças e ataques terroristas, e este arsenal tecnológico e digital influencia os valores fundamentais defendidos pela legislação europeia: direitos, liberdades e garantias. Assim, intensifica-se o desafio de estabelecer o equilíbrio entre a proteção destes valores fundamentais e a capacidade de combater e prevenir os ataques e as ameaças terroristas. Trata-se de um desafio simultaneamente político, religioso, militar e social. Face a este dilema, atualmente, observamos mudanças que surgem como resposta às consequências nefastas destes acontecimentos. Por exemplo, o paradigma de

atuação das agências policiais evoluiu de uma guerra tradicional ao terrorismo para uma guerra mais tecnológica, que vai além das políticas de segurança já existentes. No entanto, estas soluções tecnológicas, mais uma vez, levantam outros desafios ao nível das liberdades, privacidade e proteção de dados. Desta forma, afigura-se imprescindível alcançar o equilíbrio entre a privacidade individual e a segurança social na Era do Mercado Digital (Gonçalves, 2017; Drewer & Miladinova, 2017).

Todos estes conceitos, acontecimentos, teorias e factos convergem num círculo vicioso, que surge como difícil de quebrar, mas que, agitando a nossa consciência, deve soltar a nossa voz. Mergulhados numa realidade digital e tecnológica, estes conceitos serão entendidos e explorados à luz de teorias sociológicas e criminológicas. Pretende-se explorar paradigmas como os da segurança, da proteção de dados, da vigilância e da privacidade, e debater a complexidade das temáticas à luz das expectativas sociais. Além disso, exploram-se e apresentam-se estudos empíricos acerca destes temas – *Big Data,* dados, investigação criminal, vigilância, segurança e direitos humanos – em contexto não europeu. Estuda-se o futuro, no presente – contexto europeu, acontecimentos sociais contemporâneos e realidades atuais.

ORGANIZAÇÃO DOS CAPÍTULOS

A obra encontra-se estruturada em quatro partes. A primeira, com o título *O Big Data: um mundo heterogéneo*, condensa três fragmentos teóricos importantes na definição, delimitação e estruturação do tema. É feita uma contextualização do objeto de estudo *Big Data*, como forma de enquadramento do mesmo, explicação da sua relevância contemporânea e apresentação dos debates sociais e mediáticos emergentes. Posteriormente, uma definição conceptual e operacional do mesmo permite apresentar o modo como o objeto será abordado. Nesta secção encontram-se as definições consensuais da literatura científica relevante para o tema em apreço. Em terceiro lugar, descreve-se sumariamente o fator que impulsionou as controvérsias em torno do tema – o caso de Edward Snowden – e as revelações que determinaram a forma como se aborda o fenómeno e condicionam, consequentemente, o seu estudo. É feita uma análise da evolução do caso e das reações a este – desde artigos e notícias de imprensa a documentários –, procurando explorar em que condições as revelações operaram, que efeitos produziram e que consequências tiveram. As revelações de Snowden foram determinantes para a instauração de um novo paradigma de estudo e investigação. Posteriormente, surge uma reflexão acerca das questões

éticas e de direitos humanos do *Big Data*. Estas questões, inseridas sob o contexto legal europeu, são debatidas e problematizadas. Na terceira secção, surge o enquadramento temático do *Big Data* como meio de exercer a vigilância. São apresentados estudos acerca das transformações observadas na vigilância tradicional após a emergência dos grandes dados, inseridos numa Era Digital. Confere-se particular relevância ao processo de datificação e digitalização das técnicas tradicionais de investigação criminal, descrevendo as mudanças e processos contextuais, sociais e políticos que permitiram criar as condições para a emergência e expansão do *Big Data*.

A segunda parte – *Big Data e investigação criminal* – visa abordar a relação entre *Big Data* e investigação criminal, isto é, a técnica enquanto ferramenta promissora no combate e prevenção da criminalidade (organizada e transfronteiriça). Nesta secção apresentam-se as conexões identificadas pela literatura entre as duas temáticas e os estudos empíricos realizados em Departamentos Policiais. Apesar de escassos, estes estudos revelam conclusões antagónicas: em determinados Departamentos Policiais, o *Big Data* não é uma técnica usada, mas reside no imaginário policial e está em desenvolvimento; por outro lado, noutros Departamentos, o *Big Data* encontra-se em fase precoce de implementação, provocando mudanças no quotidiano policial e/ou reforçando técnicas tradicionais de investigação policial.

A terceira parte da obra – *Traçar expectativas: as narrativas sobre o Big Data* – constitui uma súmula da recolha de dados e da sua exploração, interpretação e atribuição de significações. É nesta secção que se inserem todas as sintetizações das informações e conclusões obtidas na análise das entrevistas. Em primeiro lugar, após a sintetização das escolhas metodológicas e apresentação dos participantes no estudo, é explanada a lente conceptual analítica – Sociologia das Expectativas – que serviu de base à análise, compreensão e interpretação dos discursos dos entrevistados. As Ciências Sociais têm enfatizado a importância de se estudarem as expectativas acerca dos desenvolvimentos técnico-científicos com o objetivo de nortear a sua expansão e legitimação. Em seguida, apresentam-se os dados empíricos e as diferentes expectativas dos profissionais de cooperação transnacional, profissionais laboralmente conectados e inseridos na partilha transnacional de dados na União Europeia, geneticistas forenses, juristas, professores/investigadores, investigadores criminais, profissionais que trabalham em empresas especializadas em proteção de dados e privacidade e demais pesquisadores com perfis diferenciados com conhecimentos relevantes no que concerne ao *Big Data* e à partilha transnacional de dados de ADN enquanto técnica promissora de prevenção e repressão da criminalidade.

Destaca-se a pluralidade e diversificação de expectativas de cada profissional, variando consoante a sua posição laboral e formação académica. Tal comprova que o meio contextual dos profissionais e o seu histórico formativo influenciam a sua visão sobre os fenómenos, condicionando os seus discursos. De uma forma geral, os dados destacam que a maioria dos profissionais entrevistados possui expectativas promissoras quanto à implementação do *Big Data* como ferramenta de investigação criminal, surgindo, no entanto, nuances distintas e características de cada grupo profissional. De forma a uniformizar e simplificar a complexidade dos resultados obtidos foram criadas temáticas gerais, constituindo categorias analíticas que condensam as diferentes perspetivas consideradas: i) (des)conhecimento e descrição técnico-profissional; ii) desenvolvimento, expansão e antecipação de aplicação do *Big Data* na investigação criminal; iii) perceção dos riscos e perigos do *Big Data*; e iv) pareceres éticos e de direitos humanos.

Por fim, a *Conclusão* reúne os pontos mais importantes deste caminho de produção de conhecimento científico. Este capítulo é um *lugar* de reflexão sobre o percurso realizado, um *espaço* de pensamento e reconsideração sobre as diferentes opções teóricas e metodológicas adotadas e um *intervalo de tempo* onde se podem analisar as vibrações emergentes das diferentes perspetivas dos profissionais, enquanto teias complexas que caracterizam as semelhanças e dissemelhanças que surgem entre os discursos.

PARTE I

O *BIG DATA*: UM MUNDO HETEROGÉNEO

CAPÍTULO 1. MAPEAMENTO DE CONCEITOS, PRÁTICAS E TEORIAS

1.1. DESENHO CONCEPTUAL: O QUE É O *BIG DATA*?

O *Big Data* refere-se a conjuntos de dados que são recolhidos, analisados, convertidos em algoritmos[1], categorizados numericamente e identificados por via de um índice para, posteriormente, extrair informação que informe e oriente as políticas criminais (Gonçalves, 2017; Chan & Moses, 2017; Lyon, 2014; Drewer & Miladinova, 2017; Lefèvre, 2017; Matzner, 2016; Wood, Ball, Lyon, Norris & Raab, 2006). O *Big Data* agrega e trata enormes conjuntos de dados, trabalhando com um volume infinito de informações. Estes dados são tratados e analisados a uma velocidade feroz, em tempo real; são exaustivos e abrangentes, detalhados e partilhados (Gandy Jr, 1989; Lyon, 2014; Youtie *et al*., 2016; Gonçalves, 2017). As fontes destes dados podem ser diversas: i) dirigidas (por via dos circuitos de câmaras de vigilância, por exemplo); ii) automatizadas (localização geográfica detetada pelo telemóvel, por exemplo); e iii) voluntárias (por exemplo, cedência de dados pessoais por parte dos indivíduos) (Lupton & Michael, 2017; Chan & Moses, 2015; Lyon, 2014).

Apesar de não haver consenso na literatura acerca da definição da técnica, geralmente o *Big Data* envolve (pelo menos) três considerações: o grande volume de dados que agrega; a variedade e a velocidade com que estes são processados; e o seu formato e estrutura (Chan & Moses, 2015). Noutra esteira, Boyd e Crawford (2012) apresentam uma definição do *Big Data* enquanto *fenómeno cultural, tecnológico e académico* que resulta da interação de três elementos:

---

[1] Construção matemática com uma estrutura finita, abstrata e eficaz, que cumpre uma determinada finalidade, sob certas disposições (Mittelstadt, Allo, Taddeo, Wachter & Floridi, 2016).

a tecnologia (maximização do poder computacional e da precisão algorítmica), a análise (identificação de padrões por via de um conjunto de dados) e a mitologia (crença generalizada de que grandes conjuntos de dados oferecem formas maiores de inteligência e conhecimento) (Chan & Moses, 2015). Na mesma ótica, Bartlett, Lewis, Reyes-Galindo e Stephens (2018) consideram o *Big Data* o modo indispensável de pesquisa do século XXI em toda a academia, alegando que a ciência social com dados e a computação intensiva são fenómenos contemporâneos (Halford & Savage, 2017).

Ancorados numa lente geral, diríamos que o *Big Data* é um fenómeno que está a mudar fundamentalmente o que sabemos e o que fazemos – ele consiste em tudo o que envolve capturar, armazenar, partilhar, avaliar e atuar sobre informações que os seres humanos e dispositivos criam e distribuem usando tecnologias e redes baseadas em sistemas informáticos (Gandy Jr, 1989; Herschel & Miori, 2017; Lyon, 2014; Wood *et al.,* 2006). Relativamente a esta questão, o Conselho Europeu refere que a rápida evolução tecnológica e a globalização alteraram profundamente o mundo à nossa volta e trouxeram novos desafios para a proteção de dados pessoais (Gonçalves, 2017). Isto sugere que o *Big Data* é muito mais do que apenas um grande número de dados armazenados: é um fenómeno sociotécnico complexo que congrega a interação de diferentes ciências, poderes e tecnologias.

1.2. DESENHO OPERACIONAL: COMO SE MATERIALIZA O *BIG DATA*?

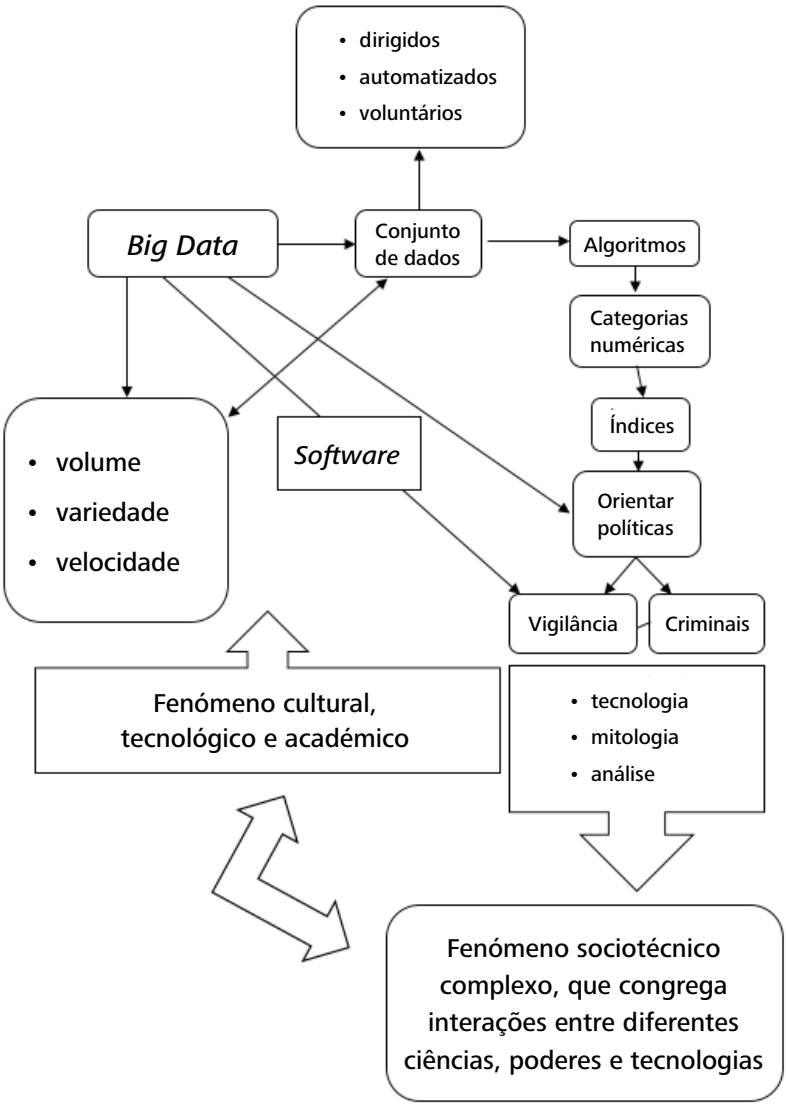
O *Big Data* depende de *softwares* para exercer a vigilância (Lyon, 2014; Matzner, 2016; Halford & Savage, 2017), o que significa que a vigilância automatizada se torna uma possibilidade crescente (Gandy Jr, 1989; Boyd & Crawford, 2012; Frade, 2016; Matzner, 2016). Com efeito, o *Big Data* opera numa vigilância cada vez mais inclinada para operações que se concentram no futuro. No contexto da governança neoliberal, esta antecipação, coloca maior ênfase numa vigilância para gerir as consequências do que numa pesquisa para compreender as causas de problemas sociais como o crime e a desordem (Lyon, 2014).

O *Big Data* faz uso prático destes grandes dados para formular teorias acerca do sujeito que é observado e vigiado, ainda que não diretamente. No fundo, é uma relação, mediada digitalmente, entre os sistemas de vigilância e o indivíduo – ao invés da tradicional relação de proximidade interativa (Ball, Di Domenico & Nunan, 2016; Lyon, 2014). A ação humana é codificada, ou seja, convertida em símbolos com significado especial (Aas, 2006), como se cada passo humano fosse uma senha comunicada por via de uma linguagem

digital (Costa, 2004). Através desses códigos e símbolos, a atividade humana e a identidade individual são armazenadas em bases de dados oficiais e, por exemplo, partilhadas com outros centros policiais (Gandy Jr, 1989; Aas, 2006). O corporal fundido com o tecnológico prevalece, hoje em dia, como a principal fonte de informação (Souza, 2010) – vista, inúmeras vezes, como a verdade. Estamos sob um poder que traduz a vida em padrões de informação, que nos leva para outros níveis de abstração (Aas, 2006).

Assim sendo, com o advento do *Big Data*, paralelamente ao desenvolvimento da tecnologia (Matzner, 2016), criou-se uma relação *humano-algoritmo* que moldou as formas como os seres humanos são tratados e classificados (Lyon, 2014; Wood *et al*., 2006). Através de *softwares*, o comportamento humano é convertido em algoritmos e, sob um número, é representado graficamente para que seja codificado e interpretado (Matzner, 2016; Wood *et al*., 2006; Kubler, 2017). O *Big Data* apoia-se em tecnologias que seguem uma lógica proativa e reconstrutiva: têm como objetivo recolher informação que permita identificar indivíduos onde a sua presença corporal e os seus atos seriam invisíveis para qualquer tecnologia de observação direta no local (Williams & Johnson, 2004; Machado & Granja, 2018). O *Big Data* converte o mundo em dados, codificando e registando movimentos da sua globalidade – suspeitos e não suspeitos, a população como um todo (Andrejevic & Gates, 2014). A forma como os dispositivos se aproximam das pessoas e/ou das redes desafia-nos a (re)pensar a proximidade e a intimidade (Van der Velden, 2015). Atualmente, a vigilância exerce-se não apenas sobre um indivíduo, mas sobre a globalidade (Fróis, 2007; Matzner, 2016), com vista a garantir a "segurança pública nacional" e a prevenir ações terroristas cada vez mais iminentes (Souza, 2010; Drewer & Miladinova, 2017; Wood *et al*., 2006; Wittendorp, 2016).

Figura 1. Definição conceptual e operacional do *Big Data*.



CAPÍTULO 2. O ROSTO DAS REVELAÇÕES: EDWARD SNOWDEN

Em junho de 2013, Snowden, analista de sistemas, ex-administrador dos sistemas da Agência Central de Inteligência (CIA) e ex-contratado da Agência Nacional de Segurança (ANS) dos Estados Unidos da América (EUA), denunciou a arquitetura de opressão projetada pela ANS. Segundo as suas declarações, a ANS usava dispositivos como telemóveis e computadores para ativar a sua localização geográfica e identificar e vigiar os indivíduos; acedia aos *sites* e endereços consultados pelos cidadãos, convertendo esses dados em algoritmos e inserindo-os nas suas bases de dados – sem que houvesse consentimento por parte dos indivíduos. Com quase cinco biliões de registos de telemóveis recolhidos pela ANS por dia, era possível aceder à localização desses indivíduos. Além disso, também eram analisados padrões de comportamento para revelar mais informações pessoais e relações entre diferentes indivíduos (Guzik, 2009; Lyon, 2014; Wright & Kreissl, 2015; Gonçalves, 2017; Matos, 2018). Provou-se que a ANS detinha numerosos programas que podiam ser classificados em algumas categorias de tecnologia de vigilância como escutas telefónicas, codificação, exploração, ferramentas de análise e bases de dados (Van der Vlist, 2017). A partir desse momento, a máquina da vigilância teve um rosto e uma identidade que ficaram conhecidos por todos aqueles que tinham sido vigiados (Van Dijk, 2014; Lyon, 2015; Matzner, 2016; Gonçalves, 2017; Bakir, Feilzer & McStay, 2017; Matos, 2018). Comprovou-se a codificação da vida e das relações sociais em algoritmos, bem como o registo destes códigos em bases de dados, partilhadas, escrutinadas e analisadas (Gandy Jr, 1989; Ball *et al*., 2016; Wittendorp, 2016). Face a estas revelações, apesar de as agências governamentais procurarem minimizar os efeitos dos metadados (Lyon, 2014), o tema da vigilância – principalmente a tecnológica, digital e

informática – nunca mais foi abordado da mesma forma, visto que a população tomou conhecimento de que era controlada nas mais diversas facetas da sua vida social, nomeadamente via redes sociais, localização geográfica, histórico de chamadas e inúmeros sectores que são "privados" (Van Dijck, 2014; Wright & Kreissl, 2015; Matos, 2018). Na conceção de Van der Vlist (2017), as tecnologias de rede de vigilância perderam a inocência após o forte impacto crítico destas revelações. No fundo, fora este o objetivo de Snowden: desbloquear o debate público sobre a invasão da privacidade (Van Dijk, 2014; Young, 2017). E estas revelações desencadearam indignações e ansiedades (Guzik, 2009).

Após a quebra deste silêncio, as práticas de vigilância sofreram alterações. Desde logo, verificou-se uma mudança para uma vigilância em massa, direcionada para o uso maciço de metadados (Lyon, 2015; Young, 2017). Os dados recolhidos são depois agregados, sintetizados e analisados. Sem que o saibam, as pessoas são constantemente vigiadas, controladas e supervisionadas, podendo ser consideradas suspeitas por terem estado num determinado local (Bakir *et al*., 2017). Os governos invocam a defesa da "segurança nacional" para legitimar a vigilância cada vez mais penetrante nos interstícios sociais, como uma política de ligação de pontos que visa prever os fenómenos criminais antes que estes ocorram (Lyon, 2014; Young, 2017). Estes fenómenos potenciam uma *transparência forçada*, em que, secretamente, para maximizar a segurança, se exige uma total visibilidade dos cidadãos (Bakir *et al*., 2017). Muitos são os autores que referem a dimensão estatal da vigilância, em resultado de uma racionalidade implacável expressa em procedimentos burocráticos. Esta condição cultural constrangedora, ajuda indubitavelmente a explicar a razão pela qual a vigilância é continuamente auto-aumentada. Surge, assim, uma governabilidade do crime através da tecnologia – sendo o crime a maior preocupação social e estatal, legitimam-se todas as práticas de segurança adotadas. Na perspetiva de Coll (2014), as declarações sobre privacidade feitas por governos e empresas podem ser consideradas como uma ferramenta de poder e governança ao serviço do capitalismo informacional. Inicialmente, a liberdade, definida como um recurso fundamental, converte a privacidade numa pré-condição para uma economia florescente no contexto da sociedade da informação (Coll, 2014). A tecnologia começa então a ser percecionada como uma ferramenta para a segurança, o controlo e a prevenção da criminalidade.

É importante compreender a vigilância no seu contexto social, político-económico e cultural (Lyon, 2015; Young, 2017; Cinnamon, 2017). Em consonância com a literatura, e face às revelações de Snowden, a máquina opressora alegou que se tratava de um mecanismo que previa proteger as

populações, acrescentando que não se poderia exigir uma percentagem máxima de segurança sem um valor mínimo de invasão da privacidade. Esta última seria consequência das garantias de segurança e proteção pública – como se tivesse de haver submissão a esta política de escrutínio da vida privada para que as ameaças criminais não atingissem as populações, à semelhança do que sucederia em situação de escravidão (Orwell, 2009; Coll, 2014). Todos aqueles que viram facetas da sua vida serem observadas, analisadas, partilhadas e convertidas em números deveriam aceitar a situação, tornando-se *vítimas* de um sistema que alegava a sua proteção e segurança máximas. A aceitação livre do escrutínio realizado pelas Agências de Segurança Estatais dar-lhes-ia o estatuto de agentes do Sistema por auxiliarem as investigações. Desta forma, houve uma obrigação de cooperar com o Sistema, aceitando a sua hipocrisia burocrática – a criação de Códigos e Leis que visam, protegem e defendem o direito à privacidade e, noutra esfera política, o direito estatal de invadir esta privacidade e aceder à esfera individual para garantir a segurança global – e tornando-nos *vítimas* potenciais da Instituição e cidadãos completamente controlados por esta (Costa, 2004; Orwell, 2009).

Houve, pelo menos, três atores importantes neste acontecimento: agências de governo, corporações privadas e, embora de forma involuntária, os cidadãos (Lyon, 2014). A datificação acabou por se converter num meio legítimo de acesso para compreender e monitorizar o comportamento humano – tornando-se numa oportunidade revolucionária de pesquisa para o investigar (Van Dijk, 2014). Consequentemente, a temática do *Big Data* em particular, e da vigilância em geral, surge mais acentuada e desenvolvida a partir deste momento, porque emerge quando as pessoas tomam conhecimento de que estão a ser vigiadas e de que os seus dados estão a ser alvo de invasão (Lyon, 2004; Gonçalves, 2017). Desde então, têm-se multiplicado os debates em torno do *Big Data* enquanto ferramenta de vigilância invisível; atualmente, na *Era dos Grandes Dados*, os autores discutem as consequências ligadas à imersão humana nesta sociedade tecnológica e digital (Taylor, 2017; Boyd & Crawford, 2012; Wood *et al*., 2006). Esta temática está estreitamente ligada aos estudos sobre a vigilância (*surveillance studies*), que se intensificam logo após o 11 de Setembro, em simultâneo com o progressivo interesse por áreas relacionadas com a vigilância, como privacidade, direitos civis e tecnologia (Fróis, 2007, 2015; Lyon, 2015; Matzner, 2016; Wood *et al*., 2006; Costa, 2004; Souza, 2010). É um tema que desde sempre suscitou um intenso debate, mas que se complexifica proporcionalmente à sofisticação e evolução dos meios disponíveis para se materializar a vigilância (Fróis, 2015).

A partir do caso Snowden, na Europa, surgiu um conjunto de preocupações sobre o crescimento dos sistemas de vigilância global que recolhem e recuperam quantidades incalculáveis de dados, com riscos potencialmente graves, não só para a proteção de dados e privacidade, mas, em última análise, para as liberdades e o sistema democrático em geral. Este caso teve o efeito de chamar a atenção pública e política para o fenómeno emergente do *Big Data* e para os riscos e incertezas que ele acarreta. Na sua revisão dos programas de vigilância após o caso, a Comissão Europeia deduziu que as atuais práticas de vigilância reforçadas pelo progresso tecnológico representam uma reconfiguração da inteligência tradicional, facilitando o acesso a uma escala muito maior de plataformas para extração de dados do que a vigilância do passado, implicando assim uma mudança na própria natureza dessas operações (Wright & Kreissl, 2015; Gonçalves, 2017; Matos, 2018).

2.1. DESAFIOS ÉTICOS E DE DIREITOS HUMANOS

Com o crescimento desta forma de fazer Ciência, crescem também as disputas éticas e colocam-se questões sobre este método. Nas últimas décadas, com a densificação dos sistemas tecnológicos, de forma proporcional, densificaram-se também os debates em torno das implicações sociais, económicas, políticas e éticas decorrentes da circulação e partilha de informação em grande escala e da recolha massiva, utilização e partilha de dados pessoais (Matos, 2018). Se o *Big Data* permite que rapidamente se capturem, analisem e explorem informações, também pode permitir o acesso a dados que comprometem a privacidade do indivíduo. E isso pode acontecer deliberada ou inadvertidamente (Herschel & Miori, 2017), na tentativa necessária de encontrar um equilíbrio entre os direitos civis e a necessidade coletiva de segurança (Matos, 2018), visto que a mudança para uma vigilância omnipresente perturba as condições prévias de uma democracia baseada na liberdade dos seus cidadãos (Bakir *et al*., 2017). Desta forma, o *Big Data* promove uma discussão de questões éticas relacionadas com a partilha e uso dos grandes dados (Herschel & Miori, 2017).

Polarizado o debate social sobre o tema, algum tempo depois, os jornais anunciavam e instalavam o caos opinativo: o jornal *The New York Times* documentava os *Oito (não, nove!) problemas com o Big Data*, alegando que se estava a cometer um grande erro com o seu uso, visto que se verificava uma atmosfera controversa no plano dos direitos humanos. Gonçalves (2018) iniciou também uma peça noticiosa, referindo que não há dúvidas de que o *Big Data* desafia o direito fundamental à proteção dos dados pessoais e os

princípios reconhecidos pela legislação europeia e inclusive pela Carta dos Direitos Fundamentais.

Esta ferramenta representa um desafio para o respeito pelos seguintes princípios: consentimento (os dados pessoais devem ser processados apenas se o seu titular tiver dado consentimento prévio e explícito nesse sentido); finalidade (os dados pessoais só devem ser coligidos para fins específicos, explícitos e legítimos e não devem ser processados de modo incompatível com esses fins); minimização (o processamento dos dados deve restringir-se ao mínimo indispensável). Estes princípios tornam-se extraordinariamente difíceis de cumprir devido à automação inerente à mineração, análise e reutilização de imensos conjuntos de dados (Gonçalves, 2018). A recolha de dados é, muitas vezes, automática e passiva (Nunan & Di Domenico, 2013, *cit. in* Ball *et al.*, 2016), criada por sensores automáticos (por exemplo, telefones) e as inferências sobre o comportamento humano são extraídas da análise desse fluxo de dados. No entanto, surgem questões acerca deste funcionamento, visto que tal recolha de dados compromete a privacidade individual e tem implicações nas relações sociais mais amplas (Boyd, 2010, *cit. in* Ball *et al.*, 2016; Coll, 2014). No fundo, permite uma extensão da vigilância comercial e governamental em diferentes aspetos da esfera privada (Ball *et al.*, 2016). Aqui levanta-se a questão da privacidade individual e da exposição à vigilância relativamente ao fluxo de dados autónomo que o indivíduo não pode controlar (Ball *et al.*, 2016; Coll, 2014; Bartlett *et al.*, 2018).

Neste sentido, surgem situações reais exemplificativas desta erosão de direitos. Por exemplo, no que toca ao caso de Bansky, o artista britânico que procura ocultar o seu nome real do domínio público. A Polícia fez uso da metodologia de *Big Data*, enquanto prática, para aceder à real identidade deste autor, mapeando a cidade. Ou seja, usou o perfil geográfico – uma técnica estatística de inferência, tradicionalmente usada para crimes em série como a violação e o assassinato – para encontrar o suspeito. Analisou os padrões espaciais das obras de arte de Bansky em Londres e Bristol, pesquisou pistas eleitorais dos antigos endereços do artista, da sua esposa e de lugares que provavelmente frequentou. As obras públicas de Bansky foram mapeadas e assim a Polícia identificou uma residência; ao entrar nela – invadindo-a –, verificou-se que não correspondia à morada de Bansky. E tudo isto é questionado. Os agentes policiais alegam que se trata de técnicas que conseguem antecipar e prevenir comportamentos terroristas; no entanto, estas técnicas absorvem a ética na sua prática (Metcalf & Crawford, 2016). Se, por um lado, como já referido, a grande análise de dados pode ser uma força de racionalização com potencial

para aumentar a eficiência e melhorar a precisão da predição, por outro lado, o uso de análises preditivas tem potencial para aprofundar padrões de desigualdade existentes (Brayne, 2017).

Consequentemente, toda esta panóplia tecnológica levanta questões em múltiplos planos: a fiabilidade dos dados obtidos; as conclusões retiradas da análise desses dados; o surgimento de uma sociedade panótica, que em tudo se assemelha à do *Big Brother*, por via de uma vigilância permanente, invisível e presente em todos os interstícios sociais (Matzner, 2016; Orwell, 2009; Coll, 2014). Ou seja, por um lado, o *Big Data* é visto como uma ferramenta poderosa para abordar vários tipos de sociedade, oferecendo o potencial de novos conhecimentos em áreas tão diversas quanto a pesquisa sobre saúde, o terrorismo e as mudanças climáticas. Por outro lado, é visto como uma manifestação preocupante de *Big Brother*, permitindo invasões de privacidade, diminuindo liberdades civis e aumentando o controlo estatal e corporativo (Boyd & Crawford, 2012; Lyon, 2014; Orwell, 2009; Coll, 2014). Não obstante, importa que nos questionemos até que ponto um grande número de dados pode revelar a qualidade das informações recolhidas sobre os indivíduos; no âmbito criminal, visto que a aplicação do *Big Data* serve ações preditivas, é necessário refletir acerca do número de ofensores estatísticos que devem ser conotados e analisados para que se desencadeie uma ação policial e, ainda, estudar o processo de agregação dos dados (Lefèvre, 2017; Metcalf & Crawford, 2016; Ball *et al*., 2016; Boyd & Crawford, 2012). Há que considerar também que o uso de uma grande quantidade de dados pessoais levanta as questões da privacidade individual e social e da proteção de dados (Drewer & Miladinova, 2017; Coll, 2014). Com a densificação dos mecanismos de vigilância invisíveis e omnipresentes, pode verificar-se o desrespeito de direitos, nomeadamente o do consentimento (Lyon, 2014).

Acresce que uma consequência prática do *Big Data* é a sua dependência, cada vez maior, da análise baseada em algoritmos; e isto representa, por um lado, um afastamento relativamente ao segmento intuitivo que traça um perfil abstrativo do indivíduo; por outro lado, uma aproximação a análises preditivas e a modelos quantitativos continuamente ajustados para prever comportamentos humanos individuais (Ball *et al*., 2016). Além disso, a presença de algoritmos converte a prática numa realidade eticamente desafiadora, devido à complexidade da análise e da tomada de decisão (Mittelstadt *et al*., 2016). De facto, trata-se de um "grande volume de dados"; no entanto, o que este termo sugere – grande quantidade, em termos numéricos de fontes de informação – não corresponde ao que representa. No fundo, refere-se à abrangência

de cobertura da vida contemporânea, a omnipresença e ao conhecimento de um registo quase completo de vidas individuais, que remove a necessidade de decisões *a priori* sobre o início da atividade da vigilância (Ball *et al*., 2016). Ainda no que concerne aos dados, importa descortinar de onde vêm, como devem ser interpretados e qual a direção a tomar, evitando incluí-los em estudos sem nexo (Boyd & Crawford, 2012). Para além disso, enquanto conjuntos grandes de dados que podem ser modelados, geralmente são reduzidos ao que pode ser ajustado num modelo matemático. No entanto, fora do contexto, os dados perdem significado e valor. A capacidade de representar as relações entre as pessoas num gráfico não significa que a correlação obtida transmita exatamente a natureza e realidade dessa mesma relação (Boyd & Crawford, 2012). Em termos de recolha de dados, os níveis de permissão anteriormente disponíveis apenas em ambientes políticos rigorosamente controlados estão agora disponíveis universalmente (Ball *et al*., 2016). Contudo, o *Big Data* não se refere apenas a conjuntos de dados muito grandes; ele consiste também em ferramentas e procedimentos usados para os manipular, analisar e gerar uma mudança computacional no pensamento e pesquisa (Boyd & Crawford, 2012).

Consequentemente, surgem questões importantes. Os dados de pesquisa em larga escala ajudarão a criar melhores ferramentas, serviços e bens públicos? Ou será que vão inaugurar uma nova onda de violações de privacidade? Será a análise de dados uma ajuda para entender comunidades *online* e movimentos políticos? Ou será usada para investigar manifestantes e suprimir o seu discurso? Dada a expansão do *Big Data* como um fenómeno social e técnico, é necessário interrogar criticamente os seus pressupostos e preconceitos (Boyd & Crawford, 2012; Coll, 2014). Obviamente, tudo isto suscita um enorme número de questões éticas e deontológicas relativas ao fluxo de informação, ao acesso a uma grande variedade de dados e a uma enorme vigilância exercida sobre o ser humano. E é sempre crucial considerar a perspetiva do observado e do vigiado – de que forma o indivíduo se apercebe de que está a ser alvo de uma observação criteriosa (Ball *et al*., 2016)? Por isso, importa ter em conta que o *Big Data* não é autoexplicativo – ou seja, não é porque o *software* denota uma correlação que esta existe simplesmente; é preciso procurar as causas, perceber o fenómeno e aceder a ele, até porque mesmo os elementos considerados mais reais sofrem mudanças e alterações (Boyd & Crawford, 2012; Chan & Moses, 2015; Amoore, 2011; Mittelstadt *et al*., 2016).

Para além disso, grande parte do entusiasmo em torno do *Big Data* provém da perceção de que este oferece acesso fácil a enormes quantidades de dados. Mas quem tem este acesso? Para que fins? Em que contextos? E com

que restrições? Enquanto a explosão da pesquisa usa conjuntos de dados dos meios de comunicação social, as fontes sugerem que o acesso é tudo menos direto; na verdade, apenas as empresas de redes sociais têm acesso a grandes dimensões de dados sociais – especialmente dados transacionais. Algumas empresas restringem o acesso aos seus dados; outras vendem o privilégio de acesso por uma taxa; e outras ainda oferecem pequenos conjuntos de dados a investigadores sediados em universidades. Isso produz desigualdades consideráveis no sistema: quem tem poder económico pode produzir um tipo diferente de pesquisa de quem não tem poder monetário (Boyd & Crawford, 2012). Como se não bastasse, há também questões sérias envolvidas na ética da recolha de dados *online*. O processo de avaliação da ética de pesquisa não pode ser ignorado simplesmente porque os dados são aparentemente públicos. Os pesquisadores devem perguntar-se – e também os seus colegas – sobre a ética do processo de recolha, análise e publicação de dados (Boyd & Crawford, 2012). O acesso aos dados gera uma série de outras questões que se prendem com o facto de estarmos imersos numa era económica informativa em que os dados pessoais dos cidadãos passam a ser vistos como moeda de troca, como forma de garantir o 'acesso gratuito'. Assim, o acesso, processamento e troca de dados pessoais converte-se num negócio (Lawless & Williams, 2010; Wright & Kreissl, 2015; Coll, 2014; Matos, 2018); gradual e consequentemente, emerge um modelo de negócio em que as empresas rastreiam e controlam os dados pessoais dos indivíduos (Lupton & Michael, 2017; Gonçalves, 2018). Tal como Araújo (2011) refletira acerca do tempo, abstraindo o objeto-estudo de reflexão: o *Big Data* converteu-se num objeto político de poder que, paralelamente ao tempo, tem valor económico e está sujeito a relações de poder. A este propósito, Lawless & Williams (2010) afirmam que recentemente se tem verificado uma forma cada vez mais mercantilizada das ferramentas de previsão científica forense. Adotou-se uma racionalidade económica destes fenómenos, capazes de prever e combater a criminalidade, sendo que o Estado neoliberal também promove esta visão mercantilizada (Lawless & Williams, 2010).

Não obstante as considerações tecidas, Brayne (2017)[2] enfatiza as repercussões do uso dos grandes dados no aumento das desigualdades sociais. A autora observou que no Departamento de Polícia de Los Angeles era usado um modelo de cálculo de perfis de risco que identificava os indivíduos que mais provavelmente desenvolveriam condutas criminais. Brayne (2017) afirma que, se for identificado um número considerável de indivíduos perigosos todos

[2] Estudo descrito detalhadamente na Parte II da presente obra.

residentes no mesmo bairro, este bairro fica sob vigilância constante da polícia sem que sequer seja necessário confirmar o seu grau de perigosidade. E caso se trate de um bairro onde residam minorias étnicas as desigualdades podem ser (ainda) maiores e mais difíceis de detetar, já que a vigilância é invisível, mas atua para as perpetuar (Brayne, 2017).

Na mesma esteira, Richards e King (*cit. in* Herschel & Miori, 2017) observam que grandes conjuntos de dados estão a ser utilizados para realizar importantes previsões. No entanto, enquanto isso acontece e se expande, os indivíduos não têm noção de que os seus dados estão a ser recolhidos e partilhados. Os autores constatam que o *Big Data* agrega conjuntos de dados pessoais de toda a natureza – histórico de chamadas, localização histórica, conexões de redes sociais, histórico de pesquisas, histórico de compras e reconhecimento facial – e que essas informações estão (já) nas mãos dos governos e corporações policiais (Herschel & Miori, 2017). Desta forma, uma vigilância sem precedentes pode, em última instância, contribuir para a expansão e perpetuação da identificação global da população, visto que se sustenta na ideia de que, se não existe nada a temer, todos os indivíduos podem integrar as bases de dados. No entanto, as vigilâncias primária e secundária enfatizam e potenciam um posterior contacto com as instâncias policiais. Assim, a inserção, sem critérios nem restrições, da população nas bases de dados pode ter consequências nefastas –, por exemplo, levar à condenação de indivíduos inocentes (Brayne, 2017). Zedner (2016) refere precisamente este aspeto: a ideia de que há uma discriminação ainda maior sobre os grupos já marginalizados que pode levar ao enfraquecimento da segurança. E toda esta panóplia de reflexões acentua-se com o advento do *Big Data*, que exige que se reflita sobre a justiça, no meio do arsenal tecnológico dos grandes dados. Onde pairam os direitos dos cidadãos? Como se define o titular dos dados de cada pessoa (Cinnamon, 2017)?

Todos estes dilemas e questões refletem-se na legislação, na preocupação eminente em proteger os direitos e as garantias dos cidadãos, para que as consequências da expansão destes métodos sejam menos lesivas para os indivíduos. Quanto a isso, existe a perceção generalizada de que o direito não acompanha a tecnologia; ou seja, alegadamente a legislação é intrinsecamente incapaz de fazer face aos progressos tecnológicos. No entanto, neste caso, ciente dos desafios decorrentes dos avanços tecnológicos neste domínio, o Conselho Europeu abriu o caminho para a reforma da proteção de dados e incluiu as tecnologias do *Big Data* na categoria das aplicações digitais a serem abordadas relativamente à definição de novos princípios orientadores. No entanto, esta inclusão e consideração não exclui o facto do *Big Data* ser perspetivado como

uma mudança fundamental, até mesmo uma mudança tectónica no uso de dados pessoais, e estes desenvolvimentos tecnológicos têm consequências na sociedade. Isto também se aplica nas tecnologias do *Big Data*, com todas as suas capacidades algorítmicas e preditivas e a capacidade de influenciar a tomada de decisões, afetando a vida das pessoas em domínios críticos. Contradizendo o que se poderia esperar de uma reforma concebida para regulamentar esta ferramenta no Estado de Direito, a Comissão Europeia define o novo regime de proteção de dados como um "facilitador de serviços do *Big Data* na Europa". Na verdade, a legislação de proteção de dados também significa maximizar o potencial das tecnologias digitais, tendo em conta as consequências económicas e sociais destas últimas (Gonçalves, 2017).

A respeito da proteção de dados e do seu consequente debate após a imersão na era do *Big Data*, em fevereiro de 2016, o Conselho da Europa iniciou o processo de elaboração de orientações específicas sobre a proteção de informações pessoais neste âmbito (Mantelero, 2017). Na última década, os efeitos do desenvolvimento de novas tecnologias e mudanças no ambiente digital criaram um contexto completamente novo. Os dados são abundantes, e em alguns casos a informação é constantemente reutilizada; as análises permitem extrair informações preditivas a partir de conjuntos de dados e os algoritmos são cada vez mais precisos, tornando-se parte dos processos de tomada de decisão. Ao contrário das orientações anteriormente adotadas pelo Conselho da Europa, que lidavam com contextos ou questões específicas, estas últimas diretrizes focam-se no uso de uma determinada tecnologia (*Big Data*) e não são específicas do setor. Além disso, uma vez que há um número crescente de aplicativos de análise do *Big Data* que geram novas questões em termos de proteção de dados, as orientações não podem definir um conjunto de disposições que levem em conta todas as potenciais implicações do uso dos grandes dados. Por estas razões, as diretrizes são necessariamente gerais, mas podem ser complementadas por outras orientações sobre a proteção dos indivíduos em campos específicos de aplicação do *Big Data*. No âmbito da investigação criminal, a União Europeia possui uma diretiva (EU 2016/680) que prevê a proteção de dados no contexto do *Big Data*; no entanto, continua a haver um vazio legal relativamente à forma como as agências policiais e demais atores da investigação e repressão do crime podem fazer a análise algorítmica de dados. Não existe um quadro legal que defina como devem ser tomadas as decisões baseadas em análises do *Big Data*, nem acerca da ética do uso desta tecnologia no contexto da aplicação da lei. E na ausência desses critérios legais, não é possível aos agentes de investigação criminal fazer uso da técnica. Apesar

desse limite, as diretrizes representam um importante passo na regulação do uso do *Big Data*, já que constituem a primeira orientação internacional neste campo, onde questões significativas surgem em relação ao paradigma tradicional de regulamentação da proteção de dados. Um conjunto de disposições das diretrizes diz respeito ao papel da intervenção humana em grandes decisões apoiadas por dados – é um tópico importante, que levanta novas preocupações acerca da liberdade individual, processos de decisão e responsabilidade. De acordo com estas diretrizes, o uso do *Big Data* deve "preservar a autonomia da intervenção humana no processo de tomada de decisões", o que levanta (ainda mais) questões acerca da subjetividade da opinião da pessoa responsável pela decisão final. Daqui se conclui que o uso do *Big Data* é um processo, que resulta dos conhecimentos do ser humano e que é criado e influenciado por este. Por isso, à semelhança do que sucede com a legislação, todas as questões éticas e de direitos humanos que se levantem sobre o *Big Data* devem ter este aspeto em conta (Mantelero, 2017).

Concluindo, apesar de todos os esforços académicos, civis, democráticos, políticos e legislativos, parece ser difícil alcançar um equilíbrio entre a defesa da segurança dos cidadãos e as potenciais ameaças às suas liberdades, garantias e direitos (Machado & Santos, 2016; Cinnamon, 2017; Matos, 2018). A inovação e a evolução na tecnologia são constantes e desejáveis, mas a forma como as tecnologias são usadas para monitorizar e governar os cidadãos é negociável. A população – enquanto massa humana cidadã – deve determinar as interações com a tecnologia, debatendo, se preciso, resistindo e propondo caminhos diferentes (Taylor, 2017). Se o paradigma dos dados totalizantes e massivamente ampliados requer novas formas de conceptualizar a vigilância, também exige esforços renovados para reinventar argumentos legais e intervenções que podem ser usadas para abordar e refrear essas práticas. Alguns autores questionam: se o *Big Data* agrega todas as informações existentes, porque não se preveem efetivamente ataques terroristas (Andrejevic & Gates, 2014)? No entanto, apesar das críticas maciças sobre os motivos epistémicos, éticos e de direitos, o *Big Data* continua a ser um termo popular, político, público e humano.

CAPÍTULO 3. O *BIG DATA*: UMA (NOVA) FORMA DE EXERCER A VIGILÂNCIA?

Vivemos numa época de expansão e intensificação da vigilância nos mais diversos sectores, e alguns autores afirmam que estamos sob a alçada de uma sociedade da vigilância que tem benefícios e riscos (Fuchs, 2011; Costa, 2004; Marx, 2004; Wood *et al.*, 2006). Atualmente os governos usam a vigilância para dar garantias de paz, segurança e controlo social (Lyon, 2004; Fuchs, 2011; Skinner, 2018). A vigilância é apresentada como algo geral e universal das sociedades, sendo conceptualizada como um fenómeno social positivo, autoevidente e constante (Fuchs, 2011)[3], que se foi convertendo gradualmente numa prática omnipresente nas sociedades modernas (Lyon, 1992; Brayne, 2017; Lyon, 2014).

Nos finais do século XX, o desenvolvimento tecnológico repercutiu-se nas técnicas de vigilância, transformando-as (Costa, 2004; Marx, 2004; Lyon, 2004; Wood *et al.*, 2006; Machado & Santos, 2016; Brayne, 2017). No discurso de políticos e autoridades de forças de segurança em todo o mundo, a tecnologia é frequentemente apresentada como uma ferramenta poderosa e inovadora na prevenção, dissuasão e luta geral contra o crime (Fróis & Machado, 2016). Nos últimos anos verificou-se o surgimento de redes interligadas de partilha de informação; o aumento da capacidade de processar, manusear, transmitir e armazenar dados; a transformação dos computadores em aparelhos de observação, simulação e processamento de dados; e a criação de tecnologias de localização geográfica em tempo real. Todo este arsenal tecnológico reconfigurou

[3] Entenda-se aqui a vigilância definida, de acordo com Lyon (2014), ou seja, uma atividade quotidiana racional de controlo da informação na Modernidade, tendo fins de produção e consumo nas sociedades capitalistas e respeitando a burocracia estatal.

as técnicas tradicionais de controlo da ordem pública e da criminalidade por parte dos órgãos governamentais (Gandy Jr, 1989; Graham, 1998; Marx, 2004; Fuchs, 2011), sendo as tecnologias computacionais também responsáveis pela expansão da vigilância no sentido do controlo da informação (Lyon, 1992; Wood *et al.*, 2006). O desenvolvimento e a densificação tecnológica permitiram a emergência de novas formas de (bio)poder (Foucault, 1994), reforçando, desta forma, a ideia já explanada de que a vigilância omnipresente – "panóptico tecnológico" (Cunha, 2008: 71) – visa salvaguardar o estado de segurança pública máxima (Cunha, 2008; Skinner, 2018). No entanto, o que se observa é que, a par deste desenvolvimento de novas formas de vigilância, continuam a aplicar-se as formas tradicionais e, por isso, o que se verifica é um maior arsenal de vigilância social (Marx, 2004).

Estas novas formas tecnológicas de controlar, vigiar e supervisionar os indivíduos têm a sua expansão, evolução e complexificação após o 11 de Setembro, nos Estados Unidos da América e sustentam-se nas políticas de segurança antiterrorismo (Lyon, 2004; Aas, 2006; Souza, 2010; Brayne, 2017). Estas políticas foram depois reforçadas e repensadas, na sequência dos ataques a Madrid, em 2004, a Londres, em 2005, e a Paris, em 2015 (Wittendorp, 2016; Matos, Santos & Machado, 2016). A partir destes acontecimentos, mais e maiores técnicas de vigilância começaram a expandir-se mundialmente, ao mesmo tempo que se espalhava e alargava o pânico moral (Lyon, 2004). O Conselho Europeu reuniu-se e lançou propostas que reforçavam a cooperação policial transnacional, com vista a reduzir e controlar as ameaças terroristas, bem como a gerir o sentimento de insegurança. Surgiram tecnologias que ambicionavam sustentar-se num raciocínio político e que tiveram repercussões na atuação das instituições policiais (Wittendorp, 2016). A luta contra o terrorismo estava instalada, pois a segurança era vista como um fenómeno socialmente produzido e, por isso, tinham de ser criadas condições para sua implementação (Wittendorp, 2016; Brayne, 2017). A estas políticas de segurança acrescem outras, que têm como objeto de controlo e vigia o corpo e o supervisionamento dos seus movimentos (Aas, 2006; Ferreira, 2014). Assim, a vigilância foi sendo expandida, estimulada e legitimada, no sentido da melhoria das suas operações, foi alvo de um amplo financiamento no que toca ao desenvolvimento de novos sensores e programas de dados que produzissem *estratégias de inteligência* (Brayne, 2017). Ou seja, um dos projetos tecnológicos levados a cabo pelos Estados-Membros para a mitigação do risco foi a expansão da vigilância tecnológica e digital. Este investimento na tecnologia estimulou a recolha massiva e constante de dados e a idealização de bases

de dados estruturadas e organizadas, criando uma cultura de fiscalização do crime sustentada em procedimentos de regulação, inspeção e controlo social (Queirós, 2018).

3.1. AS NOVAS TECNOLOGIAS E A NOVA VIGILÂNCIA

A tecnologia moderna possibilita um controlo técnico e uma supervisão anónima: o desenvolvimento das novas tecnologias cria condições para o aparecimento de uma *nova vigilância* (Gandy Jr, 1989; Lyon, 1992, 2004; Marx, 2004; Wood *et al*., 2006; Fuchs, 2011). No entanto, esta vigilância de dados, ou seja, a supervisão dos cidadãos com base em dados *online*, difere essencialmente da vigilância tradicional, na medida em que esta última prevê a monitorização vigilante dos indivíduos com fins específicos, enquanto a primeira inclui o controlo contínuo dos dados individuais sem fins predefinidos, nem declarados (Van Dijk, 2014). É uma vigilância cada vez mais incentivada por sensores, computadores, aplicativos de *software,* redes e outras tecnologias, e que evolui, progressivamente, de alvos individuais para um olhar persistente e indiferenciado sobre populações inteiras (MacWillie, 2018). Assim, surge uma vigilância descentralizada e em rede, operada por diversos agentes, dispersos e heterogéneos, e potenciada pela disponibilidade das redes digitais (Gandy Jr, 1989; Graham, 1998; Costa, 2004; Fuchs, 2011; Aas, 2006). Vigiar converte-se numa atividade anónima, fluida e móvel, em que as organizações que vigiam estão cada vez mais omnipresentes, mas ocultas perante os vigiados (Lyon, 2004, 2014; Marx, 2004; Costa, 2004); estes últimos, por sua vez, tornam-se cada vez mais transparentes sob o olhar vigilante (Costa, 2004; Lyon, 2014; Wood *et al*., 2006; Ferreira, 2014). Esta dualidade acentua-se com o advento do *Big Data* (Wood *et al*., 2006). Não há um ponto geográfico único de acesso à informação; ela pode ser consultada a partir de qualquer lugar. Da mesma forma, não há uma base de dados eletrónica única para vigilância, mas muitas dispersas que podem ser usadas em conjunto por atores poderosos com o objetivo de realizar uma pesquisa interligada de dados (Graham, 1998; Fuchs, 2011; Aas, 2006; Wood *et al*., 2006; Prainsack, 2019). Emerge, assim, uma sociedade que opera uma vigilância pós-Snowden, baseada em algoritmos (Bakir *et al*., 2017). Consequentemente, as tecnologias da vigilância têm migrado de métodos geográficos centralizados e temporalmente descontínuos para métodos geograficamente descentralizados e temporalmente contínuos (Gandy Jr, 1989; Lyon, 2004; Fuchs, 2011).

Esta panóplia tecnológica, em que os dados podem ser consultados a partir de qualquer lugar, visto estarem disponíveis a partir de qualquer ponto

digital, dificulta o estabelecimento de limites sobre quem e de que forma estes dados podem ser acedidos (Prainsack, 2019). Aliás, reflexo disso são as definições contemporâneas de vigilância que não distinguem entre relações sociais e forças tecnológicas da vigilância (Fuchs, 2011) e que não especificam se a vigilância é uma tecnologia ou uma relação social (Lupton & Michael, 2017; Lyon, 1992, 2014; Wood *et al*., 2006). O *Big Data* tem o poder de se materializar numa vigilância estruturalmente especulativa: dados que aparentemente não se relacionariam com um objetivo estratégico particular podem produzir correlações imprevisíveis, mas úteis. Além disso, o *Big Data* visa acumular um arquivo que pode ser pesquisado e classificado de forma retrospetiva. Este arquivo tem aspetos positivos e negativos: positivos na medida em que permite a consulta repetida de imagens, factos e aspetos necessários à intervenção; negativos, visto que vincula os indivíduos às cenas do crime, reconstruindo os seus movimentos, identificando-os e, eventualmente, capturando-os (Andrejevic & Gates, 2014).

As circunstâncias mudaram e a vigilância sofreu alterações, principalmente no que toca à sua margem de ação e abrangência (Costa, 2004; Fróis, 2007; Lyon, 2014). Por exemplo, existem câmaras de vigilância em todos os lugares – públicos e privados – que captam imagens sobre a totalidade dos indivíduos (Lyon, 2004; Wood *et al*., 2006; Machado & Santos, 2016; Fróis, 2007), e há um crescente desenvolvimento da recolha e armazenamento de informação com potencial uso político e militar (*intelligence*) e o aperfeiçoamento do *data mining* (Machado & Santos, 2016; Matos, Santos & Machado, 2016). Assim, os *maiores alvos* desta vigilância contínua são os cidadãos: os seus dados são analisados, processados e armazenados pelos sistemas de vigilância (Aas, 2006; Maciel & Machado, 2014; Lyon, 2014), que evidenciam modos de controlo social e de submissão do sujeito que são, em muitos aspetos, semelhantes ao que Foucault (1999) caracterizou como panótico, na medida em que colocam os indivíduos em estado de submissão pelo facto de a vigilância ocorrer de modo universal e permanente. Este tipo de vigilância é pré-construtiva, recolhendo e arquivando informação sobre qualquer pessoa de forma aberta, dirigida e indiscriminada (Machado, Queirós, Martins, Granja & Matos, 2018). Trata-se de uma vigilância que se exerce como resposta à evolução do nível de sofisticação das tecnologias, enquanto elementos essenciais da sociedade global, capitalista e financeira (Araújo *et al*., 2015).

3.2. RECONFIGURAÇÕES SOCIAIS: O QUE MUDOU?

O interesse contemporâneo da vigilância centra-se nos computadores (Gandy Jr, 1989; Lyon, 1992, 2004; Marx, 2004), através dos quais se gera muita informação, por via da partilha transnacional de dados (Lyon, 2004; Williams & Johnson, 2004). Essa informação deixa rastos infinitos, podendo ser retomada, reanalisada e alvo de (re)tratamento sempre que os órgãos responsáveis pelo seu armazenamento assim o desejarem. O trabalho exercido sobre estas informações e estes dados consiste em análises complexas que inter-relacionam pessoas, eventos e interações e que requer, consequentemente, um alargamento da rede de controlo social. Estas alterações no controlo social e na vigilância – técnicas e materializações – são vistas, por inúmeros autores, como uma resposta às transformações técnicas e sociopolíticas que ocorreram nos últimos tempos, com consequências (não intencionais) na privacidade social e individual. E uma vez que este aprimoramento técnico já faz parte dos interstícios sociais (Ferreira, 2014), importa estudá-lo. Desde que se começou a estudar o tema, um dos focos de controvérsia girava em torno das consequências potencialmente danosas da sua faceta tecnológica. A contínua sofisticação e gradual aptidão para aglutinar e intercetar dados, identificar padrões comportamentais, conceber perfis, categorizar e compartimentar pessoas, grupos e comunidades, pode levar a que esta prática seja desvirtuada. Problematiza-se a capacidade da vigilância tecnológica, que vai para além do que é declarado por governos e polícias. Prova-se a forma como esse controlo omnipresente pode ser prejudicial para o cidadão, na medida em que interfere e condiciona direitos civis, como a liberdade e o direito à privacidade; e revela-se como os sistemas de vigilância, ao materializarem-se numa lógica de diferenciação social, cooperam para a classificação, discriminação e segregação de pessoas e grupos (Fróis, 2015). Além disso, o modelo *Big Data* parece um pouco extravagante: mobiliza a promessa da profusão de dados, transmitindo a ideia de que a tecnologia emerge para controlar tudo sobre todos em todos os momentos, e armazena esses dados em formato legível. O problema deste modelo reside, necessariamente, no facto de ficar aquém do seu objetivo – não há garantias de que os dados recolhidos sejam abrangentes ou representativos. Assim, os críticos referem que se trata de uma abordagem antiquada dos dados: uma ênfase excessiva no seu conteúdo, ao invés de na sua eficácia funcional (Andrejevic & Gates, 2014).

Kierkegaard (2008) refere que, fruto das mudanças sociais e contextuais, a abolição das fronteiras entre os Estados-Membros levaram a desafios na

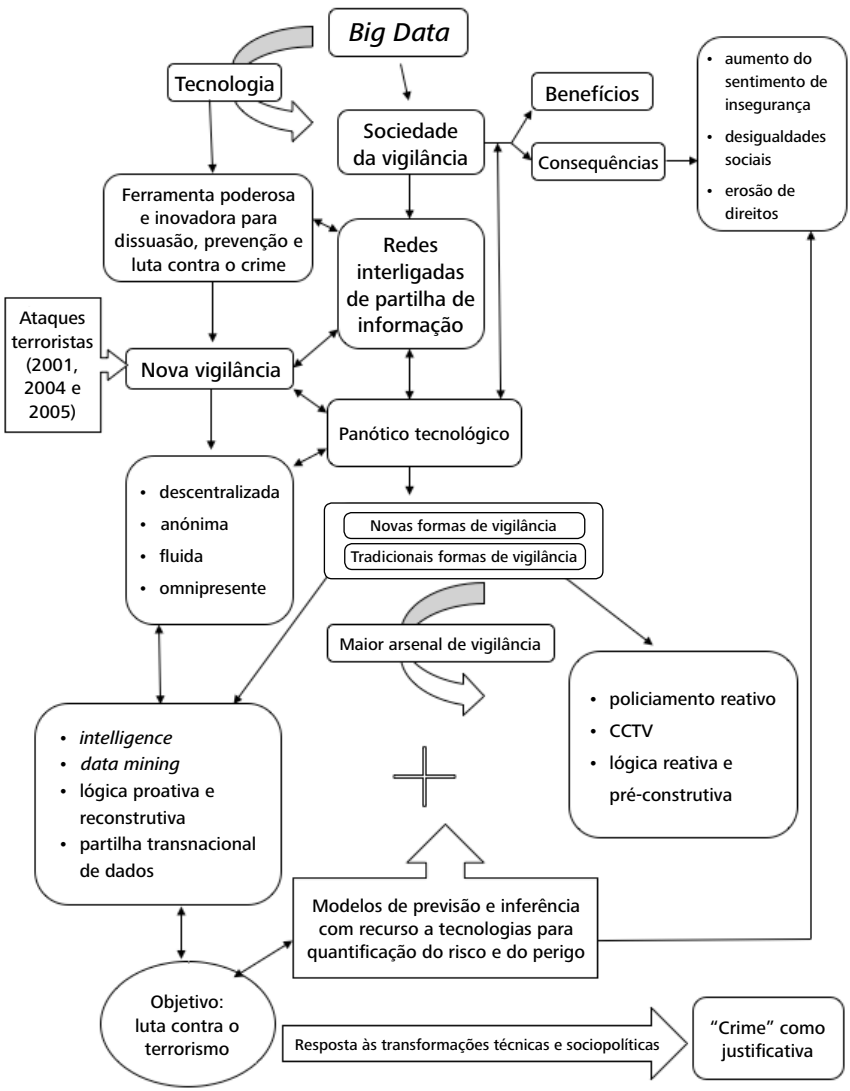
segurança. No sentido de melhorar esta última, como já referido, iniciativas recentes concentraram-se no intercâmbio de informações, ao abrigo do princípio da *disponibilidade*. Segundo este princípio, deve ser feito um pleno uso das novas tecnologias e também deve haver recíproco acesso a bases de dados nacionais; ao mesmo tempo, o Sistema que defende este princípio – Sistema de Prüm – visa promover o intercâmbio de informações entre as autoridades responsáveis pela aplicação da lei. Skinner (2018) também explora estas mudanças sociais enquanto molas propulsoras do desenvolvimento tecnológico como ferramenta ao serviço da vigilância e da sua expansão. As novas tecnologias sociais monitorizam corpos, os seus movimentos e atividades, e são um investimento enorme nos EUA, já que constituem parte de um fenómeno internacional mais amplo – as tecnologias como soluções para questões complexas de segurança. Um exemplo claro reside na génese de múltiplos sistemas computorizados conectados de recolha, armazenamento e análise de dados que permitem a partilha de informações pessoais importantes. Estas novas formas de *fazer* Ciência não servem apenas fins de controlo e vigilância, mas atribuem também aos cidadãos um rótulo, colocando-os em grupos especiais de vigilância (Skinner, 2018).

Consequentemente, o "crime" foi sendo gradualmente eleito como uma justificativa legitimadora dos discursos políticos sobre o risco, que visavam reforçar, aprofundar e aprimorar a vigilância (Wright & Kreissel, 2015). Ou seja, há uma intensificação de discursos que visam reforçar as estratégias máximas de segurança (Jasanoff & Kim, 2009) de que resulta uma gradual transformação das sociedades disciplinares em sociedades de segurança (Maciel & Machado, 2014). Assim, procura-se alcançar e garantir a segurança através da gestão do risco e da iminência das ameaças (Weber, 2016; Johnson & Williams, 2007; Skinner, 2018). Há uma centralização na gestão da ameaça criminal através da previsão do risco (Bakir *et al*., 2017; Machado *et al*., 2018), sendo que, muitas vezes, as noções de risco vêm acompanhadas de ideários de vigilância expansiva sobre os indivíduos. Neste panorama, a apropriação da tecnologia e a criação de projetos que visam a dissipação do risco não calculável por parte dos Estados decorrem de diferentes práticas, políticas e aceitação dos cidadãos (Jasanoff & Kim, 2009). Ou seja, o rumo atual das modalidades de vigilância criminal assenta, cada vez mais, em modelos de previsão e inferência a partir de populações e grupos considerados suspeitos, com recurso a tecnologias que, através da quantificação do perigo, visam controlar, prever e identificar possíveis comportamentos e atos criminosos (Machado & Santos, 2016; Machado *et al*., 2018).

Desta forma, esta maneira de exercer a vigilância pode ser perspetivada de vários ângulos. No entanto, alguns autores defendem que se trata de uma *vigilância nova*, outros consideram que se trata de um *fenómeno de vigilância eminentemente tecnológico*, outros ainda afirmam que é uma *vigilância adaptável* que procura dar resposta às ameaças externas; por fim, surgem autores que dizem que se trata apenas de um *produto da Modernidade* e de todas as suas transformações (Lyon, 2004). Inicialmente a vigilância consistiu numa forma de reforçar o poder do próprio Estado e hoje é um meio de controlo social (Lyon, 2004; Williams & Johnson, 2004). Consequentemente, o *Big Data* intensifica certas tendências de vigilância associadas à tecnologia da informação e às redes, e está, portanto, implicado em configurações fluidas. Isto é considerado de três formas: i) as capacidades do *Big Data* (incluindo metadados) intensificam a vigilância expandindo conjuntos de dados interconectados e ferramentas analíticas; a dinâmica de influência existente, a gestão de riscos e o controlo aumentam a velocidade e alcance através de novas técnicas, especialmente análises preditivas, e a ameaça criminal é gerida precisamente através desta previsão do risco; ii) o *Big Data* opera uma mudança qualitativa nas práticas de vigilância, resultando em consequências; e iii) o comportamento ético torna-se mais urgente como um modo de crítica (Lyon, 2014).

Concluindo, vivemos numa sociedade em que a vigilância está profundamente inscrita e onde o seu estudo se torna cada vez mais difícil, precisamente porque também é cada vez mais árdua a tarefa de distinguir entre práticas de vigilância e rotinas diárias. Como afirma Van der Vlist (2017), a vigilância tornou-se algo com que vivemos diariamente.

Figura 2. Transformações que operam na vigilância com o surgimento do *Big Data*



PARTE II

BIG DATA E INVESTIGAÇÃO CRIMINAL

CAPÍTULO 1. NOTAS HISTÓRICO-ESPACIAIS DO *BIG DATA* NA INVESTIGAÇÃO CRIMINAL: UMA REALIDADE ANTIGA?

A inserção e inclusão do *Big Data* na previsão e repressão da criminalidade não reside só nos tempos contemporâneos (Brayne, 2017). Os bancos de dados computorizados não são uma realidade nova. O Escritório dos Censos dos EUA implementou o primeiro equipamento de processamento automatizado do mundo em 1890 – a máquina de cartão de perfuração (Anderson, 1988, *cit. in* Boyd & Crawford, 2012) – e os bancos de dados relacionais surgiram em 1960 (Fry & Sibley, 1974, *cit. in* Boyd & Crawford, 2012). Em 1928 a Escola de Chicago elaborou um modelo que previa a probabilidade de reincidência no momento da determinação da liberdade condicional dos agentes. Nos Tribunais dos EUA as práticas de quantificação de *grandes dados* foram incluídas nos anos 1970/1980 por via de diretrizes de sentença. A Justiça Criminal, nas últimas três décadas, experimentou a implementação de práticas atuariais por via do uso de critérios numéricos de gestão do risco para prever e estimar probabilidades de risco criminal (Brayne, 2017).

Em 1994, foi implementado o *CompStat* – um modelo de gestão do risco que identificava padrões de crime, quantificando e incentivando atividades policiais, e direcionava os seus recursos. Assim, pode constatar-se que os EUA possuem o SJC mais maleável e flexível destas progressivas tendências de vigilância – desde o crescente encarceramento à expansão do policiamento. Em 1994 foram desbloqueados fundos monetários que visaram a contratação de mais de cem mil agentes policiais, e em 2002 foi regulamentado o financiamento de Agências Locais de Aplicação da Lei. Recentemente, os fundos federais foram direcionados na melhoria e ampliação do uso da tecnologia ao serviço da lei – na prevenção e repressão da criminalidade (Brayne, 2017). Ou seja, o que se

verificou nas últimas três décadas e, consequentemente, criou condições para a génese e expansão de uma análise por via do *Big Data*, foi uma transformação revolucionária no SJC: uma mudança para uma *justiça atuarial*[1] (Feeley & Simon, 1992) em que os profissionais utilizam critérios-resultado de gestão do risco (Lyon, 2004; Amoore, 2011) para estimar probabilidades de risco criminal (Andrews & Bonta, 2010; Brayne, 2017; Amoore, 2011). Promove-se assim a ideia de que é possível, e necessário, prever o risco dos indivíduos, categorizá-los e antecipar os fenómenos criminais.

Além disso, geralmente, nos casos de análise de dados no âmbito da aplicação da lei, os *softwares* fazem suposições que implícita e explicitamente dependem de teorias criminológicas – por exemplo, teorias que foram desenvolvidas para calcular e analisar o intervalo de tempo necessário de histórico criminal para prever a reincidência criminal. Ou seja, a Criminologia e as Ciências Sociais Humanas adaptam-se ao uso prático do *Big Data* e ao aparecimento de novas técnicas e metodologias em geral (Chan & Moses, 2015).

No âmbito criminal, verifica-se uma tendência crescente do uso do *Big Data* como ferramenta preventiva que orienta estratégias policiais e decisões de justiça criminal (Chan & Moses 2017). Na prática, esta técnica materializa-se em estratégias de análise de risco de indivíduos e prevenção criminal sobre lugares, intervalos de tempo e/ou pessoas; a partir destas estratégias de análise, criam-se perfis que combinam, com precisão, potenciais infratores a crimes específicos (Chan & Moses, 2015; Brayne, 2017). As ações policiais são, também, eminentemente preventivas, preditivas e orientadas para o problema, visto que se materializam após um trabalho de pesquisa e investigação que sinaliza locais, pessoas e intervalos de tempo. Estas análises baseiam-se em bancos de dados, modelos de regressão estatísticos, análises de risco e uso de técnicas avançadas de manuseamento de dados (Chan & Moses, 2015; Amoore, 2011), com o objetivo de relacionar padrões de dados e metadados com os comportamentos reais ou potenciais dos indivíduos para produzir informações sobre quem são e o que fazem (Van Dijk, 2014; Amoore, 2011). O acesso a uma maior quantidade de dados e, consequentemente, a uma multiplicidade de informação por via de bases de dados, permitirá o acesso a uma realidade mais objetiva e próxima do combate e prevenção de certas situações.

---

[1] Uma das limitações importantes existentes na literatura acerca desta temática é o atuarialismo, documentado, definido e tratado, *a priori* da génese dos grandes dados. Ainda que o paralelismo e a inter-relação possam ser estabelecidos e considerados, deve ressalvar-se esta limitação (Brayne, 2017).

No fundo, é reunir diferentes bancos de dados, criados em campos potencialmente diferentes e especializados (Dimeglio, C., Kelly-Irving, M., Lang, T., & Delpierre, C., 2018; Prainsack, 2019).

Tendo por base o desenvolvimento crescente dos grandes dados no seio da investigação criminal, procura-se explorar em que meios este se desenvolve e de que forma, como surgiu, que estudos se debruçam sobre esta relação e que efeitos tem nas práticas de prevenção e combate ao crime.

1.1. PROJETO *TOTAL "TERRORISM" INFORMATION AWARENESS*

Recentemente, exemplo da aplicação do *Big Data* ao combate à criminalidade, mais concretamente, ao terrorismo, foi o Projeto TIA – *Total "Terrorism" Information Awareness*, que propunha, de forma clara, capturar a *assinatura-informação* das pessoas de modo a dar a possibilidade ao Governo de vigiar potenciais terroristas envolvidos em vários tipos de crimes contra o Estado (Costa, 2004; Guzik, 2009; Lyon, 2015). A estratégia deste projeto baseia-se na vigilância e controlo de determinados indivíduos, na recolha e armazenamento do maior número possível de informação capturada para, posteriormente, através do uso de *softwares* e de técnicas de análise humana, detetar as atividades desses indivíduos. Toda a informação capturada, registada e analisada é passível de partilha e cruzamento com dados de outras fontes que, de forma mais eficaz e eficiente, permitem delinear as ações daqueles indivíduos. O TIA possibilita a construção do perfil total dos indivíduos, resultado do cruzamento das suas ligações telefónicas (origem, destino, data e duração), despesas registadas nos cartões de crédito e demais informações relevantes. O Projeto foi criado para produzir uma visão dos padrões comportamentais da população, tendo como objetivo auxiliar os analistas na predição de ações terroristas (Costa, 2004; Guzik, 2009; Lyon, 2015).

1.2. SISTEMA ECHELON

Todas estas tecnologias são relativamente recentes quando perspetivadas sob a ótica do Sistema Echelon – um dos mais famosos sistemas de vigilância, desenvolvido após a Segunda Guerra Mundial pela ANS dos EUA. Atualmente, o sistema tenta capturar e analisar, virtualmente, todas as chamadas de telefone e de *fax*, mensagens e *e-mails* enviados de qualquer ponto do planeta. Este sistema possui estações de interceção de sinais em todo o mundo que capturam o tráfego de comunicações via satélite, micro-ondas, celular e fibra

ótica, processando estas informações em computadores de alta capacidade que incluem programas de reconhecimento de voz e de caracteres. O Echelon marca as mensagens, grava-as, transcreve-as e analisa-as para investigações futuras. No fundo, pretende aceder ao conteúdo de mensagens transmitidas por diversos meios e trocadas pelas mais diversas instâncias, desde pessoais a Governos, organizações internacionais e organismos privados. Inicialmente, foi concebido para espionagem militar e diplomática; atualmente debruça-se sobre tráfico de droga, branqueamento de capitais, terrorismo e criminalidade económica e organizada (Costa, 2004).

1.3. EURODAC

Na mesma linha de pensamento do Sistema Echelon, Tsianos e Kuster (2016) descrevem também o Eurodac, o mais antigo banco de dados biométrico europeu de larga escala, a operar desde 2003. Este banco de dados recolhe e processa impressões digitais de requerentes de asilo, pessoas que atravessam a fronteira europeia de forma irregular e outras que se encontram ilegalmente em território da UE. O Eurodac é um instrumento para gerir a migração e a mobilidade irregular. A sua aplicação combina a tecnologia de identificação biométrica com o processamento de dados informáticos, sendo considerado parte da *fronteira inteligente* da Europa: uma fronteira difusa, que não pode ser geograficamente localizada mas que depende de locais virtuais e de instituições de controlo e vigilância, conectados através de redes de dados digitais. A introdução do termo técnico *fronteira inteligente* ou *fronteira tecnológica da Europa* processou-se em simultâneo com o de *reengenharia programática* e o redimensionamento da gestão de fronteiras após o 11 de Setembro. O Eurodac é um instrumento que possibilita a desterritorialização da fronteira europeia externa, alargando-a potencialmente a todo o espaço Schengen. O saudável funcionamento desta fronteira é possível por, recentemente, esta se ter conjugado com programas algorítmicos que lhe permitem a identificação de casos relevantes (Tsianos & Kuster, 2016).

CAPÍTULO 2. ESTUDOS EMPÍRICOS

Gradualmente, o *Big Data* expandiu-se enquanto estratégia auxiliar para a tomada de decisões, pois começou a ser visto como uma ferramenta com potencial para melhorar a eficiência e a responsabilidade. Concebido na perspetiva de melhorar a predição de comportamentos criminais, auxiliaria a lei a implementar recursos mais eficientes, ajudaria a prevenir crimes e, desta forma, reduziria a taxa de criminalidade (Brayne, 2017). Em 2014, o Presidente dos EUA, Barack Obama, referiu que o *Big Data* podia ser uma ferramenta poderosa para a aplicação da lei e que possuía potencial para fortalecer substancialmente a segurança nacional (Chan & Moses, 2017). O debate em torno deste tema tornou-se mais evidente a partir do momento em que se verificou um crescente uso de *softwares* analíticos de dados ao serviço do policiamento preventivo (Chan & Moses, 2015) e que a tecnologia se tornou um recurso muito valioso (Chan & Moses, 2017).

Os estudos empíricos que se debruçam sobre a aplicação prática, as expectativas e perspetivas do *Big Data* são ainda escassos, o que se reflete numa falta de informação acerca da relação (in)existente entre vigilância, grandes dados e consequências sociais da interseção destas duas forças (Brayne, 2017; Bartlett *et al*., 2018). Não obstante, existem alguns autores que mapearam as representações, expectativas e reconfigurações que atualmente o *Big Data* desencadeia.

2.1. LOS ANGELES

Brayne (2017) realizou um estudo qualitativo de dois anos e meio no Departamento de Polícia de Los Angeles (LAPD). Através da realização de observações e entrevistas, a autora verificou que com o advento do *Big Data*

houve mudanças no LAPD[2]. Por via de uma descrição empírica, Brayne (2017) refere que a análise dos grandes dados facilita e amplifica práticas de vigilância prévias e ancora transformações fundamentais nas atividades de vigilância. Conclui que as diferenças mais notórias foram: i) o uso crescente de pontuações de risco que cotam indivíduos; ii) uma atividade policial mais preventiva; iii) a supervisão sistemática de um elevado número de pessoas; iv) o acesso a informações pessoais através de dados e o cruzamento destes dados; v) a inserção mais facilitada de indivíduos nos bancos de dados da lei; vi) a fusão de sistemas policiais anteriormente separados, facilitando a disseminação da vigilância entre muitas e diferentes instituições policiais; e vii) a emergência da chamada "análise de rede", pela qual os agentes policiais , no seu mapa digital, interligam todos os indivíduos pontuados com alto risco, analisando depois as ligações que esses indivíduos têm a lugares e pessoas.

Concretamente, a primeira mudança descrita por Brayne (2017) residiu na quantificação do risco individual dos cidadãos, através de um novo sistema de pontos: a Operação LASER (*Los Angeles Strategic Extraction and Restoration Program*). Esta Operação relaciona lugares e indivíduos potencialmente perigosos, por via de modelos baseados em evidências. Ou seja, indivíduos-autores de crimes violentos são relacionados entre si, sendo esta uma estratégia que permite reduzir as taxas de criminalidade. Mediante e conforme o tipo de ilícito criminal que consta no Certificado de Registo Criminal de cada indivíduo, a cotação é proporcional à gravidade do ato cometido. Assim, criam-se índices de pontos que sinalizam estes indivíduos e que são partilhados pelos diferentes Departamentos e Setores da Agência Policial, por via da inserção destes cartões informativos dos cidadãos num Sistema Informático. No que concerne a esta mudança, descrita pelos entrevistados, Brayne (2017) refere que os oficiais afirmavam que, quanto maior fosse a cotação atribuída a um indivíduo, maior era a probabilidade de este ter um contacto futuro com a Polícia. No entanto, dados os escassos recursos policiais humanos e monetários, os indivíduos sobre os quais recaía uma vigilância mais intensa eram aqueles que tinham sido mais pontuados, independentemente de atualmente cometerem ou não atividades criminosas – o que leva à criação de desigualdades. Ou seja, estes

[2] Terceiro maior Departamento Policial dos Estados Unidos, com 9947 oficiais, 2947 funcionários civis, abrange uma área de quase quatro milhões de pessoas. Foram realizadas entrevistas e observações com 75 indivíduos e realizadas entre uma e cinco entrevistas de acompanhamento com uma subamostra de 31 indivíduos. Além disso, foram feitas observações em carros de patrulha e helicóptero (Brayne, 2017).

oficiais revelaram-se conscientes do contexto organizacional restritivo – no que toca aos recursos – da Polícia.

A segunda mudança verificou-se no paradigma de atuação policial: de um policiamento reativo para um policiamento preventivo. Inicialmente, o policiamento era essencialmente reativo, ou seja, os agentes policiais apenas atuavam em determinado incidente após este acontecer. No entanto, no início da década de 1980, estas estratégias revelaram-se ineficazes e, por isso, adotaram-se estratégias preventivas que visavam atuar antes de os incidentes sucederem. O mesmo acontecera no LAPD: em 2012, este começa a usar um *software* – *PredPol* – que, com base em algoritmos, elabora modelos que preveem a ocorrência de um facto criminal, tendo em conta o histórico espácio-temporal de incidentes criminais. Ou seja, se em determinado local se registou um crime, a probabilidade de se voltar a registar é maior; assim, esse local fica registado no algoritmo e classificado como *hotspot* criminal. Este *software* foi utilizado nas diferentes secções do LAPD (Brayne, 2017).

Por fim, as duas últimas mudanças representam transformações práticas nas atividades de vigilância policial. Uma delas consiste na expansão e utilização da *análise de rede* que relaciona um maior número de indivíduos – que (não) estabeleceram contacto com a Polícia. Ou seja, inicialmente, eram conotados apenas indivíduos que já tinham tido contacto com a Polícia; posteriormente, foram considerados indivíduos que se revelavam potencialmente perigosos, independentemente de um contacto prévio com a Polícia; começaram assim a ser ligadas atividades diárias, familiares diretos de indivíduos que já tinham histórico criminal e locais frequentados por estes familiares (Brayne, 2017).

A outra mudança prende-se com a criação de um Sistema de Dados Institucionais Integrado – a digitalização dos registos policiais e a sua consequente proliferação facilitam a partilha, conexão e consulta destes a partir de qualquer Secção do LAPD, por parte de qualquer agente policial, a partir das bases de dados policiais. Assim, a Polícia aumentou o leque de acesso a dados que podem ser relevantes para a sua ação (Brayne, 2017).

2.2. AUSTRÁLIA

Chan e Moses (2017) realizaram um estudo na Austrália, sob um projeto de pesquisa intitulado de *Big Data Technology and National Security*, no *Data to Decisions Cooperative Research Centre* (D2DCRC). Os autores verificaram, através de 31 entrevistas semiestruturadas a 38 funcionários responsáveis pela aplicação da lei, agentes de supervisão, agentes policiais, legisladores, informáticos

e membros da sociedade civil relevantes nas organizações na Austrália, que: i) o *Big Data* era maioritariamente associado a um grande volume de dados que só a tecnologia poderia manusear; ii) se tratava de uma técnica com uma capacidade analítica mais avançada, particularmente no contexto da previsão e análise de rede; e iii) mais de metade dos entrevistados não utilizava o *Big Data* (Chan & Moses, 2017). Todos os entrevistados enfatizaram a capacidade analítica da técnica, sendo que os agentes policiais tenderam a concentrar-se na riqueza dos dados e nas suas vantagens para a investigação. Os entrevistados pertencentes a organizações políticas perspetivavam o *Big Data* como uma técnica inovadora que criava oportunidades para uma abordagem mais proativa da segurança, baseada em inferências sobre tendências e padrões. Ou seja, apesar das promessas do *Big Data* para melhorar a eficiência e eficácia das atividades policiais e de agentes de segurança, muito pouco se sabe sobre como o *Big Data* é entendido ou imaginado por esses agentes. Este estudo permitiu concluir que, para os entrevistados, o *Big Data* é uma nova técnica de segurança que é contestada (Chan & Moses, 2017).

Os autores concluíram que, apesar de a polícia estar consciente do potencial da tecnologia para estratégias policiais mais inteligentes, não tem tempo nem recursos para beneficiar desse potencial.

2.3. FRANÇA

Kubler (2017) realizou um estudo empírico na Polícia francesa, logo depois de o estado de emergência ter sido ativado pelo Estado francês pela quarta vez, na sequência de ataques terroristas. O autor (2017) refere que numa era em que o policiamento depende fortemente, e cada vez mais, de *softwares* para agregar informações, é importante entender o potencial dessas tecnologias para influenciar a vigilância e o policiamento. Assim, por via da análise de testemunhos policiais recolhidos no âmbito de uma investigação realizada pela Assembleia Nacional Francesa, conclui que, desde 2005, a Polícia francesa usa o *IBM's computer program – i2 Analyst's Notebook* – para agregar informações e criar narrativas criminais. Esta tecnologia é um *software* de policiamento que organiza e visualiza dados, visando conectar rapidamente suspeitos com crimes, encontrar o maior número de associações entre ambos e classificá-los com base no seu nível de importância para a investigação em apreço. O *i2 Analyst's Notebook* é fruto da junção de dois programas de análise criminal – SALVAC e ANACRIM – e fazem parte de um plano mais amplo criado para dados *online* do governo, de forma centralizada e facilmente compartilhável.

O ANACRIM foi recentemente atualizado para ajudar a automatizar a recuperação de informações de todas as bases de dados da polícia para que possa ser mais fácil integrar no seu *software* de análise de rede todos os dados. O *i2 Analyst's Notebook* é um ambiente de análise de inteligência visual que pode otimizar o valor de grandes quantidades de informações recolhidas por agências governamentais e empresas. Com um *design* intuitivo e contextual, permite aos analistas agrupar, analisar e visualizar rapidamente dados de diferentes fontes, reduzindo o tempo necessário para descobrir informações-chave em dados complexos. Uma das funções mais importantes deste programa é importar dados de múltiplos conjuntos e produzir visualizações que ajudam a revelar conexões que de outra forma não seriam feitas. Este tipo de *software* permite que as agências policiais tenham acesso a grandes quantidades de dados públicos, ao mesmo tempo que economiza recursos e tempo. No fundo, reduz o tempo necessário para reconhecer informações fundamentais para a criação de uma rede sólida e oportuna que permite a identificação, previsão, prevenção e combate ao terrorismo e a atividades criminais e fraudulentas (Kubler, 2017). No programa, os algoritmos *datamine* fornecem visualizações da informação relevante; no entanto, é necessário um *analista humano* que decida se se deve, ou não, tomar uma medida ou atuar sobre o alvo. O nível em que este *software* constitui policiamento preditivo é baseado em vários fatores, sendo que os algoritmos mais conhecidos são uma tentativa de substituir o papel do analista na avaliação dos resultados do processo de mineração de dados (Kubler, 2017; Mittelstadt, Allo, Taddeo, Wachter & Floridi, 2016).

2.4. EUROPOL

Drewer e Miladinova (2017) descrevem a aplicabilidade prática do *Big Data* nas atividades da Europol. Os autores (2017) analisaram documentos de alterações legislativas, regulamentos e protocolos. No que concerne ao contexto da aplicação da lei, os grandes dados estão relacionados com uma nova abordagem de recolha de informações que facilita, estrategicamente, as investigações, visto que as fontes de recolha de dados são diversas e múltiplas. Numa escala mais ampla, a Europol, em algumas áreas específicas, particularmente o cibercrime e o terrorismo, reconhece os grandes dados como importantes, já que permitem obter um melhor perfil, que é usado por criminosos e polícias para identificar alvos potenciais. Portanto, a Europol reconhece atualmente a importância de um modelo proativo, ágil e adaptável, englobando o uso de novas tecnologias e grandes dados para apoiar o trabalho da aplicação da lei

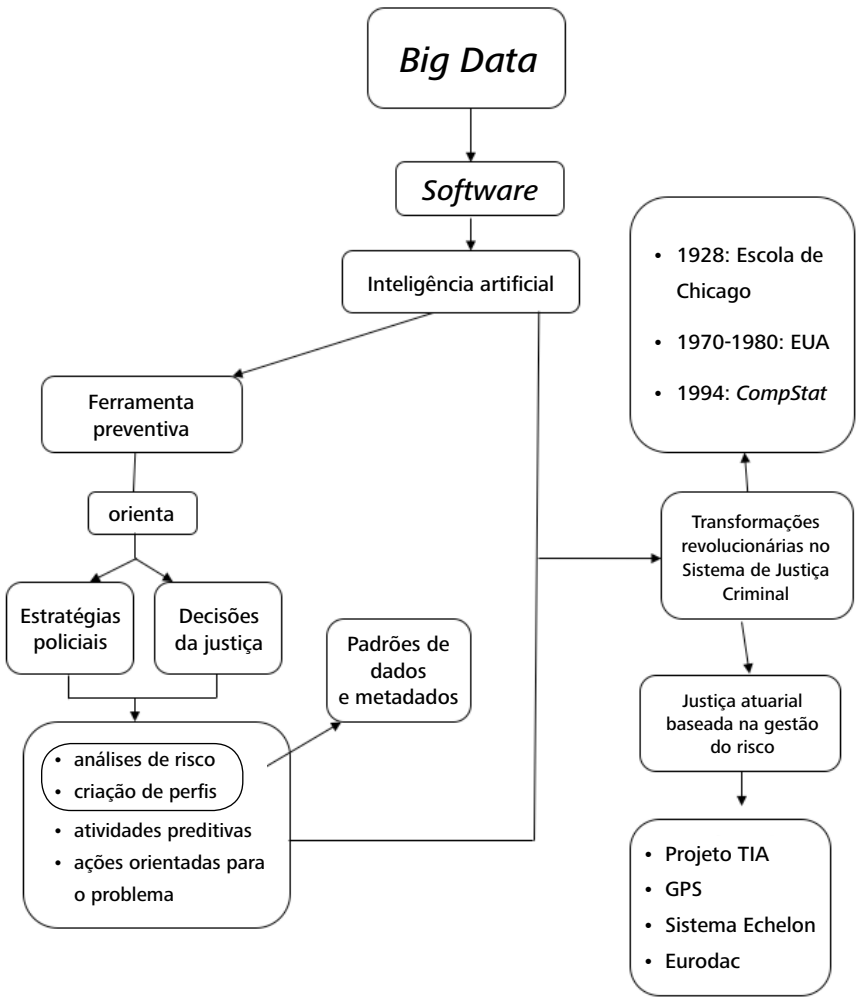
(Drewer & Miladinova, 2017; Wood *et al*., 2006; Johnson & Williams, 2007). No que toca à atividade policial transnacional, este tipo de vigilância tem um sentido preventivo e preditivo – congrega dados, analisa as suas informações e procura desencadear ações que travem atos criminais transfronteiriços e terroristas (Lyon, 2004; Johnson & Williams, 2007; Van Dijk, 2014; Wood *et al*., 2006; Matos *et al*., 2016).

No contexto do novo processamento de dados, mais avançado e modernizado, a Europol enfrenta um fluxo de grandes quantidades de informação que podem ser usadas para analisar informações estratégicas e melhor compreender a dinâmica do crime. A Europol aumentou as suas capacidades e tem agora a oportunidade de construir uma plataforma de informação capaz de facilitar uma operação mais eficaz e eficiente, bem como implementar respostas estratégicas a ameaças de segurança transfronteiriças. Atualmente, a Europol prevê uma tecnologia inovadora habilitada que interconecta mais de quinhentas Agências Policiais da Europa, utilizando constantemente os seus bancos de dados, por via de canais de comunicação, e oferecendo recursos rápidos e seguros para armazenar, pesquisar, visualizar e vincular informações. Consequentemente, também definiu altos padrões de privacidade, proteção e segurança de dados, visto que estas inovações exigiram uma adaptação regulamentar e legislativa por parte da Europol que, desde maio de 2017, realizou várias alterações no seu Regulamento. Estas alterações visaram salvaguardar medidas de segurança e mecanismos que devem ser usados para garantir uma maior proteção de dados pessoais. Qualquer processamento de dados pessoais dentro da Europol deve ser explicitamente permitido e os requisitos aplicáveis devem ser claramente definidos dentro do mandato atual da Agência – Decisão do Conselho de 6 de abril de 2009. Assim, como o trabalho efetivo da polícia exige alta inteligência, os desafios legislativos na área da aplicação da lei levaram à criação de um quadro jurídico que equilibra os interesses fundamentais de liberdade, proteção de dados e segurança. O conjunto de regras definido na Decisão do Conselho da Europol tem em conta as necessidades operacionais da agência e o direito do indivíduo à proteção adequada dos seus dados (Drewer & Miladinova, 2017).

Estando no cerne da inteligência criminal, a Europol reconhece que o rápido ritmo de desenvolvimento da tecnologia da informação exige a melhoria contínua de processos e ferramentas analíticas, tais como sistemas operacionais e mapeamento criminal. Especificamente, o Regulamento da Europol aborda o futuro da privacidade digital num contexto de tecnologia neutra. Este documento ajusta as regras existentes na direção e no ritmo das mudanças, incluindo também desafios exponenciais colocados por grandes quantidades de

dados. Sob o novo regime jurídico, a Europol passa a estar mais bem equipada para detetar rapidamente as ligações entre investigações e *modus operandi* de diferentes grupos criminosos, verificar *cross-matches* de dados e ter uma visão clara das tendências, garantindo um alto nível de proteção de dados pessoais (Drewer & Miladinova, 2017).

Figura 3: Relação entre *Big Data* e criminalidade.



PARTE III

TRAÇAR EXPECTATIVAS: AS NARRATIVAS SOBRE O *BIG DATA*

CAPÍTULO 1. CONSIDERAÇÕES METODOLÓGICAS

Considerando a escassez de trabalhos europeus direcionados para o *Big Data* enquanto potencial ferramenta de investigação criminal, esta obra assenta num estudo exploratório qualitativo. O seu principal objetivo é apresentar os resultados da análise qualitativa das expectativas sociais acerca da potencial aplicabilidade do *Big Data* nas investigações criminais europeias, por parte dos pontos de contacto nacionais em rede transnacional de cooperação policial e judiciária, bem como de geneticistas forenses e *stakeholders* de diferentes áreas (ética e regulação, investigação criminal, pesquisa universitária, empresas privadas e organizações não governamentais).

Considerando a complexidade do fenómeno objeto de estudo e a transversalidade da questão, e visando enfatizar as conceções subjetivas expressas pelos entrevistados – enquanto profissionais-atores sociais –, esta investigação ancora-se num pendor metodológico essencialmente qualitativo. Este método de investigação permite elaborar conceitos teóricos mediante a realidade onde estes surgem, em relação permanente com a empiria (Charmaz, 2009). Sendo assim, utilizando a *Grounded Theory* (Strauss & Corbin, 1990; Clarke, 2005) como âncora epistemetodológica – a construção de teorias a partir dos dados (Charmaz, 2009) –, nesta obra pretendeu-se realizar paralelismos constantes, articulações, ligações e comparações entre a recolha e a análise e entre a teoria e a realidade empírica (Strauss & Corbin, 1990). Visou-se, desta forma, apreender a pluridimensionalidade que caracteriza este objeto de estudo e privilegiar as perspetivas socialmente construídas pelos diferentes atores sociais que se relacionam profissionalmente com a temática empírica. Esta opção teórico-metodológica, aliada ao condicionalismo de apenas ser possível subtrair dos estudos pré-existentes conceitos sensibilizadores (Charmaz, 2009: 34) e não

quadros teóricos e hipóteses de trabalho gerais, implica adotar estratégias de pesquisa que possibilitem associar a construção de hipóteses e a elaboração de conceitos teóricos às circunstâncias específicas da realidade empírica localmente situada. Assim, adotaram-se os pressupostos da *Grounded Theory* (Strauss & Corbin, 1990), que viabilizam a construção de teorias "fundamentadas" nos próprios dados com elevado valor heurístico (Charmaz, 2009).

Este estudo valoriza, assim, a representatividade sociológica de cada caso, cada um afigurando-se como generalizável a enunciações teóricas e não a populações ou universos (Yin, 1994: 10; Nunes, 1992: 247-248; Brandão, 2010: 45-46). Além disso, dada a fragmentação do conhecimento existente acerca do fenómeno sob análise no contexto transnacional de dados de ADN, a abordagem indutiva foi considerada a mais apropriada por permitir que os dados informem a teoria (Elo & Kyngäs, 2008). Desta forma, o fenómeno foi estudado num contexto específico e adequado ao seu uso (Cho & Lee, 2014) – sob a alçada do Projeto EXCHANGE, que explora as dimensões sociais, culturais, éticas, regulatórias e políticas do uso de tecnologias de ADN para uso forense na União Europeia (UE).

De forma a concretizar este intuito, foram tecidas questões que visaram nortear o trabalho desenvolvido, bem como completar e densificar a temática em estudo: que expectativas em torno das potenciais aplicações do *Big Data* constroem os diferentes profissionais? Quais as suas perceções de potenciais benefícios e riscos? Que expectativas são coletivamente partilhadas entre diferentes categorias profissionais e quais as específicas de diferentes olhares profissionais? Estas questões emergem da empiria e das necessidades científicas a explorar sobre o tema. Tendo em conta que existe um conhecimento empírico diminuto neste campo em contexto europeu, este é um estudo exploratório que foi delineado de forma gradual. Foi utilizada, para o efeito, uma técnica de recolha de dados: análise de conteúdo qualitativo de entrevistas realizadas a profissionais com perfis diferenciados e com conhecimentos relevantes no que concerne ao *Big Data* e à partilha transnacional de dados de ADN.

A conceção do guião de entrevista foi elaborada pela coordenadora do Projeto EXCHANGE, Professora Doutora Helena Machado, em estreita colaboração com membros doutorados da equipa. Em simultâneo, a coordenadora elaborou um formulário de consentimento informado e um folheto informativo para que os entrevistados tivessem conhecimento da investigação em apreço. Finalmente, todas as entrevistas foram realizadas por diferentes membros da equipa, sendo a sua transcrição realizada por profissionais externos. Desta forma, o meu papel pessoal e individual cingiu-se à análise de conteúdo das

entrevistas sob a perspetiva de representações sociais sobre o *Big Data* com o objetivo de realizar a investigação que esta obra descreve. Todas as entrevistas foram conduzidas em língua dominada pelo entrevistado/a, maioritariamente em inglês (87), mas também português (28), espanhol (3) e alemão (6). As entrevistas em espanhol e alemão foram traduzidas, por profissionais externos, para inglês. Para efeitos de redação do livro, todos os extratos de entrevistas em inglês foram traduzidos para português. Estas traduções procuraram respeitar integralmente o conteúdo e sentido expressos pelo entrevistado/a.

A questão acerca do *Big Data* foi colocada aos entrevistados enquadrada no tema dos desenvolvimentos tecnológicos que têm vindo a ocorrer nos meios de investigação criminal e dos seus potenciais riscos e benefícios na prática. Os participantes foram questionados a este respeito, procurando-se em primeiro lugar perceber se tinham algum tipo de conhecimento sobre a temática. Posteriormente, foram exploradas questões relativas às suas expectativas e perspetivas no que toca aos benefícios e desvantagens da sua prática, com o objetivo também de entender se já tinham manuseado a técnica. Ou seja, para entender a forma como os entrevistados concebiam o *Big Data,* procurou-se explorar as suas representações sociais acerca da capacidade e valor daquele para o seu trabalho. Concretamente, foi perguntado aos participantes: i) qual a sua opinião relativamente aos usos potenciais que o *Big Data* proporciona?; ii) qual a sua opinião em relação aos potenciais riscos e benefícios associados ao uso de informações biométricas disponíveis e futuras como o *Big Data*? As respostas transcritas a estas questões foram analisadas minuciosamente e constituem o objeto de estudo – o fragmento empírico que possui os dados para formar a teoria.

Após uma delimitação do eixo temático – o *Big Data* –, foram estabelecidos *recortes discursivos* que caracterizam as regularidades do discurso apresentado pelos entrevistados, sempre confrontados com os sentidos heterogéneos destas similaridades (Caregnato & Mutti, 2006). O seu discurso foi analisado tendo em conta a linguagem adotada (Brown & Michael, 2003). Assim, foi feita uma análise categorial temática (Caregnato & Mutti, 2006), sendo que as categorias foram criadas em consonância com os temas que emergem na literatura científica sobre o tema em apreço; no entanto, resultam de uma análise de conteúdo qualitativo das entrevistas.

1.1. PERFIL DOS PARTICIPANTES

O material empírico analisado foram 124 entrevistas a 140 entrevistados alocados laboralmente em 25 países da UE (Alemanha, Áustria, Bélgica, Bulgária, Chipre, Eslováquia, Eslovénia, Espanha, Estónia, Finlândia, França, Holanda, Hungria, Irlanda, Letónia, Lituânia, Luxemburgo, Malta, Noruega, Polónia, Portugal, Reino Unido, República Checa, Roménia e Suécia), realizadas no âmbito do Projeto EXCHANGE (2015-2019). Estas entrevistas foram realizadas a órgãos de cooperação transnacional, profissionais diretamente conectados com a partilha transnacional de dados na UE, atores envolvidos na implementação e operacionalização da rede para o intercâmbio de dados de ADN, geneticistas forenses, juristas, professores/investigadores, investigadores criminais, profissionais que trabalham em empresas especializadas em proteção de dados e privacidade e demais pesquisadores com perfis diferenciados e com conhecimentos relevantes no que concerne ao *Big Data* e à partilha transnacional de dados de ADN. Estes entrevistados são perspetivados enquanto atores sociais, membros de comunidades científicas e, por isso, inseridos numa teia de significados específicos pertencentes a essa mesma comunidade, ainda que porosa e permeável ao exterior (Albert, Laberge & Hodges, 2009).

Concretamente, foram considerados Pontos de Contacto Nacionais (*National Contact Points)* – NCP, nos *steps* 1 e 2 – profissionais designados por cada Estado-Membro, regidos pelas normais aplicáveis do Direito nacional (Decisão 2008/615/JHA), com a função de fornecer dados, segundo o Regulamente da UE do Sistema de Prüm[1]. O relatório mais recente sobre o progresso da implementação do Sistema Prüm, datado de setembro de 2019, indica que há 25 Estados-Membros da UE operacionais ao nível da troca de ADN. No âmbito do presente estudo, foram conduzidas 37 entrevistas com 48 profissionais em 22 países operacionais no âmbito do Sistema de Prüm. Estes profissionais forenses, que desenvolvem funções como NCP, são figuras centrais no regime de Prüm, visto que protagonizam as atividades diárias que permitem o intercâmbio transnacional de dados e são cruciais nos processos de tomada de decisão. Em particular, os NCP têm como responsabilidade organizar e implementar os procedimentos e conexões necessários para: realizar trocas automáticas de dados com outras bases de dados (enviando e recebendo informações);

[1] O Sistema de Prüm foi assinado em 27 de maio de 2005 em Prüm. Este Sistema define um quadro legal que visa o desenvolvimento da cooperação entre os Estados-Membros da UE, no domínio da luta contra o terrorismo, a criminalidade transfronteiriça e a imigração ilegal (Machado & Granja, 2018).

realizar testes com outros países parceiros; e gerir e reportar correspondências de ADN. Estes papéis e as consequentes responsabilidades dos NCP de Prüm podem variar entre países, de acordo com diferentes estruturas organizacionais e legislação nacional (Machado & Granja, 2018). Para além disso, estes países atribuem a custódia das bases de dados nacionais de ADN a entidades distintas, desde autoridades judiciais até forças policiais (Santos, 2017, *cit. in* Machado & Granja, 2018). Portanto, os indivíduos que operam como NCP podem ter formação profissional e educacional variada, podendo desempenhar funções laborais em diversos contextos, desde laboratórios forenses a forças policiais (Machado & Granja, 2018). Conforme mencionado anteriormente, cada país tem autonomia para definir como atribuir diferentes papéis e responsabilidades aos profissionais forenses que atuam como Pontos de Contacto Nacionais do Sistema Prüm. Como tal, em alguns países apenas uma pessoa é responsável pelas operações associadas à troca transnacional de dados de ADN, enquanto noutros duas ou mais pessoas podem estar envolvidas nessas tarefas.

Para além das entrevistas conduzidas com NCP, foram também analisadas as entrevistas realizadas a 48 profissionais da genética forense na UE. Tendo em consideração a diversidade da comunidade de genética forense (Cole, 2013), a amostra agrega: geneticistas forenses, indivíduos que trabalham em casos criminais e são empregados por um laboratório forense; e investigadores, indivíduos empregados por universidades ou institutos forenses cuja ocupação profissional principal é a pesquisa de laboratório cientificamente controlada com aplicação no âmbito da Ciência Forense. Não raras vezes, estas posições são intermutáveis na medida em que profissionais da genética forense na UE podem, simultaneamente, desenvolver atividades de investigação e aplicar Ciência Forense em casos criminais específicos.

Por fim, foram ainda analisadas as entrevistas realizadas a 44 *stakeholders* (profissionais de Organizações Não Governamentais de Direitos Humanos ou Ciência, entidades de supervisão/regulação, órgãos de investigação criminal, empresas privadas e meios de comunicação social, bem como professores universitários/investigadores e legisladores) relevantes no âmbito de Prüm nos cinco países-alvo de estudo de caso do Projeto EXCHANGE: Alemanha, Países Baixos, Polónia, Portugal e Reino Unido[2].

[2] A análise destes estudos de caso visou, por um lado, compreender as diferenças entre países com diferentes posicionamentos em relação à troca transnacional de dados ADN, no contexto de Prüm. Por outro lado, almejou explorar as formas como diferentes legislações nacionais, recursos científicos variados, distintas tradições ao nível da utilização de Ciência Forense e desiguais posicionamentos ao nível do envolvimento político com o Sistema Prüm produzem configurações específicas que

impactam de forma significativa as dimensões sociais, culturais, éticas, regulatórias e políticas do uso de tecnologias de ADN para aplicação forense na UE. Os critérios que conduziram à escolha destes cinco países foram os seguintes: i) operacionalidade no âmbito do Sistema Prüm ou não; data de entrada na UE e situação relativa ao espaço Schengen; ii) legislação nacional relacionada com o banco de dados de ADN forense e possível impacto de Prüm; iii) nível de desenvolvimento da base de dados de ADN e proporção da população incluída; iv) condições socioeconómicas distintas e capacidade operacional, logística, tecnológica e científica desigual dos laboratórios e das tecnologias de informação e comunicação; v) debates públicos (existentes ou inexistentes) relativos à expansão do uso de tecnologias de ADN para fins de investigação criminal (fonte: Projeto EXCHANGE 2015-2019).

CAPÍTULO 2. LENTE CONCEPTUAL ANALÍTICA: SOCIOLOGIA DAS EXPECTATIVAS

Considerando que o *Big Data* é uma técnica ainda em fase emergente de aplicação no campo da investigação criminal europeia, prevalecem expectativas comummente partilhadas entre diferentes grupos profissionais e, em simultâneo, expectativas diferenciadas consoante o posicionamento profissional. Tal como prevê a *Sociologia das Expectativas*, cada profissional tem uma expectativa diferente da inovação e do progresso científicos, de acordo com a sua área académica de formação e o seu lugar contextual de trabalho. As expectativas são afirmações sobre o futuro e os elementos construtivos deste, que permitem reunir visões sobre o progresso inovador tecnológico, em fase embrionária e precoce de implementação. Assim, à margem deste olhar analítico, as diferentes expectativas dos entrevistados, posteriormente reunidas, foram analisadas e delas emergiram subjetividades – objetos de estudo e interpretação. O *Big Data* é uma técnica em fase de desenvolvimento no âmbito da investigação criminal europeia, não se encontrando ainda validada cientificamente para ser implementada como ferramenta de investigação criminal no seio da partilha transnacional de dados de ADN na UE – contexto onde os profissionais entrevistados estão inseridos.

A linha de interpretação dos dados é a da *Sociologia das Expectativas*, na medida em que o *Big Data* se rodeia de expectativas de uso orientadas para o futuro na área laboral dos entrevistados. Deve ter-se em conta que este eixo teórico central se relaciona com tecnologias que estão em fase embrionária de aplicação, onde se insere o *Big Data* no seio da investigação criminal. Nos últimos anos, um número crescente de estudos das Ciências Sociais ressalva a importância das expectativas em torno da inovação científica e tecnológica, sendo que inúmeros autores se têm debruçado sobre o papel das expectativas

na formação e mudança científica tecnológica (Borup, Brown, Konrad & Van Lente, 2006; Brown, Kraft & Martin, 2006). Em termos científicos dos desenvolvimentos tecnológicos, os atores contínua e explicitamente referem-se ao que será possível no futuro (Van Lente, 2012), e as Ciências Sociais tornam-se cada vez mais permeáveis ao desenvolvimento tecnológico inovador e à forma como este se tem processado (Hedgecoe & Martin, 2003). Além disso, a análise destas expectativas será tanto mais importante quanto mais embrionária for a fase de desenvolvimento da tecnologia, por estar repleta de incertezas (Brown, Rip & Van Lente, 2003; Selin, 2008) – como é o caso do *Big Data*.

As expectativas desempenham um papel construtivo no desenvolvimento da Ciência e das novas tecnologias. As antecipações de hoje são fundamentais para a construção do futuro (Brown *et al.*, 2003), e a compreensão dos diferentes discursos ancorados neste eixo teórico possibilita o acesso à coconstrução de visões científicas particulares do desenvolvimento tecnológico, nomeadamente o *Big Data*, pelo que se analisa a evolução desta tecnologia em termos sociológicos. A literatura acerca da *Sociologia das Expectativas* argumenta que a compreensão de expectativas, promessas, antecipações, visões e esperanças sobre fenómenos em desenvolvimento é central para aceder aos conceitos e ao seu futuro. A própria antecipação tem valor económico e epistémico, sendo que as reivindicações especulativas são fundamentais para os processos dinâmicos que criam novas redes sociotécnicas (Tutton, 2011). Visa-se, precisamente, entender a consciência futura a partir de previsões. A exploração deste tempo futuro permite o aprofundamento da responsabilidade pelo que se encontra em desenvolvimento, apesar de provocar tensões, fruto da incerteza. Assim, permite-nos entender como é que diferentes comunidades usam o conhecimento antecipatório e qual o papel das expectativas em torno das inovações (Selin, 2008).

As expectativas orientadas para temáticas futuras podem criar novas oportunidades, orientar atividades e fornecer uma estrutura, legitimando interesses, promovendo investimentos e antecipando potenciais riscos e benefícios. Tendo em conta que a probabilidade de a inovação se efetivar de forma isolada é diminuta, avaliar estas perspetivas num ambiente dinâmico, incluindo vários entendimentos e expectativas orientados para a futura aplicabilidade da técnica, pode aumentar a probabilidade de esta, posteriormente, se tornar uma prática regular. Assim, ancorados nesta perspetiva teórica e na ideia de que as expectativas e perspetivas dos atores mobilizam recursos, é importante analisar a sua dinâmica para compreender a mudança científica e tecnológica (Borup *et al.*, 2006).

Uma expectativa é uma mudança ou criação de uma nova realidade e é mobilizadora de algo; é uma afirmação sobre o futuro e pode ser positiva (promissora) ou negativa e variar a nível de conteúdo e/ou modalidade, desde esboços abstratos e abrangentes sobre o futuro (macro) a elementos detalhados (micro) (Van Lente, 2012). Em termos de conteúdo, as expectativas podem dizer respeito a aspetos sociais e, a maior parte das vezes, a um misto de aspetos. As modalidades podem variar entre declarações completamente permissivas, sem qualquer tipo de resistência, e argumentos subtilmente organizados para refutar as previsões (Van Lente, 2012). Estas expectativas – enquanto afirmações sobre alguma coisa e não descrições verdadeiras ou falsas (Van Lente, 2012) – servem, frequentemente, para unir ou mediar diferentes limites, dimensões e níveis, sendo fundamentais na coordenação de diversas comunidades e grupos de atores e podendo mediar vários níveis (micro, meso e macro) de organização. E toda esta panóplia de expectativas é maleável no tempo, enquanto resposta e adaptação a novas condições e/ou problemas emergentes (Borup *et al*., 2006).

Concretamente, as *expectativas tecnológicas* são representações em tempo real de futuras situações capacitárias tecnológicas, aproximando-se do *Big Data* e das expectativas que os diferentes profissionais têm acerca desta tecnologia. O *Big Data* surgiu, enquanto aparato tecnológico, como uma *promessa* para a resolução (mais) eficaz de casos criminais transfronteiriços, mas subsiste, no contexto em análise, enquanto ideal sem concretização prática, como um *futuro desejado* por alguns profissionais (Borup *et al*., 2006; Brown & Michael, 2003). Ancorados na perspetiva tecnológica, os estudos referem que existe uma noção culturalmente generalizada de que a tecnologia continuará a oferecer possibilidades para o progresso, apesar de este pertencer a um tempo futuro e de a distância temporal suscitar incerteza e receios (Borup *et al*., 2006; Brown & Michael, 2003; Selin, 2008). Além das variabilidades temporais – influenciadoras no processo de construção de expectativas em torno da tecnologia científica e sua evolução –, também as variabilidades sociais influenciam estas expectativas. Borup *et al*. (2006) e Van Lente (2012) referem que a tecnologia emergente dá origem a diferentes expectativas consoante a posição social dos atores: as expectativas surgem como mais autoritárias para aqueles que detêm pouca influência sobre o resultado da inovação tecnológica – o público, por exemplo; por outro lado, para aqueles que se situam no limbo entre estas duas posições – os intimamente ligados ao campo da prática da expectativa e os que conhecem o campo, mas não estão inseridos nele – por exemplo, as partes públicas interessadas, a incerteza é relativamente menor: aceitam a promessa desse futuro próximo expectável, mas não têm uma perceção dos detalhes que possam impedir a concretização

da expectativa (Van Lente, 2012). Um elevado senso de confiança reflete um distanciamento face às incertezas, e normalmente são as expectativas expressas por investigadores que conduzem as investigações sobre a área em que se insere a inovação tecnológica científica (Borup *et al*., 2006).

Para aqueles que se distanciam da prática quotidiana da (possível) técnica, as incertezas são diferentes, pois surgem ligadas ao desconhecimento; ou seja, pelo facto de não ser claro para estes profissionais o que a expectativa implica, eles são indiferentes à prática e, por isso, não têm certezas sobre ela. Por sua vez, os atores envolvidos diretamente no trabalho de campo sobre a tecnologia em evolução revelam expectativas contraditórias, ou seja, confiantes e cautelosas – estas últimas revelam uma flexibilidade interpretativa que impede que se proceda a uma padronização social devido ao facto de, frequentemente, dentro da mesma comunidade científica, se verificarem assimetrias no acesso às informações em que se baseiam as expectativas. Por exemplo, muitas vezes, as incertezas presentes nos discursos de geneticistas forenses que trabalham nos laboratórios são invisíveis para membros da comunidade política (Borup *et al*., 2006). Ou seja, para aqueles profissionais que dominam o campo onde o futuro próximo poderá acontecer, as incertezas são realidades quotidianas porque os atores conhecem os detalhes, as condições e as suposições (Van Lente, 2012; Lucivero, Swierstra & Boenink, 2011). Nesta linha de reflexão sobre a posição laboral e as expectativas, Selin (2008) refere que os cientistas sociais têm uma visão distinta de todos os outros profissionais: questionam quem está legitimado a fazer o quê, avaliam a questão dos que ganham e dos que perdem e desmistificam o que está a ser comprado, vendido e negociado. Os cientistas sociais tendem para o cultural, para as condições políticas e económicas, a partir das quais surgem os estudos futuros.

Brown e Michael (2003) propõem alguns parâmetros para uma melhor e mais próxima compreensão das expectativas em torno de inovações científicas: i) *grau de novidade técnica*, ou até que ponto uma inovação é percecionada como relativamente nova ou já enraizada nas práticas; ii) *incertezas organizacionais* que geram a procura de diferentes tipos de linguagem orientada para o futuro e discursos que refletem uma abstração temporal; e iii) *localização dos atores* relevantes dentro da rede de desenvolvimento da inovação – a sua posição laboral. Neste sentido, existem duas formas de construir expectativas acerca das inovações tecnológicas: uma delas prende-se com *retrospecting prospects* e a outra com *prospecting retrospects*.

A primeira consiste na lembrança de futuros passados ou da forma como o futuro foi em tempos representado, enquanto a segunda se debruça sobre a

forma como essas perspetivas são implantadas em tempo real contemporâneo, para construir o futuro. No que concerne ao *Big Data*, este objeto analítico aproxima-se da perspetiva dos autores na medida em que é uma tecnologia com atividade orientada para o futuro. Este exercício constante de aproximação do futuro, projetando-o no presente, cria uma pressão contemporânea e obriga a um exercício de abstração temporal constante. Este exercício – de projeção futura – torna ainda mais aparente a dualidade de discursos de risco e oportunidade. Ou seja, exacerba-se e compartilha-se a incerteza, ao invés de criar conhecimento sobre o futuro; apesar de estas representações sobre o futuro serem elementos importantes na construção do presente, o fracasso de muitos futuros projetados gera riscos e incertezas. Ainda que as expectativas assentem, em grande parte, em discursos, elas dependem da prática, que muitas vezes é inexistente; e os profissionais estão conscientes deste desfasamento entre a expectativa e o presente material. Por isso, é importante compreender esta dinâmica: aceder à forma como os atores contextualizam o futuro nas suas narrativas presentes, referindo futuros passados que fracassaram. Há uma tendência generalizada para explicar o sucesso e o fracasso indicando as propriedades da tecnologia. O que acontece é que uma excessiva centralização nas capacidades inovadoras tecnológicas potencia o esquecimento da vertente social – o palco onde os impactos são vividos e produzem consequências. O que devemos praticar é um exercício mútuo de covalorização da evolução da tecnologia e do desenvolvimento social. E tal exercício deve situar-se entre a inovação e as promessas deste desenvolvimento tecnológico, sem esquecer a necessidade de intervir em momentos-chave para que este desenvolvimento inovador não se torne prejudicial (Brown & Michael, 2003).

Lucivero *et al.* (2011) referem que, para interpretar as expectativas, importa explorar a forma como estas são construídas, como e porquê pelos diferentes públicos – profissionais –, situados em determinado ponto no tempo e no espaço e inseridos num fundo específico. Importa também mostrar algumas das características estruturais das expectativas. Paralelamente às conclusões alcançadas pelo estudo dos autores, encontramos, no que respeita ao *Big Data* e às expectativas expostas, três tipos de abordagem: i) abordagem sobre as características e funcionamento da tecnologia; ii) abordagem sobre a forma como a tecnologia será adotada pelos profissionais e integrada na prática atual; e iii) abordagem sobre como a tecnologia irá abordar um problema social ou uma necessidade. Além disso, os autores referem que as tecnologias futuras se baseiam no sucesso anterior da técnica, evidências científicas, dados científicos ou comunicação social pública. Neste caso, visto que o *Big Data* – enquanto

técnica embrionária na investigação criminal europeia – se encontra em fase de desenvolvimento, não existem evidências científicas, nem dados científicos e, por isso, a maior parte das expectativas é fruto do modo como é interpretado o que a comunicação pública social afirma sobre o *Big Data*.

Todo o futuro é previsto, pelo que surgem visões pessimistas e otimistas, esperança nos avanços futuros, bem como medos e imaginações sobre os riscos que estão intrinsecamente associados à mudança científica e tecnológica – são imaginários sombrios que fornecem material de ficção científica para melhor compreensão dos fenómenos. Normalmente, surgem discursos que antecipam a deceção, a incapacidade de corresponder às expectativas e cumprir as promessas, e igualmente a possibilidade de as esperanças não se realizarem. Qualquer visão, se manuseada e comunicada por atores confiáveis e suficientes, pode tornar-se realidade (Tutton, 2011; Gardner, Samuel & Williams, 2015). Tutton (2011: 416) refere que declarações sobre eventos e fenómenos do futuro que englobam termos como *pode, deve, poderia, espera, planeia, antecipa, acredita, estima, prevê, pretende, potencial, continuar* (ou a negação dos mesmos) são expectativas. Por vezes, os profissionais expressam uma notável ambivalência sobre o potencial impacto do uso de *Big Data*, tornando-se céticos nas suas alegações promissórias. Esta ambivalência representa a linguagem intelectual que permite aos profissionais navegar e trabalhar dentro de um terreno intermediário de um campo em evolução. Assim, surgem profissionais cautelosamente esperançosos e construtivamente céticos. Há os que consideram que o *Big Data* pode ser prejudicial e os que referem ser inútil; os que demonstram esperança e antecipações positivas, os que ficam entusiasmados com a perspetivas que os avanços tecnológicos lançariam sobre a prevenção e repressão criminal e os que demonstram desapontamento com essas novas tecnologias maravilhosas e desconforto com a propaganda especulativa em torno delas. Normalmente, em diversas áreas, os projetos de inovação surgem como resultado de um dinamismo omnipresente complexo que entrelaça expectativas positivas e negativas: é uma interação de promessa, esperança, otimismo, incerteza, pessimismo e ambivalência (Gardner *et al*., 2015). De facto, numa era em que a facilitação da tecnologia é uma realidade, existem perspetivas que demonstram que a implementação da técnica é um desafio ético (Gardner *et al*., 2015). Ainda assim, as expectativas tecnológicas são influentes no desenvolvimento de novos artefactos e conhecimentos. Muitas vezes estas expectativas resumem-se a incertezas tecnocientíficas (Pollock & Williams, 2010). Nesta perspetiva, o objetivo é identificar e analisar a robustez das expectativas expressas no que concerne ao *Big Data*.

CAPÍTULO 3. TRAÇAR EXPECTATIVAS: RESULTADOS EMPÍRICOS

Face à análise qualitativa de conteúdo das entrevistas, surgem teias complexas de significações e expectativas acerca do *Big Data*, potencialmente aplicável à investigação criminal europeia. Tendo em conta a diversidade de participantes profissionais considerados, bem como a heterogeneidade de narrativas registadas, foram criadas categorias temáticas gerais, que visam balizar, tematicamente, as significações expressas pelos entrevistados. São elas: i) (Des)Conhecimento e descrição técnico-profissional; ii) Desenvolvimento, expansão e antecipação de aplicação do *Big Data* na investigação criminal; iii) Perceção sobre os riscos e perigos do *Big Data*; e iv) Pareceres éticos e de direitos humanos.

3.1. (DES)CONHECIMENTO E DESCRIÇÃO TÉCNICO-PROFISSIONAL

Um número considerável de entrevistados não detém qualquer tipo de conhecimento acerca do *Big Data*:

> "*Big Data* não sei o que é." [NCP1]

> "Mas eu não sei… O que é *Big Data*? O que quer dizer com *Big Data*? Eu não tenho a certeza, não sei o que isso é…" [NCP1]

> "Que tipo de dados?" / "*Big Data* […]… Dados em massa?" [NCP2]

> "Não [nunca ouvi falar do *Big Data*]." / "E não sei, eu posso tentar perguntar aos meus colegas […], talvez eles trabalhem com alguma coisa relacionado com isso [*Big Data*], eu não sei." [NCP2]

> "Não sei o que é o *Big Data*." [Professor/investigador]

Este desconhecimento reflete, para além da inexistência de conhecimentos relacionados com a temática, o facto dos profissionais entrevistados não manusearem a técnica no seu contexto laboral.

Relativamente a este desconhecimento, Bartlett *et al.* (2018) e Halford e Savage (2017) referem que muitos profissionais não possuem ainda controlo sobre os dados nas Ciências Sociais. E, por isso, o seu conhecimento, apesar de existente, é extremamente reduzido. No entanto, o desconhecimento não é linear, variando consoante as diferentes posições laborais ocupadas pelos entrevistados. Os Pontos de Contacto Nacionais a desempenhar funções no *step* 1 (NCP1) em Institutos Forenses, quando questionados, revelam total desconhecimento sobre este aparato tecnológico e enfatizam a noção de que é algo que ultrapassa a sua margem de ação profissional. Além disso, não demonstram expectativas perante uma (possível) aproximação laboral com a técnica, nem perspetivam a possibilidade de desenvolvimento desta:

> "Eu não estou familiarizado com isso [*Big Data*], e acho que não vou estar num futuro próximo." [NCP1]

Os NCP1 revelam-se, a nível profissional e laboral, distantes relativamente aos desenvolvimentos tecnológicos que operam ao nível da investigação criminal. Sob a ótica de Borup *et al.* (2006), Lucivero *et al.* (2011) e Van Lente (2012), este discurso dos NCP1 pode classificar-se como autoritário, na medida em que estes profissionais não possuem conhecimentos detalhados sobre a temática do *Big Data* e procedem a uma desresponsabilização quanto à sua influência no desenvolvimento da técnica.

Noutra linha de reflexão, ao contrário do que se verifica nos NCP1, os Pontos de Contacto Nacionais alocados ao *step* 2 (NCP2) e os geneticistas forenses têm um (des)conhecimento ténue sobre esta temática, que lhes permite, de forma vaga, aceder ao conceito. Sabem explicitar as técnicas do *Big Data* e a forma como poderiam operar na prática investigativa, e um dos NCP2 que desempenham funções laborais numa Divisão de Cooperação Policial Internacional reconhece o potencial deste aparato tecnológico; no entanto, alega que não lhe parece ser a solução mais procurada pelos agentes policiais para a resolução de casos criminais, visto que crê tratar-se de uma técnica que exige recursos inexistentes na Polícia:

> "Eu penso que isso [*Big Data*] é uma solução [para resolução de casos criminais] que a Polícia tem e pode usar neste momento. […] Mas eu acho que eles não fazem uso dessas técnicas porque não têm dinheiro, […] eu nem sei quanto é que isso custa." [NCP2]

Os NCP2, na ótica conceptual teórica de Van Lente (2012), são uma categoria profissional que procede a esboços detalhados sobre o *Big Data,* na medida em que se aproximam, a um nível *micro,* das características singulares e particulares da técnica, nomeando-as. Na mesma linha de pensamento, um membro pertencente a uma Companhia Privada de Ciências Forenses refere que, além de se tratar de uma técnica possível e desejável no futuro, o *Big Data* é uma prática desafiadora a vários níveis: complexidade, economia, qualificação e mudança nas práticas policiais. Estes fatores tornam a técnica difícil de manusear, colocar em prática e implementar nos modos contemporâneos da investigação criminal. Não obstante, o entrevistado afirma também que, a ser implementado, o *Big Data* exigiria que o sistema policial se adaptasse a esta nova realidade, o que conduziria a alterações profundas nas suas atividades e dinâmicas:

> "[…] quanto mais complexo é o trabalho que temos para fazer, mais desafiador é implementar mudanças. […] os nossos cientistas precisam de ser extremamente qualificados e dispensam muito do seu tempo a lidar com as questões [relativas ao *Big Data*] que a defesa poderá levantar, em torno da contaminação, transferência e integridade do material de ADN, onde há apenas pequenas quantidades que foram identificadas [com interesse e relevo para a investigação]. […] algumas dessas questões e desses problemas ampliam-se e exacerbam-se quando se lida com as complexidades do sequenciamento da próxima geração no Sistema de Justiça Criminal […]." [Membro de Companhia Privada]

Estes entrevistados demonstram uma perspetiva dual: por um lado, o desejo expresso da aplicação material futura da técnica, devido ao seu potencial inovador e eficaz – esta é uma das características dos discursos sobre expectativas futuras, o desejo aliado à perspetiva positiva de aplicação da tecnologia (Brown & Michael, 2003). Por outro lado, num discurso cautelosamente esperançoso (Gardner *et al.*, 2015), salientam os obstáculos práticos à sua efetiva materialização, como a complexidade, os recursos económicos e mudanças que requer. Ou seja, segundo o último NCP2 e o membro pertencente a uma Companhia Privada parece haver uma desadequação laboral no que toca ao *Big Data* e às

atividades policiais. Cole (2013) referiu também o aspeto da demarcação que é feita relativamente à Polícia (*boundary work*) por parte de outros profissionais: consideram que a Polícia possui uma racionalidade diferente e, por isso, além de serem resistentes à mudança e de não possuírem condições para o desenvolvimento do *Big Data*, estes agentes também tiveram uma formação profissional que não os preparou para a inovação. Esta demarcação simbólica que é feita pelos NCP2 relativamente aos agentes policiais é evidente: um NCP2 reconhece o potencial positivo e eficaz do *Big* Data, mas considera que as forças policiais não possuem o arsenal económico e formativo, nem abertura inovadora para implementar a técnica; atribui à Polícia falta de conhecimento especializado e de investimento económico (Cole, 2013). Esta constatação assemelha-se às conclusões do estudo empírico de Chan e Moses (2017): os profissionais (re)conhecem a técnica, no entanto, não se aproximam da sua aplicabilidade prática, por considerarem que não existem recursos humanos, nem académicos, nem económicos que permitam a efetivação material do *Big Data*. Além disso, acrescentam o facto de estes polícias terem necessidade de expandir o seu conhecimento acerca da Genética Forense. Existe uma tendência geral, entre os restantes profissionais, de se demarcarem das racionalidades policiais, caracterizando estas últimas como diferentes, resistentes à implementação de novas técnicas como o *Big Data* e à inovação (Cole, 2013).

No entanto, ainda que NCP2 e geneticistas forenses desconheçam o fenómeno enquanto realidade laboral prática, partilham perspetivas da sua aplicação futura, ao contrário do que acontecera com o NCP1. NCP2 e geneticistas forenses revelam, sob a ótica de Van Lente (2012), declarações permissivas (com algumas resistências) acerca da aplicabilidade futura do *Big Data.* Por exemplo, genéticos forenses a desenvolver investigação em Universidades, apesar de se distanciarem do conhecimento da temática, consideram que o *Big Data* permite o tratamento de dados a uma grande velocidade:

> "Eu acho que a única coisa de que estou ciente neste momento é do sucesso rápido dos dados para fazer as coisas muito rápido num ambiente de imigração […] eu não estou muito conectado com isso [*Big Data*], para ser honesto." [Geneticista forense]

Apesar de o desconhecimento surgir como resposta nas diferentes camadas profissionais, existem diferenças ténues, não como fronteiras estanques, mas graduais: os NCP1 não se aproximam da temática e não revelam perspetivas futuras de aproximação, nem conhecimento. Na ótica conceptual de Van Lente (2012), os NCP1 tecem esboços abrangentes e abstratos, a um nível *macro*,

acerca do *Big Data*. Ao contrário, os NCP2 e os geneticistas forenses, apesar do conhecimento reduzido, completam as suas respostas com perspetivas de evolução e desenvolvimento da técnica do *Big Data*, aproximando-se dos detalhes da técnica, projetando-os e defendendo a sua implementação, e ainda adotando um discurso promissor quanto à temática (com elementos repletos de detalhes caracterizadores da técnica) (Van Lente, 2012).

Em contrapartida, vários são os operadores do Sistema de Prüm que revelam possuir conhecimentos acerca das questões técnicas e procedimentais por via das quais o *Big Data* se materializa. Os vários setores laborais considerados apresentam consensualmente conhecimentos que se ligam à operabilidade do *Big Data*, apesar de alguns profissionais surgirem com respostas mais relacionadas com os aspetos técnicos e outros apresentarem conhecimentos mais descritivos. A globalidade das expectativas inclui considerações técnico-descritivas sobre o fenómeno dos grandes dados. Estes discursos, transversais a qualquer categoria profissional, apelidam-se de *modernistas* e caracterizam-se por incluírem considerações acerca da conceptualização do objeto de estudo – neste caso, o *Big Data* –, da definição da temática, dos seus benefícios e aspetos positivos. Assim sendo, nesta temática geral, os discursos dos diferentes profissionais entrevistados podem caracterizar-se como *modernistas* (Stevens, Wehrens & de Bont, 2018).

Os NCP1 a desempenhar funções profissionais num Instituto de Neurologia e Genética consideram que o *Big Data* é uma técnica possibilitada pelo surgimento do sequenciamento em massa, visto que neste momento tudo é possível:

> "[O *Big Data*] poderia ser possível com o sequenciamento paralelo em massa. Agora tudo é possível." [NCP1]

Ou seja, com o desenvolvimento tecnológico atual, é possível que se criem condições para que surjam novas técnicas e novos métodos de investigação, mais sofisticados e complexos. Quanto ao sequenciamento em massa possibilitado pela digitalização em massa, Brayne (2017) refere que toda esta panóplia digital e tecnológica facilita a fusão e a partilha de registos através de instituições, tornando o armazenamento e o processamento mais fáceis, ao mesmo tempo que os dados se tornam mais eficientes para analisar e pesquisar e a análise em rede – o *Big Data* – conhece novos desenvolvimentos.

Noutra esteira, a caracterização do *Big Data* como um método preventivo e/ou proativo é evidente nas perspetivas dos NCP1. Alguns destes NCP1, a exercer funções laborais num Instituto de Criminalística ou a desempenhar

funções policiais, perspetivam a capacidade preventiva do *Big Data* como um dos maiores benefícios destes métodos – a capacidade de prever comportamentos e/ou prevenir situações danosas à sociedade, por via da agregação de dados, converte esta prática num método desejável para antecipar ações e encontrar suspeitos:

> "[…] se nós tivéssemos esses métodos [*Big Data*], poderíamos prever comportamentos [criminais] […]." [NCP1]

> "[…] graças a esse sistema, poderíamos, por exemplo, apreender os suspeitos ou perpetradores [de atos criminais]." [NCP1]

> "[O *Big Data*] deve ser preditivo e [deve ser visto como] uma ajuda [para a resolução e investigação de atos criminais] […]." [NCP1]

Sob uma perspetiva mais teórico-descritiva, sem enquadrar o *Big Data* na sua prática laboral e projetando esboços abrangentes e abstratos (Van Lente, 2012), os NCP1 tecem considerações eminentemente teóricas acerca do tema. Ou seja, consideram que uma das valências positivas do *Big Data* é a sua potencial capacidade para efetuar previsões que permitam, se eficazes, resolver de forma válida casos e investigações criminais. No que concerne a esta capacidade preventiva da técnica e quanto à forma como pode o *Big Data* operar na prática, Chan e Moses (2017) consideram que é precisamente desta valência proativa que as agências policiais fazem uso: prevenção criminal por meio de previsões sobre lugares, intervalos de tempo e/ou pessoas. Esta é uma perspetiva profissional pertencente apenas aos NCP1. Apesar de os NCP2 considerarem também esta valência preditiva e proativa das ferramentas do *Big Data*, não reconhecem, de forma semelhante, a centralidade desta característica, referindo-a apenas quando caracterizam a atividade da Europol. Desta forma, é evidente que, apesar de os NCP1 terem determinados conhecimentos relacionados com a temática, as suas perspetivas são ténues; e ainda que definam a técnica de forma descritiva, como o seguinte entrevistado, na sua maioria, consideram que se trata de uma ferramenta de investigação analítica ou de inteligência útil que não encontra facilidades na sua aplicação:

> "[O *Big Data*] pode ser uma fonte analítica ou de inteligência útil, mas não temos recursos para a usar […]." [NCP1, S03]

No que toca a este conceito, Brayne (2017) define *técnicas de inteligência* como o conjunto de atividades que se realizam antes da ocorrência de um facto criminal – incluindo desde logo a recolha dos dados, a identificação de suspeitos e de locais, de atividades e de pessoas – e que se materializam na intervenção preventiva. O mesmo entrevistado NCP1 enfatiza a ideia de que se trata de uma prática que exige recursos que o seu país não tem, à semelhança do que Chan e Moses (2017) referem: a maior parte dos órgãos policiais reconhecem o valor e a capacidade do *Big Data*; no entanto, consideram não ter tempo nem recursos para beneficiar da utilidade deste aparato tecnológico (Chan & Moses, 2017). Ou seja, os NCP1 demarcam-se do *Big Data*, apontando amplos obstáculos à sua aplicação e apresentando argumentos para refutar as previsões favoráveis à implementação desta técnica (Van Lente, 2012).

Em contrapartida, distinguindo-se dos restantes grupos profissionais considerados, os NCP2, a desempenhar funções profissionais na Direção Geral da Polícia Criminal e responsáveis pelas ligações operacionais, revelam ser detentores de um conhecimento particular ligado a esta temática que se prende essencialmente com a descrição específica e exemplificação prática do *Big Data.* Estes profissionais têm expectativas positivas e promissoras sobre a temática em apreço, revelando conhecimentos detalhados sobre a mesma, aproximando-se da sua realidade expectável e tecendo declarações permissivas quanto à sua implementação (Van Lente, 2012). Além disso, a um nível *micro*, um NCP2 descreve a técnica como facilitadora da partilha e (inter)conexão transnacional de dados e exemplifica a sua aplicabilidade prática no âmbito das atividades da Europol. Ou seja, estes profissionais nomeiam os benefícios e potencialidades do *Big Data* em contexto de partilha de um grande volume de dados. Um dos entrevistados NCP2 perspetiva o *Big Data* como uma ferramenta benéfica, precisamente por ter capacidades e funcionalidades que lhe permitem manusear um grande número de dados para serem partilhados e conectados com outras informações, por conjugar diferentes amostras e por conseguir alcançar resultados práticos, úteis e eficazes – benéfico para o contexto da investigação criminal:

> "[…] a longo prazo, eu vejo muitos benefícios se tivermos a possibilidade de conectar esses bancos de dados […]. [O *Big Data*] pode ser muito útil às vezes […]." [NCP2]

Na mesma linha de reflexão – heterogeneidade de informação e dados –, Lyon (2014) refere que é esta diversidade informativa que permite criar conhecimentos através do poder estatístico dos grandes números, que ajudam a

compreender os detalhes fragmentados das vidas individuais e auxiliam em determinadas funções policiais.

O mesmo entrevistado afirma que, em termos práticos, o *Big Data* é uma técnica que permite conectar todas as diferentes informações da Polícia, possibilitando a conexão com os casos de Prüm, ao nível do *step* 2. E tendo em conta as funcionalidades e capacidades do *Big Data*, acredita que esta técnica se trata de um projeto amplamente implementado pela Europol, enquanto entidade que prevê o intercâmbio de dados e informações com vista à prevenção e combate à criminalidade. Ou seja, na perspetiva deste profissional, o *Big Data* é utilizado como mecanismo de troca de informação privilegiado no contexto europeu, ainda que esteja em evolução. Portanto, a Europol, na perspetiva deste NCP2 a desempenhar funções profissionais na Direção-Geral da Polícia Criminal, responsável pelas ligações operacionais, legitima atualmente a importância de um modelo proativo, ágil e adaptável, englobando o uso de novas tecnologias e grandes dados para apoiar o trabalho da aplicação da lei:

> "Eu penso que esse projeto [*Big Data*] é utilizado e desenvolvido pela Europol, segundo o que eu ouvi. Eles [Europol] estão a fazer algo assim [*Big Data*] para conectar todas as informações da polícia sobre atos criminais, informações operacionais, para se conectarem com os casos de Prüm através de *matches* […]." / "[…] e sim, eu penso que a Europol está a lidar com isto [*Big Data*]." [NCP2]

Este NCP2 perspetiva o *Big Data* como uma técnica que possibilita a partilha de informação e a conexão de dados transnacionais, sendo um projeto implementado pela Europol para cumprir fins de investigação e prevenção criminal. Tal como mencionado pelo NCP2, no que concerne às atividades da Europol, Drewer e Miladinova (2017) referem que os grandes dados estão relacionados com uma nova abordagem de recolha de informações que facilita, estrategicamente, as investigações, visto que as fontes de recolha dos dados são diversas e múltiplas. Os autores, retomando a perspetiva do entrevistado, alegam que, numa escala mais ampla, e em algumas áreas específicas, particularmente o cibercrime e o terrorismo, a Europol reconhece os grandes dados como vetores importantes pois permitem obter um melhor perfil, que é usado tanto por criminosos como por polícias para identificar alvos potenciais. A Europol aumentou as suas capacidades e tem agora a oportunidade de construir uma plataforma de informação apta a facilitar uma operação mais eficaz e eficiente, bem como implementar respostas estratégicas a ameaças de segurança transfronteiriças (Drewer & Miladinova, 2017).

Esta perspetiva reflete o que Hedgecoe e Martin (2003) referiram quanto à forma como se concebem visões que legitimam os conceitos, justificando-os em bases racionais para o trabalho futuro – aqui foi exemplificada uma entidade com legitimidade (uma base racional) que faz uso do *Big Data* e que reflete a sua importância e utilidade futura na partilha transnacional de dados e na recolha de um número elevado de dados informativos variados. Os NCP2 são uma categoria profissional que revela, frequentemente, a consciência da utilidade futura da técnica, precisamente porque, desempenham funções laborais onde, futuramente, a técnica poderá vir a ser aplicada. Este discurso é designado *instrumental,* na medida em que os profissionais apontam exemplificações práticas da técnica no seu campo laboral e utilizam conceitos discursivos intimamente relacionados com a temática em apreço (Stevens *et al*., 2018). Este é um tipo de discurso frequentemente adotado pelos NCP2.

Além disso, relativamente à descrição técnico-característica do *Big Data,* excetuando os NCP2, todos os outros grupos profissionais considerados associam e referem nas suas respostas a questão dos algoritmos. Ou seja, nesta linha de representações e expectativas profissionais reflexivas de um conhecimento acerca da técnica, surgem entrevistados que perspetivam as técnicas do *Big Data* enquanto realidades algorítmicas que agregam dados do *Facebook*, biométricos, de ADN e de impressões digitais, considerando que facilitam e possibilitam a agregação de muitos e diferentes dados informativos:

> "[O *Big Data*] tem direções a seguir, algoritmos, informações recolhidas por via do *Facebook*, de dados biométricos, de ADN, de impressões digitais, de qualquer coisa, e, a partir daí, podem criar-se perfis corretos de indivíduos […]." [NCP1]

Na mesma linha reflexiva, no que respeita ao conhecimento técnico do *Big Data*, alguns geneticistas forenses a desempenhar funções laborais num Instituto Nacional de Polícia Científica referem tratar-se de um método que permite a tomada de decisões tendo por base algoritmos. Estes algoritmos potenciam a precisão das decisões e das posteriores ações para atingir os resultados pretendidos. Assim, o *Big Data* permite, por via dos algoritmos que inclui, a tomada de decisões mais precisas, segundo o seguinte geneticista forense:

> "Eu acho que [usar informações biométricas como o *Big Data*] é benéfico, porque aumenta a precisão do algoritmo […]." [Geneticista forense]

Ou seja, na perspetiva deste último entrevistado, trata-se de uma técnica que aumenta a precisão dos algoritmos. É um método que sofistica e aumenta a potencialidade das suas ferramentas, agregando variáveis raras, e que de outra forma não seria possível. Na aceção deste geneticista forense, o *Big Data* é uma técnica que permite a partilha de dados entre os diferentes países europeus, sendo esta uma valência que facilita a construção de um algoritmo completo, tornando o trabalho de equipa mais relevante e os resultados mais completos e precisos. Mittelstadt *et al.* (2016) referem que esta é a perspetiva dominante acerca dos algoritmos: há uma conceção generalizada de que os algoritmos facilitam a tomada de decisões, com grande potencial, e substituem análises e decisões humanas.

Também os restantes profissionais, nomeadamente os professores/investigadores com formação superior especializada em Economia Política, fazem esta associação entre *Big Data* e algoritmos – esta metodologia ancora-se em dados passíveis de serem convertidos em algoritmos, que facilitam a tomada de decisões.

> "A questão do *Big Data*, das decisões com base em algoritmos […]." [Professor/investigador]

Quanto à capacidade de o *Big Data* se sustentar em dados e algoritmos para descobrir padrões e associações, Gonçalves (2017) explica que esta ferramenta usa, precisamente, estes algoritmos para auxiliar na análise destes paradigmas, compreender e englobar o valor total dos dados para informar decisões, permitindo a identificação de padrões entre diferentes fontes e conjuntos de dados.

A possibilidade de as decisões de justiça criminal se ancorarem nos resultados alcançados por via do *Big Data* é abordada por Chan e Moses (2017) quando afirmam que o objetivo do uso e manuseio desta técnica é precisamente orientar estratégias policiais. Gonçalves (2017) acrescenta que existe uma expectativa em torno do *Big Data* enquanto ferramenta útil que pode levar a melhores e mais informadas decisões, tal como o último entrevistado referira.

Ainda sobre a descrição técnica do *Big Data*, o mesmo professor/investigador com especialização académica em Economia Política, define a técnica como a "ciência dos dados que oferece grande potencial de extrair informações a partir de *links*, com correlacionamento entre diferentes bases de dados", demonstrando familiaridade com os conceitos intimamente ligados à técnica e ao tema:

> "[…] aquilo que hoje se chama ciência dos dados […] este conceito de *data science* oferece – segundo […] os *data scientists* e quem está envolvido nestes setores das tecnologias de informação, engenharias informáticas, inteligência artificial, etc. – oferece grande potencial […] de extrair informação a partir de *links*, com correlacionamento entre diferentes bases de dados […]." [Professor/investigador]

Ou seja, este entrevistado perspetiva o *Big Data* como uma área científica metodológica que permite a recolha, o tratamento e o manuseamento de um grande volume de dados – de diferentes fontes –, correlacionando diferentes informações com vista a alcançar resultados úteis. Este entrevistado utiliza uma linguagem intimamente relacionada com a temática em apreço e, por isso, o seu discurso é *instrumental* (Stevens *et al.*, 2018).

A questão do armazenamento e manipulação de um grande volume de dados, geralmente associada ao *Big Data*, é também mencionada pelos NCP2 a desempenhar funções profissionais na Direção-Geral da Polícia Criminal, responsável pelas ligações operacionais. Concretamente, estes entrevistados referem que o número elevado de dados que esta técnica conjuga potencia a complexidade do seu aparato enquanto fenómeno. Desde logo, pela inovação tecnológica que constitui o *Big Data*, grande parte dos profissionais perceciona-o como um verdadeiro desafio no campo sociológico da investigação – não só um desafio, mas uma mudança que desperta interesses, receios, precauções e cautelas. Assim, a generalidade dos profissionais ressalvou sempre a sua preocupação pelo facto de englobar e trabalhar com um volume considerável de informação:

> "[*O Big Data*] são muitos dados […]." / "Como é que se armazenam tantos dados de ADN […]." [NCP2]

Também os NCP1, apesar de não revelarem conhecimento acerca da possibilidade de correlacionamento das bases de dados, nem nomearem conceitos intimamente ligados ao *Big Data* – não adotando, pois, um discurso *instrumental* (Stevens *et al.*, 2018) –, referem as questões ligadas à aplicabilidade contemporânea desta técnica. Concretamente, a tecnologização da vigilância, que se reflete, na sua ótica, na materialização do *Big Data* em ferramentas de controlo e vigia dos indivíduos. Tal como é possível compreender por via do discurso do seguinte NCP1 a exercer funções laborais na Polícia:

> "O *Big Data* é […] usado, por exemplo, para vídeo-informação […]." [NCP1]

Este NCP1 estabelece uma ligação entre o *Big Data* e as novas tecnologias, restringindo a materialização desta ferramenta aos meios audiovisuais. A recolha de informação por via de câmaras de vigilância é mencionada por vários autores e define-se como um tipo de vigilância pré-construtiva, em que o processo de recolha se desenvolve de forma aberta, percetível, dirigida, indiscriminada e a qualquer pessoa (Matzner, 2016; Fróis, 2007, 2015; Machado *et al*., 2018). Lupton e Michael (2017) apontaram a mesma dimensão pública no seu estudo acerca das perceções públicas sobre a vigilância através dos dados – é frequente esta associação entre dados, vigilância e circuitos de videovigilância. A associação entre *Big Data* e tecnologia, bem como a identificação desta técnica como uma ferramenta eminentemente tecnológica são frequentemente apresentadas nas respostas dos NCP1 a desempenhar funções laborais policiais num Instituto Forense. As respostas evidenciam-no: um dos entrevistados, quando questionado sobre o *Big Data*, direciona a sua resposta precisamente nesse sentido – os potenciais benefícios de manusear os dados por via desta técnica materializam-se na tecnologia, que é uma ferramenta que apresenta vantagens e desvantagens, consoante a finalidade do seu uso.

> "[O *Big Data*] depende de como é que os dados são usados. Porque qualquer tecnologia tem dois lados, um bom e um mau […]." [NCP1]

Esta perspetiva vai no sentido do que Matzner (2016) refere: o *Big Data* é uma ferramenta estritamente tecnológica que depende de *softwares* informáticos para se materializar. Além disso, o seu pendor eminentemente tecnológico maximiza o poder computacional e a precisão algorítmica para reunir, analisar, vincular e comparar grandes conjuntos de dados (Boyd & Crawford, 2012).

Apesar de os geneticistas forenses entrevistados também se referirem ao *Big Data* como uma tecnologia, caracterizam-no como uma prática presente na *internet*, sendo esta última fruto das novas tecnologias. Estes entrevistados, a desempenhar funções em Institutos Forenses, consideram que o *Big Data* é um *submundo* para o qual, diariamente, os cidadãos contribuem quando realizam pesquisas *online* reveladoras de preferências e interesses pessoais. Um dos entrevistados afirma que a *internet* facilita a cedência pessoal de dados informativos privados, o que beneficia o *Big Data*, apesar de este ter objetivos e finalidades distintos. Assim, na perspetiva do entrevistado seguinte, qualquer cidadão já contactou com técnicas do *Big Data*. Aliás, geneticistas forenses a exercer funções laborais num Instituto Forense questionam até que ponto nos devemos tornar tão restritivos em ceder informações para gerar os grandes

dados no âmbito da investigação criminal, quando o fazemos diariamente sem levantar questões:

> "Quando acedemos à *internet*, […] basta fazermos uma pesquisa, toda a gente faz isso [ceder dados pessoais]. Não vejo por que razão não haveremos de fazer em outro âmbito, como por exemplo no criminal. Porque é que havemos de ter tantos pruridos quando estamos a falar de resolver um crime [, quando] somos confrontados com essa troca de informação entre as bases de dados todas e mais algumas [diariamente na *internet*] […]." [Geneticista forense]

Os geneticistas forenses perspetivam que o *Big Data* já se encontra enraizado na sociedade e, por isso, o seu discurso é do tipo *pragmático* (Stevens *et al*., 2018). Ball *et al*. (2016) enfatizam precisamente este aspeto: as normativas que circulam neste meio da vigilância são inscritas através de práticas sociais de manipulação e o uso de dispositivos e características que alimentam grandes infraestruturas de dados. Assim, os dados vão sendo recolhidos e armazenados, alimentando as grandes infraestruturas de dados, nas configurações do dia a dia (Ball *et al*., 2016). Neste processo gradual, a vigilância converteu-se numa característica inevitável da vida quotidiana (Lyon, 2004; Brayne, 2017), estando presente em todos os setores sociais. Assim, a vigilância surge como um requisito de participação no mundo (Brayne, 2017): ao enviar um *e-mail*, realizar uma chamada telefónica ou uma pesquisa na *internet*, o indivíduo eterniza os seus rastreios digitais (Lyon, 2004; Brayne, 2017), tal como o último geneticista forense entrevistado referiu.

Os geneticistas forenses a desempenhar funções laborais em Institutos Forenses também consideram que a recolha de informação está já enraizada na malha social, tratando-se de uma forma muito importante de abordar o contexto social, ainda que acarrete benefícios e malefícios:

> "É claro que a análise de *Big Data* vai ser… Vai ser, não, já está a ser muito importante para a sociedade, com todos os benefícios e malefícios que isso tenha." [Geneticista forense]

Todas as perceções e expectativas dos entrevistados que desenham o *Big Data* enquanto realidade em desenvolvimento convergem no sentido prático de que esta técnica se envolveu nas malhas do tecido social, convertendo-se num elemento da sociedade que, sob a ótica de alguns profissionais, é alimentado e potenciado pelos cidadãos, responsáveis pela expansão, uso e implementação

da técnica. Concretamente, o que o seguinte NCP2 entrevistado, a desempenhar funções laborais num Centro de Análise Celular de ADN, pretendeu explicitar foi que, de facto, os cidadãos contribuem no dia a dia (in)diretamente para esta rede de dados crescente, cedendo as suas informações e sem o questionarem. Exemplificando, o mesmo entrevistado sustenta esta opinião debruçando-se sobre a realidade dos EUA:

> "Se formos aos EUA, cedemos a nossa identificação através da íris e das impressões digitais, e ninguém se importa se esse tipo de informação está numa rede de *Big Data* nos EUA. E não sabemos qual o tipo de administração que os EUA vão ter […] eu acho que 98% [das pessoas] não leem esse tipo de informação [sobre como os seus dados informativos serão tratados após a sua cedência voluntária] […]." [NCP2]

Ou seja, a recolha de informação é uma prática globalmente (re)conhecida, alimentada pela sociedade, sobre a identidade individual de cada cidadão, que cede os seus dados informativos sem questionar a sua finalidade, o lugar onde a informação vai ser armazenada ou os fins para que será usada. Consequentemente, a rede do *Big Data* expande-se de forma inquestionável. Assim, enraizou-se na sociedade a ideia de que integrar esta rede é uma obrigação social, pelo que os diferentes dados pessoais cedidos se acumulam em grandes bases de dados. O armazenamento desta informação permite que a rede dos grandes dados vá evoluindo. Quanto a esta cedência "pública" dos dados pessoais, por parte dos cidadãos, às empresas e bases, Boyd e Crawford (2012) e Frade (2016) alegam que o que se verifica é que o quotidiano é cada vez mais transparente para as grandes organizações, e as organizações envolvidas na vigilância são cada vez mais invisíveis para aqueles cujos dados são obtidos e usados. Ou seja, os cidadãos estão constantemente a transmitir informação – por exemplo, por via dos seus movimentos –, enquanto, de forma permanente, uma vigilância invisível regista e armazena essa informação sem o consentimento daqueles.

Assim, estes entrevistados – pertencentes a todas as categorias profissionais, excetuando os NCP1 – revelam uma visão de *normalização* da implementação da técnica do *Big Data* no quotidiano: esta é vista apenas como um avanço na tecnologia e, por isso, aparece contextualizada na vida quotidiana, neutralizando os riscos advindos da sua prática, implementação e propagação no que concerne à privacidade e proteção de dados.

A globalidade das expectativas profissionais aqui apresentadas culmina no consenso geral de que o *Big Data*, por se materializar por via da tecnologia

digital, por agrupar algoritmos numéricos, por congregar um grande volume de dados convertíveis em informações capazes de orientar ações de prevenção criminal, é um fenómeno com um alto nível de complexidade. Assim, a maior parte dos entrevistados que têm alguma informação acerca desta ferramenta perspetiva-a como uma questão, simultaneamente, geral e complicada: um fenómeno complexo e múltiplo.

> "Uau. Isso [o *Big Data*] é muito complexo." [NCP1]

> "Isso é uma questão muito complicada." [NCP2]

> "[…] é uma questão muito difícil quando se trata de grandes dados […]." [Membro de uma Organização Não Governamental de Direitos Humanos]

À exceção dos genéticos forenses e dos *stakeholders*, que revelaram perspetivas maioritariamente ligadas às questões preditivas e preventivas da técnica, bem como, aos algoritmos, os restantes profissionais entrevistados consideram tratar-se de uma questão complicada e desafiante a vários níveis: grande volume de dados, dificuldade em colocar em prática, exigência de mudanças policiais, formativas, contextuais e económicas. À semelhança do que Brown e Michael (2003) referem, é expectável que a emersão de uma tecnologia inovadora numa rede já constituída signifique mudanças dessa mesma rede, induzindo a alterações radicais – novas relações institucionais e profissionais. Quanto à aceção generalizada de que o *Big Data* é, efetivamente, um sistema tecnológico complexo e aparatoso, Chan e Moses (2017) acrescentam que, devido a este nível de complexidade, só a tecnologia o pode manusear; daí a sua dependência de *softwares* e redes digitais.

Por outro lado, vários são os operadores do Sistema de Prüm que possuem representações sociopolíticas quando questionados acerca do *Big Data*. Concretamente, os NCP1 foram a única categoria profissional que não abordou o *Big Data* sob uma perspetiva política. As restantes categorias profissionais, apesar de se distanciarem umas das outras, apresentam perspetivas políticas acerca do fenómeno. Nomeadamente, os NCP2 a desempenharem funções laborais num Centro de Análise Celular, tal como os geneticistas forenses, consideram que o uso e a expansão do *Big Data* são influenciados pelos ditames políticos, tornando-se dependente destes; aos decisores políticos cabe implementar, ou não, o uso desse tipo de sistema tecnológico. O *Big Data* e a sua crescente utilização constituem então um fenómeno político, sendo a Política

a responsável pela legitimação e expansão desta técnica. Assim, o *Big Data* depende dos interesses da Política (em termos de custo-benefício económico) e de quem administra:

> "Podemos ter uma opinião, mas é o mundo político que vai decidir o que vai acontecer com isso [*Big Data*]." [NCP2]

> "Eu acho que isso é uma questão política. Os políticos são quem decide se deve ser permitido ou não [o uso de *Big Data* para fins de investigação criminal], e uma vez que os nossos políticos concordaram, a lei de ADN foi ajustada […] e discutida enquanto decisão política e só poderá ser novamente redirecionada se um novo sistema político decidir que não é aconselhável usá-lo [*Big Data*] […]." [Geneticista forense]

Muitas vezes o desenvolvimento das tecnologias vem responder a ideários políticos, sendo direcionado pela política (Van Lente, 2012). Consequentemente, segundo a perspetiva de um NCP2 a desempenhar funções profissionais na Direção-Geral da Polícia Criminal, responsável pelas ligações operacionais, o *Big Data* é considerado uma arma analítica, sendo a política responsável pelo desenvolvimento e expansão destas técnicas digitais,:

> "[O *Big Data*] é uma arma analítica e Prüm também está sob essa alçada analítica…" [NCP2]

Ou seja, este é um tema com interesses políticos, muitas vezes baseados e sustentados em ideias de lucro económico-nacional, segundo a ótica dos geneticistas forenses e de NCP2. Além disso, é frequente os dados serem propriedade do Estado – enquanto entidade máxima de poder político –; por isso, depende da sua atitude a evolução das atividades em torno destes dados. Nestes casos, a entidade responsável pelo uso e manuseamento dos dados é o Estado, pelo que também é ele o responsável pelo seu (não) desenvolvimento. As autoridades estatais decidem sobre o acesso e uso dos dados (Prainsack, 2019). Assim sendo, na ótica destes entrevistados, o futuro do *Big Data* depende da opinião, dos valores, dos interesses e das decisões políticas; por outras palavras, defende-se uma responsabilização da Política no que toca ao *Big Data* (Van Dijk, 2014).

As perspetivas dos NCP2 a desempenhar funções laborais em Centros de Análise Celular e dos geneticistas forenses a desempenhar funções laborais

em Institutos Forenses assemelham-se no que respeita à dimensão política do fenómeno *Big Data*, visto que ambos os grupos defendem a ideia de que esta técnica está dependente da Política, por se tratar de algo negociável, passível de gerar lucro económico. Os dados, que podem ser comprados, convertem-se facilmente num negócio:

> "Aliás, [o *Big Data*] é um grande negócio […]. Quem consegue fazer bases de dados tem um grande negócio na mão, não é? Atualmente vendem-se e compram-se bases de dados exatamente porque as empresas querem isso." [Geneticista forense]

> "[O *Big Data*] é um grande negócio." [NCP2]

A propósito desta conceção de os dados serem parte de um negócio maior, Coll (2014) refere que a garantia de privacidade dos cidadãos se tornou um elemento fundamental para o bem-estar da economia social e, por isso, tudo aquilo que possa constituir uma ameaça deve ser feito de forma regulada, estrita e, muitas vezes, paga. Tendo em conta que a privacidade se tornou um bem de consumo, a partilha de dados é paga, como que se se tratasse do fluxo de um negócio. E a partir do momento em que a privacidade é vista como algo passível de ser manipulado pela economia, a proteção de dados e/ou o acesso a estes são vistos, consequentemente, como um negócio. Também Ball *et al.* (2016) apontam este aspeto da comercialização e negociação dos dados por parte das empresas. Os autores defendem que, a partir do momento em que as empresas perceberam que tinham muitos recursos de informações sobre padrões de compra de clientes, vendas, negócios, parceiros de cadeia e análises de sentimentos a partir de dados das redes sociais, começaram a utilizar isso a seu favor, como forma de rentabilizar o seu negócio. Assim, embora o uso de dados pelas corporações tenha sido um meio para atingir um fim, um mecanismo para gerar mais valor dos clientes, para algumas empresas (*Google*, *Facebook*, *Twitter* e outras), o armazenamento de dados tornou-se um "fim em si mesmo" (Ball *et al*., 2016; Gonçalves, 2018). Nesta linha, Boyd e Crawford (2012) afirmam que o Mercado perceciona o *Big Data* como uma oportunidade para segmentar publicidade e otimizar ofertas, e Lawless e Williams (2010) também mencionam esta *mercantilização* dos mecanismos de provisão científica. Por isso, são tecidas considerações comerciais acerca do tema em apreço, adotando-se uma racionalidade económica. Ainda segundo Lawless e Williams (2010), os geneticistas forenses – e, neste caso, também NCP2 – têm uma visão comercial do tema, o que denota a sua *identidade epistémica mutável*, já que

o uso do conceito de *negócio* reflete o facto de estes profissionais se tornarem permeáveis às estruturas discursivas em desenvolvimento.

Todos os outros profissionais, excetuando os NCP1, partilhando a mesma perspetiva política, isto é, partindo de um olhar politizado, assumem direções distintas. Como já explanado, os geneticistas forenses ancoram-se no facto de as decisões políticas – determinantes na expansão e desenvolvimento da técnica – influenciarem a negociação das bases de dados. Os NCP2 consideram que o poder político é o responsável máximo pelas decisões que se prendem com a implementação (ou não) do *Big Data*. E os professores/investigadores, nomeadamente os pertencentes ao ramo do Direito/Economia Política, ancorados também numa expectativa sociopolítica, consideram que é a atitude do sistema político que determina (ou não) os fins a que se destina a aplicação e materialização desta técnica. Ou seja, apesar de associarem o *Big Data* à Política, perspetivam-na como determinante na criação e expansão de conflitos e interesses económico-políticos, o que pode pôr em causa a proteção dos direitos, liberdades e garantias dos cidadãos.

> "[…] acho que o problema [do *Big Data*] […] tem mais a ver com a atitude do sistema político, pelo menos na Europa, que faz correr o risco de, perante estes dilemas e conflitos de valores, de interesses e de direitos, fazer prevalecer o interesse económico […] em desfavor das pessoas." [Professor/investigador]

Ainda assim, sob uma perspetiva sociopolítica, é notória a preocupação com as questões ligadas à proteção de dados, pelo facto de esta técnica englobar o uso massivo de informações do foro pessoal e privado dos cidadãos. O facto de ser algo passível de manipulação política exacerba os efeitos cautelares a adotar na recolha, armazenamento e partilha de um grande número de dados. Esta questão é maioritariamente enfatizada por membros do Corpo de investigação criminal:

> "[…] estamos agora num sistema político democrático em que se respeitam os direitos, liberdades e garantias. Agora imagina que […] o sistema político muda e passa a ser autoritário, e temos uma ditadura […] é sempre assustador se pensarmos quanto mais informação o Estado tiver…" [Corpo de investigação criminal]

Ou seja, é evidente a expectativa profissional de investigadores criminais de que o rumo a adotar na expansão e utilização destas técnicas será definido pelos governantes – expoentes máximos que ditam como se deve prosseguir

nesta área (semelhante ao que se analisara nos NCP2 e geneticistas forenses). O seguinte entrevistado perspetiva o *Big Data* como uma técnica possível devido à estruturação do sistema político vigente, ou seja, à defesa dos direitos, liberdades e garantias. Ainda assim, este entrevistado, ligado à investigação criminal, considera que tudo depende de quem detém este poder sobre os dados, pois é ele quem os manuseia:

> "[…] se a parte pública, quem está à frente dos desígnios de determinado Estado, já não estiver […] preocupada com a arquitetura dos direitos, liberdades e garantias, e estivermos perante um Estado autoritário, então isso aí […] estamos perante uma partilha de dados que pode ser apocalítica para o próprio." [Corpo de investigação criminal]

Os profissionais salientam, neste âmbito, a dificuldade política em encontrar um equilíbrio entre a defesa dos direitos, liberdades e garantias e a proteção de dados. E esta dificuldade reflete que quem detém o poder de equilibrar os dois setores detém o poder de garantir a proteção de dados. Esta entidade é a Política, segundo a ótica do seguinte entrevistado, professor/investigador em Sociologia do Direito:

> "[…] do ponto de vista legal, acredito que a legislação de facto seja sempre uma legislação de equilíbrio [entre a evolução tecnológica enquanto meio para assegurar o bem comum e o direito à privacidade] […]. Porque há realmente valores de referência, de Estado de direito. […] eu acho que isso depende […] da conjuntura política, e sobretudo de como é que evoluem as ações… […] E não é só, portanto, [uma questão] de grande segurança, securitária, no que diz respeito às bases de dados, também vai depender muito de como é que [feita] esta articulação com estas grandes empresas […]." [Professor/investigador]

Ou seja, o fenómeno do *Big Data* e a possibilidade futura do seu uso, manuseamento e aplicabilidade prática, no âmbito de investigações criminais, são percecionados (sobretudo por professores/investigadores de Direito, Economia Política e Sociologia do Direito, administrativos de entidades consultivas e jurídicas e juristas) como uma hipótese política, fruto de tensões e interesses políticos – entidades máximas que definem legalmente o conceito, o que pode e/ou deve abranger e sob que ditames deve atuar. São estas condições, exigências e interesses políticos que determinam se o *Big Data* pode expandir-se ou se deve ser eliminado enquanto técnica de investigação.

Concluindo, todos os profissionais entrevistados revelaram expectativas contraditórias, sob a ótica de Van Lente (2012), Borup *et al*. (2006) e Lucivero *et al*. (2011), ou seja, reflexões duais entre os benefícios da técnica do *Big Data* e os obstáculos práticos laborais à sua materialização. Concretamente, e de forma sucinta, os NCP2 usam uma linguagem intimamente próxima da temática, dominando o tema e descrevendo-o a um nível micro; os NCP1 tecem referências aos potenciais obstáculos práticos que podem surgir com a aplicação do *Big Data* no seu quotidiano profissional; e os geneticistas forenses e demais profissionais entrevistados tecem considerações informativas sobre o fenómeno, demonstrando, por via de um discurso consciente, na fase embrionária do fenómeno, as suas questões políticas e os dilemas que emergem fruto das tensões tecnológicas.

3.2. DESENVOLVIMENTO, EXPANSÃO E ANTECIPAÇÃO DE APLICAÇÃO DO *BIG DATA* NA INVESTIGAÇÃO CRIMINAL

No que concerne às expectativas futuras dos profissionais relativamente aos potenciais benefícios do uso do *Big Data* no âmbito das investigações criminais, surgem perspetivas comuns a todas as categorias profissionais, ainda que com diferenças ténues a considerar.

A perspetiva geral é de que esta é uma técnica em desenvolvimento, com previsões de se expandir e emancipar enquanto ferramenta tecnológica útil no combate à criminalidade. Em consonância com o que Borup *et al*. (2006) referiram, existe uma expectativa generalizada de que a tecnologia é um elemento fundamental para criar e oferecer oportunidades de progresso. Nomeadamente, os NCP1 consideram que se trata, precisamente, de uma técnica em desenvolvimento, que exige que os bancos de dados de ADN se adaptem a estas novas tecnologias. Ou seja, é percetível que o *Big Data* é uma realidade presente à qual os bancos de dados genéticos têm de se ajustar para que a sua utilização tenha efeitos válidos e eficazes:

> "[…] precisamos de adaptar as bases de dados de ADN para as novas tecnologias, para a próxima geração de sequenciamento […]." [NCP1]

Verifica-se uma consciencialização de que se trata de um fenómeno em desenvolvimento, que requer meios sofisticados e avançados para poder integrar-se e desenvolver-se. Desde logo, os NCP1 a desempenharem atividade profissional em Institutos Forenses (também responsáveis pela administração

da base de dados de ADN nacional) defendem efetivamente o desenvolvimento dos bancos de dados de ADN para que se tornem maleáveis à disseminação de ferramentas de investigação inovadoras como o *Big Data*. Além disso, consideram que, para que esta técnica se materialize, desenvolva e seja crescentemente usada, devem tornar-se mais sofisticados os meios pelos quais ela opera; ou seja, deve investir-se na criação e no desenvolvimento de ferramentas que permitam a sua expansão e efetivação. Nesta linha de reflexão, é percetível a necessidade, nomeada pelos entrevistados, de sofisticar e desenvolver meios que acompanhem e permitam o desenvolvimento destas técnicas eminentemente tecnológicas:

> "[…] tem de se assegurar um maior desenvolvimento dos bancos de dados, do nível de processamento e dos servidores, para serem capazes de processar essa grande quantidade de dados diferentes [*Big Data*]." [NCP1]

Há também profissionais que consideram o *Big Data* como um método que precisa do desenvolvimento crescente de bancos de dados para processar a informação de forma veloz e ser efetivamente útil. Por exemplo, atualmente, nos aeroportos, importa que a informação e os resultados sejam processados no imediato. Esta necessidade de urgência e velocidade de produção de provas exige meios adequados. Em geral, este tipo de discurso, adotado maioritariamente pelos NCP1, é definido como *crítico-interpretativo*, por incluir considerações críticas acerca das limitações e obstáculos atuais da técnica (Stevens *et al*., 2018).

As expectativas dos NCP1 sobre o facto de os meios de investigação atuais carecerem de desenvolvimento informático para que o *Big Data* se materialize, encontra eco nas restantes categorias profissionais entrevistadas. Estas também consideram que se trata de uma ferramenta mais avançada do que o contexto em que opera e, por isso, exige que o meio onde se insere se desenvolva para que se possa efetivar – por este motivo, não tem ampla aplicação atualmente e é necessário o desenvolvimento dos profissionais, das áreas, dos métodos e do meio.

Ou seja, há uma conceção generalizada de que se trata de uma técnica em desenvolvimento na esfera da investigação criminal: um conhecimento em construção. Sob a lente conceptual de Borup *et al.* (2006), Lucivero *et al*. (2011) e Van Lente (2012), estas expectativas revelam muita confiança no desenvolvimento das tecnologias e são comuns em investigadores académicos, tendo em conta que as suas variabilidades sociais influenciam as suas expectativas. São profissionais que não lidam com o manuseamento prático da técnica, antes se distanciando dela profissionalmente. Este tipo de expectativa é comummente

defendido por professores/investigadores, como os seguintes entrevistados geneticistas forenses a desempenharem funções de professor/investigador em contexto universitário:

> "O conhecimento é muito limitado neste momento [para associar informações genéticas a padrões de comportamento, como o *Big Data* prevê] […]."/ "[O *Big Data*] é um longo caminho, uma forma diferente de operar, mas sob uma perspetiva genética não sabemos ainda muito sobre isso." [Geneticista forense]

> "Até agora, não existe teoria por trás [do *Big Data*]." [Geneticista forense]

Para além dos geneticistas forenses inseridos profissionalmente em meios académicos, um geneticista forense a desempenhar funções num Instituto de Medicina Legal defende a mesma perspetiva:

> "Acho que ninguém neste momento tem conhecimento para gerir [o *Big Data*] […] essas informações todas juntas. Talvez no futuro, mas precisamos de ferramentas informáticas mais poderosas. […] Ninguém é capaz de gerir todas essas informações. […] Reconheço os benefícios [do *Big Data*], mas precisamos de esperar mais um pouco." [Geneticista forense]

Tal como Chan e Moses (2015) referiram, o *Big Data* é uma técnica que deve sempre ancorar-se na teoria. Contudo, pouco ainda se teorizou sobre esta ferramenta de investigação criminal, pelo que a sua utilidade enquanto metodologia de investigação é limitada.

Na mesma linha de reflexão, como já foi explanado, outros entrevistados professores/investigadores na área do Direito aproximam-se das perspetivas dos geneticistas forenses:

> "Provavelmente haverá [técnicas do *Big Data* enquanto ferramentas de investigação], só que nós não temos ainda tecnologia suficiente […] será necessário ainda um grande salto tecnológico […]. E no campo da investigação criminal isso é evidente […]." [Professor/investigador]

Os geneticistas forenses e os professores/investigadores consideram que o *Big Data* poderá trazer vantagens enquanto prática de investigação; no entanto, atualmente não existe um contexto de atuação inovador e sofisticado que permita beneficiar desta ferramenta. Estes entrevistados consideram que

o *Big Data* é resultado da evolução técnica e científica, e que envolve uma multiplicidade de questões, rápidas e inevitáveis, ao mesmo tempo que processa tudo a uma velocidade enorme. É uma realidade-fruto da tensão de duas ideias – a inevitabilidade (da evolução tecnológica) e a regulação (a necessidade de ser regulamentado pelo Direito). O equilíbrio entre estes dois ideais complexifica o pensamento sobre o tema. Neste sentido surgem expectativas contraditórias (Van Lente, 2012; Lucivero *et al*., 2011; Borup *et al*., 2006) que refletem a complexidade das expectativas. À semelhança dos entrevistados NCP1, também os entrevistados ligados a Comissões de Proteção de Dados consideram que o meio contextual e operacional de investigação deve sofisticar-se para que esta técnica se operacionalize:

> "[É] um bocadinho a ideia de alguma inevitabilidade [o uso potencial do *Big Data*] associada àquilo que tem sido – nos últimos tempos então com alguma força – o resultado da evolução técnica e científica […]." [Membro da Comissão de Proteção de Dados]

Na mesma linha de reflexão, um membro pertencente ao corpo de investigação criminal afirma:

> "O paradigma da investigação criminal tem de se adequar [à inovação dos meios e ferramentas de investigação do *Big Data*]. E tem de usar as ferramentas e os meios que são contemporâneos do momento em que intervém." / "[…] nós alterámos o paradigma da nossa relação através dos dados. […] produzimos muito mais rasto e temos mecanismos que registam muito mais a nossa pegada de presença do que antes." [Corpo de investigação criminal]

Desta forma, há uma clara preocupação por parte destes profissionais – geneticistas forenses, autoridades consultivas, diretores de instituições e membros do corpo de investigação criminal – em realçar a desadequação dos meios atuais de investigação face ao desenvolvimento e expansão do *Big Data.* Quanto à conceção de que é necessário que os meios de investigação das Ciências Sociais e Humanas se sofistiquem para criar condições para que o *Big Data* opere, Bartlett *et al.* (2018) referem que, efetivamente, falta uma base em metodologias matemáticas computacionais, pelo que deve ser feito um investimento na computação e nos seus meios e no acesso aos grandes dados para que os profissionais desta área possam manusear estas técnicas. No fundo, constata-se que existe uma representação social maioritária segundo a

qual a tecnologia evolui com a mesma velocidade que penetra os interstícios sociais, levando a metamorfoses que obrigam os variados contextos a dispor de recursos suficientemente maleáveis para se adaptarem às novas realidades com que são confrontados. No que diz respeito a esta temática, Matzner (2016) defende que só o crescente desenvolvimento da tecnologia permite a efetivação e utilização do *Big Data*, visto que ambos se desenvolveram de forma paralela e o último depende do primeiro para ter condições de atuação.

Os NCP2, apesar de partilharem esta visão consensual, não nomeiam esta relação entre a evolução dos meios tecnológicos e digitais de investigação e a necessidade de desenvolvimento do contexto de investigação atual de modo a integrar o *Big Data*. Ou seja, os NCP2 a desempenhar funções laborais em Unidades Policiais reconhecem que o *Big Data* é possível devido ao desenvolvimento e evolução tecnológica, mas não mencionam a necessidade de o contexto atual de investigação inovar para que se possa utilizar esta técnica. Apenas referem que há uma evolução no sentido do uso do *Big Data*:

> "Nós todos estamos a caminhar em direção ao *Big Data.* A cada 18 meses, os computadores evoluem…" [NCP2]

Na mesma linha de pensamento, os geneticistas forenses consideram que o uso massivo destas ferramentas tecnológicas e digitais de agregação de muitos e variados dados para atingir fins de investigação se tornou prática habitual e tornar-se-á cada vez mais iminente:

> "Eu sei que isso [*Big Data*] existe e foi feito, sei que os seus métodos estão a ser desenvolvidos […] e acho que temos poucas oportunidades para evitar o uso disso." [Geneticista forense]

> "Neste momento, sem nos darmos conta, já muitas empresas o fazem [utilizar o *Big Data*]." [Geneticista forense]

Adotando um discurso *pragmático* (Stevens *et al*., 2018), estes geneticistas forenses consideram que o *Big Data* é uma técnica que já está presente na malha social. Quanto à crescente utilização desta técnica, Chan e Moses (2017) referem precisamente esta ideia: fruto da disseminação dos tentáculos do *Big Data* nas várias facetas sociais, também no âmbito criminal se verificará, consequentemente, uma tendência crescente do seu uso como ferramenta preventiva que orienta estratégias policiais e decisões de justiça criminal.

Não obstante o reconhecimento da necessidade de desenvolvimento das ferramentas de investigação e das bases de dados para se adaptarem a este novo aparato tecnológico, os NCP1 a desempenharem atividade profissional num Instituto de Criminalística (também responsáveis pela administração da base de dados de ADN nacional) e os NCP2 a desempenhar funções laborais numa Unidade Policial consideram que se trata de uma técnica atualmente em crescente uso e com potencial para vir a desenvolver-se cada vez mais no futuro:

> "Eu acho que isso [*Big Data*] está a ser usado cada vez mais agora […]." [NCP1]

> "Nós caminhamos em direção ao *Big Data* […]." / "Hoje, com o *Big Data,* todas as pessoas sabem que estão a ser vistas e controladas […]." [NCP2]

Além disso, o *Big Data* é perspetivado, não só como um fenómeno em evolução, mas também como um caminho a seguir com interesse científico; ou seja, os NCP1 a desempenhar funções laborais em Unidades Policiais e os NCP2 a desempenhar funções de *design* e gestão do *software* informático da base de dados consideram que, enquanto técnica de investigação a ser utilizada potencialmente no futuro, poderá conduzir a conclusões interessantes:

> "Eu acho que estamos a caminhar nessa direção [do uso crescente das técnicas do *Big Data*]. […] E eu acho que pode ser interessante […] ter mais informações sobre as coisas […]." [NCP1]

> "Não só para resolução de casos criminais, [mas] até por nós [e para outros fins, o *Big Data*] poderia ser interessante […]." [NCP1]

> "Seria interessante [o uso de diferentes dados para investigação criminal], mas isso tudo capturado…" [NCP2]

Esta última perspetiva considera o *Big Data* como uma direção interessante a seguir (até para fins que não a investigação criminal) enquanto técnica que, na ótica dos entrevistados, permite obter (ainda) mais informação. Ou seja, os NCP percecionam as técnicas do *Big Data* como ferramentas promissoras a serem desenvolvidas e implementadas no seu contexto laboral. Estes profissionais são unânimes na crença de que, no futuro, o *Big Data* permitirá obter resultados que atualmente não são possíveis:

> "Eu acho que isso será o futuro, o *Big Data* permitir-nos-á ver coisas que não podemos ver hoje em dia." [NCP1]

O *Big Data* é uma técnica em expansão e no futuro permitir-nos-á possivelmente obter dados informativos que ainda não temos, facilitando a sofisticação e complexificação de determinados meios de investigação que conduzirão a resultados eficazes. Ou seja, é uma técnica potencialmente benéfica para desenvolvimentos futuros no conhecimento científico. Brown *et al.* (2006) afirmam que este aspeto está comummente presente nas expectativas: a facilidade de pensar que o *novo* oferece oportunidades nunca antes concebidas, estigmatizando e minimizando o anterior. Este ponto de vista vai ao encontro do que Boyd e Crawford (2012) referiram quando definiram o *Big Data*: um fenómeno cultural, tecnológico e académico que pode ser mitológico, na medida em que existe uma crença generalizada de que grandes conjuntos de dados oferecem uma forma mais elevada de inteligência e conhecimento que pode gerar *insights* antes impossíveis.

Quando questionados acerca da possibilidade de aplicação e inserção do *Big Data* nas práticas de investigação criminal, todos os profissionais considerados afirmam que se trata de uma técnica atualmente sem uso prático no seu contexto laboral:

> "Definitivamente, eu não uso [o *Big Data*]." [NCP1]

> "No contexto de ADN, eu não vejo isso [o *Big Data* a ser utilizado]." [NCP2]

> "Eu não consigo ver essa realidade [cruzamento de informações biométricas com associações de padrões de comportamento ou determinadas doenças] a acontecer, sinceramente." [Geneticista forense, N07]

Ou seja, todos os profissionais entrevistados não lidam, em contexto prático de trabalho, com a técnica. Este aspeto deve-se ao facto destes profissionais estarem inseridos num contexto de partilha transnacional de dados específico: a troca de perfis de ADN. Neste campo, o *Big Data* não encontra aplicação prática. No entanto, perspetivam a sua inserção no meio laboral, enquanto ferramenta de investigação benéfica e desejável:

> "Neste momento eu não uso isso [*Big Data*] na minha prática profissional." / "Posso vir a usar no futuro. Mas neste momento eu não uso, nem discuto, nem convivo com isso […]." [NCP1]

> "Eu não acho que isso [*Big Data*] seja usado no quotidiano policial [enquanto ferramenta de trabalho], talvez no futuro. Mas, em geral, acho que no trabalho de investigação, como pesquisa, sim [deve ser desenvolvido]." [NCP2]

> "Essas informações [do *Big Data*], até agora, não são usadas pela polícia, mas não posso eliminar a ideia de que haverá um banco de dados grande onde elas serão armazenadas e usadas. No entanto, até agora, não é possível." [Professor/investigador]

É frequente que as novas tecnologias, os novos conceitos, as inovações e os progressos científicos sejam expressos em termos de desejos futuros, tal como evidenciam as palavras dos últimos entrevistados (Brown & Michael, 2003). Os profissionais reconhecem estas técnicas como práticas de trabalho desejáveis, com probabilidade crescente de serem implementadas no seu contexto laboral, expressando perspetivas de que esta ferramenta se venha a desenvolver, consolidar e implementar. A vigilância automatizada torna-se uma possibilidade crescente (Boyd & Crawford, 2012; Frade, 2016; Matzner, 2016).

Apesar de a maioria dos profissionais entrevistados considerar possível e desejável, a longo prazo, a inserção destas ferramentas tecnológicas nas metodologias de investigação, apenas os NCP reconhecem a potencial utilidade prática e os benefícios de tal inserção no combate e prevenção da criminalidade transfronteiriça. Esta evidência vai ao encontro das conclusões apresentadas por Machado e Silva (2015): os profissionais que desempenham funções onde futuramente a técnica poderá vir a ser aplicada – como é o caso dos NCP relativamente ao *Big Data* – estão mais conscientes da sua utilidade. Os restantes entrevistados – geneticistas forenses e outros atores envolvidos no domínio da partilha transnacional de dados de ADN – têm uma perspetiva mais generalizada da questão em apreço. Ou seja, evidenciam perspetivas positivas acerca do *Big Data*, aliadas ao desejo da sua materialização futura – expectativas comuns e mencionadas por Brown e Michael (2003) como reflexões presentes no âmbito de uma tecnologia em desenvolvimento. Concretamente, os geneticistas forenses consideram que refletir sobre a possibilidade de utilizar o *Big Data* no âmbito da investigação criminal significa que existe um progresso cientifico útil a ser realizado, nomeadamente para a área forense, que pode fazer uso destas valências para alcançar de forma mais eficaz os seus resultados.

> "Naturalmente, esperamos pelo progresso científico […] e a questão é: até que ponto isso [*Big Data*] pode ser útil para as ciências forenses? […] para encontrar perpetradores […]." [Geneticista forense]

> "O *Big Data* só nos pode ajudar no conhecimento […]. Agora somos capazes de identificar uma pessoa como sendo pertencente àquela população líquida. Isso só é possível porque temos todos esses dados, também de estudos clínicos e de outras áreas, que nos permitem fazer essa separação e é por isso que [o *Big Data*] é útil. […] E para, por exemplo, identificar genes que dão as características faciais ou outras características do corpo […]. [É útil para isso], mas nada mais a respeito da perícia forense." [Geneticista forense]

Os geneticistas forenses perspetivam o *Big Data* como uma técnica útil neste sentido, mas não preveem outro tipo de utilidade no âmbito forense. Além disso, também reconhecem o interesse técnico-científico do tema e da ferramenta, com várias utilizações que poderão ajudar a inferir sobre a imputabilidade criminal.

> "Do ponto de vista técnico-científico, isso [*Big Data*] poderá ter interesse […] para se poder inferir porventura da imputabilidade ou não de algum indivíduo quando cometeu determinado ato criminal, por exemplo." [Geneticista forense]

Estes profissionais também consideram que a utilidade prática do *Big Data* reside na possibilidade de atingir resultados mais eficazes e demonstram uma visão positiva sobre a técnica e a sua inserção nas metodologias de investigação, quando o seu uso tem como finalidade a investigação criminal:

> "Eu acho que há determinadas investigações que se justificam quando nós estamos […] a tentar solucionar ou a tentar investigar […] um crime." [Geneticista forense]

Ou seja, visto que são profissionais que trabalham com os dados recolhidos pela polícia, se o número de dados recolhidos for maior, o objetivo a alcançar com a investigação criminal pode ser mais eficaz e célere. Desta forma, na sua perspetiva, caso seja possível integrar uma técnica que permita agregar um conjunto maior e mais variado de dados, ela será útil para a atividade profissional de investigação (Cole, 2013).

Os membros de Companhias Privadas de Ciências Forenses também consideram o *Big Data* vantajoso no âmbito da investigação criminal:

> "[O uso do *Big Data* nas Ciências Forenses] pode ser muito vantajoso […]." [Membro de companhia privada]

Os NCP2 a desempenharem funções policiais numa Diretoria de Cooperação Policial Internacional da UE também consideram estas ferramentas úteis, considerando que a resolução eficaz de um caso criminal é sempre difícil, qualquer informação que surja e ajude a resolvê-lo afigura-se vantajosa.

> "No final do dia qualquer informação [que auxilie na resolução de casos criminais] sobre qualquer coisa é útil […]." [NCP2]

Se o *Big Data* é um mecanismo de investigação que agrega, armazena, configura e codifica informação que pode ser vantajosa para a resolução de casos criminais, na ótica deste entrevistado, a sua utilização é benéfica.

No entanto, apesar de as expectativas dos entrevistados membros de entidades privadas se aproximarem das apresentadas pelos NCP (os três extratos profissionais nomeiam vantagens à inserção e aceitação de uma inovadora metodologia tecnológica de investigação), estes últimos fazem uma reflexão diferente: colocam a possibilidade de colocar o *Big Data* ao serviço do combate e prevenção da criminalidade transfronteiriça.

Concretamente, os NCP1 veem o *Big Data* como uma ferramenta útil, usada em situações de combate a ações terroristas:

> "Devido aos recentes acontecimentos terroristas, eu não teria muitas hesitações em usar os grandes volumes de dados para combater o terrorismo." / "No combate ao terrorismo eu não teria qualquer reserva […]." [NCP1]

Nestes casos, os NCP1 defendem o uso de técnicas do *Big Data* sem qualquer reserva, desde que tal permita aceder aos perfis corretos de indivíduos autores de ações terroristas. Neste sentido, e para estes fins, consideram que se trata de um mecanismo promissor para prever comportamentos futuros danosos para a sociedade.

> "Eu gostava que tivéssemos esses métodos [de *Big Data*] e que pudéssemos prever comportamentos [criminais] […]." [NCP1]

Os professores/investigadores de Sociologia do Direito também referem os benefícios de utilizar esta técnica enquanto meio e ferramenta de combate e prevenção do terrorismo. Entendem que existem certos ilícitos penais perante os quais o *Big Data* se afigura de extrema importância, dado o contexto do cometimento do crime e os seus impactos sociais: crimes

transfronteiriços, com perímetro considerável e nos quais a pesquisa, interconexão e partilha transnacional de dados se afigura útil, dada a dimensão do fenómeno criminal:

> "Do ponto de vista da eficiência da investigação criminal, ela [técnica do *Big Data*] pode trazer de facto benefício, sobretudo à criminalidade transnacional, a grande criminalidade hoje: […] a criminalidade económica, o tráfico de seres humanos […] tem de haver armas efetivas neste combate com esta dimensão de eficiência, transnacional." [Professor/investigador]

Esta visão é partilhada pelos NCP: quando questionados acerca das expectativas relativamente à futura implementação das técnicas do *Big Data* no âmbito da investigação criminal, de uma forma geral, afirmam que a perspetivam no âmbito da criminalidade transfronteiriça e transnacional. Os NCP2 a desempenharem funções laborais em Unidades Policiais consideram que a recolha e a agregação de informação múltipla e heterogénea por esta ferramenta tecnológica são úteis para combater o terrorismo:

> "Temos de usar a tecnologia e unir todas as informações que temos para combater o terrorismo […]." [NCP2]

Concluindo, é possível constatar que os NCP reconhecem a eficácia desta técnica para combater o terrorismo, enquanto manifestação de uma criminalidade transnacional e transfronteiriça cada vez mais iminente. No que toca aos tipos de crimes e circunstâncias criminais, as perceções dos NCP sobre o *Big Data* concentram-se na criminalidade grave. De acordo com Drewer e Miladinova (2017) e Van Dijk (2014), o *Big Data* é uma ferramenta projetada para prevenir e combater ações terroristas transnacionais e transfronteiriças.

Ainda assim, estes profissionais, principalmente os alocados no *step* 2 e a desempenharem funções profissionais num Centro Internacional de Cooperação em Matéria de Aplicação da Lei, apesar de reconhecerem a possível utilidade da aplicação do *Big Data* no combate à criminalidade transfronteiriça e transnacional, defendem que se trata, em casos muitos específicos e criteriosos, de uma solução de exceção. Apesar de ser uma técnica compatível com Prüm, só é permitido o seu uso para dados históricos que não tenham outra forma de serem comparados:

> "Eu vejo que isso [*Big Data*] funciona com Prüm, mas acho que é uma solução de exceção só permitida para dados históricos que não possam ser comparados de outra forma a não ser dessa […]." [NCP2]

Ou seja, o *Big Data* é uma técnica de *ultima ratio*, apenas considerada quando todos os outros mecanismos não se afigurem suficientes; é um último recurso que poderá ter maior utilidade se articulado com outras ferramentas e/ou tecnologias. Para além disso, os NCP2 entendem que é uma técnica que compara, em massa, dados sem qualquer tipo de critério, nem razão pré-estabelecida que sustente e justifique a utilidade e eficácia do seu uso, tornando-se assim um tipo de pesquisa não abrangido, nem permitido legalmente, pela legislação europeia:

> "Mas a legislação europeia não suporta essa comparação de dados em massa, sem critérios determinados […]." [NCP2]

Concluindo, os NCP consideram que se trata de uma técnica de investigação potencialmente benéfica, quando utilizada em situações de prevenção do terrorismo.

De uma forma geral, todas as categorias profissionais entrevistadas demonstram expectativas positivas, mas cautelosas quanto à implementação do *Big Data* como método de investigação criminal, sendo que os NCP2 se destacam pelo minucioso conhecimento e transposição prática que operam no sentido de abstrair e colocar os grandes dados ao serviço do combate e prevenção da criminalidade transfronteiriça e organizada. Não obstante, todas as restantes visões profissionais são consensuais relativamente aos benefícios, utilidade e potencialidade da técnica, se usada corretamente para fins de investigação criminal; partilhando igualmente a perceção do *Big Data* como fenómeno em desenvolvimento.

3.3. PERCEÇÃO DOS RISCOS E PERIGOS DO *BIG DATA*

No que concerne aos potenciais riscos e perigos que o uso do *Big Data* pode acarretar para a investigação criminal, os vários profissionais revelam preocupações e motivos pelos quais esta ferramenta se pode converter numa prática arriscada e perigosa do ponto de vista da qualidade e validade dos resultados atingidos. Esta perspetiva é partilhada pelas diferentes categorias profissionais – a preocupação com a forma como esta técnica se materializa na prática

e como opera com os dados que são o seu objeto de estudo. Ainda assim, os NCP enfatizam mais as questões ligadas à ética e aos direitos humanos do que propriamente as questões procedimentais que podem colocar em causa a validade científica da ferramenta.

NCP2 e geneticistas forenses tecem algumas considerações acerca das questões procedimentais. Tal como Cole (2013) afirma, as perceções dos geneticistas forenses relativamente aos resultados que esta técnica permite obter baseiam-se em resultados probabilísticos, amplos e geradores de incerteza, o que leva a um sentimento de insegurança por parte destes profissionais.

Os NCP1 a desempenharem atividade profissional num Instituto de Criminalística (e, em simultâneo, a exercerem funções de administração da base de dados de ADN nacional) consideram que o *Big Data* deve ser perspetivado e utilizado enquanto ferramenta auxiliar na predição de determinados padrões e não como algo que nos informa fidedignamente sobre algo – não devendo, portanto, ser visto como a única, principal e exclusiva forma de obter resultados e conclusões:

> "[O *Big Data*] deve ser visto mais como uma ajuda preditiva e não como algo que fornece informações 100% fidedignas." [NCP1]

Estes profissionais referem que esta é uma prática que pode facilmente converter-se numa ferramenta de investigação perigosa, que acarreta riscos:

> "Acho que há sempre um nível de risco. Porque esses métodos [*Big Data*] não são exatos." [NCP1]

Estas narrativas indicam que é necessário que se tenham em consideração os riscos que esta troca massiva de dados informativos acarreta, desde logo no que toca ao número de pessoas que têm acesso aos dados conectados, e à forma como os trata, manuseia e para que fins os usa. Segundo Prainsack (2019), este debate tem-se intensificado, precisamente porque se trata de dados digitais que estão disponíveis a partir de qualquer ponto geográfico e porque esta digitalização complexifica o estabelecimento de limites sobre quem tem acesso aos dados, a forma como os controla e com que fim os utiliza.

Este aspeto é frequentemente nomeado pelos NCP1 e pode revelar a falta de confiança que estes profissionais têm nas instituições que são responsáveis pela recolha massiva destes dados e pelo seu posterior uso para fins que não os previstos inicialmente (Machado & Silva, 2016; Lupton & Michael, 2017).

A este propósito, os NCP1 a desempenharem funções laborais num Instituto Forense referem que todas as tecnologias – tal como o *Big Data* – acarretam riscos e desvantagens; no entanto, as consequências dos riscos do uso desta técnica dependem da forma como esses riscos são administrados: caso sejam bem geridos, os benefícios do uso do *Big Data* podem ser maiores que as suas desvantagens.

> "[O *Big Data*] é como qualquer tecnologia: tem riscos e depende de como administramos esses riscos. Eu acho que, se o fizermos corretamente, os benefícios serão maiores que os riscos." [NCP1]

Segundo Borup *et al*. (2006) e Brown e Michael (2003), estas dinâmicas de antecipação de riscos futuros são comuns quando se abordam expectativas de uma tecnologia em desenvolvimento: são receios que estão intimamente ligados ao *Big Data*, visto que a técnica é muito recente e, por isso, geradora de medos e incertezas.

Nesta linha de reflexão, geneticistas forenses a desempenhar funções laborais como investigadores num Instituto Médico Informático e de Estatística referem que os usos potenciais do *Big Data* são uma "promessa não cumprida", porque não foi possível obter os resultados esperados aquando da sua implementação e/ou idealização. Ou seja, o facto de o *Big Data* envolver, englobar e considerar um grande número de dados não permite prever a globalidade das soluções. Pode prever-se o risco de algumas e determinadas situações, mas não é possível prever tudo:

> "O *Big Data* é uma promessa não cumprida. O *Big Data* pode, claro, fazer algo. Pode gerar talvez algumas hipóteses, mas […] não pode resolver certas questões que temos. […] não é verdade que quantos mais dados tivermos, mais coisas podemos resolver. Podemos prever o risco, por exemplo, para uma doença ou qualquer outra coisa, ou para um traço psicológico […]. O *Big Data* pode ajudar a dar algumas ideias, por exemplo, mas não pode realmente resolver tudo como prometido." [Geneticista forense]

Borup *et al*. (2006) e Brown e Michael (2003), referem que é comum que as expectativas em torno de inovações tecnológicas e científicas sejam descritas como *promessas*. As perspetivas em torno de técnicas em desenvolvimento surgem como desejos para o futuro, que não encontram ainda idealização prática para se efetivarem, mas despertam expectativas, interesses e receios. É frequente que a *promessa* inicial das tecnologias dê, *a posteriori*, lugar à desilusão, porque

a expectativa reside no imaginário; no entanto, quando se efetiva, encontra obstáculos à sua materialização. É o que sucede com o *Big Data,* visto que não se trata de uma técnica validada cientificamente para poder proceder a previsões globais e válidas, como seria esperado, na ótica dos geneticistas forenses. Desta forma, deixa de se acreditar no poder tecnológico da técnica, e a expectativa positiva dá lugar à deceção. Consequentemente, o *Big Data* pode facilmente converter-se numa ferramenta arriscada e desvantajosa caso seja perspetivada como capaz de realizar previsões certas e globais.

Na mesma ótica dos riscos e desvantagens do uso desta técnica, os NCP2 a desempenharem funções laborais num Instituto Forense referem-se ao *Big Data* como sendo uma "técnica muito perigosa", visto que permite que algumas entidades acedam a dados e informações pessoais e que os utilizem para fins que não os previstos aquando da sua recolha. Logo, é um tipo de pesquisa e de investigação que pode prejudicar os titulares dos dados se estes forem usados para outros fins; mesmo a mera consulta desses dados é um risco que pode pôr em causa o benefício das pessoas em questão.

> "Acho que [o *Big Data*] é muito perigoso, nomeadamente para entidades que têm fins lucrativos […], que têm acesso ou poderão ter acesso a uma série de informações que em conjunto lhes poderão dar informação que poderá prejudicar efetivamente as pessoas […]." [NCP2]

Relativamente a estas expectativas, o Conselho Europeu tem como um dos seus princípios a limitação da finalidade; ou seja, os dados pessoais só podem ser recolhidos para fins específicos, explícitos e legítimos e não podem ser processados de forma incompatível com esses fins (Gonçalves, 2017). Esta decisão política parece ser antecipatória do *Big Data* (Brown & Michael, 2003), pelo que esta expectativa culmina(rá) em alterações legislativas.

Ainda inseridos no conhecimento dos riscos e perigos do uso das técnicas de *Big Data*, os NCP2 a desempenharem funções de *design* e gestão do *software* informático da base de dados revelam a sua opinião e conhecimento acerca da combinação de diferentes dados de diferentes fontes, considerando que esta potencial técnica de investigação pode levar a correlações falsas. A obtenção de resultados errados não é o que se pretende quando se faz investigação e combinação de dados para chegar a resultados eficazes:

> "A análise [com o *Big Data*] pode obter um mau perfil. Nós não queremos maus perfis. Porque podemos descobrir que pessoas com certos *loci* têm maior probabilidade

> de desenvolver condutas criminais. Há criminosos com certas características… Mas essa correlação não pode ser tomada como real e única […]." [NCP2]

Os NCP2 revelam receio na medida em que, quando se manuseia um grande número de informação e de dados, a probabilidade de erro é maior e, por isso, é importante que as correlações estabelecidas sejam calculadas com cuidado. Desta forma, tendo em conta que o *Big Data* pressupõe, desde logo, a inclusão de um grande número de dados a serem trabalhados, ele converte-se numa prática *arriscada*. Partilhando da mesma perspetiva, os geneticistas forenses a desempenharem funções laborais como investigadores num Instituto Médico Informático e Estatístico referem o mesmo aspeto: o grande volume de dados não reflete a sua qualidade, sendo necessária a validação dos mesmos para que as conclusões a que se chega sejam fidedignas. Assim, este procedimento de validação dos dados e a sua posterior verificação orientam estes profissionais para situações clássicas de técnicas e ferramentas tradicionais de investigação, que não são inovadoras:

> "Neste momento, [o *Big Data*] não foi bem-sucedido e há várias razões para isso. Uma é que a qualidade do *Big Data*, geralmente, é má. Portanto, não se pode obter um bom resultado a partir de dados que estão incorretos. Então, estes dados têm de ser eliminados e ficamos sem dados. Em muitos casos é necessário validar os dados. Se tivermos muitos dados, claro, obteremos muitos resultados presumidos. No entanto terá de [se] aceder aos conjuntos de dados menores para os validar, depois validar os modelos obtidos, e assim sucessivamente […]. Isto remete-nos para situações clássicas de investigação […]." [Geneticista forense]

Esta é uma posição frequentemente defendida pelos investigadores (Selin, 2008): a avaliação das questões procedimentais da técnica. Mittelstadt *et al.* (2016) referem que esta é uma das preocupações éticas públicas mais evidenciadas pela literatura: a ênfase dada ao poder que é atribuído às correlações provenientes de algoritmos e inferências. Quanto a estas advertências, Boyd e Crawford (2012) e Chan e Moses (2015) defendem que importa ter em conta que o *Big Data* não é autoexplicativo; ou seja, não é porque o *software* denota uma correlação entre um grande volume de dados, que ela existe simplesmente; é preciso procurar as causas, perceber o fenómeno e aceder a ele (Amoore, 2011). Nenhuma conclusão deve sustentar-se apenas em correlações, pois estas podem ser apenas coincidências – mesmo os tipos de crime mais previsíveis são suscetíveis de mudanças sociais mais amplas (Chan & Moses, 2015). Assim, é

preciso que haja consciência de que, como qualquer outro método, o *Big Data* requer uma visão holística dos fenómenos. Além disso, importa enfatizar que, apesar do grande volume de dados que o *Big Data* contém, estes não devem ser vistos como representativos, nem tão-pouco como explicativos da globalidade dos fenómenos; antes devem ser interpretados como qualquer outro tipo de dados, independentemente do tamanho da amostra. Todas estas precauções se exacerbam quando o *Big Data* se aplica à prevenção da criminalidade, pois, uma vez que visa informar políticas criminais, deve orientar estas decisões de forma eficaz e válida, sustentado em dados fidedignos (Boyd & Crawford, 2012).

Ancorados no mesmo ponto de vista da antecipação dos potenciais riscos do *Big Data*, os geneticistas forenses, os NCP1 a desempenhar atividade profissional num Instituto de Criminalística e os profissionais membros de Organizações Não Governamentais de Direitos Humanos enfatizam as questões da fragilidade científica da técnica. Os entrevistados conjugam dois tipos de representações sociais: a preocupação da junção ímpar de um grande volume de dados; e as exigências minuciosas procedimentais que esta junção requer para que os resultados atingidos sejam válidos e úteis (Williams & Johnson 2004; Albert *et al*., 2009). Os entrevistados consideram que a conexão de um número elevado de dados heterogéneos não é possível, visto que a informação não possui critério de semelhança e/ou identificação. Na sua perspetiva, a agregação de dados completamente distintos não pode ser considerada, uma vez que se estão a englobar informações distintas sob o mesmo chapéu interpretativo, e isso pode acarretar riscos e perigos sérios na recolha, tratamento e posterior interpretação dos dados:

> "Há uma grande quantidade de dados [no *Big Data*], e estes devem ser processados de forma correta, porque são dados muito diferentes uns dos outros. São dados provenientes de diferentes fontes e formas…" / "O manuseamento dos dados deve ser muito cuidadoso […]." [NCP1]

> "Temos dados biométricos […] [e] impressões digitais que são procurados em bancos de dados de ADN. E esta pesquisa é feita em separado, porque os dados não estão conectados, não é permitido conectá-los." / "São muitos dados numa só mão." [Geneticista forense]

> "Estou particularmente preocupado/a com o uso nacional de dados biométricos devido à qualidade dos dados. A qualidade dos dados já é um grande problema; com o *Big Data* a preocupação é ainda maior […]." / "Eu refiro-me não só à

> qualidade dos dados, mas também à qualidade dos testes realizados e dos procedimentos que levam aos resultados." [Membro pertencente a uma Organização Não Governamental de Direitos Humanos]

Esta é uma visão frequentemente expressa por investigadores sociais, que usualmente tecem avaliações sobre o progresso científico (Selin, 2008). Na mesma linha de reflexão, os geneticistas forenses a exercer funções laborais em Institutos Forenses consideram que o *Big Data* é uma amostra numerosa e constituída por dados heterogéneos, os quais, caso não sejam manuseados com cuidado, colocam em causa a sua própria qualidade, bem como a qualidade dos procedimentos que levam aos resultados, fragilizando as conclusões que daí se possam retirar e considerar. O risco iminente prende-se com a natureza heterogénea dos dados recolhidos, acompanhada do elevado número de dados. Ou seja, segundo as seguintes respostas, trata-se de uma amostra diversa e numerosa, a partir da qual os resultados obtidos facilmente podem constituir falácias e erros. Geralmente, não é isto que se pretende, já que se pode colocar em risco a própria técnica e a investigação em causa:

> "Acho que esse cruzamento [de muitos e diferentes dados] pode ser perigoso. Porque pode induzir-nos em erro, podemos cair em falácias […] é preciso muito cuidado com esse género de coisas […]." [Geneticista forense]

> "A ideia de conectar dados da saúde com o comportamento criminal [o *Big Data*] é demais. Quando se cruzam demasiadas linhas, é exagero, surgem sempre muitos problemas éticos. Se são dados relativos à saúde, devem ser usados apenas para esse fim – a saúde, nada mais." [Geneticista forense]

Esta sensibilidade face aos dados e aos procedimentos de recolha que têm influência na validade do processo de investigação é, frequentemente, referida pelos geneticistas forenses. Estes profissionais têm como audiência última do seu trabalho e dos resultados alcançados o Tribunal, enquanto instância máxima decisória e de poder. Consequentemente, estão sujeitos a processos legais que aplicam o princípio do contraditório, podendo aqueles profissionais serem obrigados a envolver-se no processo, caso os resultados obtidos apresentem falhas, fracassos ou erros. Assim, a sua credibilidade e a do laboratório onde desempenham funções laborais podem ser fragilizadas (Cole, 2013). Conscientes desta realidade, estes profissionais demonstram uma sensibilidade profissional e técnica quanto às questões procedimentais e

de validade científica do *Big Data*. Porém, esta visão não pertence apenas aos geneticistas forenses; também os entrevistados professores/investigadores em Antropologia se referem às técnicas do *Big Data* como práticas minuciosas que devem ser manuseadas de forma criteriosa para não levarem a resultados e conclusões duvidosas:

> "Há estudos que podem ser [feitos assim]… há relações interessantes que podem ser feitas, mas tem de haver um grande controlo, e a validade do estudo é uma questão que tem de ser vista […]." [Professor/investigador]

No que toca a estas questões éticas da recolha de dados e sua posterior agregação, Boyd e Crawford (2012) afirmam que é crucial que nos questionemos acerca da origem dos dados e de como estes devem ser interpretados – caso os dados sejam incluídos em estudos sem nexo, perderão o seu valor e significado *a posteriori*. Ball *et al.* (2016) referem também que estas questões éticas que se levantam devido à expansão e agregação de grandes e diferentes dados em configurações de interesse público estão ligadas a preocupações sobre a privacidade individual e a exposição à vigilância precisamente devido ao fluxo de dados autónomo que o indivíduo não pode controlar.

Vários são os profissionais que, quando abordam a temática das questões procedimentais por via das quais se materializam as técnicas do *Big Data*, consideram que se trata de uma técnica antiética e incorreta por agregar um grande número de dados sem critérios nem procedimentos cientificamente válidos. Consequentemente, carecem de valor causal. Assim, trata-se de uma técnica de investigação sem base teórica, pelo que as previsões obtidas devem ser interpretadas com precaução.

> "Até agora não há teoria por trás disso [*Big Data*]. […] Então, tudo isso se baseia em correlações. [O *Big Data*] não tem base causal, não tem teoria […]." [Geneticista forense]

Quanto ao facto de os resultados obtidos constituírem inferências e não correlações com valor causal, Amoore (2011) afirma que o que se pretende do *Big Data* é mesmo isso: a capacidade de fazer inferências a partir dos dados, de modo a que as correlações possam ser reconhecidas e partilhadas. Uma associação não implica uma conexão causal direta, mas revela, interpreta e exibe relações entre pessoas, lugares e eventos (Amoore, 2011). Por isso, esta visão dos geneticistas forenses é frequentemente reforçada pelas questões

procedimentais e de validade científica (Cole, 2013) – incerteza técnico-científica. Van Lente (2012), Lucivero *et al.* (2011) e Borup *et al.* (2006) referem que estas incertezas são comuns em grupos profissionais que exercem funções laborais próximas dos contextos (ou neles inseridos), onde a futura técnica poderá vir a materializar-se; assim, estas preocupações são resultado da posição laboral que ocupam.

No entanto, as questões ligadas às fragilidades procedimentais da técnica não se esgotam nestas considerações dos profissionais entrevistados, visto que estes também perspetivam o desrespeito pelos fins a que se destina a recolha dos dados como uma causa de preocupação do potencial uso do *Big Data*. Concretamente, os geneticistas forenses consideram perigoso o uso de dados pessoais para atingir certos fins que não os previstos inicialmente aquando da recolha legítima e consentida. Tendo em conta que os dados incluídos e considerados pertencem ao foro pessoal e privado dos cidadãos, o *Big Data* facilmente se pode converter numa *prática perigosa*:

> "Pode ser perigoso [a recolha massiva de dados informativos sobre as pessoas para fins de investigação criminal]. […] parece-me um pouco preocupante se uma determinada empresa, como agora aconteceu com este escândalo do *Facebook*, […] usar os dados das pessoas […] para atingir determinados fins, inclusivamente [os não previstos aquando da recolha dos dados]." [Geneticista forense]

> "Pode ser perigoso também." / "[Cria-se] uma determinada ideia com base em informações que […] não foram consentidas, nem conscientes para o que estava a ser feito […]." [Geneticista forense]

> "O risco é que a utilização desta informação [obtida por via do *Big Data*] extravase este âmbito. […] E que se ultrapasse este foco da investigação criminal e se vá a outras áreas. […] isso é que acho que pode ser perigoso." [Geneticista forense]

Assim, entre os geneticistas forenses, surgem representações socioprofissionais que alertam para o facto de a recolha massiva de dados se ancorar em investigações não criminais e cumprir outros fins que não os previstos inicialmente. Verifica-se uma antecipação dos riscos e dos perigos desta técnica, caso esta seja um meio para recolher e conjugar dados para fins não consentidos. São expectativas expressas maioritariamente por NCP e geneticistas forenses. Esta conclusão vai ao encontro do que Lucivero *et al.* (2011), Borup *et al.* (2006) e Van Lente (2012) referiram: as variabilidades sociais dos profissionais

influenciam as suas expectativas; assim sendo, NCP e geneticistas forenses, por desempenharem funções laborais em meios onde o *Big Data* se poderá efetivar, aproximam-se da técnica e são também mais conscientes dos riscos e perigos a ela associados, bem como dos obstáculos à sua materialização. Esta perspetiva vai também ao encontro do que Machado e Silva (2015) afirmam: os grupos profissionais detentores de maior conhecimento acerca da temática são, proporcionalmente, os que têm maior consciência dos potenciais riscos associados à mesma. É uma perspetiva defendida na sua maioria por profissionais que estão diretamente ligados a uma prática laboral que, no futuro, pode ser palco de desenvolvimento do *Big Data.* Os profissionais salientaram o carácter peculiar, minucioso e sensível dos dados que esta técnica engloba, manuseia, agrega e partilha, tornando a sua noção sobre o fenómeno um tanto mais crítica.

Estes entrevistados têm uma perspetiva "crítica", pessimista e defensora da proteção de dados e da privacidade, ressalvando a necessidade de proteção da informação sensível; ou seja, salientam a natureza sensível deste tipo de informação (Machado & Silva, 2015). No fundo, respondendo às consequências da mudança, através da antecipação de problemas sócio-éticos e de uma avaliação crítica do processo inovador, estes profissionais auxiliam a construção do futuro (Hedgecoe & Martin, 2003).

Estes discursos caracterizam-se usualmente pela avaliação dos processos e o modo como os dados são recolhidos e analisados. Tudo o que diga respeito à avaliação de questões procedimentais práticas diz respeito à narrativa discursiva que predomina nesta temática, transversal a todas as categorias profissionais (Albert *et al*., 2009). De forma geral, todos os profissionais entrevistados teceram considerações sobre os potenciais riscos e perigos da implementação do *Big Data*: desde comentários relativos à fragilidade científica da técnica, avaliações sobre o processo de recolha de dados e seu manuseamento, passando pela análise e posterior utilização, até a reflexões abstratas com pendor avaliativo sobre os riscos procedimentais e técnicos.

3.4. PARECERES ÉTICOS E DE DIREITOS HUMANOS

No que toca às questões éticas e de direitos humanos – as preocupações que se prendem com a salvaguarda dos direitos, liberdades e garantias dos cidadãos –, todos os profissionais entrevistados fizeram-lhes referência. Desde logo, ancorados no debate político e social contemporâneo que emergiu devido ao caso Snowden, que funcionou como desbloqueador do pensamento

crítico relativamente ao *Big Data*. Este caso suscitou o debate sobre o uso de dados pessoais protegidos (considerados legalmente como privados) para fins que não os idealizados inicialmente. A questão da violação da reserva da intimidade e da vida privada que se levantou depois das revelações de Snowden é mencionada por muitos profissionais, nomeadamente pelos professores/investigadores em Direito:

> "O caso de Snowden, que tornou evidente como as empresas […] não se preocupam muito em fornecer […] a informação de que dispõem aos serviços de segurança, polícias […]." [Professor/investigador, N17]

A alusão ao caso Snowden é feita por vários profissionais da área do Direito, que veem o caso como potenciador, instalador e revelador do poder negativo dos grandes dados e das consequências nefastas que podem ter a nível dos direitos dos cidadãos. Há uma preocupação generalizada na forma como os direitos dos cidadãos – nomeadamente o direito à reserva da intimidade da vida privada – são facilmente violados por via destas técnicas que visam a recolha massiva e escrutinada de dados informativos pessoais e privados dos indivíduos, muitas vezes para outros fins que não os previstos inicialmente. Por exemplo, como o último entrevistado referiu, a propósito do caso Snowden: foram recolhidos dados pessoais privados de forma massiva que, no entanto, não ficaram à disposição dos serviços de segurança ou policiais, tendo sido utilizados para outros fins. Tal como Brown e Michael (2003) mencionaram, muitas vezes as tecnologias desdobram-se em formas que não tinham sido planeadas inicialmente. Ou seja, neste caso, ao invés de cumprirem o objetivo de garantir e promover a segurança máxima dos indivíduos, os grandes dados abriram caminho para violações da privacidade. E a questão que aqui se levanta é se este exercício de reflexão sobre os fracassos do passado pode ensinar algo que permita prevenir fracassos no futuro. Brown *et al.* (2006) afirmam que, em cada história de inovação, é frequente existir uma identidade retrospetiva, tipicamente alguém associado a uma narrativa de luta, como que se esse alguém fosse a mola propulsora do desenvolvimento tecnológico científico. Neste caso, Snowden. Ou seja, projeta-se o futuro a partir de reflexões do passado que fracassaram (Brown & Michael, 2003).

Toda esta panóplia de dilemas morais e éticos, de questões que emergem fruto da facilidade com que os meios de controlo e vigilância se expandem e abrangem todas as facetas da vida social individual, leva a que os professores/investigadores em Direito assemelhem, numa perspetiva histórica,

o surgimento do *Big Data* a acontecimentos históricos de vigilância, de policiamento reativo, capazes de exacerbar o pânico moral. Estes profissionais comparam o surgimento do *Big Data* a esse tipo de policiamento político e reativo (hoje ultrapassado), na medida em que os grandes dados também intensificam a vigilância e o controlo, ao mesmo tempo que não garantem maior segurança:

> "No tempo do [policiamento político e reativo] […] as pessoas […] tinham medo de haver um polícia, […] ali a ouvir as conversas […]. E estes sistemas [como o *Big Data*]… digamos, a perceção do risco cada vez mais generalizado de que a informação que nós podemos transmitir através dos vários sistemas eletrónicos que utilizamos possa cair em mãos que nós não sabemos bem [quem são]…" [Professor/investigador]

Metaforicamente, é como se o policiamento político e reativo, a entidade responsável no passado pela vigilância, controlo e supervisão dos cidadãos, se tenha convertido no *Big Data* e as instâncias de controlo sejam cada vez mais invisíveis perante os cidadãos controlados; por sua vez, estes últimos convertem-se em entidades sempre visíveis ao olhar atento e omnipresente das instâncias de controlo. Dito de outro modo, é como se se tivesse perdido o controlo sobre quem controla e exerce a vigilância, tudo o que é feito pode ser escrutinado por figuras invisíveis e, por isso, não identificáveis. E tudo isso aumenta, exacerba e reproduz o pânico moral. Este paralelismo entre o passado e o presente e a transposição de uma realidade passada para a criação da expectativa futura são o que Brown e Michael (2003) designam de *prospecting retrospects* – a forma como as perspetivas passadas são implantadas em tempo real, para construir o futuro. Brown *et al.* (2006) enfatizam também estas noções: as expectativas constroem-se com base em heróis retrospetivos.

Quanto a esta invisibilidade – das instâncias de vigilância perante os vigiados –, Ball *et al.* (2016) referem que ela existe precisamente porque, por via do *Big Data*, construiu-se uma relação digital entre os sistemas de vigilância e o indivíduo, que substitui o lugar da tradicional relação de proximidade interativa. Assim, os cidadãos estão cada vez mais visíveis perante as instâncias de controlo e vigia, ao mesmo tempo que estas últimas estão cada vez mais invisíveis perante os cidadãos. A ideia de que se alterou o cerne das relações humanas por via dos dados é mencionada pelos membros do corpo de investigação criminal a exercerem funções laborais na Polícia Judiciária:

> "Nós alterámos o paradigma da nossa relação através dos dados. […] Nós hoje produzimos muito mais rasto e temos mecanismos que registam muito mais a nossa pegada de presença do que antes." [Corpo de Investigação Criminal]

Relativamente a esta forma de (re)ver ou "ver de novo" a forma como agimos e comunicamos hoje em dia, Boyd e Crawford (2012) mencionam que o *Big Data* surgiu como um sistema de conhecimento, o que desde logo altera os objetos do conhecimento, ao mesmo tempo que tem o poder de informar como entendemos redes e comunidades. Nesta era de digitalização das relações humanas e sociais, Frade (2016) define a contemporaneidade como o tempo do digital e das comunicações em rede, possibilitada pelo surgimento, expansão e grande desenvolvimento das novas tecnologias. Consequentemente, segundo Aas (2006), por via do uso crescente das tecnologias, a ação humana é codificada, ou seja, convertida em símbolos com significado especial. Isto vai ao encontro do que os profissionais entrevistados percecionam: a alteração do nosso agir comunicacional. A este propósito, Matzner (2016) refere que se criou uma relação humano-algoritmo que moldou a forma como os seres humanos são tratados e classificados. Além disso, verifica-se um alargamento da rede social de controlo, densificada com a emergência destas novas formas de vigiar, que podem levar a uma maior perceção do risco e um maior sentimento de insegurança, reforçando fragilidades e enfraquecendo a segurança (Zedner, 2016; Skinner, 2018; Gonçalves, 2018).

Os profissionais apontam aspetos sociais quando revelam a sua perceção acerca das expectativas que têm para o futuro desenvolvimento da técnica – há uma relação híbrida entre o meio social e as tecnologias inovadoras que nele emergem (Van Lente, 2012).

> "Temos mecanismos que registam muito mais a nossa pegada de presença do que antes. A Via Verde, o telemóvel, sempre com localização celular. […] estamos sempre a deixar marcas […]." [Corpo de Investigação Criminal]

Os membros do corpo de investigação criminal e os NCP2 entrevistados verbalizam esta visão social do *Big Data*, ao contrário dos restantes grupos profissionais. Tecem considerações acerca das alterações sociais e pessoais que a génese de uma nova forma de exercer a vigilância pode acarretar.

Nesta linha de pensamento, os professores/investigadores especializados em Direito alegam que esta forma de vigilância pode facilmente converter a sociedade num meio semelhante ao *Big Brother*. Ou seja, surge uma perspetiva

informada e uma associação clara entre técnicas do *Big Data* e uma nova forma de exercer a vigilância – tecnológica, omnipresente e digital. Além disso, a caracterização profissional desta técnica como um *Big Brother* revela o risco e o perigo advindos da sua utilização excessiva, que poderá criar níveis máximos de vigilância (observação e controlo constante) sobre os cidadãos.

> "Novos problemas que se põem; isto é, no fundo lembramo-nos sempre de *Big Data-Big Brother.* Evidentemente que os sistemas policiais são dos mais interessados nesse cruzamento, os sistemas de segurança em geral." [Professor/investigador]

O *Big Data* surge, então, como fenómeno potenciador de toda esta panóplia de previsões e observações constantes, como resposta ao desenvolvimento sem precedentes do mundo tecnológico, potenciando e exacerbando a ideia de que estamos inseridos numa sociedade *Big Brother*, à semelhança do que Lupton e Michael (2017) referiram acerca das perceções públicas sobre uma vigilância de dados: esta aceção generalizada de que uma vigilância que opera através de uma grande quantidade de dados potencia(rá) a imersão humana numa sociedade *Big Brother.* Matzner (2016), Orwell (2009) e Coll (2014) também debatem o surgimento de uma sociedade panótica, por via de uma vigilância permanente, invisível e que penetra em todos os interstícios sociais. Ou seja, por um lado, o *Big Data* é visto como uma ferramenta poderosa; por outro lado, é visto como uma manifestação preocupante do *Big Brother* (Boyd & Crawford, 2012; Lyon, 2014; Orwell, 2009; Coll, 2014).]

Consequentemente, os profissionais entrevistados consideram que esta prática não pode aplicar-se de forma desmedida, sob pena de violar direitos humanos. Defendem que se trata de um fenómeno que deve estar previsto, de forma restrita, na lei, definindo a forma como se executa, por quem e para que fins. A maior parte dos entrevistados geneticistas forenses, membros de autoridades consultivas e profissionais de investigação criminal, entende que esta deve ser uma técnica amplamente regulada e reconhece que a sociedade está a evoluir e a criar bases no sentido dessa regulamentação.

> "[O *Big Data* é] uma questão de regulação pelo Direito […]." [Membro de uma Comissão de Proteção de Dados]

> "[O *Big Data*] deve ser devidamente enquadrado segundo […] juristas […]." [Geneticista forense]

> "[O *Big Data*] deve ser estritamente regulado, mas eu penso que estamos a caminhar nessa direção […]." [Professor/investigador]

Estes entrevistados consideram que o Direito, no sentido da legislação e regulamentação, é uma área muito importante pois baliza legalmente o número e o tipo de profissionais autorizados a manusear, recolher e tratar estes dados. Há uma clara preocupação, por parte destes profissionais, sobre quem tem acesso aos dados, visto que estão disponíveis e acessíveis a qualquer pessoa, podendo ainda ser cedidos por qualquer pessoa de forma voluntária. Então, importa definir claramente quem pode ter acesso aos dados e em que termos.

Esta perspetiva revela a falta de confiança que os profissionais considerados têm nas instituições empresariais responsáveis pela recolha de dados, preocupação frequentemente verbalizada (Machado & Silva, 2016). Segundo Lupton e Michael (2017), trata-se de uma preocupação muitas vezes referida relativamente ao exercício da vigilância por via dos dados.

Concluindo, estes entrevistados fazem um exercício de reflexão dual: ponderam os benefícios deste desenvolvimento tecnológico, mas, em simultâneo, operam uma abstração temporal e social que permita considerar de que forma se deve e se pode intervir em momentos-chave para que este desenvolvimento inovador não se torne prejudicial (Brown & Michael, 2003). Ou seja, o seu discurso é cautelosamente esperançoso (Gardner *et al.*, 2015), pois defendem a regularização legal desta atividade. Estes recortes discursivos salientam a necessidade de pensar criticamente os ideais inovadores do futuro, analisando, por via de um paradigma preditivo, a melhor forma de proceder (Brown & Michael, 2003). Em simultâneo demonstram o que Van Lente (2012) referira: os profissionais com formação académica especializada em Direito remetem os seus discursos para a necessidade de desenvolver instrumentos legais que apoiem e sustentem um progresso equilibrado para as tecnologias inovadoras.

> "Para que essas situações [de cruzamento de um grande número de informações] possam passar a ser prática corrente, eventualmente terão de sair mais leis para proteção de determinados interesses dos indivíduos. […] não pode ser qualquer pessoa que possa ter acesso a esse tipo de informações e ao cruzamento de essas informações, e não pode fazer o cruzamento assim aleatoriamente […]. Tem de estar tudo muito bem regulado e muito bem definido […]." [Geneticista forense]

> "Ainda é o problema da limitação da finalidade, e o acesso a esses bancos de dados [do *Big Data*] deve ser restrito apenas a autoridades e órgãos autorizados." [Membro de Comissão de Proteção de Dados]

> "Se as pessoas socialmente e individualmente são cada vez mais abertas às suas marcas, à criação dos tais *Big Data*, à criação de situações rastreáveis, pois então o sistema de justiça […] precisa de perceber a conduta; vai ter de aceder a isso [*Big Data*]." [Corpo de Investigação Criminal]

Os profissionais entrevistados referem que o *Big Data* é potenciado pelas marcas que as pessoas vão conscientemente deixando; ou seja, os cidadãos criam as suas próprias pegadas, e estas são recolhidas, usadas e analisadas para cumprir determinados fins. No entanto, o acesso aos dados, a sua utilização e manuseio carecem de autorização judicial, o que reflete claramente a natureza privada destas informações pessoais e o modo restritivo como podem ser usadas *a posteriori*. Esta cedência voluntária de dados pessoais é definida, nas palavras de Matzner (2016), como o processo de fornecer ativamente informações pessoais (por exemplo, através de *sites* e redes), e é a forma mais disponível e utilizada de recolher dados para expandir a rede do *Big Data*. Quanto ao acesso aos dados, Boyd e Crawford (2012) alegam que importa que nos questionemos sobre quem acede aos dados, em que contextos esse acesso opera, com que restrições e que finalidades cumpre. O *Big Data* permite o acesso a uma enorme quantidade de informações produzidas por e sobre pessoas, interações e coisas. Igualmente, permite o acesso a informações pessoais através do cruzamento destes dados, podendo comprometer a privacidade dos indivíduos (Herschel & Miori, 2017).

Na mesma linha reflexiva, os NCP1 a desempenharem funções laborais num Instituto Forense enfatizam a ideia de que, quanto maior for o número de pessoas com acesso livre aos dados, maior é a probabilidade de se registarem violações de informação pessoal; daí a necessidade acrescida de se regulamentar e prever, normativamente, as instâncias e entidades responsáveis por este acesso e manuseamento:

> "Quanto maior o número de pessoas com acesso a essa informação [*Big Data*], maior será o risco de estas informações serem analisadas ou usadas sem consentimento, para outros fins." [NCP1]

Tendo em conta que se trata de um mecanismo ameaçador do ponto de vista da proteção das liberdades, garantias e direitos, e como forma de precaver

os efeitos e impactos sociais desta violação, é necessário que a tecnologia seja manuseada com validade científica, consciencialização dos seus perigos e riscos, de forma a diminuir e/ou atenuar os efeitos nocivos que pode provocar. Também os membros de Comissões de Proteção de Dados partilham esta perspetiva, defendendo uma definição legal e normativa que limite a disponibilidade dos dados recolhidos e restrinja a margem de ação dos profissionais para os manusear:

> "Tem de ser garantido que [a informação contida nas grandes bases de dados] é privada, que só é acedida em determinados contextos. […] eu acho que é essa linha que tem de ser traçada, [e] […] se calhar, uniformizada em todos os países." [Membro de Comissão de Proteção de Dados]

Nesta perspetiva, os geneticistas forenses a exercerem funções profissionais em Institutos Forenses também revelam alguma apreensão, nomeadamente no que concerne às entidades que têm acesso aos dados e aos fins para os quais estes dados são usados (Williams & Johnson, 2004; Albert *et al.*, 2009). No que respeita diretamente à temática dos direitos humanos, estes entrevistados demonstram claras preocupações ligadas ao direito à privacidade – concretamente, com as pessoas que são alvo deste escrutínio de informação e com o impacto, a nível da privacidade, que isso tem no seu quotidiano e nas suas liberdades e garantias.

> "[Utilizar os grandes dados como potencial meio de investigação criminal, genético e forense] Seria uma invasão da privacidade das pessoas […]. Ficar com essas informações quando são entidades que não estão acreditadas para o fazer ou… não concordo muito." [Geneticista forense]

Concluindo, a maior parte dos profissionais entrevistados considera que os aspetos éticos e de direitos humanos são questões a explorar e a analisar quando o tema é, precisamente, o *Big Data*, o uso desta técnica, a cedência de dados e o objetivo da recolha e tratamento das informações. Os NCP2 a desempenhar funções de *design* e gestão do *software* informático da base de dados afirmam:

> "[O *Big Data*] é uma questão ética." [NCP2]

Lucivero *et al.* (2011) referem que estas questões são frequentemente apontadas quando se trata de uma tecnologia emergente, porque as novas tecnologias alimentam-se de esperanças e medos injustificados – este arsenal

suscita discussões ao nível da ética das tecnologias. Também os professores/investigadores em Direito e membros de Comissões Biométricas com formação superior em Direito mencionam estes aspetos nas suas respostas, não incluindo questões que se prendem com quem tem acesso aos dados, mas enfatizando a necessidade de proteção de dados devido ao risco de invasão da vida privada que o uso da técnica pode implicar.

> "As implicações que isso [*Big Data*] possa ter, nos direitos, nas liberdades, não se trata apenas de direito à vida privada e à proteção de dados, mas de uma maneira mais geral dos condicionamentos das pessoas." [Professor/Investigador]

> "[O *Big Data*] levanta muitas questões sobre a privacidade." [Membro de Comissão Biométrica]

Existe uma representação socioprofissional de que o uso das técnicas do *Big Data* na recolha compulsiva e massiva de informações pessoais e privadas pode condicionar os cidadãos, restringindo os seus direitos e as suas garantias. Aliás, os professores/investigadores em Direito entrevistados verbalizam dilemas morais e éticos que conjugam os direitos defendidos pelo Estado, a finalidade das investigações criminais e o desenvolvimento tecnológico. Apesar de reconhecerem o equilíbrio político entre a garantia de segurança e a preservação do respeito pelos direitos dos cidadãos, questionam-se se, efetivamente, a investigação criminal, para cumprir os seus fins, poderá colocar em causa, de forma legítima, os direitos dos cidadãos. Não obstante, nesta linha de reflexão, consideram que, por força do desenvolvimento tecnológico, os mecanismos de investigação sofisticar-se-ão e, consequentemente, poderão colidir com direitos, liberdades e garantias fundamentais dos cidadãos:

> "Nós temos uma perspetiva na União Europeia [acerca do equilíbrio entre o direito à privacidade e o uso de dados pessoais] para assegurar o bem comum mais vocacionada para a proteção dos direitos fundamentais. […] [A UE] procura limitar, ou pelo menos garantir, que o uso de dados pessoais não é feito de forma completamente livre e irrestrita. […] mas é mesmo necessário limitar os direitos – o direito à privacidade e à proteção dos dados pessoais, para garantir uma melhor investigação criminal? […] vamos ter sempre esta tensão nos próximos anos […]. Porque o desenvolvimento tecnológico vai-nos permitir novos mecanismos que vão ser utilizados também pelo Estado neste âmbito da investigação criminal, e que vão colidir seguramente com direitos fundamentais dos cidadãos." [Professor/Investigador]

Esta é uma perspetiva frequentemente defendida pelos investigadores. É usual que, enquanto cientistas sociais, profissionais a desempenhar funções laborais de investigação académica questionem e integrem estes debates éticos (Selin, 2008).

Em contrapartida, os investigadores criminais a desempenhar funções policiais têm uma perspetiva contrária: veem o *Big Data* como um fenómeno legítimo, com potencialidade para armazenar um número considerável de informação acerca dos cidadãos, imprescindível para garantir a sua segurança. E o facto de essa informação estar armazenada nas bases de dados não constitui uma violação dos direitos, liberdades e garantias dos cidadãos, antes uma forma de salvaguardar a sua segurança.

> "Devemos ter muita informação para salvaguardar a segurança dos cidadãos, e o ter essa informação não põe em causa os direitos, liberdades e garantias do cidadão." [Corpo de Investigação Criminal]

Este último entrevistado tem uma perspetiva que tende a priorizar o bem-estar da sociedade perante os riscos de uma sociedade sob vigilância excessiva (Machado & Silva, 2016). Desta forma, o debate ético, transversal a todas as categorias profissionais, prende-se, efetivamente, em torno das questões da proteção de dados, da vigilância excessiva dos cidadãos e das potenciais ameaças que isso acarreta para os direitos civis, como a privacidade, a liberdade e a presunção de inocência (Machado *et al*., 2018). De uma forma geral, os profissionais entrevistados contribuem para este debate contemporâneo, refletindo acerca do vazio legal – percecionado como uma lacuna – existente em torno dos grandes dados e demonstrando preocupações acerca da potencial aplicação do *Big Data* na defesa dos direitos humanos. Estes profissionais fazem um exercício de reflexão dual: ponderam os benefícios deste desenvolvimento tecnológico, ao mesmo tempo que fazem uma abstração temporal e social que permita considerar de que forma se deve e se pode intervir em momentos-chave para que este desenvolvimento inovador não se torne prejudicial (Brown & Michael, 2003). Neste caso, enfatizam a necessidade de regular, legalmente, esta atividade. Estes recortes discursivos salientam a urgência de pensar criticamente os ideais inovadores do futuro e analisar a melhor forma de proceder, por via de um paradigma preditivo.

CONCLUSÃO

Este livro ambicionou explorar as expectativas dos pontos de contacto nacionais em rede transnacional de cooperação policial e judiciária, geneticistas forenses e *stakeholders* de diferentes áreas (ética e regulação, investigação criminal, pesquisa universitária, empresas privadas e organizações não governamentais), relativamente à potencial aplicação do *Big Data,* enquanto técnica, à investigação, prevenção e repressão da criminalidade. Além disso, visou também compreender os potenciais riscos e benefícios da inclusão de uma técnica inovadora de investigação criminal no contexto de partilha transnacional de dados na UE e identificar o debate contemporâneo em torno dos impactos éticos, sociais e de direitos humanos que a expansão do *Big Data* impulsiona no contexto de uma sociedade democrática. Não obstante, procurou também percecionar as transformações ocorridas na vigilância, paralelamente ao desenvolvimento contextual tecnológico e aos acontecimentos sociopolíticos que se foram desencadeando, bem como as perceções acerca desta matéria. Este conjunto de temas complexos, diversos, mas uníssonos, foram estudados de forma multifacetada, no sentido de compreender a diversidade de expectativas e perspetivas. Devido às diferentes referências tecidas pelos diversos entrevistados, os recortes discursivos foram interpretados e considerados à luz de um movimento dialético constante entre a teoria e as transcrições, de forma a adaptar e inserir os discursos na literatura e teoria emergente sobre o tema.

Tendo em conta que se trata de um tema embrionário, principalmente no que toca à sua relação com a criminalidade – mais propriamente, enquanto meio de investigação, prevenção e repressão criminal –, esta é uma obra que visa preencher parte de um vazio científico, construindo conhecimento nesta área e contribuindo para a produção de valências científicas no que toca ao

Big Data e à criminalidade; ou seja, desmontar estas temáticas complexas, atuais e importantes, entrelaçando-as e considerando-as sob a perspetiva de quem, potencialmente, as manuseará. A inexistência de literatura, em contexto europeu, que estabeleça esta relação e de trabalhos semelhantes ao aqui apresentado dificultou o processo de análise intercalar dos dados empíricos. Consequentemente, sendo o tema invisível, não existem dados que nos permitam comparar nem quantificar a temática, a não serem escassos estudos empíricos realizados em contexto não europeu. O facto de não existirem estudos na UE dificulta a consideração dos estudos existentes, visto que o contexto se altera e, consequentemente, as condições societais, legislativas e sociopolíticas também; por isso, os estudos foram descritos e analisados, mas não serviram de base de comparação.

Assim, esta obra nasceu fruto da superação de obstáculos e barreiras epistemológicas, que foram importantes para salientar a complexidade do tema e a compreensão do discurso atual, cheio de incertezas e considerações mutáveis, por se tratar de uma área emergente e em desenvolvimento.

A análise das abordagens que teorizam o *Big Data* permitiu concluir que se trata de uma temática envolvida nos debates contemporâneos da proteção de dados, da privacidade, das garantias de segurança e de liberdade. São temas que, envolvidos no seu arsenal teórico, estão, empiricamente, imbuídos de características sociais, políticas, ambientais e contextuais – tanto o crime, como os grandes dados. Além disso, a inexistência de protocolos que regulamentem a inserção da técnica nos meios de investigação criminal, a lacuna legal e a lacuna científica teórica dificultam a construção de perspetivas e expectativas por parte dos profissionais entrevistados, que muitas vezes constroem as suas significações ancoradas nas mensagens veiculadas pelos meios de comunicação social.

No que toca à análise dos recortes discursivos, independentemente dos grupos profissionais analisados, considera-se a perspetiva de que a inclusão do *Big Data* nas práticas de investigação criminal equivale a um progresso científico e tecnológico inovador e promissor do qual os profissionais beneficiariam. No entanto, apesar de esta perspetiva ser a expressa com maior frequência pelos entrevistados, existe uma consciência de que este trabalho deve ser parte de um outro maior, capaz de abarcar a complexidade dos temas e considerar a generalidade das perspetivas – meios de comunicação social e sociedade civil, por exemplo.

Tendo em conta a amostra analisada para efeitos de cumprimento do objetivo desta obra, os recortes discursivos permitem delinear linhas gerais

conclusivas sintéticas. De uma forma geral, foi notório que os NCP2 revelaram nas suas respostas mais conhecimento sobre o *Big Data* enquanto ferramenta potencialmente aplicada à investigação criminal. Paralelamente a conhecimentos *micro* detalhados, estes entrevistados expressaram expectativas promissoras, positivas, flexíveis à inclusão dos grandes dados e à sua partilha transnacional, como uma valência para a prevenção e repressão da criminalidade transfronteiriça e organizada. Assim, por via de esboços detalhados e revelando mínimas resistências à expansão e implementação de uma nova técnica no contexto da investigação criminal, os NCP2 foram o grupo profissional que adotou um discurso *instrumental* (Stevens *et al*., 2018).

Os geneticistas forenses e os NCP1, apesar de expressarem também expectativas promissoras e positivas acerca do tema, não mencionaram aspetos detalhados sobre a técnica de *Big Data*. De forma geral, mostraram-se, maioritariamente, favoráveis à inclusão da técnica, apresentando no entanto argumentos organizados para refutar previsões (Van Lente, 2012). São a categoria profissional mais autoritária, tecendo considerações sobre os obstáculos e dificuldades práticas atuais na inclusão dos grandes dados nas práticas de investigação criminal (Van Lente, 2012; Lucivero *et al*., 2011; Borup *et al*., 2006); por isso, o seu discurso é *crítico-interpretativo* (Stevens *et al*., 2018).

No que toca às temáticas categóricas criadas, em termos de conteúdo discursivo, todos os grupos profissionais mencionaram os riscos do *Big Data* – questões procedimentais e de validade científica, considerações acerca da fragilidade da validade da técnica e preocupações com o número e tipo de dados incluídos (Jasanoff & Kim, 2009; Borup *et al*., 2006; Brown & Michael, 2003; Machado & Silva, 2015). Os geneticistas forenses e os NCP revelaram maior preocupação sobre os riscos e perigos da inclusão de uma técnica inovadora contestada no seio da investigação criminal do que as restantes categorias profissionais. Isto deve-se ao facto de se encontrarem em posições e contextos laborais que, tecnicamente, os aproximam da partilha transnacional de dados, tornando-os mais conscientes das potenciais problemáticas práticas. Esta aproximação, laboral e técnica, dos grandes dados enquanto potencial ferramenta de trabalho teve efeito também em alegações contraditórias. Ou seja, os NCP e geneticistas forenses revelaram muitas vezes discursos ambivalentes entre, por um lado, as vantagens e benefícios promissores do *Big Data* enquanto ferramenta de prevenção e combate à criminalidade e, por outro, as dificuldades de inclusão da técnica e a sua fragilidade científica (Van Lente, 2012; Lucivero *et al*., 2011; Borup *et al*., 2006). Tal reflete a importância que o contexto e meio laboral dos profissionais tem na construção de argumentos

discursivos sobre temáticas e fenómenos. Além disso, estes entrevistados também teceram considerações políticas sobre o fenómeno, convertendo-o numa prática diretamente dependente das instâncias e órgãos políticos.

Não obstante, os NCP foram o grupo profissional que mencionou com maior ênfase a utilidade da técnica. Em contrapartida, não teceram tantas considerações éticas e de direitos humanos como os restantes grupos profissionais. Além disso, os entrevistados a desempenhar funções laborais como investigadores sociais foram a categoria profissional que mais vezes questionou quem tinha acesso aos dados (Selin, 2008). No geral, todos os profissionais oscilam entre o grau de novidade da técnica e as incertezas organizacionais, fazendo abstrações temporais e antecipatórias. Tudo depende da sua posição laboral e consequente localização, como já referido. Ainda assim, é importante ressalvar que se verifica um desfasamento entre o presente material e as expectativas, assentando estas nos discursos e numa prática que é inexistente (Brown & Michael, 2003).

Adotando uma lente mais concreta, no que toca às perspetivas que se prendem com os benefícios e vantagens do potencial uso do *Big Data* no combate e prevenção da criminalidade, destacam-se os geneticistas forenses e os NCP2. Estes entrevistados têm uma visão otimista do tema, na medida em que afirmam acreditar que esta tecnologia, sendo uma ferramenta de investigação criminal valiosa, potencia(rá) um combate mais eficiente e eficaz da criminalidade – esta é, normalmente, a perspetiva de cientistas forenses com formação académica superior em Medicina Legal, Bioquímica, Ciências Forenses e Biologia (Machado, 2011; Maciel & Machado, 2014). Os entrevistados com formação académica superior no campo da saúde e da genética defendem que o *Big Data* contribuiria com maior eficiência para o combate à criminalidade, demonstrando frequentemente uma visão mais positiva e otimista da ciência e da tecnologia em geral (Machado & Silva, 2015). O *Big Data* é referido por estes profissionais como uma técnica que, dada a sua potencial qualidade e eficácia nas práticas policiais de investigação criminal, poderá ter uma utilidade operacional e ocupar uma posição científica. Estes profissionais procedem, frequentemente, a uma *neutralização* dos riscos associados às liberdades civis e aos direitos humanos, no que concerne à implementação e expansão do *Big Data* enquanto técnica de investigação criminal.

Nesta linha perspetivacional, o grupo dos NCP1 adota uma posição distinta: reconhece os potenciais benefícios, mas enfatiza a necessidade de aumentar os recursos e as valências económicas e profissionais, para que esta técnica se efetive. Além disso, é a única categoria profissional que, com uma opinião

ambivalente, revelou bastantes obstáculos laborais práticos à inserção do *Big Data* no seu quotidiano profissional e resistências face ao desenvolvimento tecnológico inovador. Em contrapartida, excetuando os NCP2, todos os restantes profissionais entrevistados perspetivam o *Big Data* como uma técnica que potencia a eficácia das decisões tomadas –com base em algoritmos que aumentam a sua precisão –, convertendo esta tecnologia numa ferramenta com um grau de objetividade maior.

No que toca às representações socioprofissionais relacionadas com os desafios éticos e sociais a enfrentar com o potencial uso do *Big Data* (benefícios e riscos), nos vários grupos profissionais é possível salientar uma tipologia: os professores/investigadores com formação académica superior em Direito e nas Ciências Sociais, os membros de Comissões de Ética também com formação especializada em Direito e os membros pertencentes a Organizações Não Governamentais de Direitos Humanos mencionam o carácter peculiar, minucioso e sensível dos dados que esta técnica engloba, manuseia, agrega e partilha. Estes entrevistados têm uma perspetiva "crítica", pessimista e defensora da proteção de dados e da privacidade, ressalvando a necessidade de salvaguardar informação sensível (Machado & Silva, 2015). No fundo, respondendo às consequências da mudança através da antecipação de problemas socioéticos e de uma avaliação crítica do processo inovador, estes profissionais auxiliam a construção do futuro (Hedgecoe & Martin, 2003). Fazem, frequentemente, avaliações processuais da técnica do *Big Data,* ou seja, analisam a forma como os dados são recolhidos, analisados, processados, agregados, convertidos em algoritmos e, posteriormente, confirmam os fins para que são utilizados (Albert *et al*., 2009).

Os profissionais ligados à vertente prática da aplicação das técnicas do *Big Data* – profissionais de investigação criminal e NCP – surgem com uma orientação utilitarista e um posicionamento mais permissivo, entusiasta e expansivo. Esta perspetiva distingue as diferentes fontes de recolha dos dados e avalia a legitimidade da extração dos mesmos de acordo com distintos comentários relacionados com a prática de investigação (Albert *et al*., 2009; Van Lente, 2012). Ou seja, os profissionais que desempenham funções laborais em setores onde o *Big Data* poderá futuramente constituir-se como técnica de investigação criminal – nomeadamente, Departamentos Policiais – estão mais conscientes da (potencial) existência e utilidade desta técnica do que outros grupos profissionais (Machado & Silva, 2015).

Por fim, surgem nuances entre diversos entendimentos de atores sociais diferentemente posicionados, nomeadamente os geneticistas forenses que

oscilam entre uma vertente mais otimista e outra mais pessimista. Esta perspetiva reside no limbo entre os benefícios da (potencial) aplicação da técnica e a incerteza baseada nos riscos que a mesma (potencial) aplicação pode acarretar. Brown e Michael (2003) referem que esta incerteza é mais acutilante e verifica-se sobretudo nas expectativas dos profissionais intimamente envolvidos na produção do conhecimento, onde a sua experiência acerca das contingências de produção de conhecimento no laboratório os torna mais cautelosos (Brown *et al*., 2003). Estas expectativas residem num nível mais *macro* porque se trata de esboços abstratos e abrangentes (Van Lente, 2012).

Concluindo, é notório que existem variadas e distintas formas de observar o fenómeno do *Big Data* – repleto de sentidos, avaliações, interpretações e definições. Cada profissional possui um olhar *seu* sobre o que lhe é questionado (Borup *et al*., 2006; Kruse, 2016). Assim, emergem *culturas epistémicas* diferentes quando se analisam os diferentes discursos narrativos. Estas divergências convertem a expansão, inserção e evolução do *Big Data* num percurso desafiador e sinuoso, visto que cada profissão define esta técnica mediante a sua ótica, que corresponde à tarefa laboral que desempenha (Kruse, 2016). Estas *culturas* espelham o facto de os diferentes profissionais possuírem vários entendimentos e significações sobre as questões em apreço (Borup *et al*., 2006; Kruse, 2016), sendo a "profissão" uma variável com impacto em termos de posicionamento perante as tecnologias de vigilância. Os tipos de discurso sobre a possibilidade futura da integração de uma tecnologia inovadora no controlo, combate, prevenção e repressão da criminalidade variam de acordo com o posicionamento dos entrevistados. A maior parte das expectativas e perceções sobre o *Big Data* são especulações diferentes que não permitem situar o (potencial) desenvolvimento da técnica – este facto é uma boa oportunidade para traçar a forma como uma técnica (enquanto conjunto de tecnologias) é moldada por diferentes visões durante a sua emergência. Assim, foi possível observar como as expectativas são constituídas em redes de conhecimento e projetadas em possibilidades futuras. Como técnica embrionária, o *Big Data* é um objeto útil para pensar o modo como as expectativas se condensam em volta de questões particulares, em diferentes momentos e lugares e por diferentes razões técnicas, políticas e culturais. Continuamente, observaremos a forma como este futuro se desenvolverá.

Sob este contexto complexo, a nível prático, é importante que este estudo impulsione e estimule a reflexão. Em primeiro lugar, pensar a vigilância sob a ótica do *Big Data* leva-nos a prosseguir com várias linhas de pensamento, nomeadamente assentes no seguinte debate: será a vigilância necessária e útil?

Ou existe para ser aplicada e expandida por instituições que visam a recolha massiva e o manuseamento de dados sobre cidadãos para fins que não são compatíveis com o respeito pelas liberdades cívicas e direitos humanos? A vigilância serve efetivamente fins preventivos de ações criminais ou é uma invasão da vida privada?

É importante que se prossigam trabalhos académicos deste teor com vista a explorar, teórica e empiricamente, os fins que serve este tipo de vigilância digital, eletrónica e fluida. No fundo, a pergunta é: o que legitima estas ações é validado quando as informações obtidas são usadas?

Em Portugal, "a prevenção dos crimes […] só pode fazer-se com observância das regras gerais sobre polícia e com respeito pelos direitos, liberdades e garantias dos cidadãos" (artigo 272.º, n.º 3, da Constituição da República Portuguesa; Canotilho, 2005), "a actividade de segurança interna pauta-se pela observância dos princípios do Estado de direito democrático, dos direitos, liberdades e garantias e das regras gerais de polícia" (artigo 2.º da Lei de Segurança Interna – n. º 53/2008 de 29 de agosto). A lei não permite o controlo das atividades dos cidadãos através da sua monitorização, alegando a defesa da integridade e privacidade da pessoa.

O *Big Data* – técnica digital, informática e tecnológica – surgiu e transformou a forma como a vigilância opera nas sociedades atuais. As transformações ocorridas permitem-nos falar de um novo paradigma, mas que não elimina os paradigmas anteriores vigentes. Apesar de todas as alterações verificadas, existem vestígios da vigilância tradicional, como, por exemplo, o policiamento preventivo em diversos espaços públicos e/ou a presença física de agentes policiais a supervisionar vários lugares; ou seja, observam-se ainda práticas de vigilância física.

Não obstante, o *Big Data* surge como uma verdadeira oportunidade para os autores aplicarem e realizarem pesquisas baseadas em abordagens multidisciplinares (Albert *et al*., 2009) – cruzamento de ideias, troca de informação, inclusão de diversos conhecimentos de distintas áreas –, enriquecendo as pesquisas e investigações. No fundo, potencia a construção de redes de conhecimento internacionais, fortalecendo e enfatizando a importância da colaboração científica globalizada. A construção de conhecimento científico é cada vez mais percecionada como um caminho fruto de esforços coletivos, que envolve cruciais atividades de interação, colaboração e troca de ideias. Assim, o desenvolvimento e a inclusão de redes que permitam o acesso e a partilha deste conhecimento são aspetos fundamentais integrantes da disseminação e criação de conhecimento (Fontes & Araújo, 2013). Ou seja, não podemos

criar dois polos opostos porque a tecnologia nem é benéfica, nem maléfica, e também não é neutra. No fundo, é a interação desta com o meio social que produz consequências nos mais diversos setores. Por isso, é necessário que as políticas que defendem e elogiam o *Big Data* enquanto técnica de controlo, vigilância e prevenção da criminalidade a avaliem e concluam acerca da sua eficácia e utilidade para que se possam tomar posições e decisões mais consistentes e informadas.

O desenvolvimento do *Big Data* suscita diversas questões éticas e deontológicas no que toca ao acesso a uma grande variedade de dados e à vigilância massiva a que estão sujeitos os cidadãos. Em suma, urge debater este fenómeno à luz da defesa de princípios de transparência, confiança pública e prestação de contas por parte das agências governamentais. Questionar e problematizar o *Big Data* significa caminhar para tornar mais democráticas e coesas as sociedades tecnológicas contemporâneas. Importa saber quem tem acesso a estes dados, o que significam, como é que a sua análise é distribuída e com que finalidades. Para além disso, importa construir bases sólidas reflexivas no conhecimento da potencialidade do *Big Data* para o definir exclusivo na produção de informação e na gestão de ações e decisões.

Urge dizer que entramos numa era em que tudo é quantificado e em que as relações humanas perderam o seu sentido íntimo, porque tudo se reduz a números – é como um exercício de observação dos fenómenos por via de uma lente numérica, matemática e algorítmica, sendo esta a única visão existente. Consequentemente, o *Big Data* diminui também, cada vez mais, os limites entre o público e o privado. Neste sentido, e analisando as consequências sociais visíveis da expansão dos grandes dados, é importante ressalvar que o *Big Data* não é uma moda que se instala por um tempo definido, mas antes uma característica que define a civilização moderna, desde a revolução científica. E caso não seja regulamentada, os problemas advindos da sua intensificação massiva tendem a explodir, visto que no futuro as pessoas deixarão cada vez mais pegadas digitais passíveis de análise, revelando detalhes minuciosos sobre as suas vidas diárias. Há assim uma necessidade de reconhecimento da importância deste fenómeno, dos seus problemas e desafios. Nenhuma solução, por si só, será suficiente, tendo em conta a complexidade da técnica, profundamente entrelaçada na história e na cultura. O que é necessário é um debate público sobre a aceitabilidade ética, a desejabilidade social e os valores-chave que uma comunidade visa defender.

Para além disso, e refletindo acerca da possibilidade de desenvolver investigações futuras para melhor entender e aceder à complexidade do *Big Data*,

importa, por exemplo, explorar o estado epistemológico desta ferramenta e descortinar o lugar onde se posiciona no campo filosófico (Matzner, 2016). Conjuntamente, num exercício de reflexão dual, importa estudar as questões levantadas por Snowden – "que tipo de sociedade queremos?" (Lyon, 2014) – e, no seguimento do que Bauman (2009) afirmara acerca da fluidez das relações sociais na Modernidade, o fluxo de dados cada vez mais livre no contexto social, político e económico. Abstraindo o objeto de estudo, criam-se condições para observar paralelismos entre estas características e as atribuídas ao *Big Data*: a questão da liquidez da vigilância no âmbito social e político, ao nível do fluxo de dados. Um tema de vigilância líquida é a necessidade de práticas devidamente éticas (Lyon, 2014). Os dilemas morais que surgem ocorrem na proximidade ética e ontológica do sujeito, dentro de relacionamentos mediados digitalmente, e são experimentados como produtivos e repressivos. Além disso, ecoando a discussão anterior sobre as questões éticas como uma continuidade das práticas de vigilância do passado para o presente, as noções de normas e competitividade também estão representadas nos estudos de vigilância. Assim, é fundamental analisar de que modo lidamos com o surgimento de uma *Era de Grandes Dados*, recordando que, ocorrendo o fenómeno num ambiente de incerteza, rapidez e mudança, as decisões atuais moldarão o futuro. Com a automação aumentada de recolha e análise de dados – bem como algoritmos que podem extrair e ilustrar padrões de grande escala no comportamento humano –, é necessário perguntar quais os sistemas que estão a dirigir essas práticas e quais as suas entidades reguladoras (Boyd & Crawford, 2012).

Estas constatações tornam evidente o facto de que, realmente, o *Big Data* representa um desafio para o respeito pelo direito à privacidade. Os grandes dados comprometem regras e deveres morais porque, em muitos aspetos, tornaram-se rapidamente muito poderosos, penetrantes e essenciais para a vida quotidiana. Isso cria um desafio moral para as sociedades, porque as pessoas querem usar as próprias tecnologias que criam o *Big Data* mas também querem tentar controlar a forma como isso as afeta. Tudo isto clarifica a ideia de que o advento do *Big Data* apresenta imensas questões preocupantes, mesmo para os entrevistados, relativamente à proteção de dados. O Comité Consultivo da Convenção tenta responder a estes dilemas, e o ponto central da legislação reside no fornecimento de respostas por via de uma abordagem que se adeque às tecnologias em evolução e às mudanças sociais (Mantelero, 2017). Por um lado, o Regulamento (UE) 2016/679 tem por objetivo definir o paradigma que regula a proteção de dados para as próximas décadas e introduz várias alterações à estrutura do sistema jurídico existente. Contudo, o Regulamento continua a

basear-se, principalmente, nos princípios que inspiraram as anteriores regulamentações nacionais nos anos 1990, sem resolver os problemas relativos à aplicação destes princípios aos novos cenários do *Big Data*. Embora o debate sobre estas duas formas de regulamentar a tecnologia continue aberto, parece difícil chegar a uma resposta conclusiva: o caso da regulamentação do *Big Data* levanta muitas perguntas sobre se o Regulamento 2016/679 representa a estrutura de proteção de dados para os próximos anos (Mantelero, 2017).

Deve haver um envolvimento público em torno destas questões, e a comunidade deve desempenhar um papel mais ativo na compreensão deste tema para poder influenciar o seu uso (Chan & Moses, 2017). Considerando a crescente expansão de redes de intercâmbio transnacional de dados, bem como de grandes dados, o estudo das perspetivas e atitudes dos profissionais envolvidos nesse intercâmbio é crucial. Assim, este livro não deve ser lido nem entendido como uma declaração final sobre o assunto, antes como uma contribuição modesta para uma questão crucial de debate que se torna cada vez mais relevante, proporcionalmente à forma como o mundo se torna cada vez mais digital.

BIBLIOGRAFIA

Aas, K. (2006). The body does not lie: Identity, risk and trust in technoculture. *Crime, Media, Culture*, *2*(2), 143-158.

Albert, M., Laberge, S., & Hodges, D. (2009). Boundary-work in the health research field: Biomedical and clinician scientists' perceptions of social science research. *Minerva*, *47*(2), 171-194.

Amoore, L. (2011). Data derivatives: On the emergence of a security risk calculus for our times. *Theory, Culture & Society, 28*(6), 24-43.

Andrejevic, M., & Gates, K. (2014). Big data surveillance: Introduction. *Surveillance & Society, 12*(2), 185.

Andrews, D., & Bonta, J. (2010). *The Psychology of Criminal Conduct*. Newark: Lexis Nexis.

Araújo, E. (2008). Technology, gender and time: A contribution to the debate. *Gender, Work & Organization, 15*(5), 477-503.

Araújo, E., Cogo, D., & Pinto, M. (2015). Mobilidades, media(ções) e cultura. *Comunicação e Sociedade, 28*, 7-14.

Bakir, V., Feilzer, M., & McStay, A. (2017). Introduction to special theme veillance and transparency: A critical examination of mutual watching in the post-Snowden, Big Data era. *Big Data & Society*, *5*(1), 1-5.

Ball, K., Di Domenico, M., & Nunan, D. (2016). Big data surveillance and the body-subject. *Body & Society*, *22*(2), 58-81.

Bartlett, A., Lewis, J., Reyes-Galindo, L., & Stephens, N. (2018). The locus of legitimate interpretation in Big Data sciences: Lessons for computational social science from -omic biology and high-energy physics. *Big Data & Society*, 5(1), 1-15.

Bauman, Z. (2009). *Confiança e medo na cidade*. Rio de Janeiro: Zahar.

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662-679.

Borup, M., Brown, N., Konrad, K., & Van Lente, H. (2006). The sociology of expectations in science and technology. *Technology analysis & strategic management, 18*(3-4), 285-298.

Brayne, S. (2017). Big data surveillance: The case of policing. *American Sociological Review*, *82*(5), 977-1008.

Brandão, A. (2010). *E se tu fosses um rapaz? Homo-erotismo feminino e construção social da identidade*. Porto: Edições Afrontamento.

Brown, N., Kraft, A., & Martin, P. (2006). The promissory pasts of blood stem cells. *BioSocieties*, *1*(3), 329-348.

Brown, N., & Michael, M. (2003). A Sociology of Expectations: Retrospecting Prospects and Prospecting Retrospects. *Technology Analysis and Strategic Management*, *15*(1), 3-18.

Brown, N., Rip, A., & Van Lente, H. (2003). Expectations in & about science and technology. In *A Background Paper for the 'Expectations' Workshop of 13-14 June* (pp. 13-14). Retirado de: https://www.york.ac.uk/satsu/expectations/Utrecht%202003/Background%20paper%20version%2014May03.pdf .

Caregnato, A., & Mutti, R. (2006). Pesquisa qualitativa: análise de discurso versus análise de conteúdo. *Texto contexto enferm*, *15*(4), 679-84.

Chan, J., & Moses, L. (2015). Is Big Data challenging criminology? *Theoretical Criminology*, *20*(1), 21-39.

Chan, J., & Moses, L. (2017). Making sense of Big Data for security. *British Journal of Criminology*, *57*(2), 299-319.

Charmaz, K. (2009). *A construção da teoria fundamentada: Guia prático para análise qualitativa*. Porto Alegre: Artmed.

Cho, Y., & Lee, H. (2014). Reducing confusion about grounded theory and qualitative content analysis: Similarities and differences. *The qualitative report*, *19*(32), 1-20.

Cinnamon, J. (2017). Social Injustice in Surveillance Capitalism. *Surveillance & Society*, *15*(5), 609- 625.

Clarke, A. (2005). *Situational Analysis: Grounded Theory after the Postmodern Turn*. Thousand Oaks: Sage.

Cole, S. (2013). Forensic culture as epistemic culture: The sociology of forensic science. *Studies in History and Philosophy of Biological and Biomedical Sciences*, *44*(1), 36-46.

Coll, S. (2014). Power, knowledge, and the subjects of privacy: understanding privacy as the ally of surveillance. *Information, Communication & Society, 17*(10), 1250-1263.

Costa, D. (2004). Sociedade de controle. *São Paulo em perspectiva, 18*(1), 161-167.

Cunha, M. (2008). Disciplina, controlo, segurança: No rasto contemporâneo de Foucault. In C. Fróis (Org.), *A sociedade vigilante: Ensaios sobre privacidade, identificação e vigilância* (pp. 67-81). Lisboa: Imprensa de Ciências Sociais.

Dimeglio, C., Kelly-Irving, M., Lang, T., & Delpierre, C. (2018). Expectations and boundaries for Big Data approaches in social medicine. *Journal of forensic and legal medicine*, *57*, 51-54.

Drewer, D., & Miladinova, V. (2017). The BIG DATA challenge: Impact and opportunity of large quantities of information under the Europol Regulation. *Computer Law & Security Review*, *33*(3), 298-308.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing*, *62*(1), 107-115.

Feeley, M., & Simon, J. (1992). The new penology: Notes on the emerging strategy of corrections and its implications. *Criminology*, *30*(4), 449-474.

Ferreira, S. (2014). A sociedade da informação como sociedade de disciplina, vigilância e controle. *Información, cultura y sociedad*, (31), 109-120.

Fontes, M., & Araújo, E. (2013). (I) Mobilidades e redes científicas internacionais: Contextos e relações em mudança. In E. Araújo, M. Fontes & S. Bento (eds.), *Para um debate sobre Mobilidade e Fuga de Cérebros* (pp. 97-124). Braga: Centro de Estudos de Comunicação e Sociedade, Universidade do Minho.

Foucault, M. (1994). *História da sexualidade. A vontade de saber*. Lisboa: Relógio d'Água.

Foucault, M. (1999). *Vigiar e punir: Nascimento da Prisão*. Petrópolis: Vozes.

Frade, C. (2016). Social theory and the politics of Big Data and method. *Sociology*, *50*(5), 863-877.

Fróis, C. (2007). Knowing Me, Knowing You: a vigilância enquanto objecto de estudo etnográfico. In *Comunicação apresentada na Conferência Ethnografeast III, Ethnography and the Public Sphere*, ISCTE (Vol. 20).

Fróis, C. (2015). Dos estudos de vigilância, videovigilância e tecnologia. Reflexão sobre o estado da arte. In M. Cunha (Ed.), *Do crime e do castigo: temas e debates contemporâneos* (pp. 147-157). Lisboa: Editora Mundos Sociais.

Fróis, C., & Machado, H. (2016), Modernization and development as a motor of polity and policing. In B. Bradford, B. Jauregui, I. Loader & J. Steinberg (Eds.), *The SAGE handbook of global policing* (pp. 391-405). London: Sage Publications.

Fuchs, C. (2011). Como podemos definir vigilância? *Matrizes*, *5*(1), 109-136.

Gandy Jr, H. (1989). The surveillance society: information technology and bureaucratic social control. *Journal of Communication*, *39*(3), 61-76.

Gardner, J., Samuel, G., & Williams, C. (2015). Sociology of low expectations: Recalibration as innovation work in biomedicine. *Science, Technology, & Human Values*, *40*(6), 998-1021.

Gonçalves, M. (2018, 20 de março). Quem controla os nossos dados? *Público*, pp. 1-5. Retirado de https://www.publico.pt/2018/03/20/tecnologia/opiniao/quem-controla-os-nossos-dados-1807206

Gonçalves, M. (2017). The EU data protection reform and the challenges of big data: remaining uncertainties and ways forward. *Information & Communications Technology Law*, *26*(2), 90-115.

Graham, S. (1998). Spaces of surveillant simulation: New technologies, digital representations, and material geographies. *Environment and Planning D: Society and Space*, *16*(4), 483-504.

Guzik, K. (2009). Discrimination by Design: predictive data mining as security practice in the United States 'war on terrorism'. *Surveillance & Society*, *7*(1), 3-20.

Halford, S., & Savage, M. (2017). Speaking sociologically with big data: symphonic social science and the future for big data research. *Sociology*, *51*(6), 1132-1148.

Hedgecoe, A., & Martin, P. (2003). The drugs don't work: expectations and the shaping of pharmacogenetics. *Social studies of science*, *33*(3), 327-364.

Herschel, R., & Miori, V. (2017). Ethics & Big Data. *Technology in Society*, *49*, 31-36.

Jasanoff, S., & Kim, S. (2009). Containing the atom: Sociotechnical imaginaries and nuclear power in the United States and South Korea. *Minerva*, *47*(2), 119-146.

Johnson, P., & Williams, R. (2007). Internationalizing new technologies of crime control: forensic DNA databasing and datasharing in the European union. *Policing & Society*, *17*(2), 103-118.

Kierkegaard, S. (2008). The Prüm decision – An uncontrolled fishing expedition in 'Big Brother'Europe. *Computer Law & Security Review*, *24*(3), 243-252.

Kruse, C. (2016). *The Social Life of Forensic Evidence*. Oakland, CA: University of California Press.

Kubler, K. (2017). State of urgency: Surveillance, power, and algorithms in France's state of emergency. *Big Data & Society*, *4*(2), 1-10.

Lawless, C. & Williams, R. (2010). Helping with inquiries or helping with profits? The trials and tribulations of a technology of forensic reasoning. *Social Studies of Science*, *40*(5), 731-755.

Lefèvre, T. (2017). Big data in forensic science and medicine. *Journal of Forensic and Legal Medicine*, *57*, 1-6.

Lyon, D. (1992). The new surveillance: Electronic technologies and the maximum security society. *Crime, Law and Social Change*, *18*(1-2), 159-175.

Lyon, D. (2004). Globalizing surveillance: Comparative and sociological perspectives. *International Sociology*, *19*(2), 135-149.

Lyon, D. (2014). Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society*, *1*(2), 1-13.

Lyon, D. (2015). The Snowden stakes: challenges for understanding surveillance today. *Surveillance & Society*, *13*(2), 139-152.

Lucivero, F., Swierstra, T., & Boenink, M. (2011). Assessing expectations: Towards a toolbox for an ethics of emerging technologies. *NanoEthics*, *5*(2), 129-141.

Lupton, D., & Michael, M. (2017). 'Depends on Who's Got the Data': Public Understandings of Personal Digital Dataveillance. *Surveillance & Society*, *15*(2), 254-268.

Machado, H. (2011). Construtores da bio(in)segurança na base de dados de perfis de ADN. *Etnográfica*, *15*(1),153-166.

Machado, H., & Granja, R. (2018). Ethics in transnational forensic DNA data Exchange in the EU: constructing boundaries and managing controversies. *Science as Culture*, *27*(2), 242-264.

Machado, H., Queirós, F., Martins, M., Granja, R., & Matos, S. (2018). Vigilância genética, criminalização e coletivização da suspeição. In S. Gomes, V. Duarte, F. Ribeiro, L. Cunha, A. Brandão & A. Jorge (Orgs.), *Desigualdades Sociais e Políticas Públicas: Homenagem a Manuel Carlos Silva.* Braga: Edições Húmus (pp. 529-548).

Machado, H., & Santos, F. (2016). Culturas de objetividade, epistemologias cívicas e o suspeito transnacional. Uma proposta para mapeamentos teóricos em estudos sociais da genética forense. In C. Fonseca, F. Rohden, P. Machado & H. Paim (Orgs.), *Antropologia da ciência e da tecnologia: dobras reflexivas* (pp. 179-203). Porto Alegre: Sulina.

Machado, H., & Silva, S. (2015). Public perspectives on risks and benefits of forensic DNA databases: an approach to the influence of professional group, education, and age. *Bulletin of Science, Technology & Society*, *35*(1-2), 16-24.

Machado, H., & Silva, S. (2016). Voluntary participation in forensic DNA databases: altruism, resistance, and stigma. *Science, Technology, & Human Values*, *41*(2), 322-343.

Maciel, D., & Machado, H. (2014). Biovigilância e governabilidade nas sociedades da informação. In H. Machado & H. Moniz (Orgs.), *Bases de dados genéticos forenses. Tecnologias de controlo e ordem social* (pp. 141-166). Coimbra: Coimbra Editora.

MacWillie, J. (2018). From Keyhole to Big Brother: The Legacies of Early Cold War Surveillance. *Surveillance & Society*, *16*(2), 203-218.

Mantelero, A. (2017). Regulating big data. The guidelines of the Council of Europe in the context of the European data protection framework. *Computer Law & Security Review*, *33*(5), 584-602.

Marx, G. (2004). What's new about the "new surveillance"?: Classifying for change and continuity. *Knowledge, Technology & Policy*, *17*(1), 18-37.

Matos, S., Santos, F., & Machado, H. (2016). Criminalidade e geopolítica da ciência na União Europeia. In *Portugal, território de territórios. Atas do IX Congresso Português de Sociologia* (pp. 1-14). Lisboa: Associação Portuguesa de Sociologia.

Matos, S. (2018). Biometria e privacidade: desafios bioéticos na cooperação policial e judicial na União Europeia. In A. Sol & S. Gouveia (Eds.), *Bioética no Século XXI* (pp. 255-286). Charleston, USA: CreateSpace Independent Publishing.

Matzner, T. (2016). Beyond data as representation: The performativity of Big Data in surveillance. *Surveillance & Society*, *14*(2), 197-210.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 1-21.

Nunes, J. (1992). *As teias da família: A construção interaccional das solidariedades primárias*. (Tese de Doutoramento). Universidade de Coimbra, Coimbra.

Orwell, G. (2009). *1984*. São Paulo: Editora Companhia das Letras.

Pollock, N., & Williams, R. (2010). The business of expectations: How promissory organizations shape technology and innovation. *Social Studies of Science*, *40*(4), 525-548.

Prainsack, B. (2019). Logged out: Ownership, exclusion and public value in the digital data and information commons. *Big Data & Society*, *6*(1), 1-15.

Queirós, F. (2018). Retratos Biogenéticos no Combate à Criminalidade: Desafios Éticos e Sociais. In A. Sol e S. Gouveia (Eds.), *Bioética no Século XXI* (pp. 287-313). Charleston, USA: CreateSpace Independent Publishing.

Selin, C. (2008). The sociology of the future: tracing stories of technology and time. *Sociology Compass*, *2*(6), 1878-1895.

Skinner, D. (2018). Race, Racism and Identification in the Era of Technosecurity. *Science as Culture*, 1-23.

Souza, L. (2010). Dilemas e hesitações da modernidade tardia e a emergência da sociedade de controle. *Revista Mediações*, *15*(2), 78-99.

Stevens, M., Wehrens, R., & de Bont, A. (2018). Conceptualizations of Big Data and their epistemological claims in healthcare: A discourse analysis. *Big Data & Society*, *5*(2), 1-21.

Strauss, A. & Corbin, J. (1990). *Basis of qualitative research: Grounded theory procedures and techniques for developing grounded theory*. Newbury Park: Sage Publications.

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, *4*(2), 1-14.

Tsianos, V. S., & Kuster, B. (2016). Eurodac in Times of Bigness: The Power of Big Data within the Emerging European IT Agency. *Journal of Borderlands Studies*, *31*(2), 235-249.

Tutton, R. (2011). Promising pessimism: Reading the futures to be avoided in biotech. *Social Studies of Science*, *41*(3), 411-429.

Van der Velden, L. (2015). Leaky apps and data shots: Technologies of leakage and insertion in NSA-surveillance. *Surveillance & Society*, *13*(2), 182-196.

Van der Vlist, F. N. (2017). Counter-Mapping Surveillance. *Surveillance & Society*, *15*(1), 137-157.

Van Dijk, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197-201.

Van Lente, H. (2012). Navigating foresight in a sea of expectations: lessons from the sociology of expectations. *Technology Analysis & Strategic Management*, *24*(8), 769-782.

Weber, J. (2016) Keep adding. On kill lists, drone warfare and the politics of databases, Environment and Planning D: *Society and Space*, *34*(1), 107-125.

Williams, R., & Johnson, P. (2004). 'Wonderment and dread': Representations of DNA in ethical disputes about forensic DNA databases. *New Genetics and Society, 23*(2), 205-223.

Williams, R., & Johnson, P. (2008). *Genetic Policing: The Use of DNA in Criminal Investigations*. Cullompton: Willan Publishing.

Wittendorp, S. (2016). Conducting Government: Governmentality, Monitoring and EU Counter-Terrorism. *Global Society*, *30*(3), 465-483.

Wood, D., Ball, K., Lyon, D., Norris, C., & Raab, C. (2006). A report on the surveillance society. *Surveillance Studies Network, UK* (https://ico.org.uk/media/about-the-ico/documents/1042388/surveillance-society-public-discussion-document-06.pdf, acedido em 02-08-2018).

Wright, D., & Kreissl, R. (2015). *Surveillance in Europe*. Oxon and New York: Routledge.

Yin, Robert K. (1994). *Case study research: Design and methods*. Londres: Sage Publications.

Young, S. (2017). Slipping Through the Cracks: Background Investigations after Snowden. *Surveillance & Society*, *15*(1), 123-136.

Youtie, J., Porter, A., & Huang, Y. (2016). Early social science research about Big Data. *Science and Public Policy*, *44*(1), 65-74.

Zedner, L. (2016). Citizenship deprivation, security and human rights. *European Journal of Migration and Law*, *18*(2), 222-242.

LEGISLAÇÃO CONSULTADA

Canotilho, J. G., & Moreira, V. (2005). *Constituição da República Portuguesa. Lei do Tribunal Constitucional*. Coimbra: Coimbra Editora.

Conselho da União Europeia (2008). Decisão 2008/616/JHA do Conselho de 23 de junho de 2008. Decisão 2008/615/JAI do Conselho de 23 de junho de 2008, Jornal Oficial da União Europeia.

Europeia, U. (2016). Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016, 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 (Regulamento Geral sobre a Proteção de Dados). Online: http://eur-lex.europa. eu/legal-content/PT/TXT/HTML.

Europeia, U. (2016). Regulamento (UE) 2016/680 do Parlamento Europeu e do Conselho, de 27 de abril de 2016, relativa à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais pelas autoridades competentes para efeitos de prevenção, investigação, deteção ou repressão de infrações penais ou execução de sanções penais, e à livre circulação desses dados, e que revoga a Decisão-Quadro 2008/977/JAI do Conselho. Online: https://eur-lex.europa.eu/eli/dir/2016/680/oj?locale=pt

Decreto-lei n.º 53/2008 de 29 de agosto – Lei de Segurança Interna.

ISBN 978-989-755-523-7

Este livro apresenta uma investigação sociológica sobre a potencial aplicabilidade do *Big Data* na segurança pública e policiamento transnacionais. Especificamente, visa a definição do *Big Data*, a explicitação dos seus processos de aplicação e a análise das expectativas sociais em torno destas técnicas, por parte de um conjunto de profissionais inseridos laboralmente na partilha transnacional de dados.

A partir das narrativas de Pontos de Contacto Nacionais em rede transnacional de cooperação policial e judiciária, geneticistas forenses e *stakeholders* de diferentes áreas (ética e regulação, investigação criminal, pesquisa universitária, empresas privadas e organizações não governamentais), analisaram-se as suas expectativas relativamente ao potencial do *Big Data* enquanto estratégia de investigação criminal. Procurou-se, desta forma, compreender as facetas sociais e culturais do *Big Data*; desmistificar os desafios europeus à aplicabilidade deste fenómeno no atual panorama da investigação criminal; e, por fim, promover um debate público acerca das questões da privacidade e dos grandes dados numa era marcadamente digital, em que as fronteiras dos direitos humanos são fluidas e a garantia da segurança pública é cada vez mais urgente. Este livro é uma janela que se abre para estas reflexões emergentes, contribuindo para fomentar o debate em redor de um tema complexo e de natureza imprecisa numa Europa crescentemente digitalizada.

exchange
erc European Research Council Established by the European Commission
European Commission
Horizon 2020 European Union funding for Research & Innovation
CECS centro de estudos de comunicação e sociedade INVESTIGAÇÃO
Universidade do Minho Instituto de Ciências Sociais