

Universidade do Minho
Instituto de Ciências Sociais

Laura Carvalho de Oliveira Neiva

Big Data na investigação criminal: previsão do risco, vigilância e expectativas sociais na União Europeia

Big Data na investigação criminal: previsão do risco, vigilância e expectativas sociais na União Europeia

Laura Carvalho de Oliveira Neiva

UMinho | 2019

Esta dissertação de mestrado foi realizada 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 sedado no Centro de Estudos de Comunicação e Sociedade, Instituto de Ciências Sociais da Universidade do Minho (Portugal).

EXCHANGE

Forensic Geneticists and the Transnational Exchange of DNA data in the European Union: Engaging Science with Social Control, Citizenship and Democracy

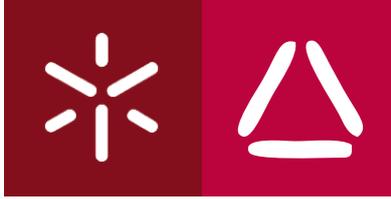
Grant agreement No 648608



Horizon 2020
European Union funding
for Research & Innovation



Março de 2019



Universidade do Minho
Instituto de Ciências Sociais

Laura Carvalho de Oliveira Neiva

**Big Data na investigação criminal: previsão
do risco, vigilância e expectativas sociais
na União Europeia**

Tese de Mestrado
Mestrado em Crime, Diferença e Desigualdade

Trabalho efetuado sob a orientação da
Professora Doutora Helena Cristina Ferreira Machado

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração da presente tese. Confirmando que em todo o trabalho conducente à sua elaboração não recorri à prática de plágio ou a qualquer forma de falsificação de resultados.

Universidade do Minho, 29 de março de 2019

Laura Carvalho de Oliveira Neiva

Agradecimentos

Esta Dissertação será, para sempre, minha e de todos aqueles que, junto a mim, a sonharam comigo. A eles, o meu sincero obrigada:

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 decorrente de investigação do mesmo.

À Professora Helena Machado, minha orientadora, por ter criado condições para que esta Dissertação fosse, hoje, o fim de um caminho que me traz novos recomeços; 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 documento. 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 que 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!

Resumo

Big Data na investigação criminal: previsão do risco, vigilância e expectativas sociais na União Europeia

O *Big Data* surge como mecanismo promissor na vigilância e previsão de risco no combate à criminalidade. O seu aparato tecnológico engloba a vigilância e observação indireta de indivíduos, recolha de um elevado número de informações potencialmente convertidas em algoritmos, com vista à produção de classificações numéricas e categorizações que visam a conceção de ações de prevenção e repressão da criminalidade. Esta investigação ambiciona o estudo deste fenómeno aplicado à segurança pública e policiamento transnacionais.

Especificamente, visa a definição de *Big Data*, a explicitação dos seus processos de aplicação e a análise de expectativas sociais em torno destas técnicas, por parte 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).

A metodologia é qualitativa e ancora-se na análise de entrevistas realizadas a estes profissionais. 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.

Palavras-chave: *Big Data*; criminalidade; cooperação transnacional; prevenção; tecnologia.

Abstract

Big Data in criminal investigation: risk prediction, surveillance and social expectations in the European Union

Big Data emerges as a promising mechanism for monitoring and predicting risk in the fight against crime. Its technological apparatus includes the surveillance and indirect observation of individuals and is based on the collection of a large number of information potentially converted into algorithms. By producing numerical classifications and categorizations such mechanism aims at the design of actions to prevent and investigate crime. This research explores the phenomenon of *Big Data* in the context of public safety and transnational policing.

Specifically, it studies *Big Data*, by explaining its application processes and analyzing social expectations around these techniques articulated by professionals engaged in the transnational cooperation for the fight against crime. These actors are those who are involved in the implementation and operation of the network for the exchange of forensic DNA data, forensic geneticists and stakeholders from different fields (ethics and regulation, criminal investigation, university research, private companies and non-governmental organizations).

The methodology is qualitative and is anchored in the analysis of interviews conducted with these professionals. The results allowed me to explore the social and cultural facets of *Big Data*, as well as to identify and contribute to the contemporary debate on issues related to privacy and data protection.

Keywords: *Big Data*; crime; transnational cooperation; prevention; technology.

Índice geral

Introdução	16
Parte I: Enquadramento teórico	24
1.1. Contextualização do objeto de estudo: <i>Big Data</i>	24
1.2. Definição concetual e operacional de <i>Big Data</i>	27
1.3. O caso de Edward Snowden	31
1.4. <i>Big Data</i> – uma nova forma de exercer a vigilância?.....	34
1.5. <i>Big Data</i> e criminalidade	44
1.6. Questões de direitos, liberdades e garantias – debate académico contemporâneo.....	55
1.7. Lente concetual empírica: <i>Sociology of Expectations</i>	62
1.8. Linhas objetivas orientadoras: Objetivos, questões de partida e hipótese de investigação	75
Parte II: Estudo empírico	80
2.1. Aspetos teórico-metodológicos: opções, bússolas e orientações.....	80
2.1.1. <i>Consideração da amostra: entrevistas</i>	82
2.1.2. <i>Análise de conteúdo qualitativo</i>	87
2.1.3. <i>Codificação e categorização</i>	88
Parte III: Resultados empíricos	94
3.1. Informação sobre <i>Big Data</i> : conhecimento, avaliação, descrição técnico-profissional e exemplificação	99
3.2. Expectativas futuras de potencial aplicação de <i>Big Data</i> : desenvolvimento, expansão e antecipação de aplicação na investigação criminal (transfronteiriça e organizada)	117
3.3. Perceção sobre os riscos e perigos de <i>Big Data</i>	130
3.4. Pareceres éticos e de direitos humanos.....	141
Parte IV: Conclusão	155
Bibliografia	169
Legislação consultada	180
Anexos	182
Anexo I – Distribuição dos/as entrevistados/as por país	182
Anexo II – Caracterização dos NCPs entrevistados em países operacionais	183
Anexo III – Caracterização dos geneticistas forenses entrevistados por país e organização onde trabalham	184
Anexo V – Tabela de sistematização de resultados 1.....	187
Anexo VI – Tabela de sistematização de resultados 2	188
Anexo VII – Tabela de sistematização de resultados 3	189
Anexo VIII – Tabela de sistematização de resultados 4.....	190

Índice de Figuras

Figura 1: Emergência e contextualização de Big Data.....	26
Figura 2: Definição conceitual e operacional de Big Data.	30
Figura 3: Transformações que operam na vigilância com o surgimento de Big Data....	43
Figura 4: Relação entre Big Data e criminalidade.....	54
Figura 5: Discursos dos diferentes entrevistados.	163

Lista de Siglas

ADN – Ácido Desoxirribonucleico

ANS – Agência Nacional de Segurança

D2D CRC – Data to Decisions Cooperative Research Centre

EUA – Estados Unidos da América

GPS – Global Positioning System

ISFG – International Society for Forensic Genetics

LAPD – Departamento de Polícia de Los Angeles

LASER – Los Angeles Strategic Extraction and Restoration Program

NCP – Ponto de Contacto Nacional

NCP1 – Ponto de Contacto Nacional alocado à etapa 1

NCP2 – Ponto de Contacto Nacional alocado à etapa 2

Projeto TIA – Projeto Total “Terrorism” Information Awereness

UE – União Europeia

Introdução

A presente investigação pretende contribuir para o debate contemporâneo que se tem aprofundado em torno dos grandes dados – *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 ainda um discurso efetivo agora – sobre as diferentes perspetivas, temporalidades, especialidades e materialidades que representam os bancos de dados, os seus objetivos e finalidades. Assim, este estudo afirma-se com o objetivo de projetar a máxima flexibilidade e permitir o possível para uma polifonia emergente.

O objetivo central que orientou este trabalho foi *aceder, compreender e analisar as perspetivas e expectativas sociais dos órgãos de cooperação transnacional, profissionais diretamente conectados com a partilha transnacional de dados na União Europeia, 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 com conhecimentos relevantes no que concerne ao Big Data e à partilha transnacional de dados de ADN*. De forma a concretizar este intuito, foram tecidas uma série de questões, que visam nortear o trabalho desenvolvido, bem como, completar e densificar a temática em estudo: que expectativas em torno das potenciais aplicações de *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 que expectativas são 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, respeitando e aceitando o que cada etapa de investigação informasse para prosseguir. Foi utilizada, para o efeito, uma técnica de recolha de dados: análise de conteúdo qualitativo de entrevistas realizadas a órgãos de cooperação transnacional, a 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.

Tratando-se de uma técnica ainda em fase emergente de aplicação na área da investigação criminal, prevalecem expectativas comumente partilhadas entre diferentes grupos profissionais; e, em simultâneo, expectativas diferenciadas consoantes o posicionamento profissional. Tal como prevê a *Sociology of Expectations*, cada profissional possui uma expectativa diferente da inovação e do progresso científicos, consoante a sua área académica de formação e o seu lugar contextual de trabalho. As expectativas são afirmações sobre o futuro e 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 destas emergiram subjetividades – objetos de estudo e interpretação. *Big Data* é uma técnica em fase de desenvolvimento no âmbito da investigação criminal, não se encontrando, ainda, validada cientificamente para ser implementada e instaurada enquanto meio de investigação criminal. Desta forma, é perspectivada como um pilar em construção, com potencial aplicação futura, enquanto materialidade subjetiva inovadora, tecnológica e científica.

Este trabalho encontra-se estruturado e dividido em quatro partes. A primeira, com o título *Enquadramento teórico* condensa seis fragmentos teóricos importantes na definição, delimitação e estruturação do tema. Nomeadamente, surge 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, urge uma definição concetual e operacional do mesmo, como forma de apresentar a forma pela qual o objeto será abordado. Nesta secção encontram-se as definições consensuais da literatura científica relevante para o tema em apreço. Em terceiro lugar, emerge a descrição sumária do fator que impulsionou as controvérsias em torno do tema – caso de Edward Snowden – revelações que determinaram a forma como se aborda o fenómeno e condicionam, conseqüentemente, o seu estudo. Delineou-se uma análise da evolução do caso e das reações a este – desde artigos, a notícias de imprensa e documentários – procurando explorar em que condições as revelações operaram, que efeitos produziram e que conseqüências tiveram, bem como,

repercussões. Concluiu-se que as revelações de Snowden foram determinantes para a instauração de um novo paradigma de estudo e investigação. Na quarta secção surge o enquadramento temático de *Big Data* como uma forma de exercer a vigilância; são apresentados estudos acerca das transformações que operaram 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, ambientais, sociopolíticos que operaram, bem como, governamentais, no sentido de possibilitar e criar condições para a emergência e expansão de *Big Data*. Em quinto lugar visa-se abordar a relação (in)existente entre *Big Data* e criminalidade – 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 (escassos) estudos empíricos encontrados e analisados. Pretende-se analisar as emergências que se têm vindo a observar na literatura científica que atribuem enfoque a esta relação, nomeadamente, em Departamentos Policiais. Apesar de escassos, os estudos revelam conclusões antagónicas: em determinados Departamentos Policiais, *Big Data* não é uma técnica que se aplique, mas que reside no imaginário policial e se encontra em desenvolvimento; enquanto que noutros Departamentos, *Big Data* encontra-se em fase precoce de implementação provocando mudanças no quotidiano policial e/ou reforçando técnicas tradicionais de investigação policial. Por fim, surgem as questões éticas e de direitos humanos consideradas e ancoradas nesta temática – grandes dados e respeito pelos direitos à privacidade e proteção de dados. É descrita a complexidade do debate e a tentativa de encontrar um equilíbrio entre garantias de segurança e de proteção de direitos. Por fim, descreve-se e enfatiza-se a importância do olhar analítico adotado para exploração e interpretação posterior dos dados empíricos – recortes discursivos – *Sociology of Expectations*. Esta lente concetual empírica, enquanto pilar basilar para interpretação, enquadramento e posicionamento dos dados, debruça-se sobre a interpretação e codificação das perspetivas, em geral, sobre inovações tecnológicas e técnicas inovadoras de progresso científico, em desenvolvimento. Assim, ancorados sobre esta perspetiva, os dados foram interpretados e apresentados no quarto capítulo. Esta primeira parte finda com a menção às questões de partida, a hipótese de trabalho e os objetivos que a presente investigação visou atingir – enquanto bússolas orientadoras do percurso de investigação.

A segunda parte – *Estudo Empírico* – apresenta e descreve o caminho de investigação seguido, especificando e delineando as diferentes opções metodológicas adotadas. Além disso, descreve o tipo de metodologia predominante – análise qualitativa de conteúdo de entrevistas; explicita o Projeto, sob o qual, esta investigação foi desenvolvida – EXCHANGE; e todo o percurso percorrido desde a escolha dos entrevistados, à análise das entrevistas. O quarto capítulo – *Resultados empíricos* – culmina, condensando, numa súpula do que foi obtido depois da recolha de dados e da sua exploração, interpretação e atribuição de significações. É neste setor que se inserem todas as sintetizações das informações e conclusões obtidas da análise das entrevistas. Aqui 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 que surgiram, por cada profissional, dependendo da sua posição laboral e formação académica. Isto comprova que o meio contextual onde os profissionais se encontram e o seu histórico formativo influencia as suas visões sobre os fenómenos, condicionando os seus discursos. De forma geral, os dados destacam que a maioria dos profissionais entrevistados possui expectativas promissoras quanto à implementação de *Big Data* como uma ferramenta de investigação criminal, no entanto surgem nuances distintas e características de cada grupo profissional. De forma a uniformizar e simplificar a complexidade dos resultados obtidos foram criadas temáticas gerais, enquanto categorias analíticas, que condensam as diferentes perspetivas consideradas: i) Informação sobre *Big Data*: conhecimento, avaliação, descrição técnico-profissional e exemplificação; ii) Expectativas futuras de potencial aplicação de *Big Data*: desenvolvimento, expansão e antecipação de aplicação na investigação criminal (transfronteiriça e organizada); iii) Perceção sobre riscos de *Big Data*; e iv) Pareceres éticos e de direitos humanos.

Por fim, o quinto capítulo é uma agregação das *conclusões* reunidas após a realização deste percurso de investigação, onde se apresentam os pontos mais importantes deste caminho de produção de conhecimento científico. Não obstante, 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

Enquadramento teórico

Parte I: Enquadramento teórico

1.1. Contextualização do objeto de estudo: *Big Data*

Inúmeros são os autores, políticos, operadores do sistema policial e meios de comunicação social que referem que a *Era do Big Data* começou (Boyd & Crawford, 2012). Em fevereiro de 2012, o *New York Times* anunciava a chegada *The Age of Big Data*; no ano seguinte o *The Economist* noticiava *Big Data* como *uma nova mina de ouro*; e o *The Guardian* identificava a técnica como a *revelação por números*. Em vários documentos oficiais é referido como o *novo petróleo para a era digital e a próxima revolução industrial*. Consequentemente, muitos são aqueles que se debruçam sobre esta possibilidade de acesso a enormes quantidades de informações produzidas por e sobre pessoas, coisas e interações. Naturalmente, surgem debates em torno dos potenciais benefícios e custos da análise de sequências genéticas, interações, registos de saúde, registos telefónicos, de governo e outros traços digitais deixados pelas pessoas (Boyd & Crawford, 2012).

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 o *mundo digital* ou o *mundo dos dispositivos de geração de dados online – Big Data*. Esse mundo é concebido em linhas estereotipadas de fluidez, mobilidade e aceleração contínua que crescem, tornando-se cada vez mais prolíferos e instalando-se como processos de rotina como, por exemplo, a proliferação de dados sociais transacionais que agora são armazenados e analisados diariamente (Frade, 2016; Youtie, Porter & Huang, 2016). Temos, assim, uma entidade dinâmica que se move continuamente e rapidamente numa única direção: para um crescimento e expansão cada vez mais rápidos, e sempre para o 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 – um tempo que ascende à medida que avança, não só em volume, mas em técnicas de eficiência (Frade, 2016). Atualmente, as sociedades evoluem, paralelamente, a um processo de globalização que requer um alto nível de controlo sobre o espaço físico e tempo, resultando na aceleração do ritmo da vida e das temporalidades organizacionais – e isto 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 como forma de 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, procurando aumentar a sua eficácia e eficiência (Van Dijck, 2014), sustentados na ideia de que a tecnologia permite a economização do tempo, contribuindo para ganhos no tempo-espaço, reforçando a sua utilização e expansão (Araújo, 2008). As Ciências Sociais foram uma dessas 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 também se foi expandido e crescendo, à medida que a digitalização em massa de informações se foi tornando uma possibilidade crescente (Brayne, 2017). E, por isso, importa antes de afirmar o que quer que seja explorar o contexto onde as realidades e existências emergem – atualmente entrelaçado com o Capitalismo – porque as capacidades tecnológicas nunca operam sozinhas. Concretamente, *Big Data* não se *auto-gerou*, mas surgiu como fruto do *capitalismo 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 faltar distâncias e eliminar tempos, através da velocidade de circulação e da híper-imposição do espaço virtual (Araújo, Cogo & Pinto, 2015). E todo este aparato móvel, digital, fluído 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.

Não obstante, o contexto europeu sofreu, também, modificações: algumas paralelas às descritas anteriormente, outras que surgiram contrárias. No entanto, todas estas mudanças se influenciam mutuamente e convergem para produzir alterações sociais que influenciam os cidadãos. Atualmente, a Europa está sob ameaças e ataques terroristas e este arsenal influencia os valores fundamentais defendidos pela legislação: direitos, liberdades e garantias; não só da União Europeia (UE), mas também dos seus cidadãos. Assim, surge um desafio: estabelecer o equilíbrio entre a proteção destes valores fundamentais, aliado à capacidade de combater e prevenir os ataques e as ameaças terroristas. Isto é um desafio, simultaneamente, político, religioso, militar e social.

Consequentemente, face a este dilema, sentiu-se a necessidade de mudar o paradigma de atuação: de uma guerra tradicional ao terrorismo, para uma guerra mais tecnológica, que vá além das políticas de segurança já existentes. No entanto, a procura de respostas tecnológicas adequadas aos ataques terroristas levanta questões ao nível das liberdades, privacidade e proteção de dados. E assim, afigura-se imprescindível alcançar o equilíbrio certo entre a privacidade individual e a segurança social na Era Do Mercado Digital (Gonçalves, 2017; Drewer & Miladinova, 2017). Neste contexto, uma das inúmeras facetas desta revolução digital é, particularmente, o surgimento do conceito dos *dados*, anteriormente explorado (Drewer & Miladinova, 2017). E 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). As mudanças que a sociedade opera devem ser perspetivadas sob a ótica do tempo e das temporalidades, daí a importância de compreender e delinear as alterações contextuais, sociais, políticas e ambientais que circundam o surgimento de *Big Data*. Os conceitos de tempo e de temporalidade estão, então, subjacentes aos processos históricos que caracterizam as mudanças sociais e, por isso, importa tê-los como foco de análise para uma compreensão holística dos (novos) paradigmas que surgem; enquanto objetos e recursos de poder (Araújo, 2012).

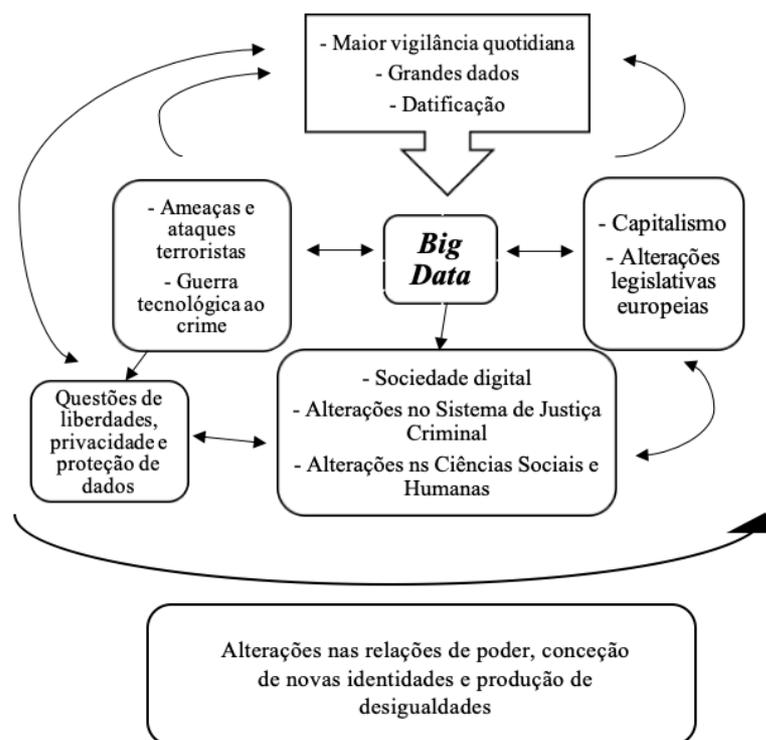


Figura 1: Emergência e contextualização de Big Data.

1.2. Definição conceitual e operacional de *Big Data*

A definição de *Big Data* não é consensual, mas a maior parte dos autores utilizam o conceito para se referirem aos conjuntos de dados que são recolhidos, analisados, convertidos em algoritmos¹, categorizados numericamente e identificados por via de um índice calculado para, posteriormente, extrair informação destes dados 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). Ou seja, frequentemente, *Big Data* surge na literatura como uma técnica que agrega e trata um enorme conjunto de dados: trabalha com um volume infinito de informações. E estes conjuntos de 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 surgir de diversas formas, entre elas: i) dirigidos (por via dos circuitos de câmaras de vigilância, por exemplo); ii) automatizados (localização geográfica detetada pelo telemóvel, por exemplo); e iii) voluntários (por exemplo, a cedência de dados pessoais por parte dos indivíduos) (Lupton & Michael, 2017; Chan & Moses, 2015; Lyon, 2014). Geralmente, a definição de *Big Data* envolve (pelo menos) três considerações: o grande volume de dados que agrega, a velocidade segundo a qual os dados são processados e a variedade dos dados, o seu formato e estrutura (Chan & Moses, 2015; Prainsack, 2019). Por sua vez, Boyd & Crawford (2012) apresentam uma definição de *Big Data* enquanto *fenómeno cultural, tecnológico e académico* que resulta da interação de três elementos: a tecnologia (maximiza o poder computacional e a precisão algorítmica), a análise (identificação de padrões por via de um conjunto de dados) e a mitologia (a crença generalizada de que grandes conjuntos de dados oferecem formas maiores de inteligência e conhecimento) (Chan & Moses, 2015). Bartlett, Lewis, Reyes-Galindo & Stephens (2018) posicionam o *Big Data* como o modo indispensável de pesquisa do século XXI em toda a academia, alegando que a ciência social com dados e computação intensiva é um fenómeno contemporâneo (Halford & Savage, 2017).

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

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

uma possibilidade crescente (Gandy Jr, 1989; Boyd & Crawford, 2012; Frade, 2016; Matzner, 2016). Além disso, *Big Data* opera numa vigilância cada vez mais inclinada em operações que se concentram no futuro – descartando o presente e o passado. No contexto da governança neoliberal, esta antecipação, provavelmente, colocará maior ênfase na vigilância para gerir as consequências, do que na pesquisa sobre a compreensão das causas dos problemas sociais como o crime e a desordem (Lyon, 2014).

Ancorados sob uma lente geral diríamos que o *Big Data* é um fenómeno que está a mudar fundamentalmente o que sabemos e o que fazemos – é 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). O Conselho Europeu, relativamente a esta questão, 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), o que surge que *Big Data* é muito mais do que apenas um grande número de dados armazenados, é um fenómeno sociotécnico complexo que congrega a interação entre diferentes ciências, poderes e tecnologias. Araújo et al. (2015) referem que as mobilidades significam, também, precisamente, complexidades, reforços de planeamento, organização e controlo, visto que a circulação dos indivíduos e da própria informação processa-se, atualmente, mediante códigos cada vez mais elaborados de vigilância, para a qual contribuem as tecnologias digitais.

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 que se cada passo humano fosse uma senha comunicada por via de uma linguagem digital (Costa, 2004). Através desses códigos e símbolos, a atividade humana, bem como, a identidade individual, são armazenadas em bases de dados oficiais e partilhadas com outros centros policiais, por exemplo (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, é evidente que com o advento do *Big Data*, paralelamente ao desenvolvimento da tecnologia (Matzner, 2016), se criou 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). 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). Os processos tecnológicos converteram-se em mecanismos pelos quais novas realidades culturais são produzidas, alterando os padrões tradicionais e criando mais e maiores possibilidades para os indivíduos negociarem a sua identidade em relação ao uso do seu tempo e do seu espaço (Araújo, 2008). E toda esta panóplia tecnológica levanta questões, desde logo, acerca da fiabilidade dos dados obtidos e das conclusões retiradas das análises desses dados; também no que toca ao surgimento de uma sociedade panótica, que em tudo se assemelha ao *Big Brother*, por via de uma vigilância permanente, invisível e penetrada em todos os interstícios sociais (Matzner, 2016), onde o Estado – entidade detentora do poder – se mantém fora do alcance dos indivíduos e estes últimos estão numa posição de constante observação (Costa, 2004). 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”, ao mesmo tempo que se luta contra ações terroristas cada vez mais eminentes (Souza, 2010; Drewer & Miladinova, 2017; Wood et al., 2006; Wittendorp, 2016).

Assim, o Estado parece ser entidade responsável pela definição de ciência e tecnologia, promovendo, mediante esta, o que deve favorecer, ou não, de apoio à sociedade. Assim, este, definindo o bem público e desmistificando o que deve usufruir de investimento, define o rumo da Ciência e os meios que o permitam, bem como, soluciona as controvérsias que possam emergir da Ciência (Jasanoff & Kim, 2009). Desta forma, converte-se numa entidade extremamente relevante para o progresso científico. É um contexto complexo e em constante mudança, que tem influência nos processos de construção de redes de conhecimento internacional, visando colaborar na produção de conhecimento científico (Fontes & Araújo, 2013), mas também na procura de soluções para a resolução dos problemas eminentes. Quanto a isso, a tecnologia pode facilitar a concretização deste objetivo – há atividades que se facilitam pelo uso da tecnologia, no entanto, esta última também provoca profundas reconfigurações nos tempos atuais de

pesquisa, trabalho e investigação, especialmente porque possibilita a realização de tarefas atualizadas a qualquer hora e em qualquer lugar: verifica-se uma colonização dos tempos e espaços (Araújo, 2008). Esta globalização das práticas científicas é potenciada pelo advento das técnicas de *Big Data* na medida em que, enquanto técnica tecnológica e digital, reforça a possibilidade de troca e partilha transnacional de saberes, culturas, conhecimentos e vidas.

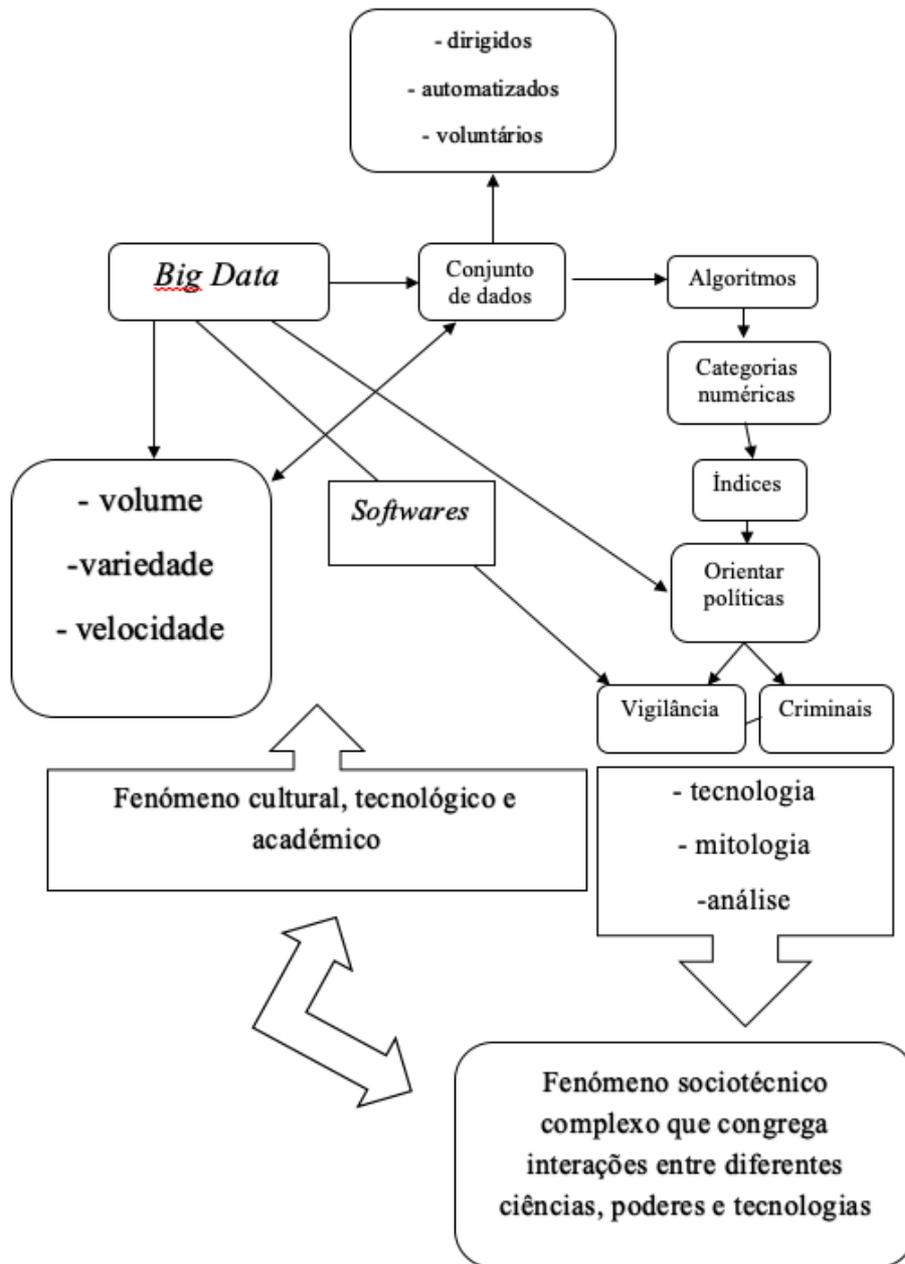


Figura 2: Definição conceitual e operacional de Big Data.

1.3. O caso de Edward Snowden

Em junho de 2013, quando Snowden deu voz à arquitetura de opressão projetada pela Agência Nacional de Segurança (ANS), nos Estados Unidos da América (EUA), a maquinaria da vigilância teve um rosto desenhado e uma identidade que ficou conhecida por todos aqueles que andavam a ser vigiados (Van Dijck, 2014; Lyon, 2015; Matzner, 2016; Gonçalves, 2017; Bakir, Feilzer & McStay, 2017; Matos, 2018). A partir desse momento, o tema da vigilância – principalmente a tecnológica, digital e informática – jamais fora falado 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 setores 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 estas revelações que tiveram um forte impacto crítico específico. No fundo, fora este o objetivo de Snowden: desbloquear o debate público acerca das invasões de privacidade que eram feitas (Van Dijck, 2014; Young, 2017). E estas revelações desencadearam indignações e ansiedades (Guzik, 2009). Por exemplo, segundo as declarações de Snowden, a ANS, usava dispositivos como telemóveis e computadores para ativar a sua localização geográfica e identificar e vigiar os indivíduos; acedia aos *sites* visitados, aos endereços acedidos e esses dados eram convertidos em algoritmos e inseridos nas suas bases de dados – sem que houvesse consentimento por parte dos indivíduos vigiados. Existiam cerca de quase cinco biliões de registos de telemóveis recolhidos pela ANS, por dia, permitindo aceder à localização destes indivíduos. Desta forma, também eram analisados padrões de comportamento para revelar mais informações pessoais e relações entre os diferentes indivíduos (Guzik, 2009; Lyon, 2014; Wright & Kreissl, 2015; Gonçalves, 2017; Matos, 2018). Ou seja, provou-se que a ANS detinha numerosos programas que podiam ser classificados em apenas algumas categorias de tecnologia de vigilância como escutas telefónicas, codificação, exploração, ferramentas de análise e bases de dados (Van der Vlist, 2017).

Após a quebra deste silêncio muitas mudanças operaram nas práticas de vigilância, que jamais teriam sido possíveis se Snowden não tivesse operado a esta rutura. Desde logo se verificou a mudança para uma vigilância em massa, direcionada e que faz uso maciço de metadados (Lyon, 2015; Young, 2017). Estes dados são, posteriormente, agregados, sintetizados e analisados e as pessoas são constantemente vigiadas,

controladas e supervisionadas, sem que o saibam e podem ser consideradas suspeitas por estarem, simplesmente, num determinado local (Bakir et al., 2017). Os governos apoiam-se nesta defesa da “segurança nacional” para legitimar esta vigilância cada vez mais penetrante nos interstícios sociais, como uma política de ligação de pontos que visam prever os fenómenos criminais antes que estes se emancipem (Lyon, 2014; Young, 2017), potenciando o fenómeno de *transparência forçada* em que, secretamente, se exige uma total visibilidade dos cidadãos para maximizar a segurança, sem o conhecimento nem consentimento destes (Bakir et al., 2017). Muitos são os autores que referem o aspeto estatal da vigilância: esta última como resultado de uma racionalidade implacável expressa em procedimentos burocráticos, esta condição cultural constringedora indubitavelmente ajuda a explicar a razão pela qual a vigilância é continuamente auto-aumentada. É importante que a compreensão da vigilância se opere no seu contexto social, político-económico e cultural (Lyon, 2015; Young, 2017; Cinnamon, 2017). Assim, surge a ideia de que se opera a governabilidade do crime, através da tecnologia – o crime é a maior preocupação social e estatal e, por isso, legitima todas as práticas de segurança adotadas. As declarações sobre privacidade feitas por governos e empresas podem ser consideradas como uma ferramenta de poder e governança a serviço do capitalismo informacional, sob a perspetiva de Coll (2014). 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). E a tecnologia começa a ser percecionada como uma ferramenta para a segurança, controlo e prevenção da criminalidade. No fundo, comprovou-se a codificação da vida e relações sociais em algoritmos, e o registo destes códigos em bases de dados, partilhadas, escrutinadas e analisadas (Gandy Jr, 1989; Ball et al., 2016; Wittendorp, 2016), sendo que as agências de governo procuraram sempre minimizar os efeitos dos metadados (Lyon, 2014). No entanto, é preciso ressaltar que o uso dos grandes dados não é mero resultado do potencial tecnológico em desenvolvimento, mas de abordagens específicas de gestão do risco em indústrias de segurança (Lyon, 2015).

Face às revelações de Snowden, a máquina opressora alegara de que se tratava de um mecanismo que previa proteger as populações, afirmando que não se poderia exigir uma percentagem máxima de segurança, sem um valor mínimo de invasão da privacidade – esta última seria consequência da garantia de segurança e proteção pública. Como que se devesse haver uma submissão a esta política de escrutínio da vida privada para que as ameaças criminais não os atingissem, à semelhança de uma atitude de escravidão (Orwell,

2009; Coll, 2014). Todos aqueles que ouviram e viram facetas da sua vida a serem observadas, analisadas, partilhadas e convertidas em números, deveriam aceitá-lo, tornando-se *vítimas* do próprio sistema que alegava a sua proteção e segurança máxima – como que se a aceitação livre deste escrutínio, por parte das Agências de Segurança Estatais, lhes desse o estatuto de agentes do Sistema por auxiliarem as investigações. Houve uma obrigação em cooperar, desta forma, 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 às esferas humanas para garantir a segurança holística – tornando-nos *vítimas* potenciais da Instituição e cidadãos completamente controlados por esta (Costa, 2004; Orwell, 2009). Assim, houve, pelo menos, três atores importantes neste acontecimento: agências de governo, corporações privadas e, embora de forma involuntária, os cidadãos (Lyon, 2014). Este é um exemplo claro do que Brown & Michael (2003) relatavam acerca dos fracassos das propriedades tecnológicas: por vezes, as tecnologias desdobram-se em formas inicialmente não previstas. Ou seja, inicialmente os Grandes Dados foram projetados e concebidos para garantir a segurança máxima e plena dos cidadãos e acabaram por violar a sua privacidade. A questão que se levanta, depois, é se podemos apreender e aprender algo com os fracassos do passado futuro. No fundo, há uma centralização nas capacidades da tecnologia e no seu papel fundamental, em detrimento de uma reflexão debruçada sobre a sociedade que lida com os impactos inevitáveis do impulso da tecnologia (Brown & Michael, 2003). No fundo, as deceções e os fracassos são características inerentes da ciência inovadora e da tecnologia (Tutton, 2011).

Não obstante, ao mesmo tempo que se operam estas reflexões, gradualmente, 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 Dijck, 2014). Assim, conseqüentemente, esta temática do *Big Data* em particular, e da vigilância, no geral, surge mais acentuada e desenvolvida a partir deste momento, porque emerge quando as pessoas se apercebem e tomam conhecimento que estão a ser vigiadas e que os seus dados estão a ser alvo de invasão (Lyon, 2004; Gonçalves, 2017). Desde então, os debates em torno do *Big Data* enquanto ferramenta de vigilância invisível são inúmeros, atualmente, na *Era dos Grandes Dados* os autores debatem as conseqüências ligadas à imersão humana nesta sociedade tecnológica e digital (Taylor, 2017; Boyd & Crawford, 2012; Wood et al., 2006). Esta temática surge estreitamente conexa ao tema geral da vigilância e aos seus estudos

(*surveillance studies*), que se intensificam logo após o 11 de setembro, em simultâneo com o progressivo interesse em áreas relacionadas com a vigilância, privacidade, direitos civis e tecnologia (Fróis, 2007; Fróis, 2015; Lyon, 2015; Matzner, 2016; Wood et al., 2006; Costa, 2004; Souza, 2010). É um tema de forte problematização desde sempre, no entanto é tanto mais complexo quanto mais aprimoradas e emaranhadas são as ferramentas disponíveis para exercer a vigilância (Fróis, 2015).

A partir deste momento, na Europa agregou-se 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 apenas para a proteção de dados e privacidade, mas, em última análise, para as liberdades e pelo sistema democrático em geral. Esta notícia teve o efeito de chamar a atenção pública e política para o fenómeno emergente do *Big Data* e para os riscos e incertezas que isso acarreta. Na sua revisão dos programas de vigilância após o caso Snowden, 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).

1.4. *Big Data* – uma nova forma de exercer a vigilância?

Vivemos numa época de expansão e intensificação da vigilância, nos mais diversos setores e, a propósito de tal constatação, surgem autores que dizem que estamos sob alçada de uma sociedade da vigilância – que tem benefícios e consequências (Fuchs, 2011; Costa, 2004; Marx, 2004; Wood et al., 2006). A verdade é que, atualmente, o Governo faz uso da vigilância para dar garantias de paz, segurança e controlo social (Lyon, 2004; Fuchs, 2011; Skinner, 2018). Esta é apresentada como um fenómeno geral e universal das sociedades, sendo concetualizada como um fenómeno social positivo, autoevidente e constante (Fuchs, 2011)², tendo-se convertido, gradualmente, numa prática omnipresente nas sociedades modernas (Lyon, 1992; Brayne, 2017; Lyon, 2014).

² Entenda-se aqui a vigilância definida, de acordo com Lyon (2014), ou seja, como uma atividade quotidiana racional de controlo da informação na Modernidade, tendo fins de produção e consumo nas sociedades capitalistas, respeitando a burocracia estatal.

Em meados do final 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) e no discurso de políticos e autoridades de forças de segurança em todo o mundo, a tecnologia é frequentemente avançada como uma ferramenta poderosa e inovadora na prevenção, dissuasão e luta geral contra o crime (Fróis & Machado, 2016). Verificou-se o surgimento de redes interligadas de partilha de informação, a capacidade de processar, manusear, transmitir e armazenar dados, os computadores transformaram-se em aparelhos de observação, simulação e processamento de dados e surgiram as tecnologias de localização geográfica em tempo real. Todo este arsenal tecnológico reconfigurou 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 – “panótico 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 e a fazer-se uso das tradicionais formas de vigilância e isto reflete uma intensificação desta – novas técnicas agregam-se às antigas e, por isso, o que se averigua é um maior arsenal de vigilância social (Marx, 2004).

Consequentemente, a tecnologia moderna possibilita um controlo técnico e uma supervisão anónima: o desenvolvimento das novas tecnologias cria condições para que se desenvolva uma *nova vigilância* (Gandy Jr, 1989; Lyon, 1992; Lyon, 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 que a vigilância de dados inclui o controlo contínuo dos dados individuais sem fins pré-definidos, nem declarados (Van Dijck, 2014). É uma vigilância cada vez mais incentivada por sensores, computadores, aplicativos de *softwares*, redes e outras tecnologias que evolui, progressivamente, de alvos individuais para um olhar persistente indiferenciado sobre populações inteiras (MacWillie, 2018). Assim, consequentemente, surge uma vigilância descentralizada e em rede, operada por diversos agentes, dispersos e heterogêneos, potenciada pela disponibilidade das redes

digitais (Gandy Jr, 1989; Graham, 1998; Costa, 2004; Fuchs, 2011; Aas, 2006). A atividade de vigiar converte-se numa atividade anónima, fluída e móvel, onde as organizações que vigiam estão cada vez mais omnipresentes, mas ocultas perante os vigiados (Lyon, 2004; Marx, 2004; Costa, 2004; Lyon, 2014); 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) – este paradoxo é acentuado com o advento do *Big Data*, como que se a recolha de dados pessoais, a vigia dos indivíduos, o seu controlo e a conversão das suas atividades em números fosse uma componente vital à sociedade contemporânea (Wood et al., 2006). Não há um único ponto geográfico de acesso à informação, ela pode ser consultada a partir de qualquer lugar. Da mesma forma, não há uma única base de dados eletrónica para vigilância, mas muitas dispersas que podem ser usadas em conjunto por atores poderosos com o objetivo de realizar pesquisa interligada de dados (Graham, 1998; Fuchs, 2011; Aas, 2006; Wood et al., 2006; Prainsack, 2019) – verifica-se o surgimento de uma sociedade transparente, que opera com 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 – onde 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 incluem uma distinção entre relações sociais e forças tecnológicas da vigilância (Fuchs, 2011), ou seja, parece não estar claro se a vigilância é considerada uma tecnologia ou uma relação social (Lupton & Michael, 2017; Lyon, 1992; Lyon, 2014; Wood et al., 2006). *Big Data* tem o poder de se materializar numa vigilância estruturalmente especulativa: dados que aparentemente não se relacionariam a um objetivo estratégico particular podem produzir correlações imprevisíveis, mas úteis. Além disso, visa acumular um arquivo que pode ser pesquisado e classificado, de forma retrospectiva. Este arquivo tem aspetos positivos e negativos: positivos na medida em que permite a consulta repetida das imagens, factos e aspetos necessários à intervenção; negativos, por sua vez, visto que vincula os indivíduos às cenas do crime, reconstruindo os seus movimentos, identificando-os e, eventualmente, capturando-os (Andrejevic & Gates, 2014).

No fundo, 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), há um crescente desenvolvimento da recolha e armazenamento de informação com potencial uso político e militar (*intelligence*) e o desenvolvimento do *data mining* (Machado & Santos, 2016; Matos, Santos & Machado, 2016). E, assim, os *maiores alvos* desta vigilância contínua são os cidadãos: os seus dados são continuamente 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 da vigilância ocorrer de modo universal e permanente. Este tipo de vigilância é pré-constitutiva, recolhendo e armazenando informação de forma aberta, dirigida, indiscriminada e a qualquer pessoa (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).

Estas novas formas tecnológicas de controlar, vigiar e supervisionar os indivíduos viram a sua expansão, evolução e complexificação após o acontecimento do 11 de setembro, como já referido, nos EUA e sustentam-se nas políticas de segurança antiterrorismo (Lyon, 2004; Aas, 2006; Souza, 2010; Brayne, 2017), reforçadas e repensadas também na consequência dos ataques a Madrid em 2004, a Londres em 2005 e a Paris em 2015 (Wittendorp, 2016; Matos et al., 2016). A partir deste momento, mais e maiores técnicas de vigilâncias começaram a expandir-se mundialmente, ao mesmo tempo que se espalhava e alargava o pânico moral (Lyon, 2004), o Conselho Europeu reuniu e lançou propostas que reforçavam a cooperação policial transnacional, com vista a reduzir e controlar as ameaças terroristas, ao mesmo tempo que procurava 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 à sua implementação (Wittendorp, 2016; Brayne, 2017). E a estas políticas de segurança acrescem-se 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 e estimulada, bem como, legitimada, tendo-se tornado alvo de interesses que visavam a melhoria das suas operações, tendo sido amplamente financiada para ser melhorada, no que toca ao desenvolvimento de novos sensores e programas de dados que produzissem *estratégias de inteligência* (Brayne, 2017). Este investimento em sistemas tecnológicos potencia a recolha massiva e constante de dados e a construção de sistemas de informação organizados e sistemáticos, criando uma cultura de controlo do crime assente num regime intensivo de regulação, inspeção e controlo social (Queirós, 2018). Ou seja, um dos projetos tecnológicos levados a cabo por Estados-membros para a mitigação do risco foi a expansão da vigilância tecnológica e digital. E enquanto a maior parte dos sistemas de vigilância operam segundo uma lógica reativa e pré-construtiva, *Big Data* apoia-se em tecnologias que seguem uma lógica proactiva e reconstrutiva: tem como objetivo recolher informação que permitirá identificar indivíduos onde a sua presença corporal e os seus atos seriam invisíveis para qualquer tecnologia de observação direta no local e tempo em que esses atos aconteceram (Williams & Johnson, 2004; Machado & Granja, 2018). *Big Data* converte o mundo em dados e codifica e regista movimentos da sua globalidade – suspeitos e não suspeitos, população como um todo (Andrejevic & Gates, 2014).

Assim, conseqüentemente, o interesse contemporâneo da vigilância centra-se nos computadores (Gandy Jr, 1989; Lyon, 1992; Lyon, 2004; Marx, 2004). Gera-se muita informação, por via da partilha transnacional de dados (Lyon, 2004; Williams & Johnson, 2004) – informação essa que deixa rastros infinitos, podendo ser retomada, reanalisada e alvo de retratamento sempre que os órgãos responsáveis pelo seu armazenamento façam questão disso. E o trabalho exercido sobre estas informações e estes dados são análises complexas que inter-relacionam pessoas, eventos e interações e que requer, conseqüentemente, um alargamento da rede de controlo social. Estas alterações no controlo social, bem como, na vigilância – suas técnicas e materializações – são vistas, por inúmeros autores, como uma resposta às transformações técnicas e sociopolíticas que ocorreram nos últimos tempos, que têm conseqüências (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 argumentação centra-se em torno das implicações potencialmente nocivas da sua vertente tecnológica que, pela sua incessante sofisticação e progressiva capacidade de aglomerar e cruzar dados, identificar padrões comportamentais, criar perfis, classificar

e compartimentar pessoas, grupos, comunidades – torna esta práticas propensas a um fenómeno de desvirtuamento da função. Problematiza-se a potencialidade da vigilância tecnológica, que vai para além do que é declarado pelos governos e polícias – demonstra-se a forma como essa presença ubíqua pode ser danosa para o cidadão, na medida em que interfere e condiciona direitos civis, como a liberdade, o direito à privacidade, à confidencialidade; e, também, revela-se como os sistemas de vigilância, ao operarem numa base de diferenciação social, contribuem 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 do excesso de dados transmitindo a ideia de que a tecnologia está a emergir 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 á quem do seu objetivo – a única maneira de garantir que nada é esquecido é reproduzir o mundo na sua totalidade em forma de dados gravados e, portanto, gravar dados sobre o processo de gravação em si e assim sucessivamente. No entanto, não há garantias de que os dados recolhidos são abrangentes ou representativos. Assim, os críticos referem que se trata de uma abordagem antiquada dos dados: uma ênfase excessiva dada ao seu conteúdo, ao invés da 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 contexto 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 pleno uso das novas tecnologias e também deve haver recíproco acesso a bases de dados nacionais; ao mesmo tempo que o Tratado, onde este princípio se abriga – 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 a sua expansão. As novas tecnologias sociais monitorizam corpos, os seus movimentos e atividades e são um investimento enorme nos EUA, constituindo-se como parte de um fenómeno internacional mais amplo – as tecnologias como correções e soluções para questões complexas de segurança. Um exemplo claro disso reside na génese de múltiplos sistemas computarizados conectados de recolha, armazenamento e análise de dados que permitem a partilha de informações pessoais importantes. Estas novas formas de *se fazer* Ciência não servem apenas fins de controlo

e vigilância, mas atribuem, também, aos cidadãos um rótulo e coloca-os em grupos especiais de vigilância (Skinner, 2018).

Conseqüentemente, o “crime” foi sendo eleito, gradualmente, 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 destes discursos que, por sua vez, visam reforçar as estratégias máximas de segurança (Jasanoff & Kim, 2009) – observa-se uma gradual transformação das sociedades disciplinares em sociedades de segurança (Maciel & Machado, 2014). Assim, visa alcançar-se e garantir-se a segurança através da gestão do risco e a eminência das ameaças (Weber, 2016; Johnson & Williams, 2007; Skinner, 2018), ou seja, 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 e do que é o risco vêm acompanhadas de ideários de vigilância expansiva sobre os indivíduos. Sob este panorama, a apropriação da tecnologia e a criação de projetos que visam a dissipação do risco não calculável por parte dos Estados possui diferentes práticas, políticas e aceitação dos cidadãos (Jasanoff & Kim, 2009). Ou seja, o rumo atual de 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 visam controlar, prever e identificar possíveis comportamentos e atos criminosos pela quantificação do perigo (Machado & Santos, 2016; Machado et al., 2018).

No entanto, existem situações em que cumprem um objetivo inesperado e contrário: reforçam o sentimento de insegurança e as desigualdades sociais. Verifica-se um aumento quantitativo dos mecanismos de vigilância que, por sua vez, reforçam as desigualdades sociais, visto que estes mecanismos não se encontram distribuídos de forma equitativa nas diferentes camadas sociais – alguns cidadãos, grupos e áreas são fiscalizadas de forma mais regular do que outras e para outros fins (Lyon, 2004; Lyon, 2014; Brayne, 2017; Taylor, 2017). No fundo, os estudos referem que se verifica uma maior intensificação da vigilância em grupos *de risco*, marginalizados e socialmente excluídos – pessoas a cumprir liberdade condicional, indivíduos com medidas sociais, emigrantes, grupos étnicos minoritários – onde os limites securitários são menores (Lyon, 2004; Lyon, 2014; Taylor, 2017; Cinnamon, 2017). Assim, a vigilância exerce-se, maioritariamente, sobre os *migrantes, terroristas e criminosos*, acentuando as desigualdades (Taylor, 2017; Skinner, 2018). Zedner (2016) refere, a este propósito, que se trata de *enfraquecer a segurança*, visto que os mecanismos de controlo recaem sobre

os indivíduos que vivem à margem da sociedade e uma maior vigilância sobre estes apenas reforça o *limbo* onde se encontram. A autora exemplifica, a propósito da lei da imigração, alegando que um excesso e concentração de controlo e vigia sobre certos grupos sociais despoja-os de cidadania, tornando-os seres humanos precários e, gradualmente, neste processo constante de diferenciação e estratificação vigilante, verifica-se uma *erosão de direitos*, reforçando desigualdades e produzindo discriminações (Lyon, 2014; Lyon, 2015; Zedner, 2016; Young, 2017; Machado et al., 2018). Paralelamente a essa crescente desigualdade social, verifica-se uma mudança, por parte do Sistema de Justiça e da investigação criminal, em basear-se no pressuposto de que determinadas populações ou grupos de indivíduos são potencialmente suspeitos – suspeição coletiva –, ao invés de considerar determinado indivíduo suspeito por uma atividade criminal (Machado et al., 2018). Estas duas categorias interpenetram-se (Cole & Michael, 2006) e reflexo disso é o explanado anteriormente: um indivíduo será mais facilmente identificado como suspeito se for membro de um grupo social ou populacional previamente considerado de risco (Cinnamon, 2017; Machado et al., 2018). No fundo, opera-se a uma classificação social, em função do risco, que só é possível por via do uso de bancos de dados pesquisáveis e técnicas de mineração de dados – isto significa que as pessoas são classificadas em certos grupos e daí podem advir (des)vantagens (Young, 2017).

Assim, importa perceber de que forma é que a identidade humana e as suas informações são traduzidas nestas instâncias digitais, visto que não reduzem apenas a informação a padrões numéricos, mas criam, também, novas conceções de identidade (Gandy Jr, 1989; Williams & Johnson, 2004; Aas, 2006; Maciel & Machado, 2014; Lyon, 2014; Kubler, 2017). Nós – cidadãos – participamos de forma ativa no controlo do qual somos alvos, por exemplo, na localização que cedemos. E a vigilância eletrónica, digital e informatizada acaba por invadir a privacidade e tornar irrelevantes todas as restrições feitas a esta última, visto que muitas lesões que são feitas a esta visam a proteção de bens maiores (Lyon, 1992). E, nesta Era Digital, segundo Gonçalves (2018) e Skinner (2018) os cidadãos tornam-se mais frágeis aos riscos, percecionando-os com maior clareza e tendo maior consciência da sua amplitude existencial. Assim, estes mecanismos densos e intensos – *novas formas difusas de tecnociência* – acabam por cumprir um objetivo inesperado: a exacerbação do risco, o aumento do sentimento de insegurança e o pânico moral (Skinner, 2018). As sociedades modernas são descritas como sociedades de risco,

marcadas por uma incerteza estrutural em relação a inúmeros fenômenos, tornando-as vulneráveis a promessas e profecias da tecnociência (Araújo, 2013).

Desta forma, esta maneira de se exercer a vigilância pode ser perspectivada de vários ângulos, no entanto, há autores que defendem que se trata de uma *vigilância nova*, outros consideram que se trata de um *fenômeno de vigilância eminentemente tecnológico*, há ainda os que afirmam que é uma *vigilância adaptável* que procura dar respostas às ameaças externas e, por fim, surgem autores que dizem que se trata apenas de *um produto da Modernidade* e de todas as suas transformações (Lyon, 2004). No fundo, inicialmente a vigilância surgiu como uma forma de reforçar o poder do próprio Estado e hoje é um meio de controlo social (Lyon, 2004; Williams & Johnson, 2004). Consequentemente, *Big Data* intensifica certas tendências de vigilância associadas à tecnologia da informação e às redes, e está, portanto, implicado em configurações frescas, mas fluídas. Isto é considerado de três formas: i) as capacidades de *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 gere-se a ameaça criminal, precisamente através desta previsão do risco; ii) *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).

No fundo, vivemos numa sociedade em que a vigilância está profundamente inscrita e, desta forma, o estudo desta torna-se cada vez mais difícil precisamente porque também é cada vez mais árdua a tarefa de distinguir entre práticas de vigilância e rotinas diárias; quando a vigilância se tornou algo onde vivemos diariamente (Van der Vlist, 2017).

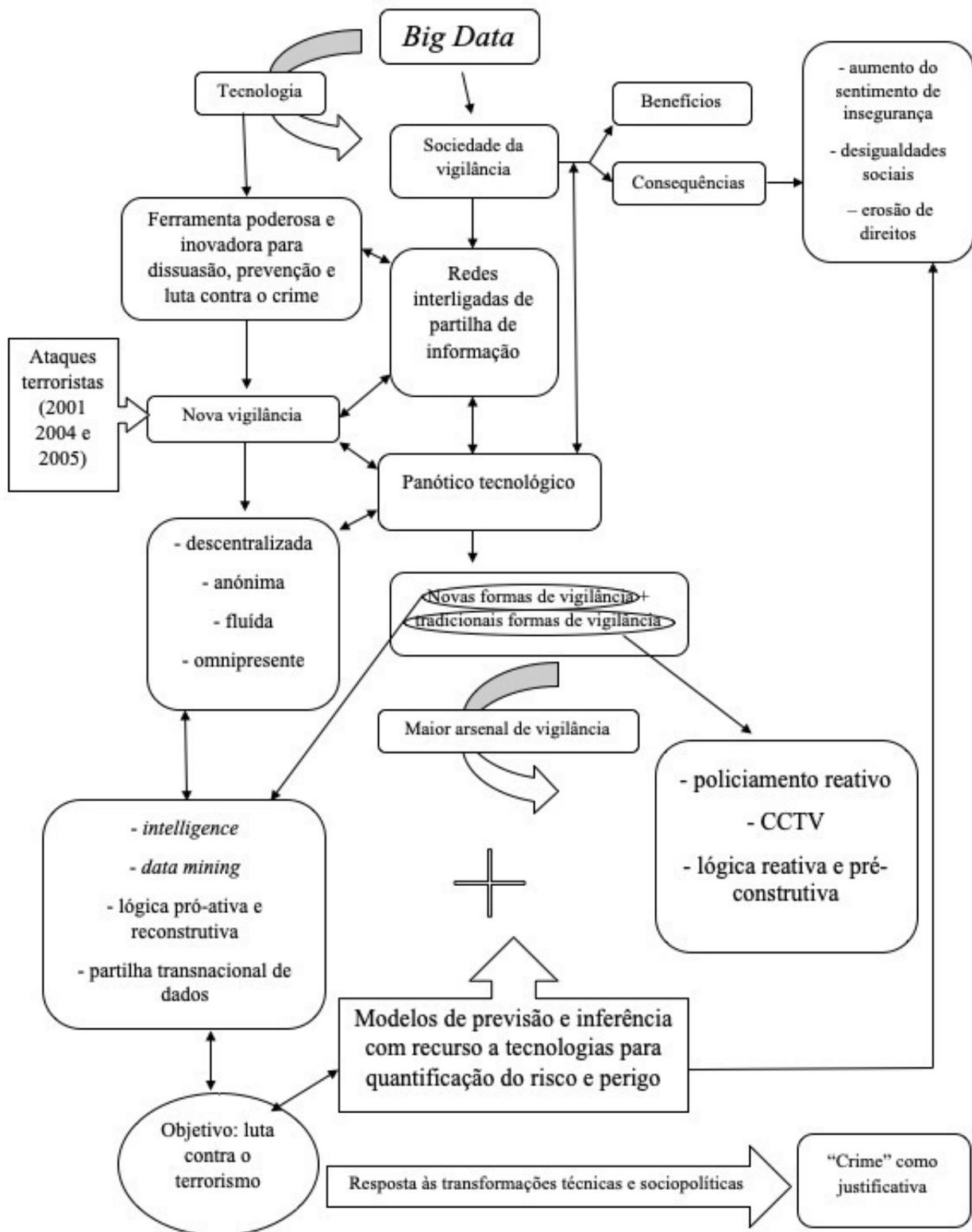


Figura 3: Transformações que operam na vigilância com o surgimento de Big Data.

1.5. Big Data e criminalidade

O *Big Data* surge e intensifica certas tendências de vigilância associadas à tecnologia da informação e às redes (Lyon, 2004; Lyon, 2014) – ancorado sob o desenvolvimento de bancos de dados informatizados de grande escala sendo que estes últimos, conseqüentemente, alimentam a produção de modelos estatísticos destinados a extrair dados, apoiados em *softwares* de inteligência artificial, que possam servir objetivos de segurança pública (Machado & Santos, 2016). É claro que, visto que o *Big Data* inclui um elevado número de dados, este expande conjuntos de dados interconectados e ferramentas analíticas, intensificando a vigilância. A gestão de riscos, o controlo e a repressão da criminalidade vão materializar-se em ações céleres e essencialmente preventivas (Lyon, 2004; Lyon, 2014; Prainsack, 2019).

No âmbito criminal, verifica-se uma tendência crescente do uso de *Big Data* como ferramenta preventiva que orienta estratégias policiais e decisões de justiça criminal (Chan & Moses 2017; Prainsack, 2019). Por exemplo, em estratégias de análise de risco de indivíduos e prevenção criminal sobre lugares, intervalos de tempo e/ou pessoas; a partir destas criam-se perfis que combinam, com precisão, os potenciais infratores aos crimes específicos (Chan & Moses, 2015; Brayne, 2017). As ações policiais são, também, eminentemente preventivas, preditivas e orientadas para o problema – reflexo claro da inclusão do *Big Data* nas estratégias dos agentes policiais, visto que todas estas ações se materializam após um trabalho de pesquisa e investigação que conota locais, pessoais e momentos, como análises de bancos de dados, de modelos de regressão estatísticos, análises de risco e uso de técnicas avançadas de manuseamento de dados (Chan & Moses, 2015; Amooore, 2011; Prainsack, 2019). Estas análises de risco relacionam padrões de dados e metadados com os comportamentos reais ou potenciais dos indivíduos para produzir informações sobre quem são e o que fazem (Van Dijck, 2014; Amooore, 2011). Também na medicina preventiva, nomeadamente na que visa prevenir e combater a violência (física, psicológica e sexual), na medida em que uma ampla gama de informação sobre o contexto de vida (social, contextual, económico e cultural) pode ser recolhido e estar relacionado com dados de saúde, por exemplo, que permita desenvolver, com sucesso, modelos de prevenção da violência. O acesso a uma maior quantidade de dados e, conseqüentemente, 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. 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).

Mas a inserção e inclusão dos *grandes dados* na previsão e repressão da criminalidade não reside só nos tempos contemporâneos (Brayne, 2017). Os bancos de dados computarizados não são novos. 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 elaborara 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 em meados de 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. Em 1994, já se implementara o *CompStat* – um modelo de gestão do risco que identificava padrões de crime, quantificando e incentivando atividades policiais, direcionando os seus recursos. Assim, pode constatar-se que os EUA possuem o Sistema de Justiça Criminal mais maleável e flexível a estas tendências crescentes de vigilância – desde o crescente encarceramento à expansão do policiamento. Em 1994 foram desbloqueados fundos monetários que visaram a contratação de mais de 100 mil agentes policiais e em 2002, regulamentou-se o financiamento de Agências Locais de Aplicação da Lei. Recentemente, os fundos federais foram direcionados na melhoria e na 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 o que, conseqüentemente, criou condições para a gênese e expansão de uma análise por via do *Big Data*, foi uma transformação revolucionária no Sistema de Justiça Criminal: uma mudança para uma *justiça atuarial*³ (Feeley & Simon, 1992) em que os profissionais utilizam critérios-resultado de gestão do risco (Lyon, 2004; Amore, 2011) para estimar probabilidades de risco criminal (Andrews & Bonta, 2010; Brayne, 2017; Amore, 2011).

³ Uma das limitações importantes existente na literatura acerca desta temática reside no atuarialismo: este foi documentado, definido e tratado, *a priori* da gênese dos grandes dados. Ainda que o paralelismo e a interrelação possam ser estabelecidos e considerados, deve ressaltar-se esta limitação (Brayne, 2017).

A ideia de que é possível e se deve prever o risco dos indivíduos e categorizá-los, para prever os fenómenos criminais.

Além disso, geralmente, nos casos de análise de dados no âmbito da aplicação da lei, os *softwares* fazem suposições que implicitamente e explicitamente dependem de Teorias criminológicas. Por exemplo, Teorias que foram desenvolvidas para calcular e teorizar o intervalo de tempo necessário de histórico criminal a considerar para prever a reincidência criminal. Ou seja, reflete o uso prático de *Big Data* e a adaptação da Criminologia e das Ciências Sociais Humanas, em geral, ao aparecimento de novas técnicas e metodologias (Chan & Moses, 2015). Gonçalves (2017) refere que, nos termos do Regulamento 2016/679, a abordagem baseada no risco representa o principal método de aplicação, atualmente. Este risco não se concentra apenas nos indivíduos ‘suspeitos’, mas sob toda a população – um risco generalizado. Assim, a densificação dos sistemas de vigilância encontra a sua justificação na necessidade de gestão do risco, visto que um sistema de vigilância é considerado mais eficaz se for capaz de irradiar os problemas antes que eles emergam (Matos, 2018).

Recentemente, exemplo da aplicação de *Big Data* ao combate à criminalidade, mais concretamente, ao terrorismo é o Projeto TIA – *Total “Terrorism” Information Awareness*, que propõe, de forma clara, capturar a *assinatura-informação* das pessoas de forma 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 vigia e controlo de indivíduos, na recolha e armazenamento do maior número possível de informação capturada para, posteriormente, através do uso de *softwares* e 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, permitam delinear as ações desses indivíduos. O TIA acaba por possibilitar a construção do perfil total dos indivíduos, resultado do cruzamento das ligações telefónicas desses mesmos indivíduos (origem, destino, data e duração), as despesas registadas nos cartões de crédito e demais informações relevantes. No fundo, o Projeto foi criado para produzir uma visão dos padrões comportamentais da população, tendo como objetivo auxiliar os analistas na predição de ações terroristas futuras (Costa, 2004; Guzik, 2009; Lyon, 2015).

Não obstante estes Projetos, criados eminentemente da necessidade de cumprir objetivos específicos – neste caso, diminuir as ameaças terroristas –, outras aplicações foram sendo criadas, por exemplo, com o objetivo de monitorizar a localização pessoal.

Estas aplicações funcionam através do GPS – *Global Positioning System* e de redes celulares e dispõem de um conjunto de tecnologias que asseguram que uma determinada rede localize o local onde a pessoa se encontra (Costa, 2004). E 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 ASN dos EUA que, atualmente, tenta capturar e analisar, virtualmente, todas as chamadas telefónicas, *fax*, mensagens e *e-mails* enviados de qualquer ponto do planeta. Empiricamente, este Sistema possui estações de interceção de sinais em todo 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, de caracteres e 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 fora concebido para espionagem militar e diplomática, atualmente debruça-se sobre tráfico de droga, branqueamento de capitais, terrorismo e criminalidade económica e organizada (Costa, 2004).

Na mesma linha de pensamento do Sistema *Echelon*, Tsianos & Kuster (2016) descrevem, também, o *Eurodac*, o mais antigo banco de dados biométrico europeu de larga escala, em operação 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 os 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 informático, sendo considerado parte da *fronteira inteligente* da Europa: uma fronteira difusa que não pode ser geograficamente localizada, mas depende de locais virtuais e instituições de controlo e vigilância, que estão conectados através de redes de dados digitais. A introdução do termo técnico *fronteira inteligente* ou *fronteira tecnológica da Europa* processou-se em simultâneo com da *reengenharia programática* e o redimensionamento da gestão de fronteiras após o 11 de Setembro. Este é um instrumento que possibilita a desterritorialização da fronteira europeia externa e, potencialmente, alargando-a a todo o espaço *Schengen*. O saudável funcionamento desta fronteira é possível por, recentemente, esta se conjugar com programas algorítmicos que lhe permitem a identificação de casos relevantes (Tsianos & Kuster, 2016).

Assim sendo, gradualmente, o *Big Data* foi-se proliferando, 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 sob uma perspectiva que melhoraria 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 referiu que *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). E o debate em torno deste tema tornou-se mais evidente a partir do momento que a popularidade em torno do uso de *softwares* analíticos de dados ao serviço do policiamento preventivo aumentou (Chan & Moses, 2015) e desde que a tecnologia se tornou um recurso muito valioso (Chan & Moses, 2017).

Os estudos empíricos que se debruçam sobre a aplicação prática, expectativas e perspectivas de *Big Data* são ainda escassos, o que se reflete numa falta de informação acerca da relação (in)existente entre vigilância, grandes dados e as consequências sociais da interseção destas duas forças (Brayne, 2017; Bartlett et al., 2018). Consequentemente, dos escassos estudos existentes, dois são exclusivamente sobre *Big Data* e a sua (potencial) aplicação no contexto de investigações criminais policiais, sendo que os restantes se prendem com breves considerações acerca de algumas implementações e alterações que foram operando em alguns Departamentos Policiais, fruto do desenvolvimento e expansão da técnica. Destacam-se, neste âmbito: um estudo qualitativo de dois anos e meio no Departamento de Polícia de Los Angeles (LAPD) de Brayne (2017); e um estudo na Austrália, no *Data to Decisions Cooperative Research Centre* (D2D CRC), de Chan & Moses (2017). Estes dois estudos revelaram duas tendências aparentemente contraditórias: se por um lado, o *Big Data* pode conduzir à mudança (Brayne, 2017); por outro lado, pode criar resistências à alteração de práticas e reforçar culturas policiais tradicionais (Chan & Moses, 2017).

Brayne (2017) observou e verificou, através da realização de observações e entrevistas, durante dois anos e meio, que com o advento do *Big Data* verificavam-se mudanças no LAPD⁴. O autor, por via de uma descrição empírica, refere que a análise

⁴ Terceiro maior Departamento Policial dos Estados Unidos, com 9947 oficiais, 2947 funcionários civis, que 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).

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 dos dados, o cruzamento destes dados; v) a inserção mais facilitada de indivíduos nos bancos de dados da lei; vi) os sistemas policiais anteriormente separados, atualmente fundiram-se, facilitando a disseminação da vigilância entre muitas e diferentes instituições policiais; e vii) a emergência da chamada “análise de rede” – onde os agentes policiais têm conotados, no seu mapa digital, todos os indivíduos pontuados com alto risco e, depois, analisam as ligações que esses indivíduos têm a lugares e pessoas. Concretamente, a primeira mudança descrita por Brayne (2017) reside na quantificação do risco individual dos cidadãos, através de um novo sistema de pontos: Operação LASER (*Los Angeles Strategic Extraction and Restoration Program*) – conota lugares e indivíduos potencialmente perigosos, por via de modelos baseados em evidências. Ou seja, indivíduos-autores de crimes violentos eram conotados e esta era uma estratégia que permitiria reduzir as taxas de criminalidade; mediante e conforme o tipo de ilícito criminal que constasse no Certificado de Registo Criminal de cada indivíduo, a cotação era proporcional à gravidade do ato cometido. Assim, criavam-se índices de pontos que vigiavam estes indivíduos e que eram 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 verbalizavam o facto de que, quanto maior fosse a conotação 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 criminais – o que leva a criação de desigualdades. Ou seja, estes oficiais revelaram-se conscientes relativamente ao contexto organizacional restritivo – no que toca aos recursos – da Polícia. A segunda mudança materializou-se numa mudança no paradigma de atuação policial: de um policiamento reativo, a um policiamento preventivo. Inicialmente, o policiamento era essencialmente reativo, ou seja, os agentes policiais apenas operavam na atuação para determinado incidente após a sua efetivação. No entanto, no início da década de 1980, estas estratégias revelaram-se ineficazes e, por

isso, adotaram-se estratégias preventivas que visavam atuar antes da ocorrência de incidentes. O mesmo aconteceu no LAPD: em 2012, este começou a usar um *software* – *PredPol* – que, com base em algoritmos, elabora modelos que prevêem a ocorrência de um facto criminal, tendo em conta o histórico espaciotemporal de ocorrência de incidentes criminais. Ou seja, se em determinado local se registou um crime, a probabilidade de se voltar a registar é maior, então esse local fica registado no algoritmo e conotado como *hotspot* criminal. Este *software* disseminou-se nas diferentes secções do LAPD (Brayne, 2017). Por fim, as duas últimas mudanças representam transformações práticas nas atividades de vigilância policial. Uma delas reside na expansão e utilização da *análise de rede* que conota um maior número de indivíduos – que (não) estabeleceram contacto com a Polícia. Ou seja, inicialmente, eram conotados apenas indivíduos que já tinham estabelecido contacto com a Polícia; posteriormente, eram conotados indivíduos que se revelavam potencialmente perigosos, independentemente de um contacto prévio com a Polícia, começando a ser conotadas atividades diárias, familiares diretos de indivíduos que já tinham histórico criminal e locais frequentados pelos familiares. Ou seja, expandiu-se a rede de vigilância, tendo-se verificado a emergência de uma linha tênue que limita, cada vez menos, a conotação de indivíduos. A outra mudança prende-se com a criação de um Sistema de Dados Institucionais Integrado – a digitalização dos registos policiais, bem como, a sua consequente proliferação facilita a sua partilha, conexão e consulta a partir de qualquer Secção do LAPD, bem como, qualquer agente policial, a partir das bases de dados policiais. Assim, a Polícia aumentou o seu leque de acessos a dados que podem ser relevantes à sua ação (Brayne, 2017).

Em contrapartida, Chan & Moses (2017) concluíram que, apesar da polícia estar consciente do potencial da tecnologia para estratégias policiais mais inteligentes, afirmam que não há tempo nem recursos para beneficiar desse potencial. Chan & Moses (2017) verificaram, através de 31 entrevistas semi-estruturadas 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 trata 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 utilizavam *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 na vantagem para a investigação que estes trazem.

Os entrevistados pertencentes a Organizações Políticas perspetivavam *Big Data* como uma técnica inovadora que cria oportunidades para uma abordagem mais proativa da segurança, baseada em inferências sobre tendências e padrões. Ou seja, apesar das promessas de *Big Data* para melhorar a eficiência e eficácia das atividades policiais e agentes de segurança, muito pouco se sabe sobre como o *Big Data* é entendido ou imaginado por esses agentes. Este estudo permitiu concluir que *Big Data* é uma técnica de segurança que, para os agentes, se trata de uma novidade que é contestada (Chan & Moses, 2017).

Kubler (2017) realizou um estudo empírico, também, na Polícia Francesa, logo depois do estado de emergência ter sido ativado pelo Estado Francês pela 4.^a vez, sucedido dos ataques terroristas. Kubler (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 tem usado 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, procurando encontrar o maior número de associações entre estes dois últimos, classificando-os com base no seu nível de importância para a investigação em apreço. O *i2 Analyst's Notebook*, surge 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. 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 de dados e produzir visualizações que ajudarão a revelar conexões que não seriam feitas de outra forma. Este tipo de *software* permite que as agências policiais tenham acesso a grandes quantidades de dados públicos, ao mesmo tempo que

economizam recursos e tempo. No fundo, reduz o tempo necessário para reconhecer informações necessárias para criar uma rede sólida e oportuna que permite a identificação, previsão, prevenção e combate ao terrorismo, 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 decisão 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). Ou seja, neste Departamento, começam a colocar-se em prática técnicas de *Big Data*, fruto das mudanças e acontecimentos sociais e contextuais.

Também Drewer & Miladinova (2017) descrevem uma aplicabilidade prática do *Big Data*, mais concretamente nas atividades da Europol, por via de uma análise documental das 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 dos dados são diversas e múltiplas. A Europol, numa escala mais ampla, em algumas áreas específicas, particularmente cibercrime e terrorismo, reconhece os grandes dados como importantes, visto 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 pró-ativo, 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 ancora 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 Dijck, 2014; Wood et al., 2006; Matos et al., 2016).

No contexto do novo processamento de dados avançado e modernizado, a Europol enfrenta um fluxo de grandes quantidades de informações que podem ser usadas para analisar informações estratégicas e melhor compreender a dinâmica do crime. As capacidades da Europol aumentaram e esta 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. A

Europol atualmente prevê uma tecnologia inovadora habilitada que interconecta mais de 500 Agências Policiais da Europa, utilizando, constantemente, os seus bancos de dados, por via de canais de comunicação, oferecendo recursos rápidos e seguros para armazenar, pesquisar, visualizar e vincular informações. Consequentemente, também definiu altos padrões de privacidade e proteção de dados, bem como, segurança destes últimos, visto que estas inovações exigiram uma adaptação regulamentar e legislativa por parte da Europol que, desde maio de 2017, operou a 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. Consequentemente, 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 sob medida, 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, a Europol e o seu regulamento abordam o futuro da privacidade digital num contexto de tecnologia neutra. O regulamento 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 está no futuro melhor equipada para detectar rapidamente as ligações entre investigações e *modus operandi* em diferentes grupos criminosos, para verificar *cross-matches* de dados e ter uma visão clara das tendências, garantindo um alto nível de proteção de dados pessoais (Johnson & Williams, 2007; Drewer & Miladinova, 2017).

Estes exemplos não esgotam a (potencial) aplicação prática do *Big Data*, intensificada logo após os ataques do 11 de setembro, época em que as Agências Federais proporcionaram financiamentos consideráveis a Agências Policiais para recolher, analisar, partilhar e armazenar grandes conjuntos de novos dados (Souza, 2010; Brayne, 2017). A partir desse momento, em Departamentos Policiais como os de Nova Iorque, Londres e Los Angeles começaram a ser implementadas estratégias policiais preditivas e

preventivas que incluem algoritmos que antecipam o local e o momento onde crimes futuros irão ocorrer; bem como, os indivíduos com maior probabilidade de desencadear condutas violentas e comportamentos de risco. Ou seja, todos estes acontecimentos potenciaram o desenvolvimento de um *policimento orientado para estratégias de inteligência*. Assim, o *Big Data* começa a expandir-se enquanto técnica que leva à tomada de decisões mais informadas, capazes e eficientes, ajudando a reduzir as taxas de criminalidade (Brayne, 2017).

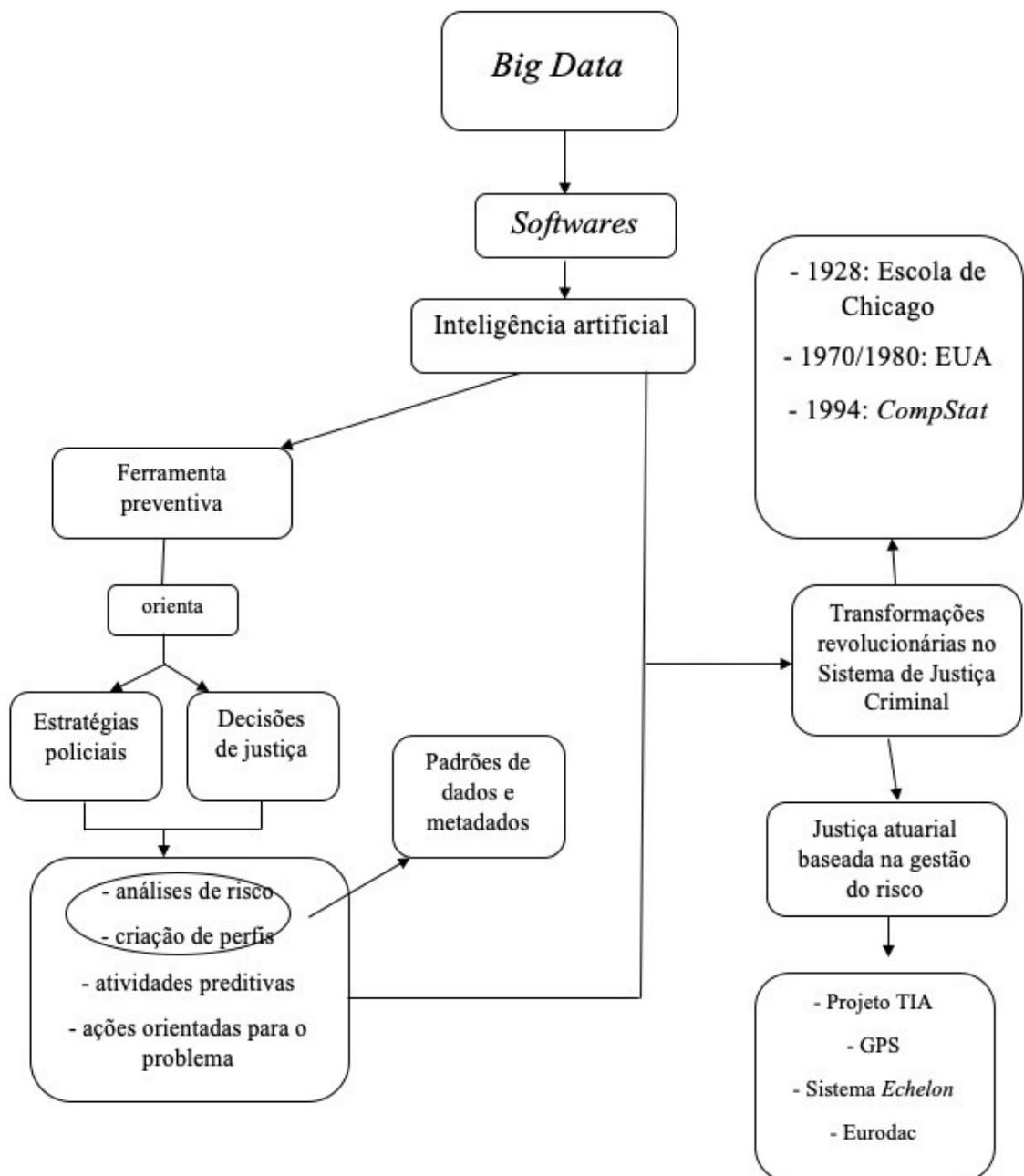


Figura 4: Relação entre Big Data e criminalidade.

1.6. Questões de direitos, liberdades e garantias – debate académico contemporâneo

Com o crescimento desta forma de se 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 tornos 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). Enquanto *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 deliberadamente 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, *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 *The New York Times* documentava os *Oito (não, nove!) problemas com Big Data*, alegando que se estava a cometer um grande erro com o seu uso, visto que se verificava uma atmosfera controversa no planeta 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. Representa um desafio para o respeito pelos seguintes princípios: consentimento (os dados pessoais devem ser processados apenas se o titular dos dados tiver dado o seu 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,

surtem 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 levantam-se questões sobre privacidade individual e exposição à vigilância relativamente ao autónomo fluxo de dados 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 Banksy*: 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 ator, mapeando a cidade. Ou seja, usaram o perfil geográfico – uma técnica estatística de inferência, tradicionalmente usada para crimes em série como a violação e o assassinato – para encontrar o suspeito. Analisaram os padrões espaciais das obras de arte de *Banksy* em Londres e Bristol, pesquisaram pistas eleitorais dos antigos endereços dele, da sua esposa e de lugares que ele provavelmente frequentou. As obras públicas de *Banksy* foram mapeadas e assim alcançaram uma residência e ao entrar na mesma – invadindo-a – esta não correspondia a *Banksy*. E tudo isto é questionado. Os agentes policiais alegam que se tratam de técnicas que conseguem antecipar e prevenir comportamentos terroristas, no entanto, são técnicas que absorvem a ética na sua prática (Metcalf & Crawford, 2016). Se, por um lado e 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 sobre as suas mais diversas arestas: a fiabilidade dos dados obtidos, as conclusões retiradas das análises desses dados, o surgimento de uma sociedade panótica, que em tudo se assemelha ao *Big Brother*, por via de uma vigilância permanente, invisível e penetrada 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, terrorismo e 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 de *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; para além disso, estudar o processo de agregação dos dados (Lefèvre, 2017; Metcalf & Crawford, 2016; Ball et al., 2016; Boyd & Crawford, 2012). E considerar, também, que o uso de uma grande quantidade de dados pessoais acarreta uma consequência óbvia: a questão da privacidade individual e social e proteção de dados (Drewer & Miladinova, 2017; Coll, 2014). Com a densificação dos mecanismos de vigilância invisíveis e omnipresentes pode verificar-se a possibilidade de se observarem direitos desrespeitados, nomeadamente o do consentimento (Lyon, 2014).

Por sua vez, 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 de estabelecer um perfil abstrativo do indivíduo e, 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 quando nos referimos a um grande número de dados, o que o termo sugere – grande quantidade, em termos numéricos de fontes de informação – não corresponde ao que representa. No fundo, o que quer referir-se é à abrangência de cobertura sobre a vida contemporânea, a omnipresença, o 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). E ainda no que concerne aos dados, importa descortinar de onde vêm estes dados, como devem ser interpretados, qual a direção a tomar e não os incluir em estudos sem nexos (Boyd & Crawford, 2012). Para além disso, enquanto conjuntos de dados grandes que podem ser modelados, os dados 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 elas transmitam o equivalente (Boyd & Crawford, 2012). Os níveis de permissão, em termos de recolha de dados, anteriormente disponíveis apenas em ambientes políticos rigorosamente controlados, estão agora disponíveis universalmente (Ball et al., 2016). No

entanto, *Big Data* não se refere apenas a conjuntos de dados muito grandes nem a ferramentas e procedimentos usados para manipulá-los e analisá-los, mas também para uma mudança computacional no pensamento e pesquisa (Boyd & Crawford, 2012).

Conseqüentemente, 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 usado para investigar os manifestantes e suprimir o seu discurso? Dada a expansão de *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 no que toca ao fluxo de informação, ao acesso a uma grande variedade de dados e uma enorme vigilância que é sediada sobre o ser humano; e é sempre crucial considerar a perspectiva do observado e do vigiado – de que forma o indivíduo se apercebe e como, de que está a ser alvo de uma observação criteriosa (Ball et al., 2016)? Não obstante, importa ter em conta que o *Big Data* não é auto-explicativo, ou seja, não é porque o *software* denota uma correlação que ela existe simplesmente, é preciso procurar as causas, perceber o fenómeno e aceder a ele, até porque mesmo os factos considerados mais reais sofrem mudanças e alterações (Boyd & Crawford, 2012; Chan & Moses, 2015; Amooore, 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 recebe acesso? Para que fins? Em que contextos? E com o que restrições? Enquanto a explosão da pesquisa usa conjuntos de dados dos *media*, as fontes sugerem que o acesso é tudo menos direto, 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 outros oferecem pequenos conjuntos de dados para investigadores baseados em universidades. Isso produz desigualdades consideráveis no sistema: aqueles com poder económico podem produzir um tipo diferente de pesquisa do que aqueles que não possuem poder monetário. Os que não têm acesso não podem reproduzir ou avaliar as reivindicações metodológicas de quem não têm acesso (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 da pesquisa não pode ser ignorado simplesmente porque os dados são aparentemente públicos. Os pesquisadores devem perguntar-se – e seus colegas –

sobre a ética do processo da recolha de dados, análise e publicação (Boyd & Crawford, 2012). E este acesso aos dados, ancora toda 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 moedas de troca, como forma de garantir o seu ‘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), emergindo, gradual e consequentemente, um modelo de negócio das empresas, assente no rastreio dos indivíduos, exercendo estas um verdadeiro domínio sobre os dados pessoais (Lupton & Michael, 2017; Gonçalves, 2018). À semelhança do que Araújo (2011) refletira acerca do tempo, abstraído o objeto-estudo de reflexão: *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) alegam que recentemente se verifica uma forma cada vez mais *mercantilizada* das ferramentas de provisão científica forense. Adotou-se uma racionalidade económica destes fenómenos – potenciais de prever e combater a criminalidade, sendo que o Estado neoliberal, também tece estas considerações mercantilizadas (Lawless & Williams, 2010).

Não obstante as considerações tecidas, Brayne (2017)⁵ enfatiza as repercussões que o uso de grandes dados tem para enfatizar as desigualdades sociais. O autor observou que no LAPD era usado um modelo de cálculo de perfis de risco que conotava indivíduos como *mais prováveis* de desenvolver condutas criminais. Brayne (2017) alega que se, por exemplo, forem conotados um número considerável de indivíduos que residam todos no mesmo bairro, aquele bairro ficará sobre vigilância constante da polícia como local onde residem indivíduos perigosos, sem que sequer se deva confirmar tais atos criminosos. Assim, criam-se aqui desigualdades; e caso se trate de um bairro onde residam minorias étnicas, as desigualdades são (ainda mais) reforçadas, sendo, simultaneamente, difíceis de detetar visto que esta vigilância é invisível, no entanto atua para as perpetuar (Brayne, 2017). Na mesma esteira, Richards & 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 reparam que o *Big Data* agrega

⁵ Artigo explanado anteriormente, nesta Dissertação, na secção “Estudos empíricos”.

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 os autores sabem que essas informações estão (já) nas mãos dos governos e corporações policiais (Herschel & Miori, 2017). Desta forma, esta vigilância sem precedentes pode, em última instância, contribuir para a expansão e perpetuação da conotação 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, a vigilância primária e secundária enfatiza e potencia um posterior contacto provável com as instâncias policiais. Assim, a inserção sem critérios, nem restrições, da população nas bases de dados pode ter consequências nefastas – levar à condenação de indivíduos inocentes (Brayne, 2017). Zedner (2016) referiu, precisamente, este aspeto: a ideia de que há uma discriminação ainda maior sobre os grupos já marginalizados, podendo levar ao enfraquecimento da segurança. E toda esta panóplia de reflexões acentua-se com o advento de *Big Data*, que exige que se reflita sobre a justiça, no meio do arsenal tecnológico dos grandes dados. Onde pairam os direitos dos cidadãos? Onde se define o titular dos dados de cada pessoa (Cinnamon, 2017)?

Todos estes dilemas e estas 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 ideia comum de um *atraso direito-tecnologia*, ou seja, a incapacidade alegadamente intrínseca da legislação para 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 protecção de dados e incluiu as tecnologias de *Big Data* como uma categoria de aplicações digitais a serem abordadas sob a protecção de dados estabelecida e os seus princípios. Esta inclusão e consideração não deixa que *Big Data* seja perspetivado como uma mudança fundamental, até mesmo uma *mudança tectónica* no uso de dados pessoais; e estes desenvolvimentos tecnológicos têm consequências na sociedade. E isso também se aplica para as tecnologias de *Big Data*, com todas as suas capacidades algorítmicas e preditivas e a sua 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 o *Big Data* ao Estado de Direito, a Comissão Europeia coloca regularmente a ênfase no novo regime de protecção de dados como um “facilitador de serviços de *Big Data* na Europa”. Na verdade, a legislação de protecção de dados também significa maximizar o potencial das

tecnologias digitais, tendo em conta as consequências económicas e sociais destas últimas (Gonçalves, 2017).

A este respeito – da protecção de dados e do seu conseqüente debate após a imersão numa *Era de Big Data* – em fevereiro de 2016, o Conselho da Europa iniciou o processo de elaboração de orientações específicas sobre a protecção de informações pessoais no contexto do *Big Data* (Mantelero, 2017). Na última década, os efeitos concorrentes 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 são capazes de extrair informações preditivas a partir de conjuntos de dados e os algoritmos são cada vez mais, tornando-se parte dos processos de tomada de decisão. Ao contrário das orientações anteriormente adotadas pelo Conselho da Europa, que lidam com contextos ou questões específicas, estas últimas Diretrizes focam-se no uso de uma determinada tecnologia (*Big Data*) e não são específicos do setor. Além disso, uma vez que há um número crescente de aplicativos de análise de *Big Data* e de policiamento preditivo que geram novas questões em termos de protecção de dados, as Orientações não podem definir um conjunto de disposições que levem em conta todas as implicações potenciais do uso dos grandes dados. Por estas razões, as Diretrizes são necessariamente gerais, mas podem ser complementadas por outras orientações sobre a protecção de indivíduos em campos específicos de aplicação de *Big Data*. Apesar desse limite, as Diretrizes representam um importante passo na regulação do uso de *Big Data*, como a primeira orientação internacional neste campo, onde questões significativas surgem em relação ao paradigma tradicional de regulamentação da protecção 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, o que levanta novas preocupações acerca da liberdade individual, processos de decisão e responsabilidade. As Diretrizes afirmam que o uso de *Big Data* deve “preservar a autonomia da intervenção humana no processo de tomada de decisões” e isto levanta (ainda mais) questões acerca da subjetividade da opinião da pessoa responsável pela decisão final. Isto leva a concluir que o uso de *Big Data* se trata de um processo humano, resultado dos seus conhecimentos, criado e influenciado por este e, por isso, todas as questões éticas e de direitos humanos que se levantem sobre ele devem ter este aspeto em conta, tal como a legislação tem (Mantelero, 2017).

Concluindo, apesar dos esforços académicos, civis, democráticos, políticos e legislativos no sentido de encontrar um equilíbrio entre a defesa da segurança dos

cidadãos e as potenciais ameaças às suas liberdades, garantias e direitos, este parece difícil de alcançar (Machado & Santos, 2016; Cinnamon, 2017; Matos, 2018). A inovação e evolução na tecnologia são constantes e desejáveis, mas as formas pelas quais as tecnologias são usadas para monitorizar e governar os cidadãos são negociáveis. 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 conceitualizar a vigilância, também exige esforços renovados para reinventar os argumentos legais e intervenções que podem ser usadas para abordar e refrear essas práticas. Autores surgem alegando que se *Big Data* agrega todas as informações existentes, porque não se prevêem efetivamente ataques terroristas? (Andrejevic & Gates, 2014). No entanto, apesar das críticas maciças sobre os motivos epistêmicos, éticos e de direitos, *Big Data* continua a ser um termo popular, político, público e humano.

1.7. Lente conceitual empírica: *Sociology of Expectations*

A linha de pensamento e, conseqüentemente, de questionamento da realidade social e interpretação dos dados prende-se com a perspectiva da *Sociology of Expectations*, na medida em que o *Big Data* rodeia-se mais de expectativas de uso orientadas para o futuro na área da vigilância e controlo da criminalidade, do que no seu uso efetivo – *Big Data* enquanto futuro possível (Van Lente, 2012). Trata-se de um *objeto real sem realidade* (Araújo, 2005: 9), na medida em que apesar de não se tratar de um objeto real, palpável, produz efeitos não materiais, mas definíveis: existe na realidade enquanto discurso que produz efeitos; é veiculado por atores que lhe pretendem dar contornos dinâmicos. (Araújo, 2005). Ou seja, este eixo teórico central relaciona-se com tecnologias e/ou técnicas que estão (ainda) em fase embrionária e relativamente precoce de aplicação, onde se insere *Big Data*. No fundo, é antecipar o surgimento evolutivo de contextos científicos tecnológicos, visto que 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). Assim, a compreensão dos diferentes discursos ancorados a este eixo teórico possibilitou o acesso à co-construção de visões científicas particulares do desenvolvimento tecnológico, particularmente *Big Data*, analisando o desenvolvimento desta tecnologia em termos sociológicos (Tutton, 2011), há uma necessidade de estudar o que não é objetivamente

fixo num tempo e num espaço, o que pode até parecer inexistente (Araújo, 2005). Tutton (2011) refere que a compreensão das expectativas, promessas, antecipações, visões e esperanças sobre fenómenos em desenvolvimento é central para aceder aos conceitos e ao seu futuro desenvolvimento. No fundo, foi o que se pretendeu com esta investigação: aceder às expectativas antecipatórias da (potencial) aplicação de *Big Data* no contexto das investigações criminais transfronteiriças. 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 desenvolvimento da responsabilidade para o 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). Compreender a extensão e a profundidade do futuro que caracterizam os indivíduos, grupos, profissionais ou classes sociais é uma das orientações temporais que define os movimentos da ação dos sujeitos no futuro; assim, é crucial aceder às ansiedades, expectativas e preocupações dos atores sociais sobre as tecnologias e a sua força definidora dos futuros da sociedade, as suas *imagens do futuro* (Araújo, 2012: 5). Apesar de que o futuro não é mais do que um tempo que conjuga o presente e o passado e, por isso, tendo por base acontecimentos passados e presentes, os indivíduos constroem representações sobre este, enquanto preocupação coletiva e categoria central no pensamento do progresso (Araújo, 2005).

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 continuamente e explicitamente referem-se ao que é (e será) possível no futuro (Van Lente, 2012) e as Ciências Sociais têm-se tornado cada vez mais permeáveis ao desenvolvimento tecnológico inovador, bem como, à forma como este se tem desenvolvido e processado (Hedgecoe & Martin, 2003), visto que o que reside no futuro está apegado, inevitavelmente, ao conceito de mudança e é esta mudança que importa estudar (Araújo, 2005). As expectativas sobre o futuro inspiram novos desenvolvimentos tecnológicos. Assim, este eixo teórico visa explorar a forma como as expectativas na Ciência são estruturadas, como emergem, como amadurecem e se tornam o centro da

atenção dramática ou, por sua vez, desaparecem silenciosamente. Não obstante, visa, também estudar a forma como estes discursos influenciam as decisões governamentais e empresariais, porque se crê que a política direciona o desenvolvimento tecnológico (Van Lente, 2012). Além disso, a análise destas expectativas é tanto mais importante quanto mais a tecnologia se encontra em fase embrionária de desenvolvimento por estar repleta de incertezas (Brown et al., 2003; Selin, 2008) – como é o caso de *Big Data*. Uma tecnologia que está a emergir e, portanto, existe principalmente na forma de visões, promessas e expectativas. E as Ciências Sociais, nas últimas décadas, tornaram-se mais conscientes da importância de avaliar a conveniência das novas tecnologias emergentes, ainda no início do seu desenvolvimento, visto que desta forma é possível antecipar desenvolvimentos quando eles ainda são maleáveis, mesmo que se encontrem ainda no desconhecido e no incerto. As tecnologias emergentes sustentam-se em hipóteses futuristas e, conseqüentemente, apresentam questões éticas altamente existenciais (Lucivero, Swierstra & Boenink, 2011). Não obstante, as expectativas são um recurso para os atores envolvidos na inovação, na medida em que podem auxiliar a legitimar, justificar e apoiar os seus argumentos, bem como, mobilizar fundos e recursos, a atenção de outros atores, permitindo a tomada de decisões, reduzindo a incerteza inerente ao desenvolvimento tecnológico. Portanto, as expectativas fazem parte de um meio que permite gerar apoios e recursos para a rede sócio-técnica emergente (Hedgecoe & Martin, 2003).

Estas expectativas orientadas para temáticas futuras podem criar novas oportunidades, orientar atividades e fornecer estrutura, legitimando interesses, promovendo investimentos e antecipando potenciais riscos e benefícios. Tendo em conta que a probabilidade da inovação se efetivar de forma isolada é diminuta, avaliar estas perspectivas num ambiente dinâmico, incluindo vários entendimentos e expectativas orientados para a futura aplicabilidade da técnica pode aumentar a probabilidade de esta, futuramente, se vir a instalar enquanto prática regular. Assim, ancorados sob esta perspectiva teórica e debruçados sobre a ideia de que as expectativas e perspectivas destes atores mobilizam recursos, é importante analisar a dinâmica destas para compreender a mudança científica e tecnológica (Borup et al., 2006), neste caso *Big Data*, enquanto realidade em progresso em que todas as expectativas formadas em seu redor constituem-se como avaliações do futuro e previsões que permitirão melhorar as capacidades de inovação dos sistemas, identificar caminhos tecnológicos e criar visões comuns de atores relevantes para a temática em apreço. Tendo em conta que para a *Sociology of*

Expectations esta orientação geral para o futuro está particularmente presente no desenvolvimento tecnológico (Van Lente, 2012), onde se insere *Big Data*.

Uma expectativa é uma mudança ou uma criação sobre uma nova realidade e é mobilizadora de algo; são afirmações sobre o futuro e podem ser positivas (promissoras) ou negativas e variar a nível de conteúdo e/ou modalidade; podendo variar de esboços abstratos e abrangentes sobre o futuro (macro) a elementos detalhados (micro) (Van Lente, 2012). As expectativas podem ter, também, um sentido de esperança sujeita a pressões e controlo político (Araújo, 2013). Em termos de conteúdo, podem dizer respeito a aspetos sociais e, a maior parte das vezes, a um misto de aspetos. As modalidades podem variar de declarações tomadas completamente permissivas, sem qualquer tipo de resistência ou a argumentos subtilmente organizados para refutar as previsões (Van Lente, 2012). Estas expectativas – enquanto afirmações sobre alguma coisa e não como descrições verdadeiras ou falsas (Van Lente, 2012) – servem, frequentemente, para unir ou mediar diferentes limites, dimensões e níveis, sendo fundamentais na coordenação de diferentes comunidades e grupos de atores, podendo, a um nível micro, meso e macro, mediar diferentes níveis de organização. E toda esta panóplia de expectativas é maleável ao tempo, enquanto respostas e adaptações a novas condições e/ou problemas emergentes (Borup et al., 2006) – *Big Data* surgiu como uma realidade contemporânea inovadora emergente que exigiu esta adaptação multifacetada, estando ligado a questões técnicas e sociais, visto que as expectativas são visões do futuro onde estas questões surgem entrelaçadas. Assim, as expectativas e imagens do futuro relativamente a *Big Data* aproximam-se do que Borup et al. (2006) descrevem: surgem desenvolvidas e reconstruídas em atividades científicas resistentes constituindo o elo entre o mundo interno e externo das comunidades científicas e seus campos de conhecimento (Borup et al., 2006). Além disso, as expectativas não descrevem apenas o futuro, mas são parte do caminho pelo qual os futuros são criados (Tutton, 2011), surgindo como reflexo e consequência das orientações temporais (Araújo, 2012).

Concretamente, as *expectativas tecnológicas* são representações em tempo real de futuras situações capacitárias tecnológicas – aproximando-se de *Big Data* e das expectativas que os diferentes profissionais possuem acerca desta tecnologia. *Big Data* é um fragmento do futuro, que paira no ideal científico e sobre o qual os entrevistados possuem expectativas, visões do futuro e imagens do amanhã, visto que se trata de uma técnica precoce, em desenvolvimento. Surgiu, enquanto aparato tecnológico, como uma *promessa* para a resolução (mais) eficaz de casos criminais transfronteiriços, mas subsiste

enquanto ideal sem concretização prática, como um *futuro desejado* por alguns profissionais (Borup et al., 2006; Brown & Michael, 2003). Borup et al. (2006) constatam que as expectativas compartilhadas têm mais força e poder para orientar futuras implementações acerca da tecnologia que potenciou o conjunto de expectativas partilhadas. Ou seja, se existem expectativas positivas partilhadas entre diferentes grupos profissionais, a probabilidade da inovação tecnológica se vir a efetivar é maior e, conseqüentemente, de potenciar o debate e a dinâmica de expectativas noutros setores. Este tipo de ambiente é fundamental em fases iniciais de construções e inovação tecnocientíficas – como o *Big Data*. Depois da fase inicial da criação de expectativas em torno das inovações tecnológicas, os estudos referem que se segue a fase decepcional que reside no facto das promessas não terem sido cumpridas e, desta forma, desacredita-se no poder tecnológico da nova ferramenta científica. As primeiras esperanças raramente são proporcionais aos resultados futuros (Borup et al., 2006).

Ancorados sob a perspectiva tecnológica, os estudos referem que existe uma noção culturalmente generalizada de que a tecnologia continuará a oferecer possibilidades para o progresso, apesar do progresso pertencer a um tempo futuro e 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 conseguinte, para aqueles que se situam no limbo destas duas posições – os intimamente ligados ao campo da prática da expectativa e os que conhecem o campo, mas não estão invertidos nele – por exemplo, as partes públicas interessadas, a sua incerteza é relativamente menor: aceitam a promessa desse futuro próximo expectável, mas não têm percepção dos detalhes que possam impedir a realização da expectativa (Van Lente, 2012). Um elevado senso de confiança reflete um distanciamento face às incertezas e normalmente são expectativas expressas por investigadores que conduzem investigações sobre a área na qual se insere a inovação tecnológica científica. Para aqueles que se distanciam da prática quotidiana da (possível) técnica, as incertezas são diferentes: surgem ancoradas ao desconhecimento, ou seja, pelo facto de não estar claro para estes profissionais o que a expectativa implica, eles são indiferentes à prática e, por isso,

incertos quanto a ela. Por sua vez, os atores envolvidos diretamente no trabalho de campo sobre a tecnologia em evolução revelarão expectativas contraditórias, ou seja, confiantes e cautelosas – esta última revela uma flexibilidade interpretativa que impede que se prossiga a uma padronização social devido ao facto de que, muitas vezes, dentro da mesma comunidade científica, verificam-se assimetrias no acesso às informações sob as quais as expectativas são baseadas. Por exemplo, muitas vezes, as incertezas provenientes de 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 se poderá executar as incertezas são realidades quotidianas porque os atores conhecem os detalhes, as condições e as suposições (Van Lente, 2012). Lucivero et al. (2011) referem o mesmo aspeto: os que trabalham nos laboratórios possuem estratégias discursivas ancoradas à materialidade da tecnologia que lá é visível – enquanto espaço onde as controvérsias e os compromissos ganham vida antes que os factos científicos sejam criados e comunicados para o mundo exterior – juntamente com as incertezas, os desafios, as controvérsias e as dúvidas. É um contexto “menos” contaminado pelas considerações estratégicas – ao contrário do que acontece, por exemplo, com académicos que fazem investigação científica – onde as considerações técnicas são as mais abordadas (Lucivero et al., 2011). Nesta linha de reflexão sobre a posição laboral e as expectativas pronunciadas, 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.

Esta linha teórica de tipo construtivista permite mapear várias dimensões relacionadas com *imaginários sociotécnicos* (Jasanoff & Kim, 2009), nomeadamente no que concerne a “mitos urbanos” em torno das técnicas – *Big Data* é o futuro, a população, enquanto massa humana, está massivamente a ser vigiada e esta técnica vai ser usada para efeitos nefastos – o Estado ancora-se no desenvolvimento da tecnologia para garantir maior segurança aos cidadãos, no entanto verifica-se uma vigilância expansiva. E surgem dimensões sociais, culturais e políticas relativas a percepções de benefícios, riscos e visões do futuro que permitem a articulação entre expectativas coletivas – perspectivas amplamente partilhadas -, e expectativas localmente/contextualmente situadas – específicas de atores sociais de determinada área profissional e/ou disciplinar. Estes

imaginários sociotécnicos são formas de vida e ordem social imaginadas, coletivamente, que se refletem na concepção e na realização de projetos científicos e/ou tecnológicos específicos de cada país (Jasanoff & Kim, 2009). *Big Data* subsiste nestes imaginários, enquanto imagem do futuro (Borup et al., 2006) e perspectiva-se este elemento como crucial na vida social e política, visto que a capacidade de o imaginar permite o desenvolvimento social. Estes não informam ou justificam determinados eventos que ocorreram no passado, o seu âmbito de atuação encontra-se no futuro, visto que os imaginários são instrumentais e futuristas (Jasanoff & Kim, 2009).

Assim, este conceito basilar pretende compreender de que modo a multiplicidade de narrativas que circulam na sociedade são apreendidas, filtradas e reinseridas na ação pública. Jasanoff & Kim (2009) defendem que os imaginários operam nas regiões pouco estudadas entre a imaginação e a ação, entre o discurso e a decisão, e entre a opinião pública pouco informada e a política estatal instrumental. E a projeção das infraestruturas de uma tecnologia, como *Big Data*, é o resultado de um conjunto de escolhas científicas, éticas e políticas. Científicas porque é necessário que funcionem corretamente para o que foram pensadas, por exemplo, a partilha e conexão de dados com outros países para prevenir e combater a criminalidade transfronteiriça. Éticas porque equacionam os potenciais riscos ou erros da tecnologia: falsos positivos, privacidade, abuso de poder. E políticas porque foram pensadas dentro de um contexto que visa atingir os mesmos objetivos: maximizar a segurança dos cidadãos. Portanto, a tecnologia (ou artefactos tecnológicos) contém em si mesma uma forte componente política que comprova que as tecnologias não podem ser pensadas fora da sua dimensão social. A ciência e a tecnologia são, em simultâneo, espaços e objetos de política (Jasanoff & Kim, 2009).

Brown & Michael (2003) propõem alguns parâmetros para uma melhor e mais próxima compreensão das expectativas em torno de inovações científicas: i) o *grau de novidade técnica* ou até que ponto uma inovação é percebida 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) a *localização dos atores* relevantes dentro da rede de desenvolvimento da inovação – a sua posição laboral. Neste sentido, existem duas formas de construir expectativas acerca de inovações tecnológicas: uma delas prende-se com *retrospecting prospects* e a outra com *prospecting retrospects*. A primeira reside na lembrança de futuros passados ou da forma como o futuro foi uma vez representado; enquanto que a segunda se debruça sobre a forma como essas perspectivas são implantadas

em tempo real contemporâneo, para construir o futuro. No que concerne ao *Big Data*, este objeto analítico aproxima-se da perspectiva dos autores na medida em que é uma tecnologia com atividade orientada para o futuro. E este exercício constante de aproximação do futuro, projetando-o no presente cria pressão contemporânea e obriga a um exercício de abstração temporal constante e 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 destas representações sobre o futuro serem elementos importantes na construção do presente, existem muitos futuros projetados que falharam e, por isso, geram riscos e incertezas. E apesar das expectativas residirem, em grande parte, em discursos, elas dependem da prática que, muitas vezes, é inexistente e os profissionais são conscientes deste desfasamento entre a expectativa e o presente material. E, por isso, é importante compreender esta dinâmica: aceder à forma como os atores contextualizam o futuro nas suas narrativas presentes nomeando futuros passados que fracassaram. Há uma tendência generalizada de explicar o sucesso e o fracasso nomeando as propriedades da tecnologia. O que acontece é que uma excessiva centralização nas capacidades inovadoras tecnológicas potencia o esquecimento da vertente social – do palco onde os impactos são vividos e produzem consequências. O que devemos apreender é um exercício mútuo de co-valorização da evolução da tecnologia, paralelamente, ao desenvolvimento social. E tal exercício deve permear-se entre a inovação e as promessas deste desenvolvimento tecnológico, a par da necessidade de intervir em momentos-chave para que este desenvolvimento inovador não se torne prejudicial (Brown & Michael, 2003).

No fundo, o que são feitas são previsões – articulações de futuros possíveis, com uma perspectiva realista, um espelho para a situação atual (Selin, 2008; Van Lente, 2012). Os entrevistados abordam *Big Data* por via de constantes paralelismos entre o que é, o que acham ser, como pode atuar e como poderia desenvolver-se. A articulação formal com o futuro acontece em situações em que as expectativas são abundantes e se realizam avaliações informais continuamente – um *mar de expectativas*. As expectativas informais circulam dentro de grupos responsáveis pela formulação de políticas e, inevitavelmente, fazem a sua legitimação e orientam o seu trabalho. Assim, a previsão baseia-se em reportórios existentes de expectativas (Van Lente, 2012).

Lucivero et al. (2011) referem que para se interpretar as expectativas importa explorar a forma como estas são construídas, como e poquê pelos diferentes públicos – profissionais – situados em determinado ponto no tempo e no espaço e inseridos num

fundo específico. Importa, também, desnudar algumas das características estruturais das expectativas. Paralelamente às conclusões alcançadas pelo estudo dos autores, aplicando ao *Big Data* e às expectativas expostas encontramos três tipos de abordagens: i) abordagens sobre as características e funcionamento da tecnologia; ii) abordagens sobre a forma como a tecnologia será adotada pelos profissionais e como será integrada na prática atual; e iii) abordagens sobre como a tecnologia irá abordar um problema social ou uma necessidade. Além disso, referem que as tecnologias futuras se baseiam sempre em algo: estado da arte (sucesso anterior da técnica), evidências científicas, dados científicos ou na comunicação social pública. Neste caso, visto que *Big Data* – enquanto técnica embrionária na investigação criminal – se encontra em fase de desenvolvimento, não existem evidências científicas, nem dados científicos e, por isso, a maior parte das expectativas são fruto do que é interpretado sobre o que a comunicação pública social ilustra acerca de *Big Data*.

Todo o futuro é previsto e, assim sendo, surgem visões pessimistas, positivas e esperanças face aos avanços futuros, bem como, medos e imaginações tenebrosas 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 decepção, a incapacidade de corresponder às expectativas e cumprir as promessas, bem como, a possibilidade das esperanças não se realizarem. Considerando que 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 englobam conceitos como *pode, deve, poderia, espera, planeia, antecipa, acredita, estima, prevê, pretende, potencial, continuar* (ou a negação dos mesmos) – estas declarações são expectativas. Por vezes, os profissionais expressam uma notável ambivalência sobre o potencial impacto do uso dos grandes dados, 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 pode ser prejudicial e os que referem ser inútil; os que demonstram esperança e antecipações positivas; os que ficam entusiasmados com a perspectivas que os avanços tecnológicos lançariam sob a prevenção e repressão criminal; e os que demonstram desapontamento por essas novas tecnologias maravilhosas e pelo

desconforto com a propaganda especulativa que surge em torno deles. Há profissionais que prevêem um futuro complexo em que a contribuição útil de *Big Data* surge: baixas expectativas do futuro. 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). Sob a alçada desta perspetiva, o objetivo é identificar e analisar a robustez das expectativas expressas no que concerne ao *Big Data*.

No que concerne às diferentes perspetivas e expectativas sociais que surgiram, face ao tema e tendo em conta estudos realizados que se debruçam sobre as perspetivas de diferentes grupos sociais relevantes em relação às tecnologias, enquanto ferramentas para combate e prevenção da criminalidade, nomeadamente Williams & Johnson (2004) e Machado & Silva (2015), é possível tecer algumas considerações. Há profissionais que, quanto a estes temas, salientam 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; surgem com uma perspetiva “crítica”, pessimista e defensora da proteção de dados e privacidade, ressaltando a necessidade da proteção de 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). Este posicionamento aproxima-se da tipologia designada, por Williams & Johnson (2004) como *exceccionalismo genético*. Nestas perspetivas são feitas, frequentemente, avaliações processuais da técnica, ou seja, explorada a forma como os dados são recolhidos, analisados, processados, agregados e, posteriormente, para que fins são utilizados (Albert et al., 2009) – expectativas a um nível micro (com elementos repletos de detalhes caracterizadores da técnica) (Van Lente, 2012). Estes profissionais fazem um exercício de reflexão dual: ponderam os benefícios deste desenvolvimento tecnológico, mas, em simultâneo, operam numa 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, nomeiam a necessidade de regular, legalmente, esta atividade. Estes recortes discursivos salientam um aspeto: a necessidade de pensar criticamente os ideais inovadores do futuro; ou seja, por via de um paradigma preditivo analisam a melhor forma de proceder. Frequentemente, esta perspetiva é defendida por profissionais que não exercem funções laborais no laboratório e, por isso, criticam a técnica (não por motivos técnicos), ancorados a considerações económicas, políticas e/ou de direitos (Brown & Michael, 2003; Brown, N., Rip, A., & Van Lente, H., 2003). Esta constatação está em sintonia com o que Van Lente (2012) refletira: os profissionais com formação académica superior em Direito ressalvam a necessidade de desenvolver instrumentos jurídicos que sustentem e apoiem o desenvolvimento das tecnologias inovadoras. Por sua vez, o contexto onde os profissionais desempenham funções laborais influencia o seu discurso acerca do fenómeno (Van Lente, 2012; Machado & Silva, 2015). Estes profissionais, mais instruídos, do ponto de vista da legislação e da regulamentação revelam uma perspetiva mais cética sobre a (potencial) aplicação da tecnologia, revelando serem mais conscientes dos riscos advindos da futura implementação da técnica (Machado & Silva, 2015; Gardner et al., 2015). Consequentemente, revelam-se mais resistentes à implementação da técnica, nomeando e referindo, frequentemente, as questões éticas e de direitos humanos já enunciadas também pela literatura, como a privacidade e a proteção de dados (Machado & Silva, 2016).

Os profissionais ligados à vertente prática da aplicação das técnicas surgem com uma orientação utilitarista, próxima do *pragmatismo biométrico* (Williams & Johnson, 2004), com 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). É notório que as expectativas destes profissionais se prendem a um nível micro porque envolvem elementos detalhados e característicos da técnica (Van Lente, 2012). Ou seja, os profissionais que desempenham funções laborais em setores onde a técnica poderá futuramente constituir-se enquanto meio de investigação criminal – nomeadamente, Departamentos Policiais – estão mais conscientes da (potencial) existência e utilidade desta técnica, quando comparados a 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 mais pessimista. Esta perspectiva reside no limbo entre ponderar os benefícios da (potencial) aplicação da técnica e a incerteza ancorada nos riscos que a mesma (potencial) aplicação pode acarretar. Brown & Michael (2003) referem que esta incerteza é mais acutilante e verifica-se eminentemente presente ancorada sob as expectativas dos profissionais intimamente envolvidos na produção do conhecimento, onde a sua experiência acerca das contingências de produção de conhecimento no laboratório, os tornam mais cautelosos (Brown et al., 2003). Estas expectativas residem num nível mais macro porque se tratam de esboços abstratos e abrangentes (Van Lente, 2012). No fundo, os profissionais que desempenham funções próximas do contexto onde a técnica tem possibilidade de se vir a materializar são mais conscientes dos riscos que a mesma acarreta para a sua prática laboral, quando comparados com outros grupos profissionais, à semelhança da conclusão a que o estudo de Machado & Silva (2015) chegara quando se debruçaram sobre as perspetivas públicas sobre os riscos e benefícios dos bancos de dados forenses de ADN tendo em conta a influência do grupo profissional, nível educacional e idade. Estas visões envolvem-se, de forma dinâmica, entre as preocupações sobre o auto-cuidado e o bem-estar coletivo da sociedade, em simultâneo com a perceção de riscos e benefícios das tecnologias de vigilância que permeiam a vida quotidiana. Estas são normatividades e subjetividades abordadas no contexto das aplicações da tecnologia genética no campo da Medicina e também reproduzidos em outros contextos sociotécnicos como o forense, tal como o estudo de Machado & Silva (2016) acerca das perceções públicas sobre a doação voluntária de ADN para a base de dados portuguesa.

Estes últimos profissionais que desempenham funções próximas do contexto onde técnica tem possibilidade de se vir a materializar – geneticistas forenses - encaixam na cultura forense designada por Cole (2013): i) a questão do tempo (quando estes projetam *Big Data* nas suas investigações criminais estão conscientes que um dos seus benefícios é, precisamente, a capacidade de processar um elevado número de dados distintos num curto intervalo de tempo, tendo em conta que a sua atividade laboral é limitada no tempo, a capacidade de *Big Data* para processar os dados a uma velocidade feroz é benéfica); ii) os dados que possuem são limitados (e, por isso, uma técnica que permita aumentar o número de dados informativos é desejável, por conseguinte pode acarretar riscos à veracidade da investigação produzida tendo em conta a forma como estes dados são recolhidos e processados); iii) os dados que possuem, além de limitados, muitos deles são degradados (assim, uma técnica que armazene um número crescente de dados

informativos que possam complementar-se aos existentes, aumentará as probabilidades de um trabalho eficaz); iv) os dados não têm como objetivo a replicação; v) o objetivo (continua a ser) é apresentar Relatórios Forenses; vi) a audiência é o Estado (caso *Big Data* se instale enquanto técnica de investigação criminal, os resultados produzidos servirão fins de prova em Tribunal); e vii) a prova forense é vista como uma ajuda à resolução do caso, tal como as técnicas de *Big Data* (que surgem como complemento às primeiras). No fundo, são considerados sob esta chuva de ideias precisamente porque todas as disciplinas forenses possuem o suficiente para serem discutidas sob os auspícios do termo *cultura forense* (Cole, 2013).

Stevens, Wehrens & Bont (2018) apresentam uma tipologia discursiva, a propósito do seu estudo sobre os discursos emergentes acerca de *Big Data* no campo da saúde. O discurso modernista – onde reside a concetualização de *Big Data*, a sua definição enquanto grande quantidade de dados. Nesta tipologia inserem-se os discursos que referem a positividade da técnica, bem como, os benefícios dos seus efeitos. Além disso, também se incluem aqui as visões que afirmam que a técnica conduzirá a ações proactivas, preditivas, preventivas e participativas na saúde centrada no paciente. Ou seja, aqui incluem-se as visões otimistas que percecionam *Big Data* como uma técnica inovadora, apesar de reconhecerem a sua não utilização no campo da saúde. Este discurso modernista esforça-se para adotar uma postura radical: mudar as formas tradicionais de produção de conhecimento. Por sua vez, o discurso instrumental preconiza que os dados são entendidos em termos de uma gama de técnicas, como reconhecimento de padrões, mineração de dados e aprendizagem automática. Neste discurso os profissionais descrevem as formas pelas quais as técnicas de *Big Data* podem auxiliar na saúde: prever resultados e aumentar a compreensão das causas. Neste discurso é demonstrado um alto valor de experimentação. Estes últimos dois discursos – modernistas e instrumentalistas – negligenciam as implicações mais amplas e potenciais efeitos sociais do uso de *Big Data*. O terceiro tipo de discurso – pragmático – caracteriza-se pela nomeação da utilidade instrumental de *Big Data* para a resolução de problemas e tomada de decisões. É um fenómeno já presente na sociedade e visa efetivar-se, cada vez mais. Este discurso centra-se, principalmente, na forma como *Big Data* deve ser implementado e descreve os passos para a sua implementação bem-sucedida. Neste discurso não se referem objeções contra o uso de *Big Data*, este é visto como benéfico sempre que auxilia a aquisição de conhecimento sobre certo tema. Semelhante ao discurso instrumentalista, o pragmático prevê uma mudança no caminho do conhecimento, visto que agora as decisões são

tomadas tendo por base *Big Data*, oferecendo mais conhecimento do que o disponível atualmente, sendo uma fonte valiosa para a tomada de decisões. O quarto discurso – cientista – refere-se a *Big Data* como uma nova tendência que inclui a recolha e análise de dados, de forma menos rigorosa que as metodologias científicas. Refere a utilidade da técnica, no entanto apenas a título exploratório: pode ser útil em algumas situações devido ao seu potencial para identificar pesquisas valiosas. Neste discurso inserem-se as críticas à técnica, bem como, é enfatizada a sua potencial perigosidade e referida a necessidade de se conjugar com outras técnicas, considerando-se as questões procedimentais éticas que a própria levanta. A este respeito, estes discursos nomeiam a necessidade de se adotarem procedimentos adequados ou corretos para a recolha e análise dos dados. Apesar das críticas, esta posição epistemológica é semelhante aos discursos modernistas e instrumentalistas na medida em que também enfatiza a ideia de que a verdade pode ser encontrada nos dados, no entanto, o discurso cientista argumenta que os dados apesar de serem informativos nunca capturam todo um domínio e as hipóteses e a teoria nunca se esgotam, presumindo-se que as evidências devam ser desenvolvidas aplicando o método científico. Por último, o discurso crítico-interpretativo, refere que a técnica apresenta os dados de forma simplista, sendo que os dados são uma representação simplista da realidade. Este discurso caracteriza-se pelo tom preocupado, pelas críticas e pela ênfase dada à incapacidade da técnica em captar adequadamente a diversidade real. Assim sendo, este discurso caracteriza-se por visões que consideram perigoso a tomada de decisões a partir de técnicas de *Big Data*, visto que estas geram conhecimento limitado e os dados devem ser manuseados com cuidado. Assim, este discurso adverte para as limitações da técnica e refere os erros que a mesma produz que podem levar a resultados falsos, prejudicando os cidadãos (Stevens et al., 2018).

Ancorados sob estas lentes analíticas, os dados empíricos serão analisados e devidamente enquadrados.

1.8. Linhas objetivas orientadoras: Objetivos, questões de partida e hipótese de investigação

As questões de partida que orientam a presente investigação são as seguintes:

- i) Que expectativas em torno das potenciais aplicações de *Big Data* constroem os diferentes profissionais?

- ii) Quais as suas percepções de potenciais benefícios e riscos?
- iii) Que expectativas são coletivamente partilhadas entre diferentes categorias profissionais e que expectativas são específicas de diferentes olhares profissionais?

A hipótese de trabalho que orienta este estudo é a seguinte: Tratando-se de uma técnica ainda em fase emergente de aplicação na área da investigação criminal, prevalecem, expectativas comumente partilhadas entre diferentes grupos profissionais; e, em simultâneo, expectativas diferenciadas consoantes o posicionamento profissional.

Parte II

Estudo empírico

Parte II: Estudo empírico

2.1. Aspetos teórico-metodológicos: opções, bússolas e orientações

Considerando a escassez de estudos e investigações direcionados para o *Big Data*, enquanto potencial ferramenta e meio de investigação criminal, esta investigação assenta num estudo exploratório qualitativo que tem como principal objetivo *aceder, compreender e analisar as perspetivas e expectativas sociais dos órgãos de cooperação transnacional, 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*. Desta forma, 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 *empíria* (Charmaz, 2009). Sendo assim, utilizando a *Grounded Theory* (Strauss & Corbin, 1990; Clarke, 2005) como âncora epistemo-metodológica – a construção de teorias a partir dos dados (Charmaz, 2009) -, nesta investigação pretendeu-se a realização de paralelismos constantes, articulações, ligações e comparações entre a recolha e análise, entre a teoria e a realidade empírica (Strauss & Corbin, 1990).

Além disso, dada a fragmentação do conhecimento existente acerca do fenómeno sob investigação, a abordagem indutiva afigura-se a mais apropriada por permitir que os dados nos informem acerca da Teoria (Elo & Kyngäs, 2008). Desta forma, o fenómeno foi estudado num contexto específico e adequado ao seu uso (Cho & Lee, 2014) – sob alçada do Projeto EXCHANGE, que explora as dimensões sociais, culturais, éticas, regulatórias e políticas do uso de tecnologias de ADN para uso forense na UE. De forma a apreender a pluridimensionalidade que caracteriza este objeto de estudo e a privilegiar as perspetivas socialmente construídas pelos diferentes atores sociais que se relacionam profissionalmente com a temática empírica, este projeto ancora-se numa metodologia de carácter eminentemente qualitativo. 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, seguindo alguns pressupostos da *grounded theory* (Strauss e Corbin, 1990), que viabilizam a construção de teorias “fundamentadas” nos próprios dados com elevado valor heurístico (Charmaz, 2009), esta investigação pauta-se por uma articulação e comparação sistemática e paralela entre recolha e análise, empiria e teoria (Strauss e Corbin, 1990). Este estudo valoriza, assim, a representatividade sociológica de cada caso, sendo que estes se afiguram como generalizáveis a enunciações teóricas e não a populações ou universos (Yin, 1994: 10; Nunes, 1992: 247–248; Brandão, 2010: 45–46).

A pesquisa do Projeto utiliza uma abordagem multi-metodológica de caráter eminentemente qualitativo de forma a apreender a pluridimensionalidade dos seus objetos de estudo e a privilegiar as perspetivas socialmente construídas pelos vários atores sociais, que inclui: i) recolha e análise de documentação em várias fontes relevantes, desde legislação e documentação oficial, até artigos científicos, correspondência aberta e *websites* relacionados com o regime de Prüm; ii) realização de entrevistas; iii) observação em encontros científicos da genética forense e outros de natureza interdisciplinar e outros contextos situacionais relevantes (por exemplo, laboratórios e outros locais de trabalho de entrevistados ligados a partilha transnacional de informação no âmbito de cooperação policial e judiciária); e iv) recolha e análise de casos criminais que envolvem o uso de tecnologias de ADN. Na presente investigação utilizou-se o material recolhido da realização de entrevistas e este estudo valoriza, assim, a representatividade sociológica de cada caso, sendo que estes se afiguram como generalizáveis a enunciações teóricas e não a populações ou universos (Yin, 1994: 10; Nunes, 1992: 247–248; Brandão, 2010: 45–46).

Este Projeto não repousa, exclusivamente, sobre a temática do *Big Data*, explorando antes, de forma geral, o intercâmbio transnacional de perfis de ADN para uso forense e policial, como uma ferramenta importante para combater o crime organizado e transfronteiriço. No entanto, um dos objetivos do EXCHANGE reside na compreensão dos imaginários culturais e narrativas dos profissionais envolvidos na operacionalização do Sistema de Prüm em relação aos desenvolvimentos tecnológicos, onde o *Big Data* se insere, com a existência das bases de dados. O guião das entrevistas abrangeu os seguintes

temas, de forma geral: visões e experiências com a implementação de Prüm ao nível nacional e europeu, opiniões sobre os propósitos e contribuições do sistema Prüm, questões éticas levantadas pela troca transnacional de dados de ADN, expectativas de desenvolvimento e inovação ao nível das tecnologias de ADN e percepções relacionadas a comunicação com o público em geral.

2.1.1. Consideração da amostra: entrevistas

Os eixos empíricos, enquanto veículos facilitadores do processo de recolha de dados, foram 124 entrevistas com 140 entrevistados baseados em 25 países diferentes: 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 União Europeia, 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 com conhecimentos relevantes no que concerne ao Big Data e à partilha transnacional de dados de ADN. A inclusão destes entrevistados não foi feita aleatoriamente, sendo que foram selecionados de forma intencionalizada, de acordo com a informação que cada um dos profissionais poderia acrescentar sobre o fenómeno em estudo, conjugada com a amostragem em bola de neve (Flick, 2002; Yin, 2011: 88-89). No fundo, foram selecionados participantes com diferentes vivências e experiências acerca do fenómeno, de modo a possibilitar a exploração das múltiplas dimensões sociais do processo em estudo (Starks & Brown Trinidad, 2007). Estes entrevistados são perspetivados enquanto atores sociais, membros de comunidades científicas e, por isso, inseridos sob uma teia de significados específicos pertencentes a essa mesma comunidade, ainda que porosa e permeável ao exterior (Albert, Laberge & Hodges, 2009).

Desta forma, os participantes-entrevistados constituem, então, um corpo profissional intimamente conexo aos contextos específicos laborais detentores de

informação acerca do tema em apreço. O principal objetivo subjacente às entrevistas foi o de captar narrativas imbuídas de significados e sentidos ilustrativos da diversidade de perspectivas relativas às dimensões sociais, culturais, éticas, regulatórias e políticas do uso de tecnologias de ADN para uso forense na UE. No que concerne aos métodos de amostragem, utilizou-se a amostragem intencional, conjugada com amostragem em bola de neve (Flick, 2002; Yin 2011: 88-89). Esta amostragem foi realizada em três níveis: 1) consulta da lista oficial de NCPs, através da consulta da lista pública providenciada nos documentos da *Working Party on Information Exchange and Data Protection* (Conselho da UE, 2015, 2017) e contactos com informantes privilegiados; 2) seleção de profissionais da genética forense na UE, selecionados por consulta de listas de membros de organizações profissionais, como por exemplo EUROFORGEN⁶ e VISAGE⁷, ambas redes de investigação composta por geneticistas forenses e outros profissionais ligados à ciência forense financiadas pela Comissão Europeia. A seleção de profissionais da genética forense na Europa contou também com a pesquisa de autores de artigos científicos na área e pesquisa de profissionais com participação em eventos de interesse, como por exemplo conferências da *International Society for Forensic Genetics (ISFG)*⁸; e 3) seleção de outros *stakeholders* dos 5 países de estudos de caso (Alemanha, Holanda, Polónia, Portugal e Reino Unido) por via da consulta de listas de membros de organizações profissionais e autores de artigos científicos na área. No caso destes últimos (*stakeholders*) o processo de amostragem foi complementado pela amostragem em bola de neve.

Concretamente, foram considerados Pontos de Contacto Nacionais – NCPs, no *step* 1 e 2 – profissionais designados por cada Estado-Membro, regidos pelas normas aplicáveis do Direito nacional aplicável (Decisão 2008/615/JHA), com a função de fornecer dados, segundo o Regulamento da UE do Sistema de Prüm. O relatório mais recente sobre o progresso da implementação do sistema Prüm, datado de novembro de 2018, indica que há 24 Estados-Membros da UE operacionais ao nível de troca de ADN. No âmbito do presente estudo, foram conduzidas 37 entrevistas com 47 profissionais em 22 países operacionais ao nível da troca de ADN no sistema Prüm. Estes profissionais forenses, que

⁶ <https://www.eurofor-gen.eu>

⁷ <http://www.visage-h2020.eu>

⁸ <https://www.isfg.org>

desenvolvem funções como NCPs, são figuras centrais no regime de Prüm, visto que protagonizam as atividades diárias que permitem o intercâmbio transnacional e são posições cruciais nos processos de tomada de decisão. Em particular, as pessoas que atuam como NCPs têm como responsabilidades organizar e implementar os procedimentos e conexões necessárias para realizar trocas automáticas de dados com outras bases de dados (enviando e recebendo informações); realizar testes com outros países-parceiros; e gerir e reportar correspondências de ADN (*hits/matches*). Estes papéis e as consequentes responsabilidades dos NCPs de Prüm podem variar entre os países de acordo com diferentes estruturas organizacionais e legislação nacional (Machado & Granja, 2018). Para além disso, os diferentes países atribuem a custódia das bases de dados nacionais de ADN a diferentes entidades, desde autoridades judiciais até forças policiais (Santos, 2017 *cit in* Machado & Granja, 2018). Portanto, os indivíduos que operam como NCPs 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 Contato 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 países duas ou mais pessoas podem estar envolvidas nessas tarefas. Nesta última situação, sempre que diferentes indivíduos demonstraram disponibilidade para participar do estudo, foram entrevistados. O guião das entrevistas abrangeu os seguintes temas: visões e experiências com a implementação de Prüm ao nível nacional e europeu, opiniões sobre os propósitos e contribuições do sistema Prüm, questões éticas levantadas pela troca transnacional de dados de ADN, expectativas de desenvolvimento inovação ao nível das tecnologias de ADN e perceções relacionadas a comunicação com o público em geral.

Para além das entrevistas conduzidas com NCPs foram, também, realizadas entrevistas a profissionais da genética forense na UE. Tendo em consideração a diversidade da comunidade de genética forense (Cole, 2013), a amostra agrega geneticistas forenses, isto é, indivíduos que trabalham em casos criminais e são empregados por um laboratório forense; e investigadores, ou seja, indivíduos empregados por universidades ou institutos forenses cuja ocupação profissional principal é 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. O guião das entrevistas abordou os seguintes temas: a organização de serviços de genética forense no país em que o/a participante esteve baseado; pontos de vista sobre a troca transnacional de dados de ADN na UE; representações sobre as perceções públicas sobre genética forense; e opiniões e experiências relativas a desenvolvimentos e inovações ao nível das tecnologias de ADN, tais como, inferência fenotípica, sequenciamento paralelo massivo e pesquisa familiar.

Por fim, foram ainda entrevistados outros *stakeholders* relevantes no âmbito de Prüm nos cinco países alvo de estudo de caso: Alemanha, Países Baixos, Polónia, Portugal e Reino Unido. 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 também 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 uso forense na UE. Os critérios que conduziram à escolha destes cinco países foram os seguintes: i) estar operacional no âmbito do sistema Prüm ou não; data de entrada na UE e situação relativa ao espaço *Schengen*; ii) legislação nacional relacionada ao 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. O guião das entrevistas abrangeu os seguintes temas: perspetivas e desafios colocados pela implementação do sistema; opiniões sobre os riscos e benefícios do sistema Prüm; e expectativas de desenvolvimento e inovação ao nível das tecnologias de ADN. Uma vez que este estudo é guiado pela *Grounded Theory* (Strauss & Corbin, 1990; Clarke, 2005), novos tópicos e conceitos abordados entrevistas foram sendo incorporados, sempre que pertinente.

As entrevistas realizadas e, posteriormente, transcritas ancoraram-se em guiões semiestruturados que procuraram captar as perspetivas edificadas e as significações enunciadas pelos entrevistados acerca do tema em apreço. Assim, foram consideradas

para esta investigação 124 entrevistas realizadas, com 140 entrevistados baseados em 25 países diferentes⁹, tendo sido, dentro destas, analisados os retratos discursivos e narrativas dos entrevistados em relação aos desenvolvimentos tecnológicos na investigação criminal, onde o *Big Data* se insere. Destes 140 entrevistados, 48 foram NCPs, com formação educacional diferenciada, a exercer funções laborais tanto em laboratórios forenses como em forças policiais¹⁰. No que concerne aos geneticistas forenses entrevistados, o número total de entrevistados foi 48¹¹. Por fim, no que diz respeito aos *stakeholders* dos estudos de caso, foram considerados 44 participantes-entrevistados¹².

A conceção do guião de entrevista foi elaborada pela coordenadora do Projeto EXCHANGE em estreita colaboração com membros doutorados da equipa. Em simultâneo, a coordenadora do Projeto 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 que a sua transcrição foi realizada por profissionais externos. Desta forma, o meu papel pessoal e individual cinjiu-se à análise de conteúdo das entrevistas sob a perspectiva de representações sociais sobre *Big Data* com o objetivo de realizar esta investigação. Todas as entrevistas foram conduzidas em língua dominada pelo entrevistado/a e entrevistado/a, tendo sido maioritariamente conduzidas em inglês (87) e as restantes em português (28), espanhol (3) e alemão (6). As entrevistas em espanhol e alemão foram traduzidas, por profissionais externos, para português. Para efeitos de redação da dissertação, todos os extratos de entrevista em inglês, espanhol e alemão considerados para esta investigação foram, por mim, traduzidos para português. Estas traduções procuraram respeitar integralmente o conteúdo e sentido expressos pelo/a entrevistado/a.

O tema do *Big Data* foi abordado nas entrevistas, perguntando 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 *Big Data*. As respostas transcritas a estas questões foram analisadas, minuciosamente, e constituem o objeto de estudo – o

⁹ Ver anexo I.

¹⁰ Ver anexo II.

¹¹ Ver anexo III.

¹² Ver anexo IV.

fragmento empírico que possui os dados para formar a Teoria.

O principal critério que orientou a recolha de extratos empíricos foi a palavra-chave *Big Data*. O primeiro passo, da análise das transcrições das entrevistas, criou condições para a constatação da existência de categorias de análise e conceitos que emergiam dos dados empíricos (Charmaz, 2009).

2.1.2. *Análise de conteúdo qualitativo*

Sob uma perspetiva qualitativa, realizou-se a análise de conteúdo de entrevistas: esta metodologia repousa sobre a materialidade linguística das condições empíricas das entrevistas, permitindo a criação de categorias rigorosas resultantes da interpretação dos dados qualitativos (Caregnato & Mutti, 2006; Cho & Lee, 2014). Este mecanismo de análise visa aceder e compreender o pensamento dos sujeitos por via do conteúdo expresso nas entrevistas. Tendo em conta que os discursos emergem das palavras dos inquiridos e as significações ancoradas aos mesmos, dependem do contexto de onde surgem, procedeu-se, desde logo, ao enquadramento, categorização e diferenciação do contexto formativo e profissional dos entrevistados (Caregnato & Mutti, 2006). Este *background* obteve-se por via de questões abertas materializadas no contexto das entrevistas semi-estruturadas, enquanto conversas intencionais que visam obter informações sobre um determinado tema, através do diálogo. As questões abertas, apesar de se inserirem no âmbito de um guião de entrevista pré-definido, caracterizam-se por potenciarem espaço para o desenvolvimento de questões e temas importantes, criando condições para o entrevistado dialogar (Bogdan & Biklen, 1994). Neste âmbito, foram evocadas perguntas que visaram aceder ao percurso, contexto e histórico académico e laboral dos entrevistados. Esta informação permitiu o acesso aos discursos e às respostas dos profissionais, perspetivando-os como uma construção do mundo social feita pelos mesmos, como algo socialmente construído, sob uma perspetiva construcionista (Caregnato & Mutti, 2006).

Após uma delimitação do eixo temático – *Big Data* – foram estabelecidos *recortes discursivos* que caracterizam as regularidades do discurso apresentado pelos entrevistados, sempre confrontados com os sentidos heterógeneos 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. Numa fase preliminar, as entrevistas foram analisadas e organizadas por via de várias leituras, estabelecimento de objetivos e elaboração de indicadores. Posteriormente, as *unidades de registo* – fragmentos discursivos recortados das entrevistas mediante o interesse científico para a investigação – permitiram a codificação dos dados. Por fim, procedeu-se à categorização dos dados, à classificação dos elementos discursivos de acordo com as suas semelhanças e diferenciações, com reagrupamento e redefinição constante, *a posteriori*.

2.1.3. Codificação e categorização

No que concerne ao processo analítico e de codificação e categorização dos dados, foi adotado um procedimento de análise comparativo permanente entre *empíria* e teoria. Assim, este processo materializou-se num movimento interativo que envolveu, em simultâneo, a recolha dos dados, a sua codificação e a construção de um texto teórico (Strauss & Corbin, 1990; Charmaz, 2009; Cho & Lee, 2014). Ou seja, as categorias de análise foram induzidas a partir dos conceitos que emergiam das entrevistas (Strauss & Corbin, 1990; Charmaz, 2009). Consequentemente, a literatura teórica e os dados empíricos foram considerados em simultâneo, permitindo uma conjugação dos mesmos, o que possibilitou o seu permanente contacto e partilha de conhecimento – como prevê a *Grounded Theory* (Cho & Lee, 2014; Strauss & Corbin, 1990).

Concretamente, numa primeira fase do processo de análise dos dados, debruçada sobre os *recortes discursivos* com interesse científico para a investigação em apreço, os conceitos enunciados em cada resposta foram sendo alvo de comparações – codificação aberta (Strauss & Corbin, 1990). Numa segunda fase, os conceitos que apresentaram similitudes foram alvo de uma codificação, tendo sido sintetizados e agrupados por categorias concetuais (Cho & Lee, 2014) – codificação axial (Strauss & Corbin, 1990) – constituindo, por fim, o núcleo de temáticas gerais criadas por via da análise dos dados. Este processo findou com uma decomposição compreensiva e interpretativa dos discursos e narrativas, tendo-se optado por delimitar os recortes discursivos mais elucidativos das expectativas e representações sociais presentes nas narrativas. Tendo em conta que o objetivo é aceder, precisamente, às *expectativas e representações sociais* atribuídas pelos entrevistados ao tema em apreço, esta metodologia qualitativa, ancorada sob a lente da *Grounded Theory* (Clarke, 2005), permite a consideração dos entrevistados enquanto

atores sociais que atribuem significados que devem ser interpretados e compreendidos à luz dos seus processos sociais e das suas interações (Cho & Lee, 2014), à semelhança do que postulava o interacionismo simbólico – cada fenómeno resulta da interação e das significações das suas partes constituintes, que interessam estudar (Starks & Brown Trinidad, 2007). Todo este processo de análise de investigação pretendeu, a partir da precisão das categorias conceituais criadas, estabelecer a generalidade dos factos (Cho & Lee, 2014).

Assim, as diferentes categorias criadas, bem como, consequentemente, subcategorias, foram agrupadas e integradas em temáticas gerais: i) Informação sobre *Big Data*: conhecimento, avaliação, descrição técnico-profissional e exemplificação; ii) Expectativas futuras de potencial aplicação de *Big Data*: desenvolvimento, expansão e antecipação de aplicação na investigação criminal (transfronteiriça e organizada); iii) Perceção sobre riscos de *Big Data*; e iv) Pareceres éticos e de direitos humanos.

Na primeira temática, concretamente, o desenvolvido, sob forma de categorias, pelos entrevistados prende-se, essencialmente, com a caracterização da técnica, bem como, a sua definição. Por sua vez, os entrevistados também referem exemplos que materializam a (potencial) aplicação do uso da técnica, nomeando a EUROPOL e os Circuitos de Videovigilância. Foram tecidas, também, algumas considerações políticas sobre o fenómeno de *Big Data*. Na segunda temática, o conteúdo restringe-se a perspetivas e expectativas sócio-profissionais que se prendem com a necessidade de *Big Data* se adaptar aos bancos de dados, tendo sido enunciadas as suas vantagens, bem como os seus benefícios, considerações acerca da sua expansão e da forma como se tem vindo a desenvolver, enquanto técnica, bem como, o seu carácter complexo, no entanto, potencialmente útil, no âmbito de investigações criminais. Na terceira temática, as narrativas tecidas concernem às questões procedimentais da técnica: o acesso, procedimentos, análise, resultados e sua posterior extrapolação. Por fim, o último tema reserva-se para perceções éticas e de direitos humanos que levantam questões ligadas aos direitos, liberdades e garantias dos cidadãos, como a proteção de dados e a privacidade, que em simultâneo com o desenvolvimento de *Big Data*, vão sendo permeáveis a estas questões.

Se por um lado, os entrevistados de diferentes áreas de atuação profissional revelam pontos em comum – efeitos das expectativas de dimensão coletiva –, por outro lado há posicionamentos distintos, ainda que subtis. Neste sentido, urge refletir acerca das articulações de distância e proximidade entre diferentes comunidades de

conhecimento – *culturas epistémicas* de diferentes categorias profissionais, como diferentes formas de pensar consoante os diferentes lugares e a diferenciada formação que estes profissionais detém (Cetina, 2009; Cole, 2013; Kruse, 2016).

Parte III

Resultados empíricos

Parte III: Resultados empíricos

A questão acerca de *Big Data* foi colocada aos entrevistados enquadrada sob o tema dos desenvolvimentos tecnológicos que têm vindo a operar nos meios de investigação criminal e que potenciais riscos e benefícios estes ancoram na prática. Os participantes foram questionados acerca da sua opinião sobre tal, sendo que em primeiro lugar procurou-se perceber se os entrevistados detinham algum tipo de conhecimento sobre a temática. Posteriormente, foram exploradas questões relativamente às suas expectativas e perspetivas no que toca aos benefícios e desvantagens da sua prática, tendo, também, como objetivo entender se já tinham manuseado a técnica. Ou seja, para entender a forma como os entrevistados-participantes concebiam *Big Data*, procurou-se explorar as suas representações sociais acerca da capacidade e valor de *Big Data* para o seu trabalho. Em determinadas entrevistas, as respostas foram consideradas irrelevantes para serem incluídas nos presentes resultados empíricos, visto que alguns dos entrevistados, além de não possuírem conhecimentos sobre *Big Data*, não se alongaram nas considerações e, por isso, não foram consideradas. Não obstante, alguns destes profissionais, que revelaram não possuir conhecimentos significativos acerca dos potenciais riscos e benefícios do uso da tecnologia e dos grandes dados para fins de investigação criminal, foram considerados por demonstrarem uma considerável representatividade acerca do desconhecimento do fenómeno. Foram criadas categorias analíticas para os diferentes profissionais considerados, no entanto, estas foram englobadas e inseridas em temáticas gerais comuns. No fundo, os diferentes profissionais surgem com diferentes posicionamentos, mas não com fronteiras estanques, logo não se afigurou necessária a divisão esvaziada por grupo profissional, por se registarem diferenças híbridas e não posicionamentos completamente distintos.

Apesar de *Big Data* não se ter emancipado, ainda, enquanto técnica válida cientificamente, isto é, considerada sob os olhos da Ciência, como suficientemente robusta, capaz, eficaz e operacional para que seja operada a sua aplicação global enquanto prática e meio de investigação criminal, esta surge nos discursos dos vários entrevistados como uma ferramenta promissora e uma tecnologia futuramente positiva. Ainda assim, após uma revisão sistemática das narrativas dos diferentes profissionais, é possível perceber-se que são extremamente evidentes as controvérsias, incertezas e dilemas que emergem, ainda, devido às diferentes perspetivas e expectativas sócio-técnicas, bem como, grau de incerteza, em torno do tema. Tratando-se de uma técnica em fase precoce

de implementação e difusão, os níveis de incerteza exacerbam-se, porque nos discursos projeta-se no futuro um paradigma em fase embrionária de aplicação e desenvolvimento. Esta abstração temporal, que as narrativas sobre *Big Data* requerem, geram fraquezas e exacerbam os riscos (Brown & Michael, 2003; Selin, 2008).

Concretamente, um número considerável de entrevistados, transversal a qualquer categoria profissional, não detém qualquer tipo de conhecimento acerca de *Big Data*. O desconhecimento ancora-se nos diferentes profissionais considerados e o reflexo desta maioria são as seguintes respostas:

“*Big data* não sei o que é.” [NCP1, N10]

“Mas eu não sei... O que é *Big Data*? O que quer dizer com *Big data*? Eu não tenho a certeza, eu não sei o que isso é...” [NCP1, W02]

“Que tipo de dados?” / “*Big data*, eu não... Dados em massa?” [NCP2, P01]

“Não [Eu nunca ouvi falar sobre *Big Data*].” / “(...) E não sei, eu posso tentar perguntar aos meus colegas que trabalham no departamento IT, talvez eles trabalhem com alguma coisa relacionado com isso [*Big Data*], eu não sei.” [NCP2, L03]

“Não sei o que é o *Big Data*.” [Professor/Investigador, N05]

Bartlett et al. (2018) e Halford & Savage (2017) referem, quanto a este desconhecimento, 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, este desconhecimento não é linear, consoante as diferentes posições laborais ocupadas pelos entrevistados. Nomeadamente, os Pontos De Contacto Nacionais a desempenhar funções no *step 1* (NCP1) em Institutos Forenses – Departamento de Análise Genética e Biológica, quando questionados, apesar de revelarem total desconhecimento sobre este aparato tecnológico, enfatizam a noção de que é algo que ultrapassa a sua margem de ação profissional, não demonstrando expectativas de (possível) aproximação laboral com a técnica, nem perspetivando potenciais possibilidades de desenvolvimento desta:

“Eu não estou familiarizado com isso [*Big Data*], e acho que não vou estar num futuro próximo.” [NCP1, T01]

Ou seja, os NCP1, revelam-se, a nível profissional e laboral, desadequados 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 *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 possuem um (des)conhecimento ténue sobre a temática que lhes permite, de forma vaga aceder ao conceito. Nomeadamente, explicitam as técnicas de *Big Data* e a forma como poderiam operar na prática investigativa, sendo que um dos NCP2 - com papel legal e estratégico na assistência mútua – a desempenhar funções laborais numa Divisão de Cooperação Policial Internacional, reconheceu o potencial deste aparato tecnológico, no entanto alegou 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 [de *Big Data*] porque não têm dinheiro, mas eu nem sei quanto é que isso [*Big Data*] custa. (...)” [NCP2, L03]

Ou seja, NCP2, e sob a lente concetual teórica de Van Lente (2012), são uma categoria profissional que procede a esboços detalhados sobre *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, com formação superior em Biologia Aplicada refere que, além do facto de se tratar de uma técnica possível e desejável no futuro, é uma prática desafiadora a vários níveis: complexidade, economia, qualificação e mudança nas práticas policiais; na medida em que é mais complexo de se manusear, colocar em prática

e implementar nos modos contemporâneos da investigação criminal. Não obstante, o entrevistado refere também que, ao implementar-se, criaria uma mudança nas práticas policiais, na medida em que exigiria que o sistema policial se adaptasse a esta nova realidade e isso 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 [de *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, D14]

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 acerca de expectativas futuras, o desejo aliado à perspetiva positiva de aplicação da tecnologia (Brown & Michael, 2003); por outro lado, ressaltam os obstáculos práticos à sua efetiva materialização, como a complexidade, os recursos económicos que exige e as mudanças que requer – um discurso cautelosamente esperançoso (Gardner et al., 2015). 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 este aspeto: a 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 de *Big Data*, também tiveram uma educação que não os preparou para a inovação. Esta demarcação simbólica que é feita, por parte dos NCP2, relativamente aos agentes policiais é evidente: NCP2 reconhece o potencial positivo e eficaz de *Big Data*, mas considera que as forças policiais não possuem o arsenal económico, nem formativo, bem como, abertura inovadora para implementar a técnica; atribuem à Polícia falta de conhecimento especializado e de investimento económico (Cole, 2013). Esta constatação assemelha-se às conclusões que o estudo empírico de Chan & Moses (2017)

alcançou: 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 de *Big Data*. Além disso, acrescentam o facto destes polícias terem necessidade de expandir o seu conhecimento acerca da Genética Forense. Existe esta tendência geral dos restantes profissionais se demarcarem das racionalidades policiais, caracterizando estas últimas como diferentes, resistentes à implementação de novas técnicas como *Big Data*, bem como, limitados a nível educacional para a inovação (Cole, 2013).

No entanto, ainda que NCP2 e geneticistas forenses desconheçam o fenómeno, enquanto realidade laboral prática, levantam perspetivas da sua aplicação futura, ao contrário do que acontecera com NCP1. NCP2 e geneticistas forenses revelam, sob a ótica de Van Lente (2012), declarações permissivas (com algumas resistências) acerca da aplicabilidade futura de *Big Data*. Por exemplo, o seguinte genético forense a desempenhar funções laborais numa Universidade, apesar de se distanciar do conhecimento da temática, considera 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 uso do sucesso rápido dos dados para fazer as coisas muito rápido num ambiente de imigração que é... eu não estou muito conectado com isso [*Big Data*], para ser honesto.”

[Geneticista Forense, D12]

Apesar do desconhecimento surgir como resposta nas diferentes camadas profissionais, existem diferenças ténues, não como fronteiras estanques, mas graduais: os NCP1 não se aproximam da temática, nem revelam perspetivas futuras de aproximação, nem conhecimento. Adotando a lente concetual de Van Lente (2012), NCP1 tecem esboços abrangentes e abstratos, a um nível *macro*, acerca de *Big Data*. Ao contrário dos NCP2 e geneticistas forenses que, apesar do conhecimento reduzido, completam as suas respostas demonstrando perspetivas de evolução e desenvolvimento da técnica *Big Data*, aproximando-se dos detalhes da técnica, projetando-os e destacando-se como defensores da sua implementação, adotando um discurso promissor quanto à temática (Van Lente, 2012).

3.1. Informação sobre *Big Data*: conhecimento, avaliação, descrição técnico-profissional e exemplificação

Vários são, também, os operadores do Sistema de Prüm, que revelam possuir conhecimentos acerca das questões técnicas e procedimentais por via das quais *Big Data* se materializa. Os vários setores laborais considerados apresentam, consensualmente, conhecimentos que se ligam à operabilidade do *Big Data*, apesar de determinados profissionais surgirem com respostas mais relacionadas com os apelos técnicos e outros apresentarem conhecimentos mais descritivos, todos tecem considerações técnico-descritivas sobre o fenómeno dos grandes dados. Estes discursos, transversais a qualquer categoria profissional, sob a ótica de Stevens et al. (2018), apelidam-se de discursos *modernistas* e caracterizam-se por incluir considerações discursivas acerca da concetualização do objeto de estudo – neste caso, *Big Data* -, a definição da temática, os seus benefícios e aspetos positivos. Assim sendo, nesta temática geral residem os discursos *modernistas* dos diferentes profissionais entrevistados considerados.

Um NCP1, a desempenhar funções profissionais num Instituto de Neurologia e Genética, referiu que crê tratar-se de uma técnica possibilitada pelo surgimento do sequenciamento em massa, visto que neste momento tudo é possível:

“Bem, quero dizer que isso [*Big Data*] poderia ser possível com... com o sequenciamento paralelo em massa agora tudo é possível. (...)” [NCP1, W02]

Ou seja, atualmente com o desenvolvimento tecnológico é possível que se criem condições para que surjam novas técnicas e novos métodos de investigação, mais sofisticados e complexos. Quanto a este 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, surgindo a análise em rede – *Big Data*.

A caracterização de *Big Data* como um método preventivo e/ou pró-ativo é evidente nas perspetivas dos NCP1. Alguns destes NCP1 a desempenharem atividade profissional num Instituto de Criminalística, em simultâneo a administrar a base de dados de ADN nacional, e num Instituto de Neurologia e Genética, bem como, um a desempenhar funções policiais (respetivamente), perspetivam a capacidade preventiva de

Big Data como um dos maiores benefícios do uso 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 prever ações e encontrar suspeitos:

“(...) se nós tivéssemos esses métodos [*Big Data*] poderíamos prever comportamentos [criminais] (...)” [NCP1, W02]

“(...) graças a esse sistema [*Big Data*] poderíamos, por exemplo, apreender os suspeitos ou perpetradores [de atos criminais].” [NCP1, X01]

“(...) isso [*Big Data*] deve ser preditivo e [deve ser visto como] uma ajuda [para a resolução e investigação de atos criminais] (...)” [NCP1, L01]

Sob uma perspetiva mais teórico-descritiva, sem enquadrar *Big Data* na sua prática laboral, 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 de *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 operar *Big Data* na prática, Chan & Moses (2017) referem que é precisamente desta valência pró-ativa que as agências policiais fazem uso: prevenção criminal por meio de previsões sobre lugares, intervalos de tempo e/ou pessoas; a partir destas criam-se perfis que combinam, com precisão, os potenciais infratores aos crimes específicos; auxiliando, assim, o encontro dos responsáveis pelos crimes. Esta é uma perspetiva profissional pertencente apenas aos NCP1 com formação superior em Engenharia Genética, a desempenhar funções laborais como Chefe de Laboratório de Genética Forense. Apesar dos NCP2 considerarem, também, esta valência preditiva e pró-ativa das ferramentas de *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 NCP1, apesar de possuírem determinados conhecimentos relacionados com a temática, as suas perspetivas são ténues e ainda que definam a técnica, de forma descritiva, como o seguinte entrevistado, consideram, maioritariamente, que se

trata de uma ferramenta de investigação analítica ou de inteligência útil que não encontra facilidades na sua aplicação:

“(...) isso [*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, de locais, de atividades e pessoas; sendo que se materializa 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 possui para a aplicar, à semelhança do que Chan & Moses (2017) referem: a maior parte dos órgãos policiais reconhecem o valor e a capacidade de *Big Data*, no entanto, consideram não ter tempo nem recursos para beneficiar da utilidade deste aparato tecnológico (Chan & Moses, 2017). Ou seja, NCP1 demarcam-se da temática, referindo amplos obstáculos à aplicação de *Big Data*, operando, sob a ótica de Van Lente (2012), à expressão sucessiva de argumentos organizados para refutar as previsões favoráveis à implementação da técnica *Big Data*.

Os NCP2, a desempenhar funções profissionais na Direção Geral da Polícia Criminal, responsáveis pelas ligações operacionais, revelam ser detentores de um conhecimento particular ligado a esta temática, distinguindo-se dos restantes grupos profissionais considerados, que se prende essencialmente com a descrição específica e exemplificação prática de *Big Data*. Estes profissionais, teoricamente, aproximando-se da tipologia de Van Lente (2012) possuem expectativas positivas e promissoras sobre a temática em apreço, quando questionados sobre tal, revelando conhecimentos detalhados sobre a mesma, aproximando-se da sua realidade expectável, tecendo declarações permissivas quanto à sua implementação. Além disso, a um nível *micro*, 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. Nomeadamente, um dos entrevistados NCP2 perspetiva *Big Data* como uma ferramenta benéfica, precisamente por ter capacidades e funcionalidades que a permitem manusear um grande número de dados para serem partilhados e conectados com outras informações, por conjugar diferentes amostras e

conseguir alcançar resultados práticos, úteis e eficazes – útil e benéfico para o contexto da investigação criminal:

“(...) a longo prazo eu vejo muitos benefícios se tivermos a possibilidade de conectar esses bancos de dados (...) isso [*Big Data*] pode ser muito útil às vezes (...)”
[NCP2, S02]

Na mesma linha de reflexão – heterogeneidade de informação e dados – Lyon (2014) referiu 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; auxiliando em determinadas funções policiais.

O mesmo entrevistado refere que, em termos práticos, *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*, refere que crê 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 da criminalidade. Ou seja, sob a perspetiva deste profissional, *Big Data* surge como mecanismo de troca de informação privilegiado no contexto europeu e utilizado, em termos práticos, ainda que esteja em evolução. Portanto, a EUROPOL, sob a perspetiva de NCP2 a desempenhar funções profissionais na Direção Geral da Polícia Criminal, responsáveis pelas ligações operacionais, legitima atualmente a importância de um modelo pró-ativo, 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, S02]

NCP2 perspetiva *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. No que concerne às atividades da EUROPOL e a sua estreita relação com o contexto de aplicação da lei, tal

como fora referido por NCP2, Drewer & 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, e retomando a perspetiva do entrevistado, alegam que a EUROPOL reconhece, numa escala mais ampla, em algumas áreas específicas, particularmente cibercrime e terrorismo, os grandes dados como vetores importantes visto que permitem obter um melhor perfil, que é usado por criminosos e polícias para identificar alvos potenciais. As capacidades da EUROPOL aumentaram e esta 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 (Drewer & Miladinova, 2017).

Esta perspetiva concecional reflete o que Hedgecoe & 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 de *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. NCP2 são uma categoria profissional que refere, frequentemente, a consciência da utilidade futura da técnica, precisamente porque, adotando a ótica de Machado & Silva (2015) desempenham funções laborais onde, futuramente, a técnica se poderá vir a aplicar. Estes profissionais, que nomeiam a vertente prática da técnica inserem-se na tipologia que Williams & Johnson (2004) criaram, mais precisamente, no *pragmatismo biométrico*: com um discurso fortemente marcado pelo entusiasmo, revelando perspetivas apologistas de inserção da técnica, com uma visão extremamente positiva. Este discurso é caracterizado, por Stevens et al. (2018) como um *discurso instrumental*, na medida em que os profissionais referem exemplificações práticas da técnica no seu campo laboral e utilizam conceitos discursivos intimamente relacionados com a temática em apreço. Este é um tipo de discurso frequentemente adotado pelos NCP2.

Além disso e inseridos sob alçada da descrição técnico-característica de *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 de *Big Data* enquanto

realidades algorítmicas que agregam dados do *Facebook*, biométricos, de ADN e de impressões digitais, sendo uma forma que facilita e possibilita a agregação de muitos e diferentes dados informativos:

“(...) [*Big Data*] têm 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, pode-se criar perfis corretos de indivíduos (...)” [NCP1, W02].

Na mesma linha reflexiva, no que respeita ao conhecimento técnico de *Big Data*, surgem geneticistas forenses a desempenhar funções laborais num Instituto Nacional de Polícia Científica, que referem tratar-se de um método que permite a tomada de decisões, tendo por base algoritmos. Estes algoritmos transformam as decisões e as posteriores ações em veículos mais precisos para atingir os resultados pretendidos. Assim, *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 *Big Data*] é benéfico, porque aumenta a precisão do algoritmo (...)” [Geneticista Forense, H03]

Ou seja, na perspetiva deste último entrevistado, trata-se de uma técnica que aumenta a precisão dos algoritmos, *estas bases que sustentam a ferramenta tecnológica*, como o próprio indica. É um método que sofisticada e aumenta a potencialidade das suas ferramentas, agregando variáveis raras, que de outra forma muito provavelmente não seria possível. Na aceção deste geneticista forense, *Big Data* trata-se de uma técnica que permite a partilha de dados entre os diferentes países europeus, sendo esta uma valência que permite a construção de um algoritmo completo; tornando o trabalho de equipa mais relevante e permitindo alcançar resultados mais completos e precisos. Mittelstadt et al. (2016) referem que esta é a perspetiva dominante acerca dos algoritmos: há uma preocupação consensual ancorada ao facto destes algoritmos facilitarem e permitirem a tomada de decisões, com grande potencial. Estes algoritmos substituem análises e decisões humanas.

Também os restantes profissionais, nomeadamente os professores/investigadores com formação superior especializada na Economia Política, surgem com esta associação entre *Big Data* e algoritmos, sendo a primeira uma ferramenta que permite, sob a sua

perspetiva, tomar decisões com base em algoritmos – esta metodologia ancora-se em dados passíveis de serem convertidos em algoritmos, que facilitam e permitem a tomada de decisões.

“A questão do *Big Data*, das decisões com base em algoritmos (...)”

[Professor/Investigador, N17]

Quanto a esta capacidade do *Big Data* se sustentar em dados e algoritmos para descobrir padrões e associações, Gonçalves (2017) explica que *Big Data* tem a capacidade de usar, 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. Esta possibilidade das decisões de justiça criminal se ancorarem aos resultados alcançados por via de *Big Data*, é abordada por Chan & Moses (2017) quando referem que o objetivo do uso e manuseio desta técnica é, precisamente, esse: 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 sob a panóplia da descrição técnica de *Big Data*, o mesmo professor/investigador com especialização académica em Economia Política, refere-se a esta como uma *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... sobretudo (...) 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, N17]

Ou seja, este entrevistado perspetiva *Big Data* como uma área científica metodológica que permite a recolha, tratamento e manuseamento de um grande volume de dados – de diferentes fontes – com o objetivo de correlacionar diferentes informações com vista a alcançar resultados úteis. O mesmo, utiliza uma linguagem intimamente

relacionada com a temática em apreço e, por isso, sob a ótica de Stevens et al. (2018) o seu discurso é *instrumental*.

Esta 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 últimos referem que o número elevado de dados que esta técnica conjuga potencia a complexidade do seu aparato, enquanto fenómeno. Desde logo, pela inovação tecnológica de que se trata o *Big Data*, grande parte dos profissionais percecionam-no 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 ressaltou sempre a sua preocupação pelo facto de englobar e trabalhar com um volume considerável de informação:

“(...) *Big Data* (...) são muitos dados (...)” / “(...) Mas, sim, claro, é grande...
Como é que se armazenam tantos dados de ADN (...)” [NCP2, S02]

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

“*Big data* é (...) usado, por exemplo, para video-informação (...)” [NCP1, O02]

Desta forma, este NCP1, estabeleceu uma ligação entre *Big Data* e novas tecnologias, restringindo a materialização desta ferramenta nos meios de comunicação audiovisuais. Esta forma de recolher informação por via de câmaras de vigilância é nomeada por vários autores e define-se como um tipo de vigilância pré-constitutiva, em que o processo de formação de informação se procede de forma aberta, perceptível, dirigida, indiscriminada e a qualquer pessoa (Matzner, 2016; Fróis, 2007; Fróis, 2015; Machado & et al., 2018). Lupton & Michael (2017) concluíram a mesma aceção 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. Esta

associação entre *Big Data* e *tecnologia*, bem como, a identificação desta técnica como uma ferramenta eminentemente tecnológica é frequentemente apresentada nas respostas dos NCP1, a desempenhar funções laborais policiais num Instituto Forense. As suas respostas evidenciam claramente isso: um dos entrevistados a desempenhar estas funções laborais, quando questionado sobre *Big Data*, direciona a sua resposta precisamente nesse sentido – os potenciais benefícios de manusear os dados por via desta técnica materializam-se por via da tecnologia, que é uma ferramenta que apresenta vantagens e desvantagens, mediante a finalidade do seu uso. Ou seja, é perceptível que percecionam *Big Data* como uma tecnologia:

“Isso [*Big Data*] depende de como é que os dados são usados. Porque qualquer tecnologia tem dois lados, um bom e um mau (...)” [NCP1, J01]

Esta perspetiva vai no sentido do que Matzner (2016) referiu: *Big Data* como 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 dos geneticistas forenses entrevistados também se referirem ao *Big Data* como uma tecnologia, fazem-no de forma menos nítida, caracterizando-o como sendo uma prática presente na *internet*, sendo esta última, fruto das novas tecnologias. Os mesmos, com especialização em Bioquímica e Medicina Legal e a desempenhar funções em Institutos Forenses, referem que *Big Data* é um *submundo* para o qual, diariamente, os cidadãos contribuem quando realizam pesquisas *onlines* reveladoras de preferências e interesses pessoais. Um dos entrevistados, pertencente à mesma área disciplinar, refere que a *internet* facilita esta cedência pessoal de dados informativos privados e o mesmo pretende o *Big Data*, apesar de que com objetivos e finalidades distintas, é o que se verifica, sendo um tema familiar à globalidade e permitindo o contacto com o mundo tecnológico, por via desta. Assim, sob a perspetiva do seguinte entrevistado, qualquer cidadão já contactou com técnicas de *Big Data*. Aliás, o geneticista forense, com formação académica superior em Bioquímica e a exercer funções laborais num Instituto Forense, questiona até que ponto nos devemos tornar tão restritivos em ceder informações para gerar os grandes dados, 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 porque 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 todos e mais algumas [diariamente na *internet*] (…)” [Geneticista Forense, N16]

Ou seja, os geneticistas forenses perspetivam que *Big Data* já se encontra enraizado na sociedade e, por isso, o seu discurso, sob a ótica de Stevens et al. (2018) é do tipo *pragmático*. Ball et al. (2016) enfatizam precisamente este aspeto: as normativas que circulam dentro deste meio da vigilância são inscritas através de práticas sociais de manipulação, uso de dispositivos e características que alimentam grandes infraestruturas de dados. Assim, os dados vão sendo recolhidos, armazenados e vão alimentando as grandes infraestruturas de dados, nas configurações do dia-a-dia (Ball et al., 2016). Sob este 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) – enviar um e-mail, realizar uma chamada telefónica ou uma pesquisa na *internet* – os indivíduos eternizam os seus rastreios digitais (Lyon, 2004; Brayne, 2017), à semelhança do que o último geneticista forense entrevistado referiu.

Estes geneticistas forenses, oriundos de áreas disciplinares como a Biologia e a Genética, a desempenhar funções laborais em Institutos Forenses, também consideram que se trata de uma análise 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, N23]

Esta ideia consensual de que a técnica se encontra (já) presente na sociedade é um discurso apelidado de *pragmático* sob a ótica de Stevens et al. (2018) e, neste caso, típico e frequentemente utilizado pelos geneticistas forenses, como já referido anteriormente.

Todas as perceções e expectativas dos entrevistados que desenham *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, com formação académica superior em Direito e a desempenhar funções laborais num Centro de Análise Celular de ADN, pretendeu explicitar foi que, efetivamente, os cidadãos, diariamente, contribuem (in)diretamente para esta rede de dados crescente, cedendo as suas informações, sem questionarem esta atividade. Exemplificando, o mesmo, sustenta esta sua opinião debruçando-se sob 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 usar (...) eu acho que 98% [das pessoas] não lê esse tipo de informação [sobre como os seus dados informativos serão tratados após a sua cedência voluntária] (...)” [NCP2, I02]

Ou seja, é uma prática globalmente (re)conhecida, alimentada pela sociedade, sob a identidade individual de cada cidadão que cede os seus dados informativos sem questionar a finalidade de tal, nem do lugar onde a informação vai ser armazenada, nem para fins será usada. Desta forma, conseqüentemente, a rede do *Big Data* expande-se, de forma inquestionável. Assim, enraizou-se, socialmente, a ideia de que integrar esta rede é uma obrigação social e, desta forma, conseqüentemente, os diferentes dados pessoais cedidos acumulam-se em grandes bases de dados, a informação fica armazenada e a rede dos grandes dados vai evoluindo, crescendo e tornando-se cada vez mais útil. Quanto a esta cedência “pública” dos dados pessoais, por parte dos cidadãos, às empresas e bases, Boyd & Crawford (2012) e Frade (2016) alegam que o que se verifica, no fundo, é que o quotidiano é cada vez mais transparente para as grandes organizações e as organizações envolvidas em vigilância são cada vez mais invisíveis para aqueles cujos dados são obtidos e usados. Ou seja, o cidadão está constantemente a transmitir informação, por via dos seus movimentos (por exemplo) e existe uma vigilância invisível que, de forma permanente, regista e armazena essa informação, sem o consentimento dos indivíduos.

Ou seja, estes entrevistados – pertencentes a todas as categorias profissionais, excetuando NCP1 – revelam uma visão de *normalização* da implementação da técnica de *Big Data* no quotidiano – é apenas um avanço na tecnologia e, por isso, esta aparece

contextualizada na vida quotidiana, neutralizando, por sua vez, os riscos advindos da sua prática, implementação e propagação no que concerne à privacidade e proteção de dados (Machado & Silva, 2016). Sob a ótica de Van Lente (2012), excetuando NCP1, todos os restantes profissionais entrevistados possuem expectativas positivas e promissoras, aproximam-se, a um nível *micro*, da técnica, operando na idealização de esboços detalhados de abstração da mesma nas suas práticas laborais e profissionais, demonstrando-se permissivos quanto à sua inclusão, com resistências mínimas.

A globalidade das expectativas profissionais aqui apresentadas culmina num consenso geral de que, por considerar e materializar-se 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, *Big Data* trata-se de um fenómeno com um alto nível de complexidade. Assim, no seio destes profissionais, surgem respostas que refletem conhecimento sobre as técnicas de *Big Data*, como já vimos; no entanto, é notório que a maior parte dos entrevistados que detinham alguma informação acerca desta ferramenta perspetivam-na como uma questão, simultaneamente, geral e complicada: um fenómeno complexo e múltiplo.

“Wow. Isso [*Big Data*] é muito complexo.” [NCP1, H01]

“Isso é uma questão muito complicada [quando questionado/a acerca de *Big Data*].” [NCP2, S01]

“(...) é uma questão muito difícil quando se trata de grandes dados (...)”
[Membro de Organização Não Governamental/de Direitos Humanos, G07]

Assim, é notório que, à exceção dos genéticos forenses, que revelaram perspetivas maioritariamente ligadas às questões preditivas e preventivas da técnica, bem como, à sua força algorítmica; 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 & Michael (2003) referiram: é expectável que a imersão de uma nova tecnologia inovadora numa rede já constituída aflore em mudanças dessa mesma rede, induzindo a alterações radicais – novas relações institucionais e profissionais. Quanto a esta aceção generalizada de que o *Big Data* é,

efetivamente, um sistema tecnológico complexo e aparatoso, Chan & Moses (2017) também verificaram que se trata de uma ferramenta que, devido ao seu nível de complexidade, só a tecnologia a pode manusear, daí a sua dependência aos *softwares* e às redes digitais. E esta complexidade advém, nas palavras de Fróis (2015), precisamente devido ao aprimoramento e emaranhamento das ferramentas disponíveis para se exercer a vigilância e o controlo.

Por outro lado, vários são os operadores do Sistema de Prüm que possuem representações sociopolíticas quando questionados acerca de *Big Data*. Concretamente, NCP1 foram a única categoria profissional que não abordou *Big Data* sob uma perspectiva política. As restantes categorias profissionais, apesar de, no cerne da resposta, se distanciarem umas das outras, no fundo, possuem todas perspectivas políticas acerca do fenómeno. Nomeadamente, um NCP2, com formação académica superior em Direito e a desempenhar funções laborais num Centro de Análise Celular, à semelhança do que se observa nas respostas dos geneticistas forenses, consideram que o uso e expansão de *Big Data* são influenciados pelos ditames políticos, tornando-se um fenómeno dependente destes, a quem cabe a decisão de implementar, ou não, a possibilidade de usar esse tipo de sistema tecnológico. Como podemos verificar pelas seguintes respostas – *Big Data* e a sua crescente utilização como um fenómeno político, sendo a Política a responsável pela legitimação e expansão desta técnica. Assim, *Big Data* é visto como um fenómeno politizado, dependente 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, I02]

“(…) Eu acho que isso [*Big Data*] 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, A03]

Tal como Jasanoff & Kim (2009) referiram: os artefactos tecnológicos não podem ser equacionados fora do seu contexto político, pois este influencia a sua evolução,

expansão e materialização. A Ciência e a tecnologia são espaços e objetos da Política (Jasanoff & Kim, 2009), sendo que, muitas vezes, o desenvolvimento das tecnologias vem, muitas vezes, responder a ideários políticos e estatais. Além disso, crê-se que a política direciona o desenvolvimento tecnológico (Van Lente, 2012).

Consequentemente, sendo a Política responsável pelo desenvolvimento e expansão destas técnicas digitais, segundo a perspetiva de NCP2 a desempenhar funções profissionais na Direção Geral da Polícia Criminal, responsável pelas ligações operacionais, isso potencia o facto de *Big Data* ser considerado uma *arma analítica*, como o próprio indica:

“(...) Isso [*Big Data*] é uma arma analítica e Prüm também está sob essa alçada analítica...” [NCP2, S02]

Ou seja, é um tema com interesses políticos, muitas vezes baseados e sustentados em ideias de lucro económico nacional. Segundo a ótica dos geneticistas forenses e NCP2, a partir do momento que as técnicas de *Big Data* possam trazer lucro económico que sustentem a sua implementação e expansão, estas serão ferramentas apoiadas pelos decisores políticos – muitas vezes órgãos máximos facilitadores das mudanças sociais. Além disso, muitas vezes os dados são propriedade do Estado – enquanto entidade máxima de poder político – e, por isso, depende da sua atitude perante esta posse, a evolução das atividades em torno destes mesmos dados. Nestes casos, a entidade responsável pelo uso e manuseamento dos dados é o Estado, logo é, consequentemente, a entidade responsável pelo seu desenvolvimento, ou não; as autoridades estaduais decidem sobre o acesso e uso dos dados (Prainsack, 2019). Assim sendo, o futuro do *Big Data* depende da opinião, dos valores, dos interesses e das decisões políticas, sob a ótica deste profissional. Parece haver uma responsabilização da Política no que toca ao *Big Data* (Van Dijck, 2014). A este respeito, Machado & Santos (2016), Machado & Granja (2018) e Machado et al. (2018) referem que um exemplo particularmente esclarecedor da motivação por trás de um modelo de policiamento baseado na tecnologia como instância de um projeto político no contexto da UE é a criação de um sistema europeu para o intercâmbio transnacional de dados genéticos forenses entre os países da UE com o objetivo de combater a criminalidade transfronteiriça e o terrorismo (e também a imigração ilegal) através do denominado Tratado de Prüm (assinado em 2005).

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

assemelham-se, na sua ótica política, do fenómeno *Big Data*, visto que ambos os extratos profissionais defendem a ideia de que esta técnica está dependente da Política, por se tratar de algo negociável, passível de gerar lucro económico. Os dados, passíveis de compra, convertem-se, facilmente, num negócio:

“(…) Aliás, é um grande negócio [*Big Data*] (...) 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, N16]

“Isso [*Big Data*] é um grande negócio.” [NCP2, I02]

Coll (2014) refere, quanto a esta conceção dos dados serem parte de um negócio maior, que a garantia da privacidade dos cidadãos se tornou um elemento fundamental para o bem-estar da economia social e, por isso, tudo aquilo que possa constituir-se uma ameaça para tal, 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 que a privacidade é vista como algo passível de ser manipulado por via da economia, a proteção de dados e/ou o acesso a estes é visto também, consequentemente, como um negócio. É evidente que se trata, no seu conceito mais linear, de uma forma dos diferentes países trocarem informações pessoais entre eles, de forma a atingir os seus objetivos (Gonçalves, 2018). Também Ball et al. (2016) referem este aspeto da comercialização e negociação dos dados, por parte das empresas. Os autores referem que a partir do momento 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 de redes sociais, começaram a utilizar isso a seu favor, como forma de rentabilizar o seu negócio. Assim, embora o uso de dados pelas corporações fosse uma vez um meio para um fim, um mecanismo para gerar mais valor dos clientes, para algumas empresas, como *Google*, *Facebook*, *Twitter* e outros, o armazenamento dos dados tornou-se um “fim em si mesmo” (Ball et al., 2016; Gonçalves, 2018). Nesta linha, Boyd & Crawford (2012) referem que o Mercado percebe o *Big Data* como oportunidade para segmentar publicidade e otimizar ofertas. Lawless & Williams (2010) também referem esta *mercantilização* dos mecanismos de provisão científica e, por isso,

tecem-se considerações comerciais acerca do tema em apreço, adotando-se uma racionalidade económica. Segundo estes últimos autores, os geneticistas forenses – e, neste caso, também NCP2 –, têm esta visão comercial do tema e isso ajuda a introduzir uma *identidade epistémica mutável*, pois o uso destes conceitos – *negócio* – reflete o facto destes profissionais se tornarem permeáveis às estruturas discursivas em desenvolvimento.

Todos os outros profissionais, excetuando NCP1, partilhando da mesma perspetiva política, assumem direções distintas, apesar do ponto de partida ser semelhante – um olhar politizado. Como já explanado, os geneticistas forenses ancoram-se ao facto das decisões políticas – determinantes na expansão e desenvolvimento da técnica – influenciarem a negociação das bases de dados; os NCP2 consideram que é, apenas, o poder político o responsável máximo pelas decisões que se prendem com a implementação (ou não) do *Big Data*; sendo que professores/investigadores, nomeadamente os pertencentes ao ramo do Direito – Economia Política –, ancorados também sob uma 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 também referirem e caracterizarem *Big Data* mediante a Política, estes últimos profissionais perspetivam-na como determinante na criação e expansão de conflitos e interesses económico-políticos, podendo isso 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 fazem correr o risco de que perante estes dilemas e conflitos de valores e de interesses e de direitos, fazer prevalecer o interesse económico (...) em desfavor das pessoas. (...)” [Professor/Investigador, N17]

Ainda assim, sob uma perspetiva sociopolítica, é notória a preocupação com as questões ligadas à proteção de dados, pelo facto desta técnica englobar o uso massivo de informações do foro pessoal e privado dos cidadãos. E o facto de ser algo passível de manipulação política exacerba os efeitos cautelares a adotar na recolha, armazenamento e partilha do grande número de dados. Esta questão é, maioritariamente, enfatizada pelos profissionais da Polícia Judiciária – membros pertencentes ao corpo de investigação criminal com formação superior em ADN e investigação criminal – tal como o seguinte extrato comprova:

“(…) estamos agora num sistema político democrático em que se respeita 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, N28]

Ou seja, é notória a expectativa profissional de investigadores criminais de que o rumo a adotar por parte da expansão e utilização destas técnicas pertence aos governadores – expoentes máximos que ditam como se deve prosseguir nesta área (semelhante ao que se analisara em NCP2 e geneticistas forenses). Por sua vez, o seguinte entrevistado perspetiva *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 pertencente à investigação criminal – Polícia Judiciária e com formação superior em ADN e investigação criminal –, considera que depende de quem detém este poder sobre os dados, pois é quem os manuseia, como podemos ver:

“(…) se a parte pública, quem está à frente dos desígnios de determinado estado já não estiver (...) preocupado com a arquitetura dos direitos, liberdades e garantias, e estivermos perante um estado autoritário, então isso aí está-se a ver mesmo que estamos perante uma partilha de dados que pode ser apocalítica para o próprio.” [Corpo de Investigação Criminal, N28]

Os profissionais reforçam neste âmbito, e conseqüentemente, também, a dificuldade política em encontrar o equilíbrio entre a defesa dos direitos, liberdades e garantias e a proteção de dados. E esta dificuldade reflete, claramente, que quem detém o poder de equilibrar os dois setores, conseqüentemente, 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:

“(…) mas 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, de facto, da conjuntura política, e sobretudo como é que evolui as ações... (...) E não é só, portanto em geral, de grande segurança, de securitária no que diz respeito às bases de dados, também vai depender

muito como é que esta articulação com estas grandes empresas (...)”

[Professor/Investigador, N26]

Ou seja, o fenómeno do *Big Data* e esta possibilidade futura do seu uso, manuseamento e aplicabilidade prática no âmbito de investigações criminais e posterior conexão de resultados com outros territórios é percecionada, maioritariamente pelos professores/investigadores inseridos em áreas disciplinares como o Direito, a 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 de definição legal do conceito, do que pode e/ou deve abranger e sob que ditames deve atuar. Sob as condições, exigências e interesses políticos, o *Big Data* pode expandir-se ou ser eliminado enquanto possível técnica de investigação. Concluindo, apesar dos esforços académicos, civis, democráticos, políticos e legislativos no sentido de encontrar um equilíbrio entre a defesa da segurança dos cidadãos e as potenciais ameaças às suas liberdades, garantias e direitos, este parece difícil de alcançar (Machado & Santos, 2016), tal como os profissionais referiram.

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 de *Big Data* e obstáculos práticos laborais à sua materialização. Concretamente, mas de forma sucinta, NCP2 com uma linguagem intimamente próxima da temática, dominando o tema e definindo-o, descrevendo-o, a um nível micro; NCP1 tecem referências aos potenciais obstáculos práticos e materiais que podem surgir com a potencial aplicação de *Big Data* no seu quotidiano profissional; geneticistas forenses e demais profissionais considerados tecem considerações informativas sobre o fenómeno, demonstrando, por via de um discurso consciente, a fase embrionária do fenómeno, as suas questões políticas e os dilemas que emergem – fruto das tensões tecnológicas¹³.

¹³ Ver anexo V.

3.2. Expectativas futuras de potencial aplicação de *Big Data*: desenvolvimento, expansão e antecipação de aplicação na investigação criminal (transfronteiriça e organizada)

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

Surge uma perspetiva geral de que se trata de uma técnica em desenvolvimento, com previsões de se expandir e reconhecer enquanto ferramenta tecnológica útil no combate à criminalidade. À semelhança do que Borup et al. (2006) referiram: existe uma expectativa generalizada de que a tecnologia é um elemento fundamental para criar e oferecer oportunidades para o progresso. Nomeadamente, NCP1 referem que se trata, precisamente, de uma técnica em desenvolvimento, que exige adaptação por parte dos bancos de dados de ADN. Estes últimos têm de se adaptar a estas novas tecnologias e ao desenvolvimento contextual e social do meio onde se inserem, bem como, os cidadãos. Ou seja, é perceptível que *Big Data* é uma realidade presente à qual os bancos de dados genéticos têm de ajustar para que da sua utilização surtam efeitos válidos e eficazes; como se pode verificar na seguinte resposta de um NCP1 a desempenhar funções num Laboratório da Polícia Científica:

“Bem, nós precisamos de adaptar as bases de dados de ADN para as novas tecnologias, para a próxima geração de sequenciamento (...)” [NCP1, G05]

Ou seja, existe uma consciencialização de que se trata de um fenómeno em desenvolvimento, que requer meios sofisticados e avançados para puder integrar-se e desenvolver-se. Desde logo, os NCP1 a desempenharem atividade profissional em Institutos Forenses, sendo, também, responsáveis pela administração da base de dados de ADN nacional, perspetivam efetivamente esse caminho: o desenvolvimento dos bancos de dados de ADN, para que se tornem maleáveis à disseminação de ferramentas de investigação inovadoras, tal como o *Big Data*. Além disso, os NCP1 a desempenharem atividade profissional em Institutos de Criminalística e responsáveis pela administração da base de dados de ADN nacional consideram, também, que para que esta técnica se materialize, desenvolva e seja crescentemente usada devem sofisticar-se 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 perceptível a necessidade sentida de sofisticar e desenvolver meios que acompanhem e permitam o desenvolvimento destas técnicas, exigentes do ponto de vista da tecnologização:

“(...) tem que 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, L01]

Neste caso, este entrevistado, considera tratar-se de um método que carece do desenvolvimento crescente de bancos de dados que lhe permita processar a informação de forma veloz, para que este seja efetivamente útil, por exemplo atualmente nos aeroportos – importa que a informação e os resultados sejam processados no imediato; e esta necessidade de urgência e velocidade de produção de provas exige meios adequados para que se efetive. Esta necessidade de eliminar tempos, de acelerar ações e funções é, também, referida por Araújo (2013) nos seus trabalhos empíricos sobre as temporalidades sociais. Atualmente, existe este senso de que a tecnologia é crucial porque permite diminuir os tempos de espera, acelerando a produção e efetivação de resultados úteis e visíveis num curto intervalo de tempo. Esta questão do tempo – a sua preciosidade e a capacidade de o acelerar é, muitas vezes, referida nos discursos dos entrevistados. No geral, este tipo de discurso, adotado maioritariamente por NCP1 é caracterizado, sob a tipologia de Stevens et al. (2018) como um discurso *crítico-interpretativo*, por incluir considerações críticas acerca das limitações e obstáculos atuais da técnica.

No que toca a esta visão profissional dos NCPs de que os meios de investigação carecem de desenvolvimento informático para que *Big Data* se materialize, as restantes categorias profissionais entrevistadas, incluem-se nesta panóplia, mas sob uma perspetiva diferente: consideram que se trata de uma ferramenta mais avançada do que o contexto onde visa operar e, por isso, exige que o meio onde se insere se desenvolva para que se possa efetivar – e, por isso, não tem ampla aplicação atualmente, carece do desenvolvimento dos profissionais, das áreas, dos métodos e do meio. Ou seja, parece ser geral a conceção de que se trata de uma técnica, de um tema, de uma área em desenvolvimento na sociedade: um conhecimento em construção (no sentido de que se sabe e se faz pouco a previsão de comportamentos criminais a partir da agregação de muitos e vários dados), mas já com alguma utilidade prática e em vias de se teorizar e constituir como ferramenta essencial de investigação. Sob a lente concetual de Borup et

al. (2006), Lucivero et al. (2011) e Van Lente (2012), estas expectativas que revelam muita confiança no desenvolvimento das tecnologias são comuns em investigadores académicos, tendo em conta que as suas variabilidades sociais influenciam as expectativas que defendem. No fundo, são profissionais que não lidam com a utilidade prática laboral da técnica, distanciam-se desta, profissionalmente – este tipo de expectativa é comumente defendido por professores/investigadores; como os seguintes entrevistados geneticistas forenses a desempenharem funções de professor/investigador em contexto universitário referem:

“(...) O conhecimento é muito limitado neste momento [para associar informações genéticas a padrões de comportamento, como *Big Data* prevê] (...) /“(...) Isso [*Big Data*] é um longo caminho, uma forma diferente de operar, mas sob uma perspectiva genética não sabemos ainda muito sobre isso [*Big Data*].” [Geneticista Forense, A02]

“(...) Até agora, não existe teoria por trás disso [*Big Data*]. (...)” [Geneticista Forense, N01]

Não obstante os geneticistas forenses inseridos profissionalmente em meios académicos, surge um geneticista forense a desempenhar funções num Instituto de Medicina Legal com a mesma perspetiva:

“(...) Acho que ninguém neste momento tem conhecimento para gerir isso [*Big Data*] (...) essas informações todas juntas. Talvez no futuro, mas precisamos de ferramentas informáticas mais poderosas. (...) Ninguém é capaz de gerir todas essas informações. (...) Precisamos de esperar mais um pouco. Reconheço os benefícios [de *Big Data*], mas precisamos de esperar mais um pouco.” [Geneticista Forense, J02]

À semelhança do que Chan & Moses (2015) referiram: é uma técnica que deve, sempre, ancorar-se à Teoria e, de facto, ainda se teorizou pouco sobre *Big Data* enquanto potencial ferramenta de investigação criminal. Por isso, esta técnica enquanto metodologia de investigação pode ser útil, mas é limitada.

Na mesma linha de reflexão, como já fora explanado, surge outro entrevistado professor/investigador na área do Direito, que se aproxima das perspetivas dos geneticistas forenses:

“(…) Provavelmente haverá [técnicas de *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, N27]

Os geneticistas forenses e os professores/investigadores consideram que o *Big Data* poderá trazer vantagens enquanto prática de investigação, no entanto atualmente não encontrará um contexto de atuação inovador e sofisticado que permita retirar deste os maiores benefícios. Consideram que *Big Data* é resultado da evolução técnica e científica, alegando que se trata de um tema que envolve uma multiplicidade de questões, rápidas e inevitáveis, ao mesmo tempo que se processa tudo a uma velocidade enorme. Ou seja, parece ser uma realidade-fruto da tensão de duas ideias – a inevitabilidade (da evolução tecnológica) e a regulação (a necessidade de ser regulamentado pelo Direito). E o equilíbrio entre estes dois ideais complexifica o pensamento sobre o tema. Surgem, neste sentido expectativas contraditórias (Van Lente, 2012; Lucivero et al., 2011; Borup et al., 2006) que refletem a complexidade das perspetivas. Estes mesmos entrevistados pertencentes a Comissões de Proteção de Dados perspetivam que o meio contextual e operacional de investigação deve sofisticar-se para que esta técnica se operacionalize, à semelhança do que os entrevistados NCP1 referiram:

“(…) Um bocadinho a ideia de alguma inevitabilidade [o uso potencial de *Big Data*] associada àquilo que tem sido, nos últimos tempos então com alguma força, o resultado da evolução técnica e científica, (...)” [Membro de Comissão de Proteção de Dados, N03]

Na mesma linha de reflexão, um membro pertencente ao corpo de investigação criminal, a desempenhar funções na Polícia Judiciária, enfatiza precisamente a mesma perspetiva:

“(...) O paradigma da investigação criminal tem que se adequar [fruto da inovação dos meios e ferramentas de investigação *Big Data*]. E tem que 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, N19]

Desta forma, é perceptível de que há uma clara preocupação e consideração por parte destes profissionais - geneticistas forenses, autoridades consultivas, diretores de instituições e membros do corpo de investigação criminal - em realçar a desajustação que existe entre a futura possibilidade de materialização do *Big Data* e os meios atuais onde esta materialização pode e deve atuar. Quanto a esta conceção de que é necessário que os meios de investigação das Ciências Sociais e Humanas se sofisticem para que se criem condições para que *Big Data* opere, Bartlett et al. (2018) referem que há, efetivamente, falta de base em metodologias matemáticas computacionais, devendo operar-se no sentido de desenvolver a computação, os meios computacionais e o acesso aos grandes dados para os profissionais desta última área, para que possam manusear estas técnicas. No fundo, estas constatações refletem que existe uma representação social maioritária que a tecnologia evolui, proporcionalmente à velocidade que penetra os interstícios sociais, levando a metamorfoses nos variados contextos, os quais têm de ter recursos para serem maleáveis o suficiente para se adaptarem às novas realidades perante as quais são confrontados. No que diz respeito a esta temática, Matzner (2016) refere que só o crescente desenvolvimento da tecnologia permite a efetivação e utilização de *Big Data*, visto que ambos se desenvolveram, de forma paralela, e o último depende do primeiro para ter condições de atuação. Neste sentido, Fróis (2015) descreve, exatamente, este dilema: a sofisticação tecnológica paralela ao aperfeiçoamento das técnicas de vigilância.

Os NCP2, apesar de partilharem esta visão consensual de que, por força do desenvolvimento e evolução dos meios tecnológicos e das ferramentas de combate e prevenção da criminalidade, o *Big Data* surge como a manifestação clara da evolução dos aparelhos tecnológicos e do mundo digital, não consideram esta relação entre evolução e sofisticação dos meios tecnológicos e digitais de investigação e o obrigatório desenvolvimento do contexto de investigação para que se torne apto à receção e inserção destas técnicas atuais. Ou seja, os NCP2 a desempenhar funções laborais em Unidades Policiais, reconhecem que *Big Data* é possível devido ao desenvolvimento e evolução

tecnológica, mas não reconhecem a necessidade de o contexto de investigação inovar para que se possa efetivar esta técnica, aliás apenas referem que há uma evolução no sentido do uso de *Big Data*:

“(...) Nós todos estamos a caminhar em direção ao *Big Data*. A cada 18 meses, os computadores evoluem...” [NCP2, H02]

Na mesma linha de pensamento, os geneticistas forenses a exercer funções laborais em Universidades e Institutos Forenses, consideram, também, que o uso massivo destas técnicas 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 eminente:

“(...) 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, A03]

“Neste momento, sem nós darmos conta, já muitas empresas o [*Big Data*] fazem. (...)” [Geneticista Forense, N16]

Sob a ótica destes geneticistas forenses, *Big Data* é uma técnica que já está presente na malha social e este tipo de discurso é nomeado por Stevens et al. (2018) como o discurso *pragmático* e é, neste caso, frequentemente citado por geneticistas forenses. Quanto à crescente utilização desta técnica, Chan & Moses (2017) referem precisamente esta ideia: fruto da disseminação dos tentáculos de *Big Data* nas várias facetas sociais, também no âmbito criminal, se verificará, conseqüentemente, uma tendência crescente do seu uso como ferramenta preventiva que orienta estratégias policiais e decisões de justiça criminal.

Não obstante o reconhecimento da necessidade do desenvolvimento das ferramentas de investigação e das bases de dados para se adaptarem a este novo aparato tecnológico, NCP1, a desempenharem atividade profissional num Instituto de Criminalística, sendo, também, responsáveis pela administração da base de dados de ADN nacional, e NCP2 a desempenhar funções laborais numa Unidade Policial,

consideram que se trata de uma técnica em crescente uso, atualmente, e com potencial para vir a desenvolver-se, de forma crescente, no futuro:

“Eu acho que isso [*Big Data*] está a ser usado cada vez mais agora (...)” [NCP1, L01]

“(...) Nós caminhamos em direção ao *Big Data* (...)” / “(...) Hoje, com o *Big Data*, todas as pessoas sabem que estão a ser vistas e controladas (...)” [NCP2, H02]

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

“Eu acho que estamos a caminhar nessa direção [do uso crescente das técnicas de *Big Data*] (...) E eu acho que pode ser interessante ir nessa direção [de *Big Data*], ter mais informações sobre as coisas (...)” [NCP1, Y01]

“(...) Não só para resolução de casos criminais, até por nós [e para outros fins, *Big Data*] poderia ser interessante (...)” [NCP1, Y01]

“Seria interessante [o uso de diferentes dados para investigação criminal], mas isso tudo capturado...” [NCP2, S01]

Esta última perspetiva reflete o *Big Data* como uma direção interessante a seguir, enquanto técnica, sob a sua ótica, que permite obter (ainda) mais informação acerca dos dados que possuímos, interessante até para fins não criminais. Ou seja, NCPs percecionam as técnicas de *Big Data* como ferramentas promissoras a serem desenvolvidas e implementadas no seu contexto laboral. E esta expectativa é unânime entre estes profissionais, pois revelam que, no futuro, poderá criar condições e possibilitar observar coisas que atualmente não somos dotados para tal:

“(...) Eu acho que isso [*Big Data*] será o futuro, *Big Data* permitir-nos-á ver coisas que não podemos ver hoje em dia.” [NCP1, J01]

É uma técnica em expansão e que, possivelmente – refere o entrevistado NCP1 a desempenhar funções laborais num Instituto Forense – no futuro, nos permitirá atingir e obter dados informativos que hoje ainda não são possíveis, facilitando a sofisticação e complexificação de determinados meios de investigação, que permitirão alcançar resultados mais eficazes. Ou seja, é uma técnica potencialmente benéfica para desenvolvimentos futuros no conhecimento científico. Brown et al. (2006) referem este aspeto nas expectativas: a facilidade de se pensar que o *novo* oferece oportunidades nunca antes concebidas, estigmatizando-se e minimizando-se o anterior. Este ponto de vista vai de encontro ao que Boyd & Crawford (2012) referiram quando definiram *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 podem gerar *insights* que anteriormente eram impossíveis.

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

“Definitivamente, eu não uso isso [*Big Data*].” [NCP1, X01]

“No contexto de ADN, eu não vejo isso [*Big Data*].” [NCP2, V01]

“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, no entanto, perspetivam a sua futura inserção no seu 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 [*Big Data*] (...)” [NCP1, O02]

“Eu não acho que isso [*Big Data*] seja usado no cotidiano policial [enquanto ferramenta de trabalho], talvez no futuro. Mas em geral, acho que no trabalho de investigação, como pesquisa, sim [deve ser desenvolvido].” [NCP2, F02]

“(…) Essas informações [de *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 são armazenadas e serão usadas. No entanto, até agora, não é possível.”

[Professor/Investigador, G02]

É frequente que as novas tecnologias, os novos conceitos, as inovações e os progressos científicos sejam expressos em termos de desejos futuros, tal como os últimos entrevistados referiram (Brown & Michael, 2003). Os profissionais reconhecem estas técnicas como práticas de trabalho aceitáveis, com probabilidade crescente de virem a ser implementadas enquanto ferramentas no seu contexto laboral; revelando perspectivas de que se venha a desenvolver, consolidar e implementar, à semelhança do que Boyd & Crawford (2012), Frade (2016) e Matzner (2016) referiram: a vigilância automatizada torna-se uma possibilidade crescente.

Apesar de ser unânime que a maioria dos profissionais entrevistados considera que será possível e desejável, a longo prazo, a inserção destas ferramentas tecnológicas nas metodologias de investigação, apenas os NCPs reconhecem a potencial utilidade prática e os possíveis benefícios de tal inserção no combate e prevenção da criminalidade transfronteiriça. Esta evidência vai de encontro às conclusões referidas por Machado & Silva (2015) no seu estudo: os profissionais que desempenham funções onde futuramente a técnica se poderá vir a aplicar – como é o caso de NCPs relativamente ao *Big Data* – estão mais conscientes da utilidade da técnica. Os restantes entrevistados – geneticistas forenses e outros atores envolvidos no domínio da partilha transnacional de dados – possuem uma perspectiva mais generalizada da questão em apreço. Ou seja, estes entrevistados possuem perspectivas positivas acerca de *Big Data*, aliadas ao desejo futuro da sua materialização concreta – expectativas comuns e referidas por Brown & Michael (2003) como reflexões presentes no âmbito de uma tecnologia em desenvolvimento. Concretamente, os geneticistas forenses com formação especializada em Biologia Genética e Medicina e a desempenhar funções laborais em Universidades e Institutos Forenses, respetivamente, consideram que refletir sobre a possibilidade de aplicar as técnicas dos grandes dados no âmbito da investigação criminal é reflexo de que existe um

progresso científico útil a ser concretizado, nomeadamente para a área forense, que pode fazer uso destas valências para alcançar de forma mais eficaz os seus resultados: encontrar os perpetradores.

“(...) Naturalmente, esperamos pelo progresso científico (...) e a questão é até que ponto isso [*Big Data*] pode ser útil para as ciências forenses? Bem, (...) para encontrar perpetradores (...)” [Geneticista Forense, A02]

“(...) *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 [*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, A05]

Ou seja, perspetivam *Big Data* como uma técnica útil neste sentido, mas não preveem outro tipo de utilidade na aplicação destas técnicas, na área forense. Além disso, também reconhecem o interesse técnico-científico do tema e da ferramenta, com várias utilidades práticas no meio, que poderão auxiliar a realização de inferências acerca da imputabilidade criminal, tornando-se uma prática útil nas investigações criminais. Como o seguinte geneticista forense, a desempenhar funções laborais num Instituto Forense refere:

“(...) 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,

N02]

Os geneticistas forenses consideram que a utilidade prática de *Big Data* opera na possibilidade de atingir resultados mais eficazes e caso o permita, demonstram uma visão positiva sobre a técnica e a sua inserção nas metodologias de investigação, visto que é justificável o seu uso, caso a sua finalidade seja, efetivamente, 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, N02]

Ou seja, visto que são profissionais que trabalham com os dados recolhidos pela polícia (que muitas vezes são limitados), se o número de dados recolhidos for maior, talvez o objetivo a alcançar com a investigação criminal seja mais eficaz e célere. De facto, os geneticistas forenses estão limitados quanto aos dados que são o seu objeto de estudo – reduzidos e, às vezes, degradados – desta forma, caso seja possível integrar uma técnica que permita agregar um conjunto maior e variado de dados, esta técnica é útil para a sua atividade profissional de investigação (Cole, 2013), sob a sua perspetiva. No entanto, é perceptível que os geneticistas forenses com formação na Biologia, Genética e Medicina possuem, apesar da sua expectativa positiva acerca de *Big Data* e da defesa da remota possibilidade de o considerar enquanto metodologia de investigação, uma perspetiva simplista, apresentado duas utilidades práticas. Na mesma linha de reflexão, surgem as expectativas de um membro pertencente a uma Companhia Privada de Ciências Forenses, que considera as vantagens do uso de *Big Data* no âmbito da investigação criminal:

“(...) Isso pode ser muito vantajoso [o uso de *Big Data* nas Ciências Forenses] (...)” [Membro de Companhia Privada, D13]

Sob esta perspetiva da antecipação de aplicação de *Big Data* no âmbito da investigação criminal, NCP2 a desempenharem funções policiais numa Diretoria de Cooperação Policial Internacional da UE, consideram que estas ferramentas são úteis, visto que a resolução eficaz de um caso criminal é sempre difícil, logo qualquer informação que surja e que ajude a resolvê-lo afigura-se útil.

“No final do dia qualquer informação [que auxilie na resolução de casos criminais], sobre qualquer coisa é útil [então, *Big Data* também o pode ser] (...)” [NCP2, W01]

Assim, se o *Big Data* é um mecanismo de investigação que agrega, armazena, configura e codifica informação que pode ser útil para a resolução de casos criminais, na

ótica deste entrevistado, a sua utilização é benéfica – tendo em conta a dificuldade da tarefa, qualquer técnica que se afigure útil, deve ser considerada para a investigação, inclusive *Big Data*.

No entanto, apesar das perspetivas e expectativas dos entrevistados-membros de entidades privadas se aproximarem, de forma conexas, às apresentadas pelos NCPs, na medida em que os três extratos profissionais nomeiam vantagens gerais e inerentes à inserção e aceitação de uma nova e inovadora metodologia tecnológica de investigação, os NCPs surgem com uma reflexão diferente: refletem acerca da possibilidade de a colocar ao serviço do combate e prevenção da criminalidade transfronteiriça. NCP1 a desempenhar funções profissionais num Instituto de Neurologia e Genética, refere-se ao *Big Data* como uma ferramenta útil, usada em situações que visem combater 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, W02]

Nestes casos, NCP1 defende o uso de técnicas de *Big Data* sem qualquer reserva nem questão, desde que permita aceder aos perfis corretos de indivíduos-autores de ações terroristas, tal como o próprio enfatiza. Neste sentido e para estes fins, considera que se trata de um mecanismo promissor para prever/antecipar comportamentos futuros danosos para a sociedade.

“(…) Eu gostava que tivéssemos esses métodos [de *Big Data*] e que pudéssemos prever comportamentos [criminais] (…)” [NCP1, W02]

Também um dos entrevistados, professor/investigador na Sociologia do Direito, refere esse aspeto: os benefícios de se utilizar esta técnica enquanto meio e ferramenta de combate e prevenção do terrorismo, ou seja, para a investigação de determinado tipo de crime. O seguinte entrevistado considera que existem certos ilícitos penais, perante os quais *Big Data* se afigura de extrema importância, dado o contexto do cometimento do crime e os seus impactos sociais; concretamente, refere-se a crimes transfronteiriços, com perímetro considerável e sob os 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 de *Big Data*] pode trazer de facto benefício, sobretudo à criminalidade transnacional, a grande criminalidade hoje, sobretudo criminalidade económica, o tráfico de seres humanos, a grande criminalidade é hoje transnacional. (...) [para] esta criminalidade de natureza transnacional (...) tem de haver armas efetivas neste combate com esta dimensão de eficiência, transnacional.” [Professor/Investigador, N26]

Esta visão é partilhada pelos NCPs: quando questionados acerca das expectativas da futura implementação das técnicas de *Big Data* no âmbito da investigação criminal, sob uma ótica generalizada, referem que perspetivam uma futura implementação destas ferramentas no âmbito da criminalidade transfronteiriça e transnacional. NCP2 a desempenharem funções laborais em Unidades Policiais, consideram que todos os tipos de informação agregados e recolhidos por esta ferramenta tecnológica são úteis, precisamente, para combater o terrorismo, por possibilitar a partilha de muita informação heterogénea:

“(…) Temos que usar a tecnologia e unir todas as informações que temos para combater o terrorismo (...)” [NCP2, H02]

Assim, é possível constatar que os NCPs reconhecem a utilidade prática eficaz do uso desta técnica para combater atentados terroristas, enquanto manifestação da criminalidade transnacional e transfronteiriça, cada vez mais eminente. No que toca aos tipos de crimes e circunstâncias criminais, as perceções dos NCPs sobre *Big Data* concentram-se na criminalidade grave. 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 de aplicação de *Big Data* à criminalidade transfronteiriça e transnacional, defendem a perspetiva de 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, P01]

Ou seja, é uma técnica de *ultima ratio*, apenas considerada quando todos os outros mecanismos não se afigurem adequados para tal, é um último recurso que poderá ter maior utilidade se articulado com outras ferramentas e/ou tecnologias. Para além disso, NCP2 referem 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 [*Big Data*], sem critérios determinados (...)” [NCP2, P01]

Concluindo, NCPs consideram que se trata de uma técnica de investigação potencialmente benéfica quando utilizada em situações de prevenção do terrorismo. No que toca ao combate a ações terroristas, também Drewer & Miladinova (2017) e Van Dijck (2014) referem que o *Big Data* surge como forma de prevenir estas ameaças transnacionais e transfronteiriças cada vez mais eminentes. De uma forma geral, todas as categorias profissionais entrevistadas demonstram-se permissivas quanto à implementação de *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 culminam num consenso que reúne os benefícios, utilidades e potencialidades da técnica, se usada corretamente para fins de investigação criminal; bem como, perceções de *Big Data* enquanto fenómeno em desenvolvimento¹⁴.

3.3. Perceção sobre os riscos e perigos de *Big Data*

No que toca aos potenciais riscos e perigos que o uso de *Big Data* pode acarretar para a investigação criminal, os vários profissionais revelam preocupações e motivos pelos

¹⁴ Ver anexo VI.

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 pela qual esta técnica se materializa na prática e como opera com os dados que são o seu objeto de estudo (Williams & Johnson 2004; Albert et al., 2009). Ainda assim, os NCPs demonstram maior preocupação ética e de direitos humanos do que propriamente ligada às questões procedimentais que podem colocar em causa a validade científica da ferramenta. NCP2 e geneticistas forenses tecem algumas considerações acerca das questões procedimentais. À semelhança do que Cole (2013) referira quanto aos genéticos forenses – as suas perceções relativamente aos resultados a que chega esta técnica baseiam-se em resultados probabilísticos, amplos e geradores de incerteza; demonstrando que estes profissionais se sentem inseguros quanto aos resultados que *Big Data* pode alcançar.

A este respeito, NCP1 a desempenhar atividade profissional num Instituto de Criminalística, desempenhando, em simultâneo, 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 ser visto como a única, principal e exclusiva forma de obter resultados e conclusões:

“(…) Isso [*Big Data*] deve ser visto mais como uma ajuda preditiva e não como algo que fornece informações 100% fidedignas.” [NCP1, L01]

Desta forma, o mesmo entrevistado refere que 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, L01]

Ou seja, a última perspetiva defende 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, a forma como os trata, manuseia e para que fins os usa. Segundo Prainsack (2019) este debate tem-se intensificado, precisamente porque nos referimos a dados digitais que estão disponíveis a partir de qualquer ponto geográfico e esta digitalização complexifica o estabelecimento de limites sobre quem tem acesso aos dados, controlo sobre estes e para que fins os utiliza,

intensificando este debate. Esta preocupação sobre quem tem acesso aos dados foi uma das conclusões que Lupton & Michael (2017) obtiveram quando se debruçaram sobre as percepções públicas acerca de uma vigilância através dos grandes dados. Este aspeto é frequentemente nomeado pelos NCP1 e pode revelar a falta de confiança que estes profissionais têm nas instituições que estão responsáveis pela recolha massiva destes dados e posterior uso dos mesmos para fins que não os previstos inicialmente. Estas conclusões são semelhantes às obtidas no estudo realizado por Machado & Silva (2016) sobre as percepções acerca da cedência voluntária de ADN para as bases de dados nacional portuguesa – a (não) cedência voluntária pode refletir-se, e ser proporcional, ao grau de confiança que a população tem nas instituições que recolhem o ADN e para que fins os usam. A este propósito, NCP1, a desempenharem funções laborais num Instituto Forense, referem que todas as tecnologias – tal como *Big Data* – acarretam riscos e desvantagens; no entanto, as consequências dos riscos do uso desta técnica dependerão da forma como esses mesmos riscos serão administrados: caso sejam bem geridos, os benefícios do uso de *Big Data* podem ser maiores que as suas desvantagens.

“(…) [*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, J01]

A referência que é feita aos riscos da técnica pelos entrevistados insere-se no que Jasanoff & Kim (2009) referiram, a propósito dos *imaginários sociotécnicos*: sendo *Big Data* uma materialidade que emerge na imaginação e nas expectativas sociais destes profissionais, quando estes fazem a abstração de a transpor na realidade prática quotidiana laboral onde se inserem, surgem obstáculos que se materializam nos riscos percecionados, expectáveis e imaginados pelos inquiridos. Ou seja, os profissionais ados têm em consideração os riscos que podem advir da utilização da técnica, prevendo-os e precavendo a sua expansão e exacerbação. Jasanoff & Kim (2009) referem-se a estes riscos como instrumentais e futuristas: residem no futuro, mas instrumentalizam-se na imaginação do presente. Estas dinâmicas de antecipação de riscos futuristas são comuns quando se abordam expectativas de uma tecnologia em desenvolvimento, segundo Borup et al. (2006) e Brown & Michael (2003): são receios que estão intimamente ligados ao *Big Data*, visto que a técnica é prematura e, por isso, geradora de medos e incertezas.

Nesta linha de reflexão, geneticistas forenses a desempenharem funções laborais como investigadores num Instituto Médico Informático e de Estatística, referem que os usos potenciais de *Big Data* são uma *promessa não cumprida*, porque não são capazes de atingir os resultados a que se propuseram aquando da sua implementação e/ou idealização. Ou seja, o facto de envolver, englobar e considerar um grande número de dados não lhe permite prever a globalidade das soluções; na perspetiva dos geneticistas forenses, pode prever o risco de algumas e determinadas situações, no entanto não é capaz de prever tudo:

“(...) *Big Data*, é uma promessa não cumprida. *Big Data* pode, claro fazer algo. Pode gerar talvez algumas hipóteses mas, por outro, pode – bem, 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 (...) *Big Data* pode ajudar a dar algumas ideias, por exemplo, mas não pode realmente resolver tudo como prometido.”

[Geneticista Forense, O08]

Borup et al. (2006) e Brown & Michael (2003) nomearam esta característica das expectativas em torno da inovação tecnológica e científica – *promessa* – as perspetivas em torno de técnicas em desenvolvimento surgem como desejos para o futuro, que não encontram, ainda, idealização prática para se efetivar, mas despertam expectativas, interesses e receios. E a desilusão ancorada após a criação da expectativa da *promessa* foi mencionada nos estudos de Brown et al. (2006) sobre as expectativas de mudança e as identidades contestadas das células-tronco do sangue. Ou seja, 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 materiais à sua materialização. Neste caso, relativamente a *Big Data*, visto que não se trata de uma técnica validada cientificamente para ter capacidades de proceder a previsões globais e válidas, como seria esperado, na ótica dos geneticistas forenses. Desta forma, desacredita-se no poder tecnológico da técnica e a expectativa positiva dá lugar à deceção. Consequentemente, *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, NCP2 a desempenharem funções laborais num Instituto Forense, referem-se ao *Big Data* como

sendo uma *técnica muito perigosa*, visto que permite o acesso a dados e informações pessoais a entidades que podem usar esses dados e informações para fins que não os previstos inicialmente, aquando da sua recolha. Logo, é um tipo de pesquisa e de investigação que pode prejudicar os titulares dos dados que estão a ser usados, se usados para outros fins, por exemplo, ou mesmo apenas a consulta desses dados é um risco, que pode pôr em causa o benefício das pessoas em questão.

“Acho que é [*Big Data*] 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, N14]

Esta é uma das preocupações dos profissionais entrevistados, à semelhança do que Machado & Silva (2016) concluíram no seu estudo quanto às perceções populacionais da cedência voluntária de ADN para as bases de dados portuguesas: a recolha do ADN para outros fins que não os previstos inicialmente. No que toca a esta perspetiva, 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). Este regime político parece ser antecipatório quanto ao *Big Data* (Brown & Michael, 2003) – esta expectativa culmina em alterações legislativas.

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

“(...) A análise [com *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, S01]

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 a que se chegam sejam calculadas com cuidado. Desta forma, tendo em conta que *Big Data* pressupõe, desde logo, a inclusão de um grande número de dados a serem trabalhados, conseqüentemente, converte-se numa prática *arriscada*. Partilhando da mesma perspetiva, os genéticos 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 qualidade destes últimos, sendo necessária a validação dos mesmos para que as conclusões a que chegam sejam fidedignas. Assim sendo, este procedimento de validação dos dados e sua posterior verificação, remete estes profissionais a situações clássicas de técnicas e ferramentas tradicionais de investigação, que não são inovadoras:

“(…) Neste momento, não foi bem sucedido [*Big Data*] e há várias razões para isso. Uma é que a qualidade de *Big Data*, geralmente, é má. Portanto, não se pode obter um bom resultado a partir de dados que estão incorretos. Então, estes dados têm que 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 que 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, O08]

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 à importância e o poder que são atribuídas às correlações provenientes de algoritmos e inferências. Quanto a estas advertências, Boyd & Crawford (2012) e Chan & Moses (2015) referem 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, *Big Data* requer, também, uma visão holística dos fenómenos. Além disso, importa enfatizar que, apesar do grande

volume de dados que *Big Data* contém, estes não devem ser vistos como representativos, nem tão-pouco como explicativos da globalidade dos fenómenos, mas antes devem ser interpretados, como qualquer outro tipo de dados, independentemente do seu tamanho amostral. Todas estas precauções se exacerbam quando *Big Data* se aplica à prevenção da criminalidade, visto que visa informar políticas e resoluções criminais, deve orientar estas decisões de forma eficaz e válida, sustentado em dados fidedignos (Boyd & Crawford, 2012).

Ancorados sob o mesmo ponto de vista da antecipação dos potenciais riscos de *Big Data*, surgem os geneticistas forenses, NCP1 a desempenhar atividade profissional num Instituto de Criminalística e os profissionais-membros de Organizações Não Governamentais/Direitos Humanos, que enfatizam também as questões da fragilidade científica da técnica. Os entrevistados conjugam dois tipos de representações sociais: a preocupação eminente devido à 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). Desde logo, consideram que a conexão de um número elevado de dados heterógeneos não é possível, visto que a informação não possui critério de semelhança e/ou identificação. Ou seja, 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 [em *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, L01]

“(...) Temos dados biométricos (...) 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, P02]

“(…) 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 *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 Organização Não Governamental/Direitos Humanos, G07]

Esta é uma visão frequentemente expressa por investigadores sociais – estes profissionais, usualmente, tecem avaliações sobre o progresso científico (Selin, 2008). Na mesma linha de reflexão, os geneticistas forenses a exercer funções laborais em laboratórios de Institutos Forenses, consideram que é uma amostra numerosa e constituída por dados heterógeneos, que caso não sejam manuseados com cuidado colocam em causa a qualidade dos mesmos, bem como a qualidade dos procedimentos que levam aos resultados, fragilizando as conclusões que daí se possam retirar e considerar. O risco eminente é o que se prende com a natureza heterogénea dos dados recolhidos, acompanhada do elevado número de dados. Ou seja, trata-se de uma amostra diversa e numerosa e, a partir desta, os resultados obtidos facilmente se podem tratar de falácias e erros – segundo a seguinte resposta – e não é o pretendido, podendo 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-nos induzir em erro, podemos cair em falácias (….) é preciso muito cuidado com esse género de coisas (….)” [Geneticista Forense, N08]

“(…) A ideia de conectar dados da saúde com o comportamento criminal [*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, X03]

Esta sensibilidade face aos dados e aos procedimentos de recolha dos mesmos que têm influência na validade do processo de investigação é, frequentemente, nomeada pelos geneticistas forenses. No fundo, estes profissionais têm como audiência última do seu trabalho e dos resultados alcançados o Tribunal, enquanto instância máxima decisória e de poder, sendo sujeitos a processos legais com a aplicação do princípio do contraditório

em que podem ser obrigados a envolver-se no processo caso os resultados obtidos apresentam falhas, fracassos ou erros. Assim, a sua credibilidade e a do laboratório onde desempenham funções laborais pode fragilizar-se (Cole, 2013). Estes profissionais, conscientes desta realidade, demonstram esta sensibilidade profissional e técnica quanto às questões procedimentais e de validade científica de *Big Data*. Porém esta visão não pertence apenas aos geneticistas forenses, também os entrevistados professores/investigadores em Antropologia, referem-se às técnicas de *Big Data* como práticas minuciosas que devem ser manuseadas de forma criteriosa para não levar a resultados e conclusões duvidosas:

“(...) Há estudos que podem ser [feitos assim: associar padrões de comportamento para mapear a prevenção de determinadas doenças]... há relações interessantes que podem ser feitas, mas tem que haver um grande controlo e a validade do estudo é uma questão tem que ser vista (...)” [Professor/Investigador, N05]

No que toca a estas questões éticas da recolha de dados e posterior agregação, Boyd & Crawford (2012) referem 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 nexos, 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, ligam-se a preocupações sobre a privacidade individual e exposição à vigilância precisamente devido ao autónomo fluxo de dados que o indivíduo não pode controlar.

Vários são os profissionais que quando abordam a temática ligada às questões procedimentais por via das quais se materializam as técnicas de *Big Data*, consideram que se trata de uma técnica antiética e incorreta por agregar um grande número de dados sem critérios nem procedimentos válidos cientificamente. Consequentemente, dada a fragilidade do método, este pode, sob a ótica dos geneticistas forenses desenvolver atividade profissional de professor/investigador universitário na área da Antropologia, estabelecer previsões e correlações, no entanto estas carecem de valor causal. Assim, trata-se de uma técnica de investigação sem base teórica e 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. [*Big Data*] não tem base causal, não tem teoria (…)” [Geneticista Forense, N01]

Quanto ao facto dos resultados obtidos se tratarem de inferências e não de correlações com valor causal, Amoore (2011) refere que o que se pretende de *Big Data* é mesmo isso: a capacidade de fazer inferências nos dados, de modo a que estas últimas possam ser reconhecidas e partilhadas. Uma associação não implica uma conexão causal direta, mas sim revela, interpreta e exhibe relações entre pessoas, lugares e eventos (Amoore, 2011). Consequentemente, esta visão dos geneticistas forenses é frequentemente reforçada com ênfase nas questões procedimentais e de validade científica (Cole, 2013) – incerteza tecnicocientífica. 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 inseridos nestes, onde a futura técnica se poderá vir a materializar e, desta forma, estas suas preocupações são resultado da posição laboral que ocupam. Em contrapartida, os mesmos autores referem que é incomum, ou mesmo raro, encontrar este tipo de reflexões, por exemplo, em grupos profissionais que se caracterizam por funções laborais relacionadas com a administração, secretariado e/ou política, pelo facto de não estarem inseridos no contexto de materialização da técnica.

Ainda assim, as questões ligadas às fragilidades procedimentais da técnica não se esgotam nestas considerações, por parte dos profissionais entrevistados, estes também perspetivam o desrespeito pelos fins a que se destina a recolha dos dados como uma causa de preocupação do potencial uso de *Big Data*. Concretamente, os geneticistas forenses em geral, 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, *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 inicialmente aquando da recolha dos dados]” [Geneticista Forense, N20]

“Pode ser perigoso também” / “(...) Vai-se (...) criar uma determinada ideia com base em informações que (...) não foram consentidas, nem conscientes para o que estava a ser feito (...)” [Geneticista Forense, N22]

“(...) O risco é que a utilização desta informação [obtida por via de *Big Data*] extravase este âmbito. (...) É que se ultrapasse este foco da investigação criminal e se vá a outras áreas. (...) isso é que acho que pode ser perigoso.” [Geneticista Forense, N02]

Assim, surgem representações socioprofissionais, por parte dos geneticistas forenses, que alertam para o facto da recolha massiva de um grande número de dados se ancorar em investigações não criminais e cumprir outros fins que não os previstos inicialmente. Ou seja, existe uma antecipação dos riscos e dos perigos desta técnica, caso esta seja um meio para recolher e conjugar dados, para fins não consentidos. Estes resultados aproximam-se dos alcançados pelo estudo de Machado & Silva (2016) relativamente à perceção populacional da cedência voluntária de ADN para as bases de dados nacionais de ADN. São expectativas expressas, maioritariamente, por NCPs e geneticistas forenses – esta conclusão vai de encontro ao 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, NCPs e geneticistas forenses, visto que desempenham funções laborais em meios onde *Big Data* se poderá efetivar, aproximam-se da técnica e são, também, mais conscientes, dos riscos e perigos a ela associados, bem como, obstáculos à sua materialização. Esta perspetiva vai, também, de encontro ao que Machado & Silva (2015) referiram: 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, maioritariamente, por profissionais que estão diretamente ligados com uma prática laboral que, futuramente, possa ser palco de desenvolvimento de *Big Data*.

Esta referência a aspetos éticos e sociais, onde se enfatiza o carácter peculiar e minucioso, pessimista e crítico da técnica, insere-se na tipologia *exceccionalismo genético* de Williams & Johnson (2004) – nestes discursos, usualmente e sob a perspetivas dos autores, tecem-se considerações acerca da avaliação do processo pelo qual a técnica se materializa – avaliação dos processos, do modo pelo qual os dados são recolhidos, a forma como são analisados, tudo o que diga respeito à avaliação de questões procedimentais

práticas que é a narrativa discursiva que predominou nesta temática, sendo esta última transversal a todas as categorias profissionais. De forma geral, todos os profissionais entrevistados teceram considerações sobre os potenciais riscos e perigos da implementação de *Big Data*, desde comentários que se debruçaram sobre a fragilidade científica da técnica, como avaliações sobre o processo de recolha de dados, o seu manuseamento, análise e posterior utilização, bem como, reflexões abstratas com pendor avaliativo sobre os riscos procedimentais e técnicos¹⁵.

3.4. Pareceres éticos e de direitos humanos

No que toca às questões éticas e de direitos humanos – a esta consideração vasta pelas preocupações que se prendem com a salvaguarda dos direitos, liberdades e garantias dos cidadãos – todos os profissionais considerados teceram referência. Desde logo, ancorados ao debate político e social contemporâneo que emergiu devido ao caso de Snowden¹⁶, enquanto desbloqueador do pensamento crítico relativamente ao *Big Data*, ou seja, a questão do uso de dados pessoais protegidos e vistos, legalmente, como privados, para fins que não os idealizados inicialmente. Esta questão da violação da reserva da intimidade e da vida privada que se levantou depois das revelações de Snowden é referida por muitos profissionais, nomeadamente pelos professores/investigadores em Direito:

“(...) O caso do 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 nomeação do caso de Snowden é feita por vários profissionais inseridos na área do Direito e estes vêem 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. Concretamente, 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, precisamente, a recolha massiva e

¹⁵ Ver anexo VII.

¹⁶ Tópico tratado na subsecção 1.3. da presente dissertação (“O caso de Edward Snowden”).

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 de Snowden: foram recolhidos dados pessoais privados, de forma massiva, no entanto não ficaram à disposição dos serviços de segurança, nem policiais, tendo sido utilizados para outros fins. Tal como Brown & Michael (2003) referiram, muitas vezes as tecnologias desdobram-se em formas e maneiras que não tinham sido planeadas inicialmente. Ou seja, neste caso, os grandes dados abriram caminho para violações da privacidade, ao invés de cumprir o objetivo de garantir e promover a segurança máxima dos indivíduos. E a questão que se levanta aqui é se este exercício de reflexão sobre os fracassos do passado pode ensinar algo que permita prevenir os fracassos do futuro? Brown et al. (2006) referem que é frequente a nomeação de uma identidade retrospectiva, que cada história de inovação possui, tipicamente, alguém acompanhado de 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 nomeando futuros passados que fracassaram (Brown & Michael, 2003).

Toda esta panóplia de dilemas morais e éticos, de questões que emergem fruto da facilidade pela qual 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* à antiga PIDE, capaz de exacerbar o pânico moral, na medida em que intensifica a vigilância e o controlo, na mesma proporção que não garante maior segurança:

“(…) No tempo do antigo regime (…) as pessoas (…) tinham medo de haver um polícia, ali um PIDE 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... (…)” [Professor/Investigador, N17]

Metaforicamente, como que se atualmente a PIDE, a entidade responsável, na altura, pela vigilância, controlo e supervisão dos cidadãos se tenha convertido no *Big Data* e as instâncias de controlo se tenham convertido, cada vez mais, invisíveis perante os cidadãos controlados e, por sua vez, estes últimos converteram-se em entidades sempre visíveis sob o olhar atento e omnipresente das instâncias de controlo. Em contrapartida, perdeu-se o controlo sob quem controla e exerce a vigilância, tudo o que é feito pode ser

escrutinado por figuras não identificáveis, devido à sua invisibilidade. 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 é o que Brown & Michael (2003) designam de *prospecting retrospects* – a forma como as perspectivas passadas são implantadas em tempo real, para construir o futuro. Brown et al. (2006) enfatizam também estas noções: as expectativas constroem-se de heróis retrospectivos, acompanhando-se de narrativas de luta, tal como se trata o exemplo dado pelo entrevistado.

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 mediada digitalmente entre os sistemas de vigilância e o indivíduo – ao invés 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. Dando a ideia de que se alterou o cerne das relações humanas por via dos dados. E esta constatação é mencionada pelo seguinte membro do corpo de investigação criminal, a exercer 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, N19]

Quanto a esta forma de (re)ver ou “ver de novo” a forma como agimos e comunicamos, hoje em dia, Boyd & Crawford (2012) referem que o *Big Data* surgiu como um sistema de conhecimento e isso já altera os objetos do conhecimento, ao mesmo tempo que tem o poder de informar como entendemos redes e comunidades. Quanto a esta digitalização das relações humanas e sociais, Frade (2016) define a contemporaneidade como o tempo do digital e das comunicações em rede – possibilitados pelo surgimento, expansão e grande desenvolvimento das novas tecnologias. Consequentemente, por via do uso crescente das tecnologias, segundo Aas (2006), a ação humana é codificada, ou seja, convertida em símbolos com significado especial. Assim, na conceção de Costa (2004), cada passo do humano é uma senha comunicada por via de uma linguagem digital. E isto vai de encontro ao que o entrevistado percecionara: alteração do nosso agir comunicacional. Aliás, Matzner (2016) refere que se criou uma

relação *humano-algoritmo* que moldou as formas como os seres humanos são tratados e classificados. Além disso, verifica-se uma extensão e 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). Ou seja, estes profissionais e adotando a lente concetual de Van Lente (2012) nomeiam 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. E isso é tido em conta pelos diferentes entrevistados. Prova disso é o discurso do mesmo entrevistado - pertencente ao corpo de investigação criminal a exercer funções profissionais na Polícia Judiciária -, que refere que o *Big Data* se materializa hoje na sociedade por via de mecanismos que registam as nossas “marcas”, as nossas “pegadas”, como por exemplo a via verde, a geolocalização e os demais dispositivos que registam a nossa presença.

“(…) 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, N19]

Os membros do corpo de investigação criminal e NCP2 entrevistados apresentam esta preocupação inigualável, quando comparados com os restantes grupos profissionais – esta visão social do fenómeno de *Big Data* – tecem considerações que se prendem com estas alterações sociais e pessoais que a génese de uma nova forma de se exercer a vigilância pode acarretar. E sob esta linha de pensamento, surge um entrevistado professor/investigador, especializado em Direito, que alega 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 de *Big Data* e uma nova forma de se exercer a vigilância – tecnológica, omnipresente e digital. Além disso, a caracterização profissional desta técnica como *Big Brother* transparece o risco e o perigo advindos da sua utilização excessiva, que poderá criar níveis máximos de vigilância sobre os cidadãos – observação e controlo constante.

“(…) Novos problemas que se põem, isto é, no fundo lembramo-nos sempre *Big Data*, *Big Brother*, evidentemente que os sistemas policiais são dos mais interessados

nesse cruzamento, os sistemas de segurança em geral. (...)” [Professor/Investigador, N17]

Big Data surge, então, sob a sua concepção, 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 & Michael (2017) referiram acerca das perceções públicas sobre uma vigilância de dados: esta aceção generalizada de que uma vigilância que opera por via de uma grande quantidade de dados potencia(rá) a imersão humana numa sociedade *Big Brother*. Quanto a este tema, Matzner (2016), Orwell (2009) e Coll (2014) também debatem sobre o surgimento de uma sociedade panótica, por via de uma vigilância permanente, invisível e penetrada 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 de *Big Brother* (Boyd & Crawford, 2012; Lyon, 2014; Orwell, 2009; Coll, 2014).

Em contrapartida, os NCP2 a desempenhar funções laborais em Unidades Policiais, alegam que a maioria social, quando questionada e/ou confrontada com este tema de *Big Data* refere que se tratará de converter a sociedade num meio semelhante ao do *Big Brother*, como os profissionais a exercer funções em Universidades e com formação académica superior em Direito perspetivam. No entanto, para os NCP2, o trabalho desenvolvido em redor desta técnica é benéfico e jamais criará condições semelhantes à do *Big Brother*, defendendo a prática desta técnica:

“(...) As pessoas vão começar a dizer “Isto [*Big Data*] é o *Big Brother*”, estão a monitorizar tudo. Mas eu não acho (...)” [NCP2, H02]

Não obstante e em seguimento do que fora explicado, um membro pertencente a uma Comissão de Proteção de Dados, profissionais de investigação criminal a exercer funções profissionais na Polícia Judiciária, professores/investigadores de Direito e geneticistas forenses a exercer funções laborais como professores universitários, consideram que esta prática não pode aplicar-se de forma desmedida, para não acarretar consequências diretas de violações de direitos humanos. Consideram que se trata de um fenómeno que deve estar amplamente previsto, de forma restrita, na lei, enquanto

entidade que deve regular a forma como se executa, por quem e para que fins. É unânime que a maior parte dos entrevistados geneticistas forenses, membros de autoridades consultivas e profissionais de investigação criminal consideram que deve ser uma técnica amplamente regulada, destacando o facto de perspetivarem que a sociedade está a evoluir e a criar bases no sentido dessa regulamentação.

“(...) [*Big Data* é] uma questão de regulação pelo Direito (...)” [Membro de uma Comissão de Proteção de Dados, N03]

“[*Big Data*] deve ser devidamente enquadrado segundo (...) juristas (...)”
[Geneticista Forense, N04]

“Isso [*Big Data*] deve ser estritamente regulado, mas eu penso que estamos a caminhar nessa direção (...)” [Professor/Investigador, G02]

Estes entrevistados consideram que o Direito, no sentido da legislação e regulamentação, é uma área muito importante no sentido de balizar, legalmente, o número e o tipo de profissionais autorizados a manusear, recolher e tratar estes dados. Há uma clara preocupação, por parte destes profissionais, sob quem tem acesso aos dados, visto que a disponibilidade dos mesmos não define nem estabelece quem pode aceder a estes, pois estão disponíveis e acessíveis a qualquer pessoa; bem como, são cedidos por qualquer pessoa, de forma voluntária. Então, importa definir, claramente, quem o pode fazer e em que termos. Esta perspetiva revela a falta de confiança que os profissionais considerados têm nas instituições empresariais responsáveis pela recolha dos dados, visto que é uma preocupação eminente e frequentemente referida por estes (Machado & Silva, 2016). Segundo Lupton & Michael (2007) trata-se de uma preocupação frequentemente referida quando se trata de apreender as perceções públicas sobre o 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 numa 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 nomeiam a necessidade de regular, legalmente, esta atividade.

Estes recortes discursivos salientam um aspeto: a necessidade de pensar criticamente os ideais inovadores do futuro; ou seja, por via de um paradigma preditivo analisam 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 se desenvolverem instrumentos legais que apoiem e sustentem um progresso legal 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, N24]

“(...) Ainda é o problema da limitação da finalidade, e o acesso a esses bancos de dados [de *Big Data*] deve ser restrito apenas a autoridades e órgãos autorizados.”
[Membro de Comissão de Proteção de Dados, G04]

“(...) 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 que aceder a isso [*Big Data*]. (...)” [Corpo de Investigação Criminal, N19]

Ou seja, os profissionais entrevistados referem que o *Big Data* é potenciado pelas marcas que as diferentes pessoas vão, conscientemente, deixando, ou seja, os cidadãos criam as suas próprias pegadas e estas são recolhidas, usadas e analisadas para obter determinados fins. No entanto, este acesso aos dados, a sua utilização e manuseio carecem de autorização judicial, o que reflete claramente a natureza privada destas informações pessoais e o caráter restritivo como podem ser usadas, *a posteriori*. Esta cedência voluntária de dados pessoais é definida, sob as palavras de Matzner (2016) como o processo de fornecer ativamente informações pessoais, por exemplo, por via de *sites* e redes; e é uma das mais frequentes formas de recolher dados para expandir a rede do *Big Data* – a forma mais disponível e utilizada. E quanto a este acesso aos dados, Boyd & Crawford (2012) alegam que importa que nos questionemos precisamente sobre isso:

quem tem acesso aos dados? Em que contextos? Com que restrições? Qual a finalidade? A verdade é que *Big Data* permite o acesso a uma enorme quantidade de informações produzidas por e sobre pessoas, interações e coisas. E, desta forma, consequentemente também o acesso a informações pessoais através dos dados e o cruzamento destes dados, podendo comprometer a privacidade dos indivíduos (Herschel & Miori, 2017).

Na mesma linha reflexiva, NCP1, a desempenharem funções laborais num Instituto Forense, enfatizam precisamente 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 destas informações serem analisadas ou usadas sem consentimento, para outros fins. (…)” [NCP1, E01]

Ou seja, é evidente que, 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 que pode provocar. Também os membros de Comissões de Proteção de Dados, partilham desta perspetiva – são apologistas de uma definição legal e normativa que restrinja a margem de ação que os profissionais possuem para manusear os dados recolhidos, mediante a disponibilidade limitada desses dados:

“(…) Tem de ser garantido que é privada [a informação contida nas grandes bases de dados], 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, N25]

Nesta linha perspetivacional, os geneticistas forenses a exercerem funções profissionais em Institutos Forenses, também revelam algumas preocupações, nomeadamente ligadas às entidades que têm acesso aos dados e aos fins para os quais estes dados são usados (Williams & Johnson, 2004; Albert et al., 2009). Concretamente, no que se liga diretamente à temática dos direitos humanos, estes entrevistados

demonstram claras representações preocupantes ligadas aos direitos à privacidade e de igualdade – ou seja, as pessoas que são alvo deste escrutínio de informações e que impactos, a nível da privacidade, é que este acesso tem no seu quotidiano e nas suas liberdades e garantias – com podemos reparar na resposta do seguinte geneticista forense:

“(...) Seria [utilizar os grandes dados como potencial meio de investigação criminal, genético e forense] 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, N21]

Concluindo, mediante a panóplia ética e de direitos humanos, a maior parte dos profissionais entrevistados considera que são questões a ter em conta, a explorar e a analisar quando o tema é, precisamente, *Big Data*, o uso das suas técnicas, a própria cedência dos dados e os fins que se visam atingir após a recolha das informações e o seu tratamento. NCP2 a desempenhar funções de *design* e gestão do *software* informático da base de dados alega que:

“(...) [*Big Data*] é uma questão ética. (...)” [NCP2, S01]

Lucivero et al., (2011) refere que estas questões são frequentemente citadas 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. Assim, conseqüentemente, paralelamente a este olhar sob um cenário irrealista emergente – como o *Big Data* – os entrevistados que referem questões éticas contribuem para transformar as promessas estratégicas sobre as tecnologias, em expectativas geralmente partilhadas, por todos os grupos profissionais entrevistados. No que toca a esta preocupação geral e unânime por parte dos entrevistados com as questões ligadas à recolha de dados, privacidade e proteção de dados, os professores/investigadores em Direito e membros de Comissões Biométricas com formação superior em Direito, nomeiam precisamente esses temas nas suas respostas, não incluindo questões que se prendem com quem tem acesso aos dados, mas enfatizando os riscos de invasão da vida privada e proteção de dados que o potencial uso da técnica pode acarretar.

“(...) 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 os condicionamentos das pessoas. (...)” [Professor/Investigador, N17]

“(...) [*Big Data*] levanta muitas questões sobre a privacidade. (...)” [Membro de Comissão Biométrica, D03]

Existe uma representação socioprofissional que o uso das técnicas de *Big Data* na recolha compulsiva e massiva de informações pessoais e privadas, pode condicionar os cidadãos, restringindo-lhe 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, as finalidades das investigações criminais e o desenvolvimento tecnológico. Apesar de reconhecerem o equilíbrio político entre a garantia da segurança e a preservação do respeito pelos direitos dos cidadãos, questionam-se, no que concerne à investigação criminal, se efetivamente esta, para cumprir os seus fins, poderá colocar em causa, de forma legítima, os direitos dos cidadãos? Não obstante, consideram, nesta linha de reflexão que, por força do desenvolvimento tecnológico os mecanismos de investigação sofisticar-se-ão e, conseqüentemente, 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. (...) onde de facto se 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, N27]

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

Em contrapartida, os investigadores criminais, a desempenhar funções policiais, possuem uma perspetiva contrária: perspetivam *Big Data* como um fenómeno legítimo, visto que possui potencialidades para armazenar um número considerável e útil de informação acerca dos cidadãos, imprescindível para garantir a sua segurança. E o facto dessa informação estar armazenada nas bases de dados não constitui uma violação aos direitos, liberdades e garantias dos cidadãos, mas 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, N28]

Este último entrevistado tem uma perspetiva que tende a priorizar o bem-estar da sociedade sob 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, liberdade e presunção da inocência (Machado & et al., 2018). De uma forma geral, todos os profissionais entrevistados contribuem para este debate contemporâneo, refletindo, também, acerca do vazio legal – percecionado como uma lacuna – existente em torno dos grandes dados; bem como, demonstrando preocupações acerca da potencial aplicação de *Big Data* no que concerne à defesa dos direitos humanos¹⁷.

¹⁷ Ver anexo VIII.

Parte IV
Conclusão

Parte IV: Conclusão

Este estudo ambicionou explorar as expectativas expressas por parte 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 de *Big Data*, enquanto técnica, à investigação, prevenção e repressão da criminalidade; desmistificar 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 de *Big Data* impulsiona no contexto de uma sociedade democrática. Além disso, perceber as transformações que a vigilância operou, paralelamente, ao desenvolvimento contextual tecnológico e aos acontecimentos sociopolíticos que se foram desencadeando; bem como, as percepções acerca disso. Este conjunto de temas complexos, diversos, mas uníssonos, foram estudados de forma multifacetada, no sentido de aceder e compreender a diversidade de expectativas e perspetivas que surgem. Além disso, devido às diferentes referências tecidas pelos diversos entrevistados, os recortes discursivos foram interpretados e considerados, sob 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.

Esta investigação, eminentemente qualitativa e interpretativa, caracterizou-se, assim sendo, por um movimento dinâmico e interativo, visando enquadrar os conceitos enunciados pelos entrevistados – o material empírico –, no arsenal científico analisado. A principal técnica de recolha de dados foi a análise qualitativa de conteúdo de entrevistas semiestruturadas, realizadas pelo Projeto EXCHANGE, que foram transcritas e posteriormente interpretadas e categorizadas. A análise categorizada dos recortes discursivos visou recolher as narrativas e significações da diversidade das expectativas e perspetivas profissionais que surgiram.

No que toca às barreiras encontradas durante este caminho de investigação – parte integrante de todo este processo – estas foram importantes para salientar a complexidade do tema e a sua compreensão atual quando o discurso efetivo não está concretizado e, por isso, preenchido de incertezas e considerações trémolas, por se tratar de uma área emergente e em desenvolvimento. Em primeiro lugar, porque se trata, efetivamente, de um tema embrionário, principalmente no que toca à sua relação com a criminalidade –

mais propriamente, enquanto meio de investigação, prevenção e repressão criminal. A inexistência de literatura que estabeleça esta relação e de trabalhos semelhantes ao aqui apresentado, dificulta o processo de análise intercalar dos dados empíricos. Consequentemente, o tema é 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, consequentemente, as condições societais, legislativas e sociopolíticas também, pelo que os estudos foram descritos e analisados, mas não serviram de base de comparação.

Não obstante, a inexistência de outro trabalho de investigação equiparável a este converte esta investigação numa contribuição única nacional e europeia para uma (melhor) compreensão do fenómeno *Big Data*, enquanto realidade social contemporânea, bem como, enquanto potencial técnica de investigação, prevenção e repressão da criminalidade organizada e transfronteiriça, sob a ótica dos profissionais que lidam, laboralmente, com a partilha transnacional de dados na UE, 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. Ou seja, é 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.

Neste sentido, este tema de investigação e este estudo almejam constituir-se parte de um vazio científico, construindo conhecimento nesta área e contribuir para a produção de valências científicas no que toca ao *Big Data* e à criminalidade – desmontar estas temáticas complexas, atuais e importantes, interlaçando-as e considerando-as sob a perspetiva de quem as manuseia. A análise das abordagens que teorizam o tema 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 sobre o seu arsenal teórico, estão embuídos, empiricamente, de características sociais, políticas, ambientais e contextuais – tanto o crime, como os grandes dados. Além disso, a inexistência de protocolos que regulamentam a inserção da técnica nos meios de investigação criminal, a lacuna legal que se verifica, de mãos dadas à lacuna científica teórica, dificulta a construção de perspetivas e expectativas por parte dos profissionais

entrevistados, que muitas vezes, constroem as suas significações ancoradas às mensagens veiculadas pelos meios de comunicação social.

No que toca à análise dos recortes discursivos, independentemente dos grupos profissionais considerados, considera-se a perspetiva promissora de que a inclusão de *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 desta perspetiva ser, em simultâneo, a expressa com maior frequência pelos entrevistados, existe uma consciencialização de que este trabalho é um fragmento-parte de um trabalho que deve ser maior, para ser 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 investigação, os recortes discursivos permitem delinear umas linhas gerais conclusivas sintéticas. De uma forma geral, foi notório que os NCP2 revelaram, por via das suas respostas, mais conhecimento sobre *Big Data* enquanto potencial ferramenta aplicada à investigação criminal. Nomeadamente, conhecimentos *micro*, detalhados, paralelamente à expressão de expectativas promissoras, positivas, verbalizando declarações permissivas à inclusão dos grandes dados e da sua partilha transnacional, como uma valência positiva para a prevenção e repressão da criminalidade transfronteiriça e organizadas. Assim, NCP2, por via de esboços detalhados e revelando mínimas resistências sobre a expansão e implementação de uma nova técnica no contexto da investigação criminal, foram o grupo profissional que demonstrou possuir maior conhecimento e familiaridade com *Big Data* (Van Lente, 2012; Brown & Michael, 2003; Borup et al., 2006), caracterizando-se o seu discurso como *instrumental* (Stevens et al., 2018). Os geneticistas forenses e NCP1, apesar de serem detentores, também, de expectativas promissoras e positivas acerca do tema, não demonstraram possuir um conhecimento tão detalhado e próximo do tema, caracterizando-se o seu discurso como *pragmático* (Stevens et al., 2018). De forma geral, demonstraram-se, maioritariamente, favoráveis à inclusão da técnica, no entanto, apresentaram argumentos organizados para refutar previsões (Van Lente, 2012). São a categoria profissional que se apresenta mais autoritária, tecendo considerações sobre obstáculos práticos e dificuldades empíricas na inclusão dos grandes dados, atualmente, nas práticas de investigação criminal (Van Lente, 2012; Lucivero et al., 2011; Borup et al., 2006), caracterizando-se o seu discurso como *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 da técnica – questões procedimentais, de validade científica, considerações acerca da fragilidade da validade da técnica e preocupações com o número e tipo de dados que inclui (Jasanoff & Kim, 2009; Borup et al., 2006; Brown & Michael, 2003; Machado & Silva, 2015). Inseridos nestas considerações cautelosas, os geneticistas forenses e NCPs revelaram maior preocupação sobre os riscos e perigos da inclusão de uma técnica inovadora contestada no seio da investigação criminal, quando comparados com as restantes categorias profissionais. Isto deve-se ao facto de se encontrarem em posições e contextos laborais que, tecnicamente, os aproxima desta partilha transnacional de dados e os torna mais conscientes para as potenciais problemáticas práticas encontradas. Esta aproximação, laboral e técnica, dos grandes dados enquanto potencial ferramenta de trabalho teve efeito, também, nas suas alegações contraditórias. Ou seja, NCPs e geneticistas forenses, muitas vezes, revelaram discursos ambivalentes, entre as vantagens e benefícios promissores de *Big Data* enquanto ferramenta de prevenção e combate à criminalidade, balanceando com as dificuldades da inclusão da técnica, bem como, a sua fragilidade científica (Van Lente, 2012; Lucivero et al., 2011; Borup et al., 2006) – o que reflete a importância que o contexto e meio laboral dos profissionais tem na sua construção de argumentos discursivos sobre temáticas e fenómenos. Além disso, também teceram considerações políticas, que se refletiram em visões politizadas sobre o fenómeno, convertendo-o numa prática dependente, diretamente, das instâncias e órgãos políticos.

Não obstante, NCPs foram o grupo profissional que mencionou, com maior ênfase a utilidade da técnica (Machado & Silva, 2015), podendo enquadrar-se o discurso de NCP2, no que toca à menção de *Big Data* como útil, sob a tipologia discursiva *pragmatismo biométrico*. Em contrapartida, visto que estes últimos profissionais não teceram considerações éticas e de direitos humanos, com tanto ênfase como os restantes grupos profissionais considerados, os professores/investigadores, geneticistas forenses, membros do corpo de investigação criminal, membros de Comissões de Proteção de Dados e todos os profissionais, excetuando NCPs, enquadram o seu discurso no *excepcionalismo genético* por ser vinculado a perspetivas éticas e de direitos humanos (Machado & Silva, 2015). Além disso, os entrevistados a desempenhar funções laborais como investigadores sociais foram a categoria profissional que mais vezes questionou quem tinha acesso aos dados, tecendo considerações avaliativas da mesma (Selin, 2008). De uma forma geral, uma súmula dos discursos dos diferentes profissionais, permite a

caracterização uníssonas de um discurso *modernista* (Stevens et al., 2018). No geral, todos os profissionais oscilam entre o grau de novidade da técnica, como Brown & Michael (2003) referiram e também sobre incertezas organizacionais, fazendo abstrações temporais com linguagem futurista. Tudo depende da sua posição laboral e, conseqüente, localização, como já referido. Ainda assim, é importante ressaltar que se verifica um desfazimento entre expectativas e o presente material, as expectativas residem nos discursos, mas dependem da prática que é inexistente (Brown & Michael, 2003).

Adotando uma lente mais concreta, no que toca às perspectivas que se prendem com os benefícios e vantagens do potencial uso de *Big Data* no combate e prevenção da criminalidade, destacam-se os geneticistas forenses e NCP2, com uma noção otimista do tema, na medida em que referem acreditar que esta tecnologia potencia(rá) o combate mais eficiente e eficaz da criminalidade, sendo uma ferramenta de investigação criminal valiosa – esta é, normalmente, a perspectiva de cientistas forenses com formação académica superior em Medicina Legal, Bioquímica, Ciências Forenses e Biologia (Machado, 2011; Maciel & Machado, 2014), ou seja, os entrevistados com formação académica superior no campo da saúde e da genética referem, com mais frequência, que *Big Data* contribuiria com maior eficiência no combate à criminalidade, demonstrando, frequentemente uma visão mais positiva e otimista da Ciência e da tecnologia em geral (Machado & Silva, 2015). *Big Data* é referido por estes profissionais como uma técnica com potencial utilidade operacional, podendo vir a ocupar uma posição científica, salientando a sua potencial qualidade e eficácia inserida nas práticas policiais de investigação criminal (Williams & Johnson, 2008; Machado & Silva, 2015). Estes profissionais, frequentemente, procedem a uma *neutralização* dos riscos associados às liberdades civis e direitos humanos, no que concerne à implementação e expansão de *Big Data* enquanto técnica de investigação criminal (Machado & Silva, 2016).

Nesta linha perspetivacional, NCP1 adota uma posição distinta: reconhece os potenciais benefícios, mas enfatiza a necessidade de se aumentarem os recursos e as valências económicas e profissionais, para que esta técnica se efetive. Além disso, foram a única categoria profissional que, ancorada sob uma opinião ambivalente, revelaram bastantes obstáculos laborais práticos à inserção de *Big Data* no seu quotidiano profissional – esta é uma das modalidades que Albert et al. (2009) apresentam no seu trabalho sobre os limites culturais dos cientistas sociais no campo da pesquisa em saúde. Além disso, é reflexo de uma das modalidades que Van Lente (2012) referiu: a criação de refutações e resistências face ao desenvolvimento tecnológico inovador. Este

posicionamento pode ancorar-se no facto de serem o grupo profissional que detém maior conhecimento sobre *Big Data* e sobre a forma como este se pode materializar – o elemento *conhecimento* como preditor da força atitudinal deste grupo profissional face às expectativas de *Big Data*. Tendo em conta que NCP1 é o grupo profissional com a visão mais oposta (Machado & Silva, 2015). Estas resistências surgem em relação a objetivos normativos: o *dever ser* da prática científica válida e emergem, temporalmente, da cultura (Brown et al., 2006). Em contrapartida, excetuando NCP2, todos os restantes profissionais entrevistados perspetivam *Big Data* como uma técnica que potencia a precisão e eficácia das decisões tomadas – ancora-se em algoritmos e estes últimos aumentam a precisão das decisões –, convertendo esta tecnologia numa ferramenta com um grau de objetividade maior. No que toca às representações sócio-profissionais relacionadas com os desafios éticos e sociais a enfrentar com o potencial uso de *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, membros de Comissões de Ética também com formação especializada em Direito e membros pertencentes a Organizações Não Governamentais/Direitos Humanos salientam 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. Surgem com uma perspetiva “crítica”, pessimista e defensora da proteção de dados e privacidade, ressaltando a necessidade da proteção de informação sensível (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). Este posicionamento aproxima-se da tipologia designada, por Williams & Johnson (2004) como *excecionalismo genético* (Williams & Johnson, 2004). Nestas perspetivas, são feitas, frequentemente, avaliações processuais da técnica de *Big Data*, ou seja, explorada a forma como os dados são recolhidos, analisados, processados, agregados, convertidos em algoritmos e, posteriormente, para que fins são utilizados (Albert et al., 2009) – expectativas a um nível *micro* (com elementos repletos de detalhes caracterizadores da técnica) (Van Lente, 2012).

Os profissionais ligados à vertente prática da aplicação das técnicas de *Big Data* – profissionais de investigação criminal e NCPs – surgem com uma orientação utilitarista, próxima do *pragmatismo biométrico* (Williams & Johnson, 2004), com 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). É notório que as expectativas destes profissionais se prendem a um nível *micro* porque envolvem elementos detalhados e característicos do *Big Data* (Van Lente, 2012). Ou seja, os profissionais que desempenham funções laborais em setores onde *Big Data* se poderá futuramente constituir-se enquanto 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 mais pessimista. Esta perspetiva reside no limbo entre os benefícios da (potencial) aplicação da técnica e a incerteza ancorada nos riscos que a mesma (potencial) aplicação pode acarretar. Brown & Michael (2003) referem que esta incerteza é mais acutilante e verifica-se eminentemente presente ancorada sob as expectativas dos profissionais intimamente envolvidos na produção do conhecimento, onde a sua experiência acerca das contingências de produção de conhecimento no laboratório, os tornam mais cautelosos (Brown et al., 2003). Estas expectativas residem num nível mais *macro* porque se tratam de esboços abstratos e abrangentes (Van Lente, 2012). No fundo, os profissionais que desempenham funções próximas do contexto onde a técnica tem possibilidade de se vir a materializar são mais conscientes dos riscos que a mesma acarreta para a sua prática laboral, quando comparados com outros grupos profissionais, à semelhança da conclusão a que o estudo de Machado & Silva (2015) chegara quando se debruçaram sobre as perspetivas públicas sobre os riscos e benefícios dos bancos de dados forenses de ADN, tendo em conta a influência do grupo profissional, nível educacional e idade. Estas visões envolvem-se, de forma dinâmica, entre as preocupações sobre o auto-cuidado e o bem-estar coletivo da sociedade, em simultâneo com a perceção de riscos e benefícios das tecnologias de vigilância – como *Big Data* – que permeiam a vida quotidiana. Estas são normatividades e subjetividades abordadas no contexto das aplicações da tecnologia genética no campo da Medicina e também reproduzidos em outros contextos sociotécnicos como o forense, tal como o estudo de Machado & Silva (2016) acerca das perceções públicas sobre a doação voluntária de ADN para a base de dados portuguesa.

Concluindo, é notório que existem variadas e distintas formas de observar o fenómeno de *Big Data* – repleto de sentidos, de avaliações, de interpretações e definições. Cada profissional possui um olhar *seu* sobre o que lhe fora 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 de *Big Data* um 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 dos 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 discursos 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 *Big Data* são especulações diferentes que não permitem situar o (potencial) desenvolvimento da técnica – este facto é uma boa oportunidade para traçar a forma pela qual uma técnica emergente (enquanto conjunto de tecnologias) está a ser moldada por diferentes visões durante a sua emergência. Assim, foi possível observar como as expectativas são constituídas em redes de conhecimento, projetadas em possibilidades futuras. Como técnica embrionária, *Big Data* é um objeto útil para pensar a forma como as expectativas se condensam em volta de questões particulares, em diferentes momentos, lugares e por diferentes razões técnicas, políticas e culturais. Continuamente, observaremos a forma como este futuro se desenvolverá.

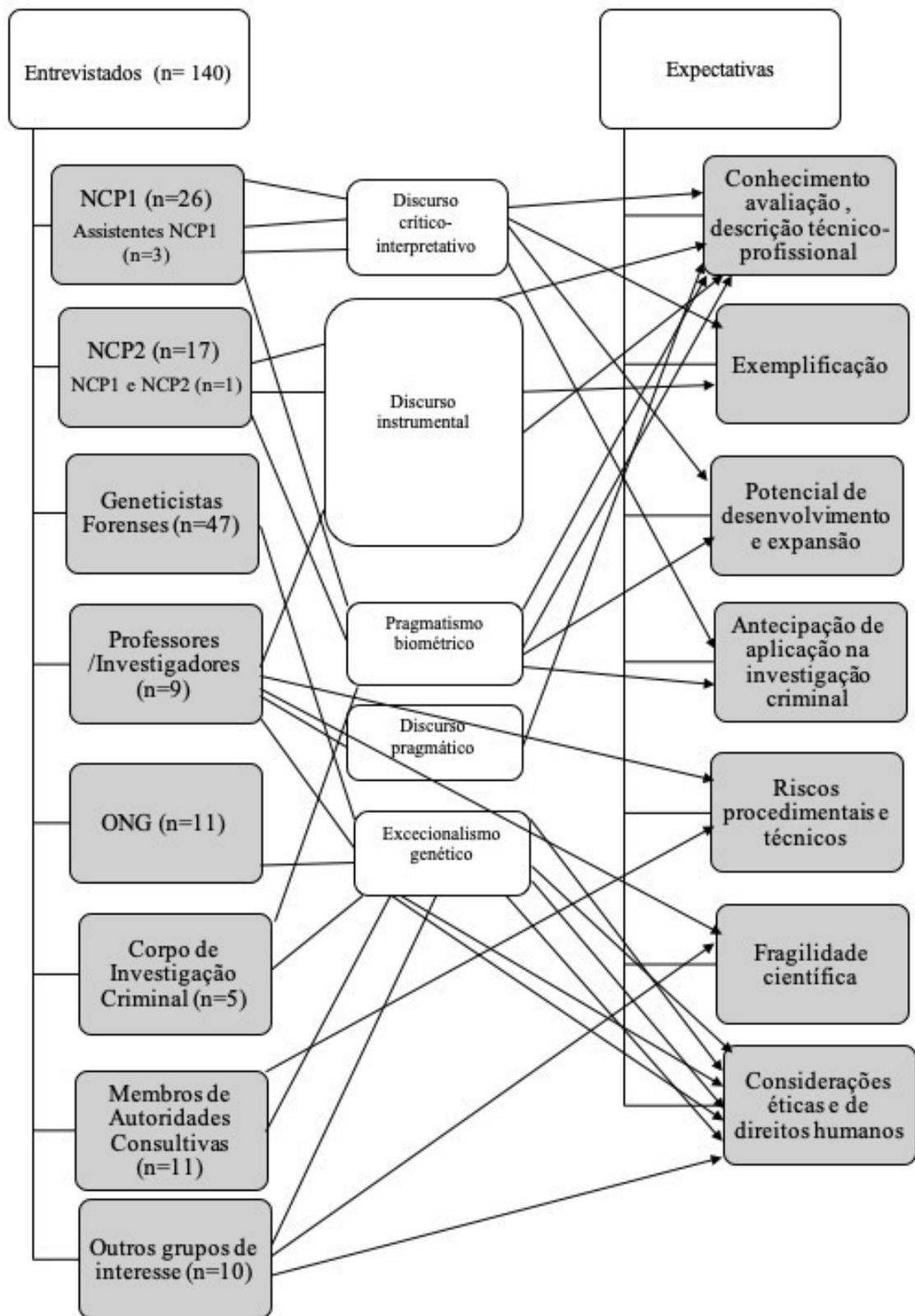


Figura 5: Caracterização dos discursos dos diferentes profissionais entrevistados.

Sob este contexto complexo, a nível prático, é importante que este estudo impulse 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 manuseamento de dados sobre cidadãos como 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 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 fluída. No fundo, 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 Constituição da República Portuguesa; Canotilho, 2005), *a atividade 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. Pois 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 isso, 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, mas também não se trata de algo neutro. 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 a 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 da 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 é que estes dados significam, como é que a análise a estes dados é distribuída e com que finalidades. Para além disso, importa construir bases sólidas reflexivas no conhecimento acerca da potencialidade do *Big Data* para se constituir exclusivo na produção de informação e na gestão de ações e decisões.

Urge dizer que entramos numa *Era* em que tudo é quantificado e as relações humanas perderam o seu sentido íntimo, porque tudo se reduz a números; à semelhança de 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 estas consequências sociais visíveis da expansão dos grandes dados, é importante ressaltar que *Big Data* não se trata de uma moda que se instala por um tempo definido, é 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 uma necessidade, conseqüente, de reconhecimento de quando se fala 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, 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 do *Big Data*, descortinar o lugar onde este se posiciona no campo filosófico (Matzner, 2016). Conjuntamente, num exercício de reflexão dual, importa estudar as questões levantadas por Snowden – “que tipo de sociedade é que nós queremos?” (Lyon, 2014). À semelhança do que Bauman (2009) refletira 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 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, ocorrem dentro de relacionamentos mediados digitalmente e são experimentados como produtivos e repressivos. Além disso, e ecoando a discussão anterior sobre 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 como é que lidamos com o surgimento de uma *Era de Grandes Dados*, visto que enquanto o fenómeno ocorre 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 são as suas entidades reguladoras (Boyd & Crawford, 2012). Esas constatações tornam evidente o facto de que, realmente, *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 também essenciais para a vida quotidiana. Isso cria um desafio moral para sociedades, porque as pessoas querem usar as próprias tecnologias que criam *Big Data*, mas também querem tentar controlar a forma como isso os 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 Comitê 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 regulatório em 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, sobre os princípios que inspiraram as anteriores regulamentações nacionais nos anos 90, sem resolver os problemas relativos à aplicação destes princípios aos novos cenários do *Big Data*. Embora o debate sobre estas duas maneiras diferentes de regulamentar a tecnologia ainda esteja aberto, parece difícil chegar a uma resposta conclusiva: o caso da regulamentação de *Big Data* levanta muitas perguntas sobre se o Regulamento 2016/679 representa a estrutura de proteção de dados para os próximos anos? Isto sugere que pode valer a pena reconsiderar toda a abordagem regulamentar da UE, para adotar uma proteção de dados mais flexível (mas não mais fraca) (Mantelero, 2017).

Deve haver um envolvimento público mais aberto 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 para o intercâmbio transnacional de dados, bem como, dos grandes dados, o estudo das perspetivas e atitudes dos profissionais envolvidos nesse intercâmbio é crucial. Assim, esta dissertação não deve ser lida nem entendida como uma declaração final sobre o assunto, mas antes como uma contribuição modesta para uma questão crucial de debate que está definida para se tornar 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. (2005). O conceito de futuro. In *Atas de Conferência “O Futuro não pode começar”* (pp. 7-49). Núcleo de Estudos de Sociologia, Universidade do Minho.

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

Araújo, E. (2011). A política de tempos: elementos para uma abordagem sociológica. *REVISTA DE CIÊNCIAS SOCIAIS-POLÍTICA & TRABALHO*, 34, 19-40.

Araújo, E. (2012). As identidades e a “política dos tempos sociais”. In *VII Congresso de Sociologia* (pp. 1-14). Associação Portuguesa de Sociologia (APS).

Araújo, E. (2013). A espera e os estudos sociais do tempo e sociedade. In Araújo & Duque (eds.), *Os tempos sociais e o mundo contemporâneo. Um debate para as ciências sociais e humanas* (pp. 9-27). Universidade do Minho: Centro de Estudos de Comunicação e Sociedade / Centro de Investigação em Ciências Sociais.

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*. Zahar.

Bogdan, R., & Biklen, S. (1994). *Investigação qualitativa em Educação: Uma introdução à teoria e aos métodos*. Porto: Porto Editora.

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 June* (pp. 13-14).

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

Cetina, K. (2009). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press.

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.

Cole, S., & Michael, L. (2006). The Social and Legal Construction of Suspects. *Annual Review of Law and Social Science* 2(12), 39–60.

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: Ensaio 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.

Felson, M., & Clarke, V. (1998). Opportunity Makes the Thief. Police Research Series Paper 98, Policing and Reducing Crime Unit. *Research, Development and Statistics Directorate. London: Home Office.*

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.

Flick, U. (2002). *Métodos qualitativos na investigação científica*. Lisboa: Monitor.

Fontes, M., & Araújo, E. (2013). (I) Mobilidades e redes científicas internacionais: Contextos e relações em mudança. In Araújo, E., Fontes, M. & Bento, S. (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). 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). Quem controla os nossos dados? *Público*, pp. 1–5.

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 (Org.), *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). 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.

Metcalf, J., & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society*, 3(1), 1-14.

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*. 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.

Starks, H., & Brown Trinidad, S. (2007). Choose your method: A comparison of phenomenology, discourse analysis, and grounded theory. *Qualitative health research*, 17(10), 1372-1380.

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 Dijck, 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>, accessed 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.

Yin, Robert K. (2011). *Qualitative Research from Start to Finish*. First Edit. Guildford Press.

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>.

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

Anexos

Anexo I – Distribuição dos/as entrevistados/as por país

Tabela 1. Distribuição dos/as entrevistados/as por país

Países	Entrevistas	Entrevistados/as
Alemanha	14	17
Áustria	1	1
Bélgica	2	3
Bulgária	1	3
Chipre	2	2
Eslováquia	1	1
Eslovénia	4	5
Espanha	4	5
Estónia	1	1
Finlândia	1	2
França	3	3
Hungria	3	3
Irlanda	1	1
Letónia	2	3
Lituânia	2	2
Luxemburgo	1	2
Malta	3	3
Noruega	2	2
Países Baixos	11	13
Polónia	12	14
Portugal	29	30
Reino Unido	16	16
República Checa	3	3
Roménia	2	2
Suécia	2	2
Outros¹⁸	1	1
Total	124	140

Fonte: Entrevistas realizadas no âmbito do projeto EXCHANGE (2015-2019).

¹⁸ Foi realizada uma entrevista a um representante da União Europeia.

Anexo II – Caracterização dos NCPs entrevistados em países operacionais

Tabela 2. Caracterização dos NCPs entrevistados em países operacionais

Formação Profissional	
Biologia, ciência forense e campos relacionados	20
Polícia	11
Direito	9
Tecnologias e sistemas informáticos	5
Outros	2
Organização	
Policial	27
Institutos Forenses	20
Tipo de trabalho	
Cooperação policial internacional	22
Laboratorial	21
Tecnologias e sistemas informáticos	4
Atividades relacionadas com Prüm	
NCP 1	26
NCP 2	17
Assistentes de NCP1	3
NCP 1 e NCP 2	1
Número total de participantes	47

Fonte: Entrevistas realizadas no âmbito do projeto EXCHANGE (2015-2019).

Anexo III – Caracterização dos geneticistas forenses entrevistados por país e organização onde trabalham

Tabela 3. Caracterização dos geneticistas forenses entrevistados por país e organização onde trabalham

País	
Alemanha	6
Eslovénia	1
Espanha	3
França	1
Hungria	1
Irlanda	1
Malta	1
Noruega	2
Países Baixos	6
Polónia	6
Portugal	14
Reino Unido	5
República Checa	1
Roménia	1
Organização	
Policia	8
Universidades / Institutos Forenses	41
Número total de participantes	49

Fonte: Entrevistas realizadas no âmbito do projeto EXCHANGE (2015-2019).

Anexo IV – Número de entrevistas realizado em cada estudo de caso

Tabela 4. Número de entrevistas realizado em cada estudo de caso (excetuando NCPs e geneticistas forenses)

Alemanha	8
Países Baixos	4
Polónia	6
Portugal	14
Reino Unido	11
Outros (Representante EU)	1
Número total de participantes	44

Fonte: Entrevistas realizadas no âmbito do projeto EXCHANGE (2015-2019).

Anexo IV - continuação

Tabela 5. Caracterização profissional dos *Stakeholders*

Organizações não governamentais Direitos Humanos	10
Entidade Supervisão/Regulação	11
Professor(a) Universitário(a)/Investigador(a)	9
Legislador	5
Órgãos de investigação criminal	5
Empresas privadas	2
Organizações não governamentais Ciência	1
Meios de comunicação social	1
Número total de participantes	44

Fonte: Entrevistas realizadas no âmbito do projeto EXCHANGE (2015-2019).

Tabela 6. Formação Educacional dos *Stakeholders*

Direito	19
Biologia e campos relacionados	11
Ciências Sociais e campos relacionados	7
Ciências Políticas	4
Sem informação	2
Outros	1
Número total de participantes	44

Fonte: Entrevistas realizadas no âmbito do projeto EXCHANGE (2015-2019).

Anexo V – Tabela de sistematização de resultados 1

Categoria Profissional/Grande categoria analítica	Informação sobre <i>Big Data</i>: conhecimento, avaliação, descrição técnico-profissional e exemplificação
NCP1	<ul style="list-style-type: none"> - Robustez - Conhecimento teórico - Desadequação laboral - complexo
NCP2	<ul style="list-style-type: none"> - Otimismo - Demarcação relativamente à polícia (<i>boundary work</i>) - Conhecimento prático-descritivo - Complexo - Visão politizada - Negócio
Geneticistas Forenses	<ul style="list-style-type: none"> - Conhecimento técnico-descritivo - Visão politizada - Negócio
Professores Universitários/Investigadores	<ul style="list-style-type: none"> - Conhecimento descritivo e técnico profundado - Conhecimento concetual - Conhecimento informático - Visão politizada
Organizações Não Governamentais (Direitos Humanos e Ciência)	<ul style="list-style-type: none"> - Complexidade - Afastamento
Membros do Corpo de Investigação Criminal	<ul style="list-style-type: none"> - Visão politizada
Membros de Autoridades Consultivas	<ul style="list-style-type: none"> - Visão politizada
Stakeholders	<ul style="list-style-type: none"> - Demarcação relaticamente à polícia (<i>boundary work</i>)

Anexo VI – Tabela de sistematização de resultados 2

Categoria Profissional/Grande categoria analítica	Expectativas futuras do potencial aplicação de <i>Big Data</i>: desenvolvimento, expansão e antecipação de aplicação na investigação criminal
NCP1	<ul style="list-style-type: none"> - Pessimismo - Obstáculos - Adaptação - Em desenvolvimento - Especificação de ilícitos penais (criminalidade transfronteiriça)
NCP2	<ul style="list-style-type: none"> - Otimismo - Utilidade - Em desenvolvimento - Conjugação com outras tecnologias - Especificação de ilícitos penais (criminalidade transfronteiriça)
Geneticistas Forenses	<ul style="list-style-type: none"> - Otimismo - Em desenvolvimento - Útil - Promessa não cumprida
Professores Universitários/Investigadores	
Organizações Não Governamentais (Direitos Humanos e Ciência)	
Membros do Corpo de Investigação Criminal	- Útil
Membros de Autoridades Consultivas	<ul style="list-style-type: none"> - Robustez e pessimismo - Em desenvolvimento - Especificação e ilícitos penais: terrorismo
Stakeholders	- Otimismo

Anexo VII – Tabela de sistematização de resultados 3

Categoria Profissional/Grande categoria analítica	Percepção dos riscos e perigos
NCP1	- Riscos procedimentais e técnicos - Fragilidade científica
NCP2	- Riscos procedimentais e técnicos
Geneticistas Forenses	- Riscos procedimentais
Professores Universitários/Investigadores	- Fragilidade científica
Organizações Não Governamentais (Direitos Humanos e Ciência)	- Fragilidade científica
Membros do Corpo de Investigação Criminal	
Membros de Autoridades Consultivas	- Riscos procedimentais e técnicos (direito)
Stakeholders	

Anexo VIII – Tabela de sistematização de resultados 4

Categoria Profissional/Grande categoria analítica	Pareceres éticos e de direitos humanos
NCP1	- Vazio legal (necessidade de regulamentação)
NCP2	- Vazio legal (necessidade de regulamentação) - Violação direitos humanos
Geneticistas Forenses	- Vazio legal (necessidade de regulamentação) - Preocupações éticas e de direitos humanos
Professores Universitários/Investigadores	- Vazio legal (necessidade de regulamentação) - <i>Big Brother</i> - <i>Snowden</i> - PIDE - Direitos, liberdades e garantias
Organizações Não Governamentais (Direitos Humanos e Ciência)	
Membros do Corpo de Investigação Criminal	- Vazio legal (necessidade de regulamentação) - Considerações acerca da proteção de dados - Sociedade de vigilância excessiva
Membros de Autoridades Consultivas	- Vazio penal (necessidade de regulamentação) - Considerações relativas à privacidade
Stakeholders	