

Instituto Superior de Ciências Policiais e Segurança Interna



Miguel Rodrigues Pereira

Aspirante a Oficial de Polícia

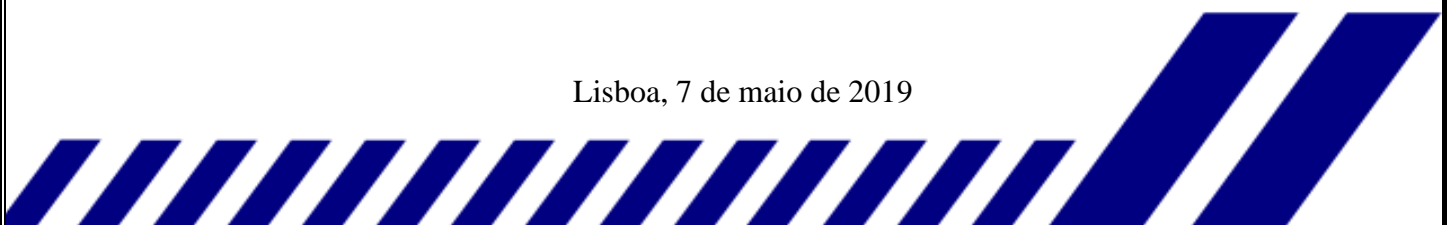
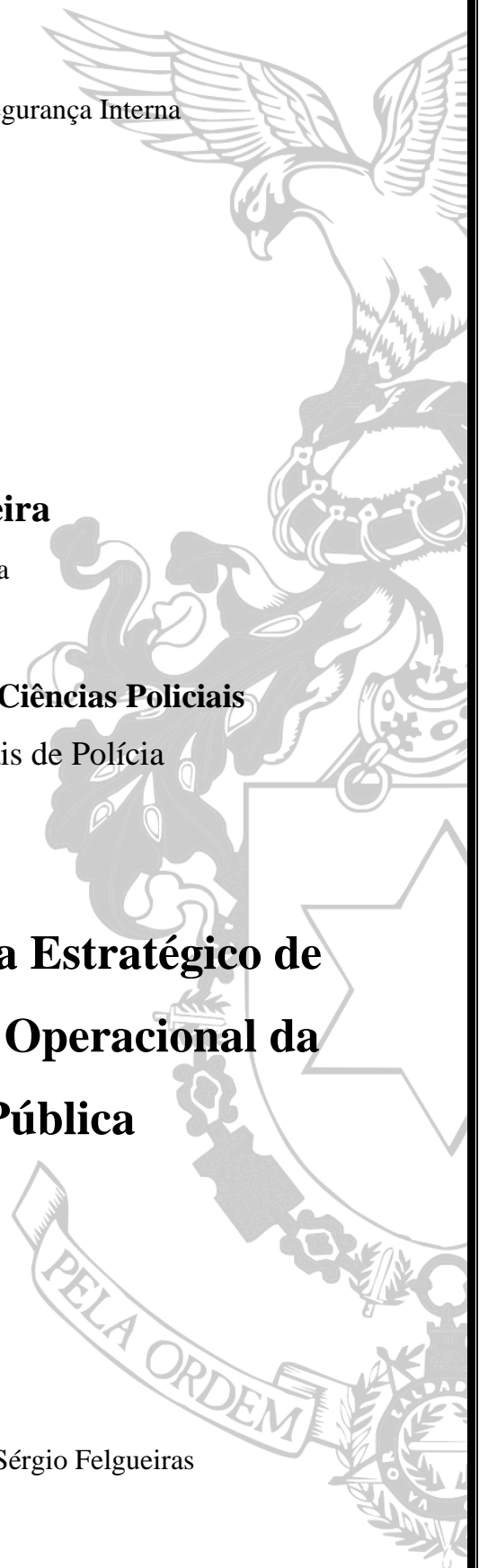
Dissertação de Mestrado Integrado em Ciências Policiais

XXXI Curso de Formação de Oficiais de Polícia

***Big Data: O caso do Sistema Estratégico de
Informação, Gestão e Controlo Operacional da
Polícia de Segurança Pública***

Orientador: Superintendente, Prof. Doutor Sérgio Felgueiras

Lisboa, 7 de maio de 2019



Instituto Superior de Ciências Policiais e Segurança Interna



Miguel Rodrigues Pereira

Aspirante a Oficial de Polícia

Dissertação de Mestrado Integrado em Ciências Policiais

XXXI Curso de Formação de Oficiais de Polícia

***Big Data: O caso do Sistema Estratégico de
Informação, Gestão e Controlo Operacional da
Polícia de Segurança Pública***

Dissertação apresentada ao Instituto Superior de Ciências Policiais e Segurança Interna
com vista à obtenção do grau de Mestre em Ciências Policiais, elaborada sob a
orientação do Superintendente, Prof. Doutor Sérgio Felgueiras.



Estabelecimento de Ensino: Instituto Superior de Ciências Policiais e Segurança
Interna

Curso: XXXI CFOP

Orientador: Superintendente, Prof. Doutor Sérgio Felgueiras

Título: *Big Data: O caso do Sistema Estratégico de
Informação, Gestão e Controlo Operacional da
Polícia de Segurança Pública*

Autor: Miguel Rodrigues Pereira

Local de Edição: Lisboa

Data de Edição: maio de 2019



Dedicatória

À minha avó, Amália Rodrigues.



Agradecimentos

Foram cinco rigorosos anos. Uma etapa difícil, apenas tornada ultrapassável pela participação ativa da minha família no meu desenvolvimento pessoal e profissional.

Em especial, um agradecimento à minha mãe, por fazer de mim o que sou hoje e por representar um modelo perfeito de pessoa e de mãe.

Ao meu pai agradeço pelo apoio que sempre me deu, e continua a dar, na prossecução dos meus objetivos por via do trabalho, da dedicação e da resiliência.

Agradeço à minha irmã por me pôr a cabeça no lugar quando ela começa a andar à deriva, o que não raras vezes acontece.

Um agradecimento especialmente vincado à Inês Nunes. Nunca as palavras serão suficientes para descrever a minha gratidão. És e sempre serás uma peça fundamental na minha vida.

Agradeço em particular ao Mário Vaz pelos pareceres informáticos essenciais ao desenvolvimento deste trabalho. Ah, e pela amizade também! Assim como ao Vitor Patrico (nome mal escrito de propósito), com o qual sempre pude contar, nos momentos mais difíceis. Para ti, um muito obrigado.

À Ana Fragoso, sem a qual, a realização desta dissertação não tinha sido possível.

Deixo ainda um agradecimento às minhas duas grandes famílias Baseau e Manhuço, que permanecerão eternamente no meu coração.

Agradeço ao meu curso, XXXI CFOP, pelo caminho percorrido lado a lado. Apesar de todos os altos e baixos, somos e seremos sempre *primus inter pares*.

Reservo um agradecimento especial ao meu orientador, Exm^o. Sr. Superintendente Sérgio Felgueiras pela disponibilidade demonstrada ao longo do desenvolvimento da presente Dissertação.

Por fim, agradeço ao Instituto Superior de Ciências Policiais e Segurança Interna, pelo desenvolvimento pessoal e profissional que me proporcionou.

Resumo

Nos últimos anos, tem-se verificado uma evolução acelerada dos sistemas e tecnologias de informação e comunicação e, conseqüentemente, a uma crescente dificuldade no armazenamento e tratamento de dados, dando origem ao *Big Data*. Este é um fator impulsionador de mudanças sociais sem paralelo, com as quais as polícias têm de lidar. Assim sendo, as complexidades do Policiamento no século XXI obrigam à adoção de soluções inovadoras e adaptadas a uma realidade em constante mutação. Neste contexto, esta investigação pretende olhar para o sistema de informação central na PSP, de modo a compreender a sua implementação e desenvolvimento. Para este efeito, recorreu-se a uma abordagem qualitativa da problemática de investigação, procedendo à análise de conteúdo de quatro documentos estratégicos referentes à implementação e desenvolvimento do SEI. Os resultados da análise permitiram-nos identificar a maioria dos componentes fundamentais do *Big Data* nos documentos constitutivos do *corpus*. Contudo, verificámos que a avaliação do desempenho das tecnologias e a objetividade da análise são duas ideias ausentes nos documentos estratégicos analisados. Esta ausência, podendo ela ser sintomática da não aplicação destas ideias na PSP, revela-se preocupante pelo facto de serem consideradas fundamentais no âmbito do *Big Data*. Finalizando, a presente dissertação possibilitou a abertura de novos caminhos de investigação, designadamente, estudar a operacionalização destas estratégias do ponto de vista do *Big Data*, verificando, concretamente, a sua aplicação.

Palavras-chave: *Big Data*; Polícia; tecnologia; informação; estratégia.

Abstract

In recent years there has been an accelerated evolution of information and communication systems and technologies and, consequently, an increasing difficulty in storing and processing data, giving rise to *Big Data*. This is a driver of unparalleled social change that police have to deal with. As such, the complexities of policing in the 21st century require the adoption of innovative and adapted solutions to the changing reality. In this context, this research intends to look at the central information system in the PSP, in order to understand its implementation and development. To this end, a qualitative approach to the research problem was used, analyzing the contents of four strategic documents related to the implementation and development of SEI. The results of the analysis allowed us to identify most of the fundamental components of *Big Data* in the constituent documents of the corpus. However, we have found that the evaluation of technology performance and the objectivity of the analysis are two missing ideas in the strategic documents analyzed. This absence, which may be symptomatic of the non-application of these ideas in PSP, is worrisome given these ideas are considered fundamental in the scope of *Big Data*. Finally, the present dissertation allowed the opening of new research paths, namely, to study the operationalization of these strategies from the point of view of *Big Data*, verifying, concretely, its application.

Keywords: *Big Data*; police; technology; information; strategy.

Índice Geral

| | |
|--|------|
| Agradecimentos | i |
| Resumo | ii |
| Abstract..... | iii |
| Índice Geral | iv |
| Índice de Tabelas | vii |
| Índice de Anexos | vii |
| Lista de Siglas e Abreviaturas | viii |
| Capítulo I – Introdução..... | 1 |
| Capítulo II – Enquadramento teórico | 3 |
| 1. Uma definição de <i>Big Data</i> | 3 |
| 1.1. Informação | 4 |
| 1.2. Tecnologia | 7 |
| 1.3. Métodos | 8 |
| 1.4. Impacto | 10 |
| 2. Soluções..... | 13 |
| 2.1. <i>Hadoop framework</i> | 13 |
| 2.2. A solução da Google..... | 13 |
| 3. <i>Big Data</i> no contexto securitário contemporâneo | 15 |
| 3.1. Policiamento Tradicional..... | 16 |
| 3.2. Policiamento Orientado para os Problemas | 17 |
| 3.3. Policiamento Orientado pela Inteligência..... | 20 |
| 3.4. Desafios e Soluções | 22 |
| 3.5. Polícia e <i>Big Data</i> | 24 |
| 4. O SEI | 28 |
| 4.1. Descrição do Sistema..... | 28 |
| 4.2. Potencialidade e desafios | 31 |
| Capítulo III – Método | 33 |
| 1. Abordagem Qualitativa | 33 |

| | |
|---|----|
| 2. <i>Corpus</i> | 34 |
| 3. Instrumento: Análise de Conteúdo | 35 |
| 4. Procedimento | 36 |
| Capítulo IV – Apresentação, Análise e Discussão de Resultados | 38 |
| 1. Impacto | 40 |
| 2. Polícia | 41 |
| 3. Métodos | 42 |
| 4. Tecnologias | 44 |
| 5. Informação | 45 |
| Capítulo V - Conclusões | 46 |
| Bibliografia | 51 |
| ANEXOS | 60 |

Índice de Tabelas

| | |
|--|----|
| Tabela 1. Distribuição das u.r. pelas categorias | 38 |
| Tabela 2. Distribuição das u.r. pelas subcategorias..... | 38 |

Índice de Anexos

| | |
|--|----|
| Anexo 1. Autorização para a realização de entrevistas | 61 |
| Anexo 2. Grelha de análise de conteúdo (com exemplos). | 62 |

Lista de Siglas e Abreviaturas

| | |
|----------|---|
| ADN | Acido Desoxirribonucleico |
| COMETPOR | Comando Metropolitano do Porto |
| CP | Carro de Patrulha |
| CRIME | <i>Consolidated Records Intelligence Mining Environment</i> |
| DL | <i>Digital library</i> |
| EUA | Estados Unidos da América |
| EUSEI | Equipa Única SEI |
| GPS | <i>Global Positioning System</i> |
| GSI | Gabinete de Sistemas de Informação |
| HDFS | <i>Hadoop Distributed File System</i> |
| IBM | <i>International Business Machines</i> |
| IoT | <i>Internet of Things</i> |
| NBD-PWG | <i>NIST Big Data Public Working Group</i> |
| OCR | <i>Optical Character Recognition</i> |
| OLSEI | Oficial de Ligação SEI |
| PESI | Plano Estratégico de Sistemas de Informação |
| POI | Policimento Orientado pela Inteligência |
| POP | Policimento Orientado para os Problemas |
| PSP | Polícia de Segurança Pública |
| RAM | <i>Random Access Memory</i> |
| RFSEI | Regulamento de Formação SEI |
| RNSI | Rede Nacional Segurança Interna |
| SARA | <i>Scanning, Analysis, Response, Assessment</i> |
| SCoT | Sistema de Contraordenações de Trânsito |
| SEI | Sistema Estratégico de Informação |
| SI | Sistemas de Informação |
| SIGAE | Sistema Integrado de Gestão de Armas e Explosivos |
| U.R. | Unidade de Registo |
| VDL | <i>Very Large Digital Library</i> |

Capítulo I – Introdução

A evolução tecnológica nos últimos anos tem ganho proporções vertiginosas. Estas tecnologias constituem hodiernamente uma ferramenta de uso fácil e recorrente, constituindo uma parte fundamental do nosso quotidiano. A comercialização destas tecnologias permite a criação de conexões entre todas as pessoas por via dos dispositivos informáticos comumente utilizados. O facto de cada um desses dispositivos se apresentar como um “sensor” e, portanto, produtor de dados, levanta um conjunto de alterações no paradigma dos sistemas de gestão de informação. É neste contexto, de uma sociedade cada vez mais familiarizada com as tecnologias e a massificação da produção da informação, que surge o *Big Data*, um termo pouco consensual, mas que abrange várias dimensões da sociedade. Quando a quantidade de informação produzida, recolhida e armazenada é de tal forma avassaladora que não pode ser tratada segundo métodos simples, surge a necessidade de criação de soluções inovadoras que não só permitam responder às necessidades que este fenómeno acarreta, mas também aproveitar as suas potencialidades.

Neste cenário acelerado e tecnológico em que vive a sociedade, onde a atuação policial é profundamente mediatizada, tem-se verificado uma complexificação e diversificação do crime e uma crescente exigência para as Polícias. Já não é somente espectável um modelo de policiamento tradicional que se resume ao mero patrulhamento e resposta às ocorrências. Fala-se em policiamento comunitário, policiamento orientado pela inteligência, policiamento orientado para os problemas, entre outros, isto é, quer-se uma polícia moderna, eficiente e, simultaneamente, próxima do cidadão, que resolva efetivamente problemas concretos das populações e não adote apenas meras medidas isoladas que apenas atenuem os sintomas.

A conjuntura económica do nosso país implica que tenha havido uma redução drástica dos recursos (Fernandes, 2014), sejam eles humanos, financeiros ou tecnológicos, disponíveis em vários setores do Estado, nomeadamente, no setor responsável pela administração interna, que tutela a PSP. Assim sendo, tem sido exigido à PSP que faça mais e melhor com menos recursos, o que significa que, incontornavelmente, tem de ser feito um esforço de racionalização dos meios ao seu dispor, por forma a acompanhar a evolução da sociedade, sem perder eficácia no cumprimento das missões que lhe estão atribuídas. O SEI foi uma das soluções concebidas pela PSP para responder às exigências da sociedade moderna, também esta cada vez informatizada.

É neste contexto que surge a pertinência da presente investigação. O SEI sempre foi desenvolvido consoante as necessidades da PSP, orientadas não só pela sua missão, mas também pelo serviço público ao cidadão. Contudo, não existem investigações prévias que exprimam se a conceção e o desenvolvimento deste sistema de informação tiveram, e têm, em conta todas as potencialidades e desafios que o *Big Data* acarreta. Deste modo, consideramos pertinente estabelecer os seguintes objetivos de investigação: fazer uma contextualização de âmbito geral do fenómeno do *Big Data*, identificando potenciais soluções; identificar desafios inerentes à complexidade da ligação entre o *Big Data* e a segurança; fazer uma caracterização do SEI; refletir sobre o processo de desenvolvimento do SEI, tendo em conta a exploração conceptual do *Big Data*. Assim sendo, a nossa problemática de investigação é “Numa perspetiva de *Big Data*, como é que o SEI foi desenvolvido?”. Para o efeito iremos, inicialmente, explorar este conceito, procurando apresentar uma definição consensual para o mesmo, identificando as suas vertentes mais relevantes e explorando potenciais soluções que tenham sido apresentadas até agora. Seguidamente, restringiremos o nosso escopo com o intuito de tentar compreender a importância do *Big Data* no contexto securitário contemporâneo e, mais concretamente, no contexto das atividades desenvolvidas pelas instituições policíacas. Iniciaremos essa análise com uma explanação sucinta dos modelos de policiamento mais prevalentes no contexto policial internacional, na expectativa de que eles possam dar uma primeira alavancagem na exposição das necessidades específicas das organizações policíacas quanto à sua estrutura, objetivos, missão e cultura. Posteriormente, faremos uma súpula dos desafios e soluções apresentadas até agora pelos diversos autores que associaram o *Big Data* à polícia, aproximando-nos, assim, da realidade em estudo. Findando o enquadramento teórico, faremos uma descrição do SEI enquanto sistema de informação da PSP fazendo, desde logo, o levantamento das potencialidades e dos desafios com os quais este sistema se depara.

Na sequência do enquadramento teórico será descrito o método utilizado neste trabalho. Optámos por uma abordagem qualitativa de índole exploratória a qual se justifica pelo contexto em que a investigação está a ser desenvolvida, isto é, na ausência de estudos prévios sobre o *Big Data* na PSP. Para a constituição do *corpus* foram escolhidos documentos de índole estratégica e feita a respetiva análise de conteúdo, constituindo este o nosso instrumento de investigação. Constarão ainda neste ponto os procedimentos desenvolvidos ao longo desta dissertação, apresentando os critérios definidos para a categorização e análise. Sequencialmente os resultados da análise serão apresentados e discutidos, através da interpretação dos mesmos no contexto da investigação.

Finalmente, das conclusões apresentadas constarão reflexões sobre os resultados obtidos e sobre todo o enquadramento teórico desenvolvido, sempre na expectativa de abrir novos caminhos para investigações futuras.

Capítulo II – Enquadramento teórico

Steve Lohr (2013), na pesquisa que fez em busca da origem do termo *Big Data*, averiguou que John Mashey, um cientista da empresa *Silicon Graphics*, foi o primeiro a popularizar este conceito através das exposições que fazia sobre o tema, apesar de não ter nenhum artigo científico publicado onde o refira. Lohr teve acesso a uma das apresentações que Mashey usava como suporte para as suas exposições, onde abordava o tema para explicar “um conjunto de problemas” associados à contínua expansão da tecnologia. Nestas apresentações, Mashey já identificava potenciais necessidades futuras que adviessem desta tendência, tais como, o aumento da capacidade de armazenamento, processamento e transmissão, recorrendo para isso a uma análise comparativa dos 10 anos anteriores.

Contudo, muito mudou desde a década de 90, altura em que Mashey popularizou o conceito. À medida que as tecnologias de informação se comercializaram e banalizaram, passando a fazer parte da hodiernidade das pessoas e empresas, este ganhou outros contornos. Naturalmente, o surgimento de novos autores que abordam este tema sob novas perspetivas vai distanciando o conceito de *Big Data* do sentido atribuído primeiramente por Mashey.

1. Uma definição de *Big Data*

Atualmente, o termo *Big Data* encontra-se disperso em ramos tão diversificados como o da informática, medicina, economia, gestão, biologia, segurança, entre incontáveis outros. Embora seja utilizado nesta quantidade vastíssima de áreas, este não as afeta a todas da mesma forma (Manyika *et al.*, 2011). Esta dispersão do conceito por áreas tão vastas tem gerado incongruências quanto ao seu significado. De facto, ao longo dos anos, e à medida que o termo foi sendo mais utilizado, vários autores propuseram as suas próprias definições, alterando ou ignorando outras que tivessem sido apresentadas até então. Os autores De Mauro, Greco, e Grimaldi (2016) defendem que, devido a este caos no estado da arte, surge a necessidade primordial de estabelecer os limites e as características deste conceito para que possa haver um desenvolvimento coerente da literatura neste campo. Para o efeito, foram

identificados quatro temas fundamentais nos quais as várias definições até então apresentadas se inserem: informação, tecnologia, métodos e impacto.

1.1. Informação

A *digitization* foi um processo com elevada importância histórica neste contexto. Karen Coley (2006), descreveu este processo dizendo que consistia na transformação de sinais analógicos em sinais digitais. Esta transformação não passa da representação de uma realidade em código binário para que o computador o possa ler e replicar essa representação. Vejamos o exemplo da Google, descrito por James Somers (2017). Em 2004 a Google iniciou um projeto de *scanning* de todos os livros a que pudesse ter acesso. O grande objetivo desta iniciativa de *digitization* em massa era simples: “facilitar a procura de livros relevantes para as pessoas – especialmente os livros que não encontrariam de outra forma, por não serem impressos” (Google, 2019). Para o efeito, esta empresa montou um dispositivo que virava as páginas dos livros automaticamente e tirava fotografias das páginas. Contudo, os computadores não compreendiam o que aquelas páginas diziam, ou seja, a única maneira de descodificar o conteúdo das páginas era colocar um humano a ler. Deste modo, havia ainda mais um passo a dar para que estas representações da realidade pudessem ser analisadas.

Assim sendo, o passo seguinte é chamado *datafication* por Mayer-Schönberger e Cukier (2013). Esta estruturação permite um tratamento superior dos dados permitindo a extração de padrões e correlações entre eles, que doutro modo seria impossível. Revisitando o exemplo na Google mencionado no parágrafo anterior, o passo seguinte ao *scanning* das páginas dos livros foi aplicar um *software* de *Optical Character Recognition* (OCR) de modo a que o computador conseguisse distinguir as letras e palavras contidas naqueles livros (Coyle, 2006). Deste modo, eles tornaram possível, de entre outros exemplos, uma análise da evolução linguística, das tendências relativas à frequência do uso de determinadas palavras, da dimensão lexical da língua inglesa e até denotaram a repressão de certas palavras no período de governo Nazi, durante a II Guerra Mundial, ao analisarem livros alemães publicados ao longo do século passado (Michel *et al.*, 2011). Abordando outro exemplo, para entendermos melhor o alcance deste processo, Mayer-Schönberger e Cukier (2013) exploraram o caso de Shigeomi Koshimizu, professor no Instituto Avançado de Tecnologia Industrial, em Tokyo. Este professor e a sua equipa decidiram começar a recolher dados da postura corporal das pessoas enquanto estão sentadas. Para o efeito, colocaram sensores em 360 pontos nos bancos dos carros de modo a medir a pressão que era exercida em cada um

deles, sendo esta pressão medida numa escala de 0 a 256. Os resultados desta recolha permitiram concluir que cada pessoa tem uma postura única. Em ambiente de testes, o sistema que foi criado e alimentado com os dados recolhidos nesta experiência conseguia distinguir as pessoas pela sua postura com uma precisão de 98%. Pode parecer uma informação inútil, mas estes dados estavam a ser implementados num sistema antirroubo para as viaturas, que só iniciariam a marcha se reconhecessem a pessoa que se sentasse no banco. Provavelmente, na ausência desta tecnologia, não só nunca seria possível desenvolver estes mecanismos, como também nunca extrairíamos valor com este elevado nível de precisão sobre a postura corporal de um ser humano sentado. Muito menos ainda pensaríamos que essa informação poderia ser revelante e útil.

Assim, esta progressão leva-nos a estabelecer a hierarquia dados-informação-conhecimento-sabedoria. Jennifer Rowley (2007) identificou os elementos essenciais de cada um destes níveis:

- *Data* ou dados, não têm significado ou valor porque não se encontram contextualizados nem interpretados, são descrições objetivas, elementares, desorganizadas e não processadas de coisas, eventos, atividades ou transações;
- *Information* ou informação, são dados contextualizados do modo a representar uma realidade, processados obtendo um significado ou organizados e interpretados de maneira a adquirir valor e propósito;
- *Knowledge* ou conhecimento, apesar de não haver consenso quanto à sua natureza, do ponto de vista da lógica desta hierarquia, podemos defini-lo com sendo a combinação de dados e/ou informação organizada e processada originando experiência, aprendizagem acumulada e especialização aplicada a um determinado problema ou atividade;
- *Wisdom* ou sabedoria, é o nível mais ambíguo desta hierarquia e muitas vezes ignorado quando se fala de sistemas de informação, contudo, a definição mais consensual para Rowley (cit. in Rowley, 2007) é: “A capacidade de aplicar o comportamento mais adequado, tendo em conta o que é conhecido (*knowledge*) e o que é considerado como bom (considerações éticas e sociais)” (p. 177).

Gu Jifa (2013) sumariza a definição de cada um dizendo: “Dados são o nível mais básico; Informação adiciona um contexto; Conhecimento adiciono a forma como se usam; Sabedoria adiciona quando e o porque de utiliza-los” (p. 714). Para De Mauro *et al.* (2016), a informação é o “combustível fundamental” (p. 124) para as aplicações de *Big Data*, que

tratam essencialmente da atribuição de propósito à informação de modo a gerar conhecimento.

Para De Mauro *et al.* (2016) o crescimento da informação disponível deve-se à crescente quantidade de dispositivos conectados à Internet e equipados com sensores digitais (smartphones, câmaras fotográficas, tablets, etc.). Se, por um lado, estes dispositivos possibilitam a *digitization*, por outro, a conexão à internet permite que esses dados sejam recolhidos e organizados segundo critérios específicos, transformando-os em informação. Assim, a crescente variedade de formatos de dados afigura-se uma característica importante dos dados que são produzidos. Podemos ter dados em formato de imagens, registos de *streaming data*¹, vídeos, dados de localização GPS, em suma, todos os dados produzidos por qualquer tipo de sensor que proceda à *digitization*. Deste modo, a multiplicidade dos tipos de dados e a sua coexistência é um dos maiores desafios associado ao manuseamento de *Big Data* na atualidade (Manyika *et al.*, 2011). Segundo Dhankhad (2019), estima-se que nos últimos dois anos tenham sido criados 90% da totalidade dos dados existentes no mundo. A maioria desses dados estão destrutturados e, portanto, tornam-se difíceis de analisar. Os sistemas, diga-se, comuns, apenas estão preparados para lidar com dados estruturados. Como tal, desta premissa surgem duas exigências: do ponto de vista da tecnologia, a de desenvolver sistemas com capacidade de armazenamento e processamento superiores que correspondam aos requisitos dos métodos de tratamento; do ponto de vista dos métodos, de conceber métodos de estruturação e análise dos dados.

Diversos autores (Beyer & Laney, 2012; Zaslavsky, Perera & Georgakopoulos, 2013; Zikopoulos, Eaton, deRoos, Deutsch & Lapis, 2011) definem *Big Data* focando-se nos atributos dos dados. Laney (2001) propôs o modelo dos 3 V's (volume, velocidade e variedade) para exprimir a necessidade de formas inovadoras no processamento da informação, apesar a associação ao conceito de *Big Data* só ter sido feito mais tarde pelos autores supracitados. Houve ainda autores que introduziram alterações a esta definição, tais como, veracidade (Schroeck, Shockley, Smart, Romero-Morales & Tufano, 2012), valor (Dijcks, 2013), complexidade e destrutturação (Intel IT Center, 2012; Suthaharan, 2014). Contudo, através da pesquisa realizada, verificamos que existem outras dimensões do *Big Data* para além dos atributos dos dados, que serão exploradas adiante nesta dissertação.

¹ Emissão contínua de dados.

1.2. Tecnologia

Neste sentido, para aprofundar o tema da tecnologia em *Big Data*, é importante ter como ponto de partida que os sistemas comumente utilizados raramente têm capacidade de processamento e armazenamento suficientes para lidar com as características deste tipo de dados. A definição introduzida pela Microsoft Research (2013) alude precisamente a esta vertente considerando que se trata de um “termo cada vez mais usado para descrever o processo de aplicação de um vasto poder de processamento – o mais recente no *machine learning* e na inteligência artificial – em conjuntos de informação em massa e altamente complexos”. Deste modo, para Microsoft Research (2013) o fator preponderante para o *Big Data* é o poder de processamento.

De Mauro *et al.* (2016) denotaram uma relação entre os repositórios de *Big Data* e as *digital librareis* (DLs) ou bibliotecas digitais. Candela *et al.* (2007) define este conceito como sendo: “Uma organização, potencialmente virtual, que de modo abrangente recolhe, gere e preserva a longo prazo conteúdo digital rico, e fornece à comunidade de utilizadores funcionalidades especializadas sobre o seu conteúdo, de qualidade mensurável e de acordo com políticas codificadas” (p.16). Os requisitos para ser considerada uma biblioteca digital de elevada dimensão ou *very large digital libraries* (VDLs), de acordo com Candela, L., Manghi, P. e Ioannidis, Y. (2012) são: ter um volume elevado de entidades ou um tamanho elevado; ter uma elevada exigência de rapidez na recolha e no processamento das entidades; e existir uma vasta variedade de entidades. Resumindo, quando as DLs têm um volume, velocidade e variedade elevados, carecem de “tecnologias e métodos especializados” para o seu manuseamento (Candela *et al.*, 2012, p.61). Deste modo, pode estabelecer-se um paralelo entre a definição apresentada de VDLs e as definições acima enunciadas de *Big Data*. Como tal, conclui-se que o desenvolvimento das tecnologias e métodos associados ao *Big Data* são necessários para continuar a desenvolver o estudo da gestão da informação destas bibliotecas, que podem servir como fonte de alimentação de dados para os sistemas de informação de *Big Data* (De Mauro *et al.*, 2016).

Outro dos problemas que surge associado à gestão de *Big Data* reside na sua transmissão, pois as redes de comunicação precisam de sustentar transferências de dados de elevada dimensão e rapidez. Por outro lado, devido à sua complexidade, estes sistemas requerem técnicas de *benchmarking* específicas para avaliar o seu desempenho (Xiong *et al.*, 2013). Este autor defende que devem ser estabelecidos sistemas de referência para permitir que esta avaliação seja feita considerando-a, portanto, fundamental no desenvolvimento e

melhoramento de sistemas posteriores, tais como o *Hadoop framework*, que será explorado adiante nesta dissertação. Através desta avaliação é que é possível determinar, não só se o sistema foi desenvolvido convenientemente, mas também se está ajustado às exigências específicas de cada organização.

Quanto ao armazenamento, a necessidade de armazenar maiores extensões de dados em dispositivos mais pequenos também tem sido uma das dificuldades emergentes com a evolução das tecnologias de informação. Segundo a Lei de Moore (2006), o número de *transistors* que se conseguem introduzir numa placa de sílica tende a duplicar a cada 18 a 24 meses, aumentando a capacidade de armazenamento dos dispositivos. Contudo, embora seja verdade que esta previsão de Moore (2006) se tem verificado, também a quantidade de dados produzida tem aumentado exponencialmente mantendo, assim, a pressão constante de evoluir os sistemas de armazenamento para que estes possam acompanhar esta tendência incremental. Apesar desta pressão, consta-se que “as ferramentas disponíveis para lidar com o volume, velocidade e variedade do *Big Data* têm melhorado imenso nos últimos anos” (McAfee & Brynjolfsson, 2012, p. 66), não estando fora do alcance financeiro das pessoas e das instituições, pois a maioria destas tecnologias é *open source*².

No decorrer da nossa pesquisa, verificámos a existência de algumas propostas de definição de *Big Data* cujo foco se centra nas necessidades tecnológicas associadas a este conceito. Para Ward e Barker (2013), *Big Data* refere-se ao “armazenamento e análise de *data sets* grandes e complexos usando uma série de técnicas incluindo, mas não exclusivamente, NoSQL *MapReduce* e *machine learning*” (p. 2). De outro ponto de vista, apesar do NBD-PWG (2014) aludir aos atributos dos dados, restringe a sua posição dizendo que *Big Data* são “vastos *data sets*, primariamente caracterizados por volume, velocidade e/ou variedade, que requerem arquitetura escalável para um armazenamento, manipulação e análise eficientes.” (p. 6). Importa sublinhar que este autor fez menção a uma característica relevante como a escalabilidade do sistema, ou seja, garantir um design da sua arquitetura que permita expansões da sua capacidade.

1.3. Métodos

Nos pontos anteriores foram identificados os problemas da velocidade de produção e da variedade de dados. Neste tema, estão englobados os métodos específicos usados no

² Está disponível em fontes abertas, ou seja, é de aquisição livre.

tratamento de dados em resposta a estes problemas. No processo de extração de valor é necessário, primeiramente, aplicar métodos de melhoria da qualidade dos dados, reestruturando-os e reorganizando-os. A qualidade dos dados mede-se através do grau de aptidão para o uso por parte dos utilizadores e dos consumidores, estendendo-se até onde o valor de utilização do *Big Data* pode ser realizado e posteriormente aproveitado (Wang & Strong, 1996; Wang, Ziad & Lee, 2002).

Uma vez que a qualidade esteja garantida, são necessários métodos para a análise dos dados e extração de conhecimento relevante para a tomada de decisão. É-nos perceptível que os métodos utilizados para o tratamento desta quantidade de dados não podem ser os métodos estatísticos comuns. Esses métodos foram compilados por Manyika *et al.* (2011) e Chen, Chiang e Storey (2012) numa lista dos procedimentos mais utilizados, tais como: *cluster analysis*, algoritmos genéticos, processamento de linguagem natural, *machine learning*, redes neurais, modelos preditivos, modelos de regressão, análise de redes sociais, análise de sentimentos, processamento de sinais e visualização de dados (De Mauro *et al.*, 2016). Contudo, tal como referido no ponto anterior, a evolução constante da tecnologia, impõe que estes métodos estejam em constante revisão, atualização, mutação e aperfeiçoamento.

Perante este contexto, cada vez mais as organizações têm na sua base sistemas de informação alimentados por dados recolhidos e organizados pelos seus funcionários. A partir desta realidade, surgem dois aspetos que devem ser considerados pelas referidas organizações: investimento em formação específica para os funcionários (Chen *et al.*, 2012; Manyika *et al.*, 2011); introdução de uma mudança cultural generalizada para que os funcionários lidem disciplinadamente com os dados (Buhl, Röglinger, Moser & Heidemann, 2013). Esta mudança cultural deve ser feita não só ao nível dos funcionários, mas também ao nível dos líderes das organizações, que não raras vezes “carecem de conhecimentos sobre o valor do *Big Data* bem como do modo como esse valor é alcançado” (Manyika *et al.*, 2011, p. 12).

Boyd e Crawford (2012), defendem que a implementação destas soluções de *Big Data* nos seus sistemas e a integração das mesmas no processo de tomada de decisão não é suficiente para garantir o aproveitamento total destas soluções. Deve então haver um processo de avaliação contínuo do desempenho das aplicações e uma constante busca por novas possibilidades de inovação. Dumbill (2013) formula a sua definição de *Big Data* dizendo que o “*Data*” pode ser considerada “*Big*” quando a capacidade dos métodos e tecnologias convencionais é excedida, sendo necessárias abordagens alternativas. Fisher,

DeLine, Czerwinski e Drucker, (2012) complementam esta perspetiva dizendo que o conceito de “*Big*”, em termos de dimensão, está ligado à Lei de Moore (2006) e, consequentemente com a capacidade armazenamento dos sistemas comerciais.

1.4. Impacto

As aplicações para a gestão de *Big Data* têm revelado uma enorme capacidade de adaptação às diferentes áreas onde são empregues, seja em investigação científica ou no setor empresarial. Isto permite a sua crescente aplicação em mais áreas da sociedade, levando a que, progressivamente, sejam integradas nos processos de tomada de decisão e resolução de problemas. Segundo Dhankhad (2019), esta é razão principal pela qual nos devemos focar na recolha e tratamento de dados, pois estes, quando processados adequadamente, eles permitem a extração de conhecimento que suportará as nossas decisões, tornando-as mais fundamentadas, independentemente da área de negócio. No entanto, é importante ter em conta que até agora, segundo McAfee e Brynjolfsson (2012), “O poder do *Big Data* não apaga a necessidade da perspetiva ou da compreensão humana” (p. 65). Deste modo, pode-se dizer que, tal como já referido, a tendência é de as empresas garantirem uma atualização constante dos seus sistemas e das capacidades analíticas para se manterem relevantes no setor onde estão inseridas.

Noutra perspetiva, a crescente produção de dados, de entre os quais dados pessoais, levanta a problemática da proteção da privacidade. Neste aspeto, este fenómeno pode acarretar consequências negativas, tais como, fugas de informação, acesso e divulgação ilegítimas de informações pessoais, etc. Esta matéria tem gerado discussões em vários países sobre a legitimidade e sobre a conveniência do uso desses dados, seja para que fim for. Os próprios consumidores dos “produtos” de *Big Data*, sabendo da possibilidade destes acontecimentos, começam a levantar entraves ao uso dos seus dados pessoais iniciando um processo de desequilíbrio entre a privacidade e a utilidade (Manyika *et al.*, 2011). Cody Agnellutti (2014) defende que enquanto essa luta decorre, perdemos o foco do potencial do tratamento destas informações. Não obstante, ressalva que é uma questão que não pode ser ignorada.

Derivado ainda deste problema, um dos aspetos a ter em consideração quanto ao *Big Data* é o segurança dos sistemas. Apesar do nos trazer grandes benefícios “a recolha, transmissão, processamento, análise, armazenamento e gestão de *Big Data* inevitavelmente aumenta o risco de fugas de informação” (Chen & Yan, 2016, p. 13). A colmatação destas

falhas de segurança, segundo Manyika *et al.* (2011), deverá ser feita por via de “mecanismos tecnológicos e políticos” (p. 11), de modo, não só a evitar que elas sejam exploradas, mas também a estabelecer responsabilidades quando tal acontece. Assim sendo, ao garantirmos a segurança e a proteção dos dados pessoais também estamos a aumentar o grau de confiança das pessoas na abordagem que o *Big Data* faz aos dados e à informação, o que significa que poderemos começar a reequilibrar a balança da privacidade e da utilidade.

Uma preocupação crescente nos setores empresariais mais dependentes da informação é o acesso à mesma. Com o objetivo de inovar e desenvolver soluções para os seus problemas, as empresas têm progressivamente recorrido a informação proveniente de múltiplas fontes externas (Manyika *et al.*, 2011). Neste sentido, há instituições que são recetivas à disponibilização sem custos da sua informação, outras estão abertas à possibilidade de venda da informação, enquanto que outras restringem completamente o acesso a essa informação. O que se verifica é que a detenção de um determinado tipo de informação pode ser um fator de valorização no mercado numa determinada área, o que gera políticas de partilha de informação distintas, consoante o tipo de informação e a própria cultura competitiva da empresa. Manovich (2012) defende que devem ser tomadas medidas reguladoras do acesso à informação por forma a permitir uma acessibilidade imparcial e impeditiva da monopolização da informação que possa determinar posições dominantes no mercado onde essas empresas estão inseridas. Segundo Boyd e Crawford (2012), a criação destas separações pode ser um potencial inibidor da evolução neste setor.

Há ainda outro impacto a ter em conta aquando da interiorização do *Big Data* no seio das empresas. De facto, quando uma empresa se torna *data-driven*³ ela é tendencialmente forçada a reconsiderar a sua estrutura e todos os processos da organização tendo em conta a disponibilidade de nova informação que, tal como exposto anteriormente, pode ser uma vantagem perante os outros atores no mercado (De Mauro *et al.*, 2016). McAfee e Brynjolfsson (2012) corrobora esta ideia dizendo que “Quanto mais as empresas se caracterizam no sentido de se tornarem *data-driven*, melhor é o seu desempenho nas medidas adotadas para atingir objetivos financeiros e operacionais” (p. 65). Contudo, Manyika *et al.* (2011) deixa o aviso de que o setor público fica mais facilmente alheio a estas evoluções devido à falta da competitividade que se verifica no setor privado. Os sistemas já implementados neste setor não raras vezes são incompatíveis com a integração sistemas de *Big Data*, o que impede a análise dos dados de modo a extrair valor dos mesmos (Manyika

³ Termo usado para as empresas/organizações que integram nas suas decisões as informações tratadas através de tecnologias e métodos de *Big Data*.

et al., 2011). Assim sendo, “os líderes das organizações e os decisores políticos deverão ter em conta o modo como as estruturas das empresas evoluem em consonância com o mundo do *Big Data* se quiserem determinar como otimizar a criação de valor ao nível das firmas individuais, dos setores e das economias como um todo” (Manyika *et al.*, 2011, p. 12).

Em suma, vários autores fizeram sobressair os efeitos do *Big Data* na sociedade. Boyd e Crawford (2012) definem *Big Data* como “um fenómeno cultural, tecnológico e académico” (p. 663) baseado em 3 elementos: tecnologia, pela maximização do poder computacional e da precisão algorítmica; análise, pela identificação de padrões em *data sets* de elevadas dimensões; mitologia, pela crença de que *data sets* de elevada dimensão fornecem uma forma superior de *intelligence*, carregando consigo uma aura de verdade, precisão e objetividade (De Mauro *et al.*, 2016). Consideremos ainda Mayer-Schönberger e Cukier (2013), que descrevem *Big Data* segundo três mudanças principais na forma como analisamos informação: mais dados, porque todos os dados são usados em vez de uma só amostra; mais desorganização, porque até os dados incompletos e imprecisos podem ser usados assim como os completos mas limitadores; correlação, porque se torna uma operação fundamental para desvendar a causalidade dos acontecimentos de modo a tomar decisões.

Por fim, revendo todos estes temas os autores De Mauro *et al.* (2016) propuseram uma definição construída com base em todas as variadas abordagens ao *Big Data*. Por conseguinte, segundo estes autores, a sua análise permitiu-lhe concluir que os conceitos nucleares deste termo são:

- Volume, Velocidade e Variedade, remetendo para os atributos dos dados e da informação;
- Tecnologias e Métodos, aludindo aos requisitos necessários ao tratamento e manuseamento deste tipo de informação;
- Valor, para descrever a transformação da informação em conhecimento que pode ser benéfico para as empresas e para a sociedade.

Conjugando todos os itens abordados ao longo dos temas com os núcleos conceptuais supramencionados, podemos definir *Big Data* como “um ativo de informação caracterizado por um volume, velocidade e variedade elevados de tal forma que requer tecnologia específica e métodos analíticos para a sua transformação em valor” (p. 131). Referir-nos-emos ainda nesta dissertação ao “fenómeno do *Big Data*” para explicar todas as mudanças que o surgimento e a proliferação destes ativos de informação acarreta tanto para a sociedade como para as organizações.

2. Soluções

2.1. *Hadoop framework*

O *Hadoop* é uma *framework*⁴ concebida pela Apache que tem como objetivo principal permitir o processamento de uma quantidade vastíssima de dados através da distribuição do processamento por vários *nodes*⁵. Tem nele incorporado um sistema de distribuição de ficheiros, *Hadoop Distributed File System* (HDFS), que permite a transferência rápida entre os *nodes* e o servidor central (Thuma, 2018). Este é um sistema completo com várias aplicações responsáveis por cada tarefa constitutiva do processo de tratamento dos dados e também pela redistribuição do processamento destas aplicações pelos *nodes*. Deste modo, este torna-se um sistema completo e versátil, facilmente adaptável às exigências de diferentes potenciais consumidores desta *framework*.

Tal como acima referido, o HDFS tem a função de distribuir o armazenamento e o processamento dos dados mal eles sejam introduzidos no sistema. Este sistema interage com o *Hadoop* de forma a que primeiro seja feita uma distribuição dos dados pelos *nodes* e só depois seja feito o armazenamento propriamente dito, tendo ainda nele incorporado um sistema de credenciação, programável pelo gestor do sistema (Dhankhad, 2019).

Este sistema funciona por blocos de informação, distribuindo-a pelos vários *nodes*, mas garantindo a redundância da mesma para garantir que esta não é desperdiçada (Zikopoulos *et al.*, 2011). O *MapReduce* é uma *framework* de análise de dados com linguagem complexa, dividida em duas tarefas distintas: *map* e *reduce*. É o “coração do *Hadoop*” (Zikopoulos *et al.*, 2011, p. 60). Esta aplicação, devido à sua complexidade, requer programadores com experiência para trabalhar nela, os quais são incomuns. No entanto, existem ferramentas que podem ser conjugadas com o *MapReduce* para simplificar o seu desenvolvimento e manutenção.

2.2. A solução da Google

A Google, como empresa detentora do maior repositório de informação do mundo, deparou-se com vários problemas na gestão dessa informação. Devido a este facto, sempre foi do seu interesse investigar a melhor maneira de tratar esta quantidade de informação.

⁴ *Software Framework* é uma plataforma concreta ou conceptual onde o código comum com funcionalidades genéricas pode ser seletivamente especializado ou sobreposto pelos desenvolvedores ou utilizadores (Technopedia, 2019).

⁵ *Node* é qualquer sistema ou dispositivo conectado a uma rede (TechTerms, 2019).

Neste sentido, aproveitou os sistemas que foram sendo criados para o efeito (*Hadoop*), melhorando-os através da integração de novas aplicações adequadas às dificuldades de manuseamento. Passaremos, então, a explicar sucintamente a função de cada uma destas componentes.

O *HBase* é uma base de dados NoSQL⁶ de suporte ao HDFS. O objetivo desta base de dados é reorganizar os dados para permitir um armazenamento mais fácil, por um lado, e por outro, permitir o acesso facilitado aos dados. Portanto, estes são introduzidos, distribuídos e armazenados no HDFS de forma a que fiquem facilmente acessíveis. A transferência de dados entre o HDFS e a *HBase* são feitos recorrendo a uma ferramenta chamada *Sqoop* (Dhankhad, 2019). Esta está concebida para importar dados de vários tipos de bases de dados, tais como, Oracle⁷ e MySQL⁸ para o HDFS (para ser feita distribuição) e exportar dados do HDFS para as restantes bases de dados. Portanto, esta ferramenta seria essencial para complementar a leitura, armazenamento e distribuição de dados dos vários sistemas em uso na PSP, tais como, o SEI, SCoT, SIGAE, etc., apresentando-se como uma possível solução para a baixa compatibilidade dos dados presentes nestes sistemas, tal como será referido adiante nesta dissertação.

O *Spark* é uma *framework*, semelhante ao *MapReduce* na sua integração com o *Hadoop*, contudo, esta foi concebida para o processamento de dados de forma muito mais eficiente (por RAM) e rápida. Sumariamente, as vantagens que esta *framework* apresenta em relação a qualquer uma das outras acima referidas é ter uma melhor capacidade de processamento de *data streams*, ser mais facilmente programável e funcionar autonomamente (Kalron, 2018). No entanto, esta tecnologia é largamente mais dispendiosa do ponto de vista monetário do que as alternativas mencionadas.

Devido à complexidade inerente ao *MapReduce* desenvolveu-se a aplicação *Pig* no sentido de permitir que seja possível manusear o *Hadoop* passando menos tempo a escrever em programas de *mapper* e *reducer* e direcionando o foco para a análise de dados no manuseamento do *Hadoop*, simplificando a linguagem (Zikopoulos *et al.*, 2011).

⁶ NoSQL é uma base de dados não relacional. No site mongoDB (2019) é explicado que este tipo de base de dados tem melhor desempenho a lidar com elevados volumes de dados estruturados, semiestruturados ou destrutturados em constante mudança.

⁷ A Oracle, ou Oracle RDBMS, é um sistema de gestão de base de dados multimodelo desenvolvida pela Oracle Corporation (Oracle, 2017).

⁸ MySQL é um sistema de gestão de base de dados relacional (DB-Engines, 2019).

O *Hive* é uma ferramenta semelhante ao *Pig* em função, contudo, está mais preparada para extração, transformação e carregamento de informação de bases de dados estruturadas e acrescenta o suporte do *Cloudera Search* (Dhankhad, 2019).

De modo a fornecer um ambiente de trabalho para o utilizador foi desenvolvido o *Cloudera Search*. Este é um motor de pesquisa com interface *userfriendly*. Segundo Olesker (2012) esta ferramenta “fornece uma integração do *Hadoop*, equipado com o HDFS, com o *MapReduce*” (p. 6).

O *Hue* é o centro de comando do sistema da Google. Permite aceder a todos as ferramentas integrados do *Hadoop*, facilitando a gestão de cada um deles e possíveis atualizações ou alterações (Dhankhad, 2019).

O sistema foi desenvolvido pela Google para solucionar o problema de tratamento e análise de uma enorme quantidade de dados. Sendo uma *open-source*, esta ferramenta não acarreta custos na sua aquisição e está construída de forma a permitir a sua utilização por outras empresas ou organizações, tendo a possibilidade de adaptar este *framework* às suas necessidades, aproveitando, assim, os seus recursos disponíveis. Para utilizar esta ferramenta, as empresas e organizações têm de verificar os seguintes pré-requisitos:

- Desenvolver o *software* próprio consoante as necessidades da empresa;
- Ter *nodes* pelos quais o processamento possa ser distribuído;
- Formar uma equipa especializada para a manutenção, desenvolvimento e permanente avaliação do sistema.

A ideia patente no artigo de Dhankhad (2019) é de que a agregação das várias ferramentas vai permitir um tratamento completo do *Big Data*, ultrapassando, deste modo, muitas das dificuldades associadas ao manuseamento de grandes quantidades de informação. Assim sendo, perante todo este enquadramento teórico, esta é um proposta de solução eficiente, e simultaneamente, adaptável aos vários ramos, para todas as questões levantadas por este fenómeno.

3. *Big Data* no contexto securitário contemporâneo

Bayley (2016) refere no seu artigo “*The Complexities of 21st Century Policing*” que o conceito de segurança e tem-se complexificado ao longo das últimas décadas e, como tal, o conceito de policiamento tradicional não é suficiente para lidar com este novo paradigma. Este autor identifica seis vertentes nas quais o policiamento se tornou mais complexo: nas suas tarefas, ao nível das exigências públicas, estratégias, tecnologias, responsabilidades e

recursos. Consequentemente, o policiamento adotou uma abordagem inteiramente nova no controlo da criminalidade, nomeadamente, para prevenir o crime de forma proativa, e não através de dissuasão reativa. Isso envolve abordar as circunstâncias, tanto iminentes quanto remotas, que promovam a criminalidade. A Polícia aceita, assim como os estudiosos, que punir criminosos pelos atos cometidos não é suficiente para garantir a segurança pública (Nagin, Solow & Lum, 2015).

Perante este panorama, também a Polícia se foi tentando adaptar não só face às mudanças aceleradas acima enunciadas, mas também face à evolução nos vários ramos da ciência. Esta nova abordagem ao *policing* enquanto vertente académica da atividade policial veio fundamentar novas decisões e procedimentos policiais com base em métodos alternativos à mera experiência profissional e ao costume. Por conseguinte desta matriz de pensamento, têm surgido novos modelos de policiamento que abordam e estruturam a atividade policial de formas distintas. Segundo Elias (2018), não há modelos de policiamento puros, “muitos países apostam cada vez mais num modelo de policiamento híbrido, mesclado, holístico e tecnológico, procurando congregiar diversos elementos de outros modelos para melhor responder à complexidade crescente das nossas sociedades” (p. 113).

3.1. Policiamento Tradicional

O modelo de policiamento tradicional tem sido o modelo utilizado pelas polícias a nível internacional desde a sua formação. As críticas feitas a este modelo têm sustento na evidência de que é ineficaz no combate à criminalidade, baseando-se numa componente predominantemente reativa. Na generalidade, as estratégias mais utilizadas neste tipo de policiamento são, segundo Fernandes (2014) e Skogan e Fryll (2004):

- Aumento do número de elementos policiais, na expectativa de que este aumento, só por si, origine uma redução da criminalidade;
- Policiamento (quase) aleatório por todas as partes da comunidade como meio de prevenção, ainda que os estudos feitos relativamente a este ponto demonstrem que esta estratégia não tem efeitos significativos na redução da criminalidade e no aumento do sentimento de segurança;
- Resposta rápida às ocorrências de modo a aumentar a probabilidade de interetar os suspeitos logo após o cometimento dos crimes;
- Investigações no seguimento dos crimes cometidos;

- Aplicação geral da coerção e das detenções, resolvendo as ocorrências singularmente e, portanto, apenas reduzindo a criminalidade a curto prazo.

Fernandes (2014) sumariza este modelo dizendo que “pode ser caracterizado pela centralidade do poder coercivo da lei penal como principal meio de resolução de incidentes, apostando na ideia da omnipresença policial como base da prevenção” (p. 175). Grandes mudanças sucederam o estudo feito pela *Kansas City Police Department* quando começaram a introduzir variações num fator da sua estratégia de policiamento, designadamente, número de polícias em patrulha, não sendo verificado nenhum impacto na criminalidade (Ratcliffe, 2016). Este foi o estudo que despoletou tantos outros que revelaram a ineficiência deste modelo de policiamento, dando, assim, início ao processo de reformulação do conceito de policiamento e ao surgimento de novas modalidades.

3.2. Policiamento Orientado para os Problemas

Herman Goldstein foi quem concebeu este modelo de policiamento, definindo-o como “um plano abrangente para melhorar o policiamento no qual a elevada prioridade associada à resolução de problemas substantivos molda a organização policial, influenciando todas as mudanças no pessoal, na organização nos procedimentos” (Goldstein, 1990, p. 32, como referido em Fernandes, 2014). Para Goldstein (1979) a profissão de polícia não se resume à mera aplicação da lei, mas sim, à resolução de vários problemas, onde a aplicação da lei é apenas “um meio para atingir um fim” (p. 242).

O policiamento orientado para os problemas (POP) tem como objetivo principal propor soluções para a ineficiência da Polícia pois esta apenas se tem limitado à resposta o mais célere possível das ocorrências. Na perspetiva do policiamento tradicional, a Polícia investe os seus recursos na resolução das ocorrências individualmente, não as interligando e, portanto, nunca identificando a sua origem. Para Goldstein (1995), as organizações policiais ainda têm uma estrutura demasiada forte e rígida, argumentando que os níveis mais baixos da hierarquia policial requerem mais flexibilidade e autonomia na tomada de decisão. Esta atuação contraria não só as missões atribuídas à polícia, mas também a própria perceção que a comunidade tem da atividade policial (Skogan & Fryll, 2004). Assim sendo, o POP está direcionado para a resolução dos problemas da comunidade, sendo essencial, para o efeito, a análise criminal. Para pôr em prática essa análise, ele propôs o modelo SARA (*Scanning, Analysis, Response, Assessment*).

Na sua conceção inicial, este modelo seria aplicado ao nível das instituições policiais (macro), estando cada uma delas a cargo da resolução dos problemas na sua jurisdição. Contudo, a implementação do POP revelou-se mais eficaz quando essa tarefa era aplicada aos comandantes de esquadra (nível micro), atribuindo-lhes, assim, a autonomia, flexibilidade e responsabilidade, que este autor defendia que deveriam ter (Fernandes, 2014). A dificuldade que esta perspetiva levantou prende-se com o panorama acima apresentado no qual as instituições policiais ainda se apresentavam com hierarquias funcionais muito rígidas e, portanto, muito formais. Algumas tentativas de implementação deste modelo não têm sido bem-sucedidas e, por vezes, apenas o são a curto prazo. Como tal, para Scott e Kirby (2012, p. 5) é necessário ter atenção às seguintes questões:

- A necessidade imperativa de resposta às ocorrências diárias pode sobrepor-se à adoção do método de atuação orientado para os problemas;
- A resposta aos problemas é focada nas detenções e na reação em vez de se focar na prevenção;
- Os níveis hierárquicos mais elevados não desempenham uma liderança ativa e não fornecem suporte organizacional suficiente para a resolução dos problemas;
- À medida que os elementos se vão habituando ao modelo de atuação do POP, vão sendo substituídos por outros menos dedicados a essa perspetiva;
- Os elementos de base vão-se tornando mais céticos com o tempo, e, portanto, reduzem o grau de atuação;
- Os elementos de base não percebem ou não aceitam o POP como constitutivo de trabalho de polícia;
- A responsabilidade pela resolução de problemas é unicamente delegada nos elementos de base, não fornecendo uma gestão adequada a esses elementos;
- A Polícia carece de compreensão quanto à forma de trabalhar eficientemente com instituições externas;
- Um ambiente político interno instável pode distanciar a organização de um desenvolvimento devido do POP.

Apesar destas dificuldades, várias avaliações que foram sendo feitas às Polícias que implementaram este modelo de policiamento revelam resultados positivos, quando comparado com o modelo tradicional anteriormente existente. De um modo geral, verificou-se uma redução da criminalidade e melhoria da perceção da comunidade em relação à Polícia. Consequentemente, foi possível continuar a atribuir maiores níveis de autonomia

aos comandantes e aos patrulheiros, o que moldou a própria estrutura cultural e organizacional em torno deste modelo de policiamento.

Por fim, Scott e Kirby (2012, pp. 3-70) fizeram o levantamento de um conjunto de passos essenciais não só com vista à implementação deste modelo de policiamento, mas também ao desenvolvimento, aperfeiçoamento e avaliação do POP nas instituições policiais. Numa fase preparatória:

- Compreender o conceito de policiamento orientado pelos problemas;
- Estudar as tentativas de implementação anteriores e, caso não tenham sido bem-sucedidas, reconhecer quais foram as falhas;
- Perceber de que modo o POP se pode relacionar com outros conceitos de policiamento;
- Identificar o enquadramento correto do POP na atividade policial da instituição em questão;
- Diagnosticar os pontos fortes e os pontos fracos da instituição policial;
- Integrar os elementos no processo de implementação dando a conhecer as motivações que levaram a esta escolha e as alterações que serão feitas;
- Gerir a mudança organizacional;
- Clarificar as funções e as responsabilidades;
- Garantir recrutamento, formação e seleção adequados.

Quanto à estruturação de instituição:

- Decidir entre uma especialização ou generalização do POP;
- Descentralizar a organização onde for possível, centralizando onde é necessário;
- Desenvolver sistemas para identificar e definir problemas;
- Desenvolver a capacidade da instituição de analisar problemas;
- Desenvolver a capacidade da instituição para avaliar problemas;
- Otimizar o desempenho dos elementos.

Quanto à aplicação do POP:

- Organizar e atribuir responsabilidades;
- Providenciar suporte administrativo para quem vai resolver os problemas;
- Gerir o período de serviço dos elementos para facilitar a resolução de problemas;
- Cultivar e gerir o estabelecimento de parcerias eficientes;
- Documentar os projetos do POP;
- Reconhecer e celebrar os projetos bem-sucedidos na resolução de problemas.

Quanto à manutenção do POP:

- Monitorizar o processo de mudança organizacional;
- Cultivar qualquer relação com outros setores que possam ser benéficos ao POP (cidadãos, empresas, tribunais, instituições governamentais, academias, comunicação social, etc.);
- Inculcar valores organizacionais e hábitos vantajosos para o POP;
- Desenvolver os sucessores.

A aplicação do POP tem revelado resultados positivos, aumentando o sentimento de segurança dos cidadãos e reduzindo a criminalidade (Fernandes, 2014). Esta é uma alternativa ao policiamento tradicional que oferece melhores condições para o cumprimento das missões atribuídas à Polícia, priorizando uma atuação preventiva, em detrimento de uma atuação meramente reativa. Contudo, esta não foi a única resposta encontrada para a mudança do paradigma social e securitário do final do século passado.

3.3. Policiamento Orientado pela Inteligência

O policiamento orientado pela inteligência (POI) surgiu num contexto semelhante ao POP: alteração do paradigma da criminalidade e maior exigência do ponto de vista da gestão de recursos (humanos, materiais e financeiros) e da satisfação dos cidadãos, da comunidade política e dos órgãos de comunicação social (Fernandes, 2014). Segundo Ratcliffe (2002), a crescente disponibilidade de informação e os novos modelos de gestão pública dos serviços policiais, potenciaram a produção de enormes quantidades de dados relativos à atuação policial, o que permitiu a sua monitorização. Este autor reporta o início da proliferação deste modelo de policiamento a dois relatórios vastamente influentes: *Helping With Enquiries – Tackling Crime Effectively*, publicado a 1993 pela *Audit Commission*, e *Policing With Intelligence*, publicado a 1997 por *Her Majesty’s Inspectorate of Constabulary*. O primeiro, focou-se essencialmente na ideia de uma Polícia *cost-effective*, ou seja, alocação do mínimo de recursos para os melhores resultados na redução da criminalidade (Ratcliffe, 2002). O segundo relatório identificou um conjunto de fatores fundamentais para a estruturação e implementação do *intelligence-led policing*, afirmando que a inteligência “permite uma compreensão clara da criminalidade, identifica quais os criminosos que estão ativos, quais os crimes que estão ligados e onde existe a probabilidade de ocorrerem crimes” (p. 1) corroborando ainda a ideia do relatório anterior quanto à aplicação de “recursos valiosos

(...) de forma eficaz contra os desafios atuais e tendências emergentes, garantindo as melhores oportunidades de intervenção positiva” (p. 1)

Desta feita, a definição de policiamento orientado pela inteligência apresentada mais recentemente por Ratcliffe (2016) é:

um modelo de negócios para a polícia que vê a inteligência criminal como uma combinação do que está mais comumente separado como análise criminal e a inteligência criminal, trabalhando numa estrutura de gestão de informação que permite aos analistas influenciar os decisores, e onde um espectro de soluções preventivas a longo prazo e orientadas para os problemas são extraídas de evidências basilares (p. 108).

expondo-se assim as duas funções primordiais do POI: “ter a capacidade de antecipar o risco e influenciar a ação” (Martins, 2013, p.18).

Em consonância com o exposto, a inteligência passa a assumir um papel central nas organizações policiais, pelo que, recolher informações e usar os resultados proporcionados pelos analistas para melhorar a atividade policial, são os novos objetivos dos dias de hoje (Treverton, Wollman, Wilke & Lai, 2011). Desta forma, com o intuito de pôr em prática a ideia de uma Polícia preventiva, quebrando com a rotina do modelo de policiamento tradicional, priorizou-se a adoção de estratégias e táticas fundamentadas na inteligência produzida pela organização policial (Peterson *et al.*, 1997). De igual modo, “A inteligência produzida pelos atores tem como fim apoiar a prossecução das suas missões e é, em regra, utilizada logo que produzida na condução das suas atividades” (Fernandes, 2014, p. 168), o que implica, automaticamente, que esta tem de ser oportuna, para cumprir os seus fins. De maneira mais sucinta, “à polícia cabe a tarefa de previsão, ou seja, antecipar a prevenção, e fá-lo através da produção de informações” (Clemente, 2010, p.159). Isto implica que o papel da análise das informações tenha uma importância acrescida (Fernandes, 2014), o que depende da existência de “analistas devidamente formados para proceder à interpretação do ambiente, e decisores influenciados que usem as informações policiais de forma a criar o referido impacto através de iniciativas estratégicas, operacionais e táticas” (Baraças, 2009, p.9). Não obstante da importância do papel dos analistas, é necessário que os decisores assumam um papel ativo neste processo, de maneira a proporcionar uma perspetiva o mais realista possível do ambiente, dito operacional, ao analista, para que este haja em

concordância, definindo prioridades de acordo com um *feedback* mútuo (Barrett, 2006; Fernandes, 2014).

O POP e o POI não são de tal forma distantes ao ponto de serem incompatíveis. Aliás, Ratcliffe (2016) é da opinião que “o POP poderia beneficiar de um maior uso da inteligência criminal e um maior foco no suspeito, enquanto o POI poderia beneficiar das capacidades estratégicas de resolução de problemas do POP” (p. 109).

3.4. Desafios e Soluções

Chan e Moses (2016) afirmam que o *Big Data* tem surgido como um potencial aliado das instituições policiais e das agências de inteligência. Contudo, segundo Crawford (2014) existem várias “ansiedades” quanto ao uso de *Big Data* para a segurança. Uma delas está relacionada, não com a quantidade de dados que as polícias têm na sua posse, mas sim com a imensidão que não está e que pode estar a deixar na penumbra uma vasta possibilidade de correlações entre eles (Crawford, 2014). Contudo, mesmo que as polícias tenham informação suficiente, quando esta não é organizada convenientemente, o seu potencial perde-se. Esta perspetiva baseia-se na crença de que quanto mais informação tivermos, mais aproximados estamos da verdade. Desde o setor empresarial ao setor público começou-se a apostar na recolha em massa de todo o tipo de dados. Contudo, mesmo após o início desta crença, continuaram a surgir inúmeras situações em que todo este manancial de informações e de agências de inteligência não conseguiram prever nem prevenir (Crawford, 2014). Resumidamente, só porque temos toda informação na nossa posse não significa que conseguiremos extrair valor dela porque: a informação pode mesmo assim não ser suficiente; pode conter ruído, ou seja, não ter veracidade, baixa qualidade; os métodos que utilizamos podem não ser os corretos para extrair o valor de que precisamos; ainda assim, o valor que extraímos pode não ser bem aplicado.

Um dos desafios predominantemente identificados por Goldstein (1990) consiste na resistência à mudança da organização. Manyika *et al.* (2011) referiu que o setor público é tendencialmente mais resistente no que concerne à reestruturação das organizações. Um dos motivos para isso pode ser o desconhecimento das possíveis vantagens que essa mudança acarreta. Babuta (2017) no seu estudo nas organizações policiais no Reino Unido, verificou que “quando novas tecnologias eram adotadas pelas Polícias, era frequente que a implementação falhasse” por não haver formação adequada, e porque os elementos policiais, muitas vezes, “escolhiam não as utilizar” (p. 39). Chan e Moses (2016) concluíram através

do seu estudo que poucos polícias na sua amostra conheciam ou valorizavam o potencial inerente ao *Big Data*. O seu estudo abrangeu ainda decisores políticos que se demonstraram leigos na matéria. Este desconhecimento pode ser inibidor do desenvolvimento de projetos de implementação do *Big Data* no setor da segurança. Ao não haver entendimento entre os técnicos informáticos, que desenvolverão o sistema, e elementos polícias, que deverão fornecer as informações necessárias e relevantes para serem implementadas nos sistemas de forma útil, estes projetos correm o risco de ficar aquém das expectativas.

De outra perspetiva, a objetividade da análise nos modelos de policiamento que a ela recorrem, tem sido um dos problemas identificados ao longo deste estudo. Segundo Ratcliffe (2016) os analistas têm a tendência de recorrer às tecnologias de informação disponíveis para analisar a informação de uma forma tão generalizada que acaba por não ser útil para os decisores. Por outro lado, quando os produtos dessa análise são em número demasiado considerável ao ponto de os comandantes receberem demasiada informação, esta deixa de ser relevante: “analistas que enviem rajadas de mails diários são vistos como mais contributivos para o barulho de que para uma solução” (Ratcliffe, 2016, p. 223). Neste sentido, é importante que os decisores percebam os produtos desta análise e as incorporem nas decisões de âmbito operacional, aproveitando, assim, o seu valor (Evans, 2012).

Outra das perspetivas da objetividade, surge de uma das preocupações inerentes ao uso de algoritmos por parte da Polícia como forma de análise de *Big Data* é a sua subjetividade. Este tipo de algoritmos, comumente usados no *predictive policing*, não são objetivos por terem o cunho humano na sua elaboração. Um dos grandes marcos do *predictive policing*, apontado até agora como a grande falha deste processo, será a forte cultura discriminatória racial presente no EUA, o que acaba por se impregnar no próprio algoritmo que é desenvolvido (Ferguson, 2017). De facto, o uso de *Big Data* pode levar a mais facilmente a abordar indivíduos que já tenham cometido crimes, ou, ainda menos objetivo, se tiver comportamentos que não estejam de acordo com o que a tecnologia de *Big Data* identifica como normal (Speranza, 2017).

Hoje em dia ainda há pouca preocupação com a qualidade dos dados (Ratcliffe, 2016). Jagadish (2015) afirma que “as previsões só podem ser tão boas quando a informação introduzida, e muito frequentemente, esta informação tem erros” (p. 1). Consequentemente, surge a complexa questão da atribuição de responsabilidades. Para Bovens, Schillemans e Goodin (2014) responsabilização envolve fornecer respostas a quem tem a legitimidade para as pedir, face às consequências. No exemplo específico da tecnologia de videovigilância, o *profiling* e a seleção social não se afiguram uma nova forma de policiar, mas sim a reflexão

do próprio perconceito humano (Ferguson, 2017). Não atentar a este fator pode levar a uma disseminação da responsabilização, ou seja, atribui-se o ónus daquela análise à tecnologia que tratou os dados e, portanto, que cometeu o erro de sinalizar aquela pessoa indevidamente, excluindo a responsabilidade dos criadores dos algoritmos (Speranza, 2017).

3.5. Polícia e *Big Data*

Nos EUA o *Big Data* tem sido integrado nas instituições policiais fundamentalmente para recolha e armazenamento de informação sobre amostras de ADN, videovigilância massificada e *predictive policing* (Joh, 2014). Contudo, existem mais exemplos de Polícias que recorrem às tecnologias e métodos de *Big Data* para aprimorar as suas estratégias e táticas de policiamento.

A *Albuquerque Police Department* tinha na sua posse uma riqueza de informações em relação a cada uso da força, que poderia incluir o registo criminal do suspeito, dados biométricos, câmaras de veículos policiais e registo áudio das ocorrências, videovigilância, relatórios de ocorrências de inúmeras revisões de procedimentos sobre o uso da força, artigos de jornais relevantes, testemunhos e os registos do tribunal. Apenas uma pequena fração dessas informações estava a ser “sistematicamente analisada para encontrar fatores de alarme, determinar quem estava em falta ou rever a política e o treino em torno do uso da força pelo polícia” (Olesker, 2012). Usando o *Hadoop*, os métodos de *Big Data* e *machine learning*, todos esses dados poderiam ter sido armazenados e analisados, encontrando novos padrões e fatores desconhecidos para explicar o aumento anómalo da violência. Ao lançar todos os dados disponíveis sobre o problema, as correlações que os analistas omitiriam sem eles, podem começar a ganhar forma (Olesker, 2012).

As Polícias de vários países já estão a integrar este conceito na sua forma de atuação e a extrair vantagens da compreensão e aplicação do mesmo. Em muitas delas a especialidade de analistas criminais, isto é, profissionais qualificados para o tratamento e análise da criminalidade, já foi integrada na estrutura da instituição. Não há, no entanto consenso quanto à própria definição da função de uma analista criminal. Segundo Fernandes (2014) argumentando que a origem desta ausência de consenso:

resulta do facto do leque de atividades desenvolvidas sob a designação de “análise criminal” ser bastante amplo, variando entre a análise estatística, para apoio, por exemplo, na distribuição de recursos (tipo e quantidade), até à análise de padrões

criminais ou das atividades de organizações criminosas, em apoio à investigação criminal (p. 138).

Ainda assim, Willey (2013a) defende que os analistas criminais têm “um papel essencial no apoio às instituições policiais a quantificar, avaliar e a responder à mudança da atividade criminal” (p. 1). Através da compreensão destes fenómenos, enriquecida pelos métodos usados em *Big Data*, estes analistas podem ajudar os polícias a perceber os fatores que influenciam a criminalidade e, assim, a conceber soluções apropriadas para prevenir e combater devidamente o crime e as incivilidades.

Exemplo disso é a *Durham Police Department*, que se serviu da tecnologia disponibilizada pela IBM para reduzir a criminalidade violenta em 50% num período de apenas 4 anos. Baseando-se em dados históricos, os analistas criaram uma base de dados inteligente que lhes permitia observar relações entre pessoas, locais e outras entidades, o que permitiu à polícia de Durham reorganizar a aplicação dos seus recursos (Willey, 2013b). Um outro exemplo desta nova realidade é a equipa de analistas da *Vancouver Police Department* que implementou, tendo por base a mesma tecnologia, um sistema por eles próprios desenvolvido, *Consolidated Records Intelligence Mining Environment (CRIME)*. Esta ferramenta permitiu-lhes perceber, através da correlação de dados, várias tendências com um grau de precisão elevado e prever onde e quando era provável que ocorressem os crimes. A avaliação do desempenho destes sistemas é positiva na medida em que implicou uma redução de 24% na criminalidade em geral e 9% na criminalidade violenta e grave, no período entre 2007 e 2011 (Willey, 2013b). Nas mãos de trabalhadores com capacidades especializadas, estes produtos, quando conjugados, podem ajudar a melhorar a resposta operacional (Evans, 2012).

Devemos ter em conta, no entanto, que esta tecnologia apenas é levada ao seu exponencial máximo perante condições muito específicas, nomeadamente, a necessidade de ser aplicada numa *smartcity*. Este conceito de cidades inteligentes nasceu associado ao desenvolvimento tecnológico. Numa primeira fase, era considerada uma cidade inteligente aquela que na gestão dos grandes temas governativos fazia uso da tecnologia (Lacinák & Ristvej, 2017). Em pouco tempo as cidades começaram a ser munidas dos mais variados dispositivos na área das energias, gestão de resíduos, sistemas de vigilância, etc. O conceito foi evoluindo para a área da sustentabilidade ambiental e era considerada inteligente a cidade que tivesse essa preocupação de monitorização de espaços verdes, automação de sistemas de rega, uso de materiais ecológicos na construção, criação de fontes de produção de energias

renováveis, entre outros⁹. Com a massificação do uso de *smartphones*, surgiram as aplicações, que abriram novos canais de interação entre entidades e comunidade, novos canais de visitação e conhecimento de lugares e cidades, bem como, novas formas de criar alertas e fazer denúncias. Atualmente o conceito abrange de forma particular a “humanização” dos sistemas, isto é, a funcionalidade da tecnologia, do *Big Data* e outros conceitos generalistas na vida dos cidadãos. É uma cidade *smart* aquela que usar a tecnologia não como um fim mas sim como um meio que contribua para felicidade dos cidadãos (Caragliu, Del Bo & Nijkamp, 2011). Mohanty, Choppali e Koungianos (2016) considera que uma *smartcity* é “um lugar onde redes e serviços tradicionais são tornados mais flexíveis, eficientes e substanciais com o uso da informação, digital e das tecnologias de telecomunicação, para melhor as suas operações em benefício dos seus habitantes” (p. 60).

Desta feita, é dentro deste contexto que surgem condições adicionais para trabalhar a componente preventiva da polícia, criando ferramentas que as conectem com a comunidade e com o cidadão, que lhes garantam “proximidade”, facilidade de comunicação, e que lhes aumentem o sentimento de segurança (Willey, 2011). Por outro lado, o sistema torna-se mais rico se for interativo, isto é, se existir fluxo de informação nos dois sentidos, cidadão-polícias, polícias-cidadão. O sistema oferece riscos, pois quanto mais sensores estão conectados ao sistema, maior é o volume a velocidade e variedade da informação. Neste sentido, as tecnologias e métodos de *Big Data* têm um papel fundamental na garantia de uma correta recolha, armazenamento, tratamento e processamento da informação.

Contudo, uma polícia tecnológica não é uma polícia desumana. a vertente humana tem de ser central. A tecnologia é apenas um meio, uma ferramenta, que necessita sempre da análise e avaliação que só a perceção humana permite fazer. Perry *et al.* (2013) reitera a ideia dizendo que as “tecnologias não devem ser vistas com um substituto (...) das estratégias policiais já concebidas” (p. 27).

Em bom rigor uma cidade não poderá ser considerada “*Smart*” se um dos fatores determinantes para a qualidade de vida dos cidadãos não for garantido: a segurança. Porque quando falamos em tecnologia, conectividade, velocidade, armazenamento de dados, falamos de redes, falamos de Internet, imediatamente nos surge a preocupação da vulnerabilidade dos sistemas. Um dos desafios é, portanto, a dos sistemas associados à *smartcity*. Nenhum sistema é inviolável, mas é possível definir métodos que garantam a sua fiabilidade e esta deve ser uma preocupação da própria governação do país. Esta realidade

⁹ Estas ideias foram defendidas publicamente, em 2016, pelo Ministro do ambiente da República da Eslováquia, tendo em conta a estratégia nacional da sustentabilidade e desenvolvimento daquele país.

levou à criação de um novo conceito, o *blockchain*, que visa criar um sistema incorruptível, inicialmente pensado para a transação de valores. O *blockchain* é uma *digital ledger* incorruptível de transações económicas que está programada para registar tudo, não só transações financeiras mas virtualmente tudo de valor (Tapscott & Tapscott, 2016). O conceito evoluiu e, atualmente, estende-se à transação de informação diversa:

Blockchain como mecanismo seguro e descentralizado dá a oportunidade de aumentar a conectividade sem comprometer a segurança de processos administrativos. Com o aumento de dispositivos IoT¹⁰ e as suas conhecidas vulnerabilidades, é essencial acoplar uma tecnologia segura. Embora a maior parte dos dispositivos IoT não possua os recursos necessários para guardar *Blockchain*, a imutabilidade do sistema associada às chaves criptográficas pode ser usada para estabelecer a identidade dos dispositivos. Os dispositivos podem apresentar credenciais para se autenticar num ecossistema IoT, com essas credenciais armazenadas e validadas posteriormente a partir de um *blockchain* público (Smart Cities, 2019)

Os problemas gerados aquando das várias tentativas de integração do *Big Data* na atividade policial são amplos. Contudo, houve soluções apresentadas que obtiveram resultados positivos na redução da criminalidade na área das polícias onde foram implantadas, mantendo o foco nas potencialidades do *Big Data*. Assim sendo, após uma análise global de todas as tendências, Ferguson (2017) acautela que para uma polícia se preparar para a introdução de tecnologias e métodos de *Big Data* na sua estrutura deve primeiro corresponder aos seguintes requisitos:

- Identificar os riscos que a tecnologia do *Big Data* tem o objetivo de reduzir;
- Garantir a qualidade da informação e a metodologia correta para a escolha da informação;
- Perceber se estas alterações terão impactos positivos na prática policial e na relação com as comunidades;
- Testar a tecnologia (responsabilizar e garantir a transparência);

¹⁰ *Internet of Things* remete para a extensão da conectividade da Internet aos dispositivos eletrónicos do dia-a-dia (Brown, 2016).

- Garantir o respeito pelos direitos dos cidadãos.

Por outro lado, Chan e Moses (2016) defendem que, apesar de não haver razões para acreditar na ineficácia do *predictive policing*, é necessário acautelar que essa eficácia ainda não foi provada. Para o efeito, será necessário compreender as limitações e as assunções impregnadas no *software* garantindo que resultados ineficientes e injustos são evitados.

Contudo, como vimos, as potencialidades do *Big Data* não se resumem ao *predictive policing*. Qualquer tratamento da vasta quantidade de informação na posse de qualquer polícia pode ser feito com recurso a estas tecnologias e métodos, seja qual for a utilidade que decidamos dar-lhe. Desta feita, as soluções apresentadas servem apenas como exemplo de possíveis aplicações do *Big Data* na Polícia.

4. O SEI

4.1. Descrição do Sistema

A definição formal apresentada pela PSP deste sistema de informação é:

O SEI é um sistema de informação cujo objetivo é assegurar a criação, manutenção e disponibilização da informação necessária e relevante à atividade operacional e de gestão da PSP, garantindo a sua atualização, coerência, integração e acessibilidade em tempo útil e de forma segura, funcionando como uma ferramenta complementar e de suporte às diversas atividades policiais de carácter operacional ou de apoio.

O SEI foi implementado em maio de 2004, fruto do desenvolvimento de um projeto para a produção de um sistema de informação na PSP que, segundo Lemos (2011), foi apressado pela realização do Euro 2004. Apesar disso, na sua opinião este foi o argumento ideal para obtenção de financiamento para o projeto em causa. Na conjectura prévia à implementação deste sistema verificava-se que vários comandos tinham os seus próprios sistemas, ou não, e havia uma completa descoordenação quanto aos procedimentos administrativos. Os processos eram desenvolvidos individualmente e a partilha de informação era praticamente nula por não haver uma base de dados centralizada (Accenture / Polícia de Segurança Pública, 2002)

Neste sentido, em 2002 foi desenvolvido o Plano Estratégico de Sistemas de Informação (PESI) no qual estava determinado que a missão do sistema de informação que viesse a ser criado seria “Assegurar a criação, manutenção e disponibilização da informação necessária e relevante à atividade operacional e de gestão da PSP, garantindo a sua atualização, coerência, integração e acessibilidade em tempo útil e de forma segura.” O SEI foi a resposta encontrada. Este sistema foi concebido para centralizar toda a informação da PSP e servir de núcleo de interoperabilidade entre outros sistemas que, entretanto, se vieram a desenvolver, nomeadamente, SIGAE, SCoT, entre outros.

Tendo por linha condutora a missão acima enunciada, o PESI estabeleceu as seguintes diretrizes para a conceção deste sistema de informação:

- Criação de um repositório único de dados;
- Estruturação da informação de forma lógica e orientada a conceitos;
- Maximização da utilização de ligações e associações tipificadas entre itens de informação;
- Potenciação da utilização de diversos tipos de informação;
- Registo da informação uma única vez;
- Definição, uniformização e operacionalização de procedimentos;
- Minimização da circulação de suportes físicos de informação;
- Racionalização dos circuitos de informação internos;
- Recurso a dispositivos de acesso alternativos;
- Apresentação da informação de forma simples, consistente e amigável;
- Alinhamento com as tendências de mercado em matéria tecnológica;
- Utilização de novos canais de interação com os cidadãos e a sociedade em geral.

Estavam assim traçados os objetivos deste sistema de informação. Contudo, Lemos (2011) refere que os objetivos do SEI não têm sido cumpridos na totalidade. Ainda assim, ressalva que o Gabinete de Sistemas de Informação (GSI), cujas competências são exploradas mais à frente nesta dissertação, tem sido fundamental para adaptar o SEI às necessidades da PSP que vão surgindo com o passar do tempo, e que inicialmente não estavam previstas nos objetivos traçados para este sistema.

Na perspetiva dos recursos humanos à disposição do sistema podemos distinguir quatro grupos: os utilizadores, coordenadores, apoio e desenvolvimento. No primeiro grupo inserem-se todos os elementos que utilizam o SEI como ferramenta de trabalho e, portanto, são os principais responsáveis pela introdução de dados de todo o género no SEI, sejam

peças, objetos, locais, ocorrências, etc., portanto, são a maior “fatia” dos recursos humanos que caracterizam este sistema e, por conseguinte, uma peça fundamental para que o SEI tenha esta vasta quantidade de informação. Todos estes elementos, atualmente, recebem formação específica para a sua função, estando distinguidos os módulos específicos ministrados enumerados no Regulamento do Formação SEI (RFSEI).

Na categoria dos coordenadores, podem-se distinguir dois tipos de coordenação: a central e a local. Na primeira a PSP dispõe da Equipa Única do SEI (EUSEI), implementada em 2011 e na dependência direta da Direção Nacional da PSP. De entre outras funções destacam-se: exercer, de forma integrada, a supervisão do SEI; estabelecer as prioridades de desenvolvimento e evolução do SEI; supervisionar a qualidade dos dados no SEI; ser o ponto de contacto com os Oficiais de Ligação SEI; colaborar na gestão do *Help Desk* nacional, em coordenação com o GSI; acompanhar a utilização do SEI pelo dispositivo.

A nível local existem os Oficiais de Ligação SEI (OLSEI), tendo por função: contribuir para a qualidade do SEI e dos seus dados; propor alterações ao SEI; gerir o *Help Desk* local, dando resposta às solicitações apresentadas, devendo reencaminhá-las para o *Help Desk* nacional quando não estejam no âmbito das suas competências;

A criação de equipas de apoio técnico especializado aos utilizadores foi uma das diretrizes mais vincadas no PESI. Deste modo, na terceira categoria enquadram-se os elementos que fornecem apoio aos elementos policiais no manuseamento do SEI. Novamente, este serviço funciona a nível local e a nível nacional. Como já referido anteriormente, o OLSEI é responsável pela gestão do *Help Desk* local, devendo este ser o ponto de comunicação com o *Help Desk* Nacional nos casos em que o problema apresentado extravasa os seus conhecimentos. Deste modo, as competências comuns a estes dois serviços são: responder a questões funcionais, questões logísticas e de suporte à infra-estrutura e configurações do SEI; esclarecer as dúvidas e problemas na utilização do SEI. Esta quarta categoria enquadram-se os elementos pertencentes ao Gabinete de Sistemas de Informação, cujas missões atribuídas pela PSP, enquanto responsáveis pelo desenvolvimento do SEI, são: corresponder às solicitações da EUSEI, no que respeita à manutenção do sistema ou novos desenvolvimentos; gerir a base de dados do SEI, no que respeita à sua manutenção, segurança e *backups*; desenvolver, na vertente técnica, os protocolos de interligação com entidades externas.

O Regulamento de Formação SEI (RFSEI) refere que:

Em termos tecnológicos, o SEI é executado através de uma aplicação do tipo *browser*, o mesmo que é utilizado regularmente para aceder à Internet, sendo o SEI uma aplicação baseada em tecnologia *Web*. O sistema divide-se em vários módulos, cada um deles com múltiplas funcionalidades, direcionadas para as várias tarefas existentes.

4.2. Potencialidade e desafios

O facto de ter sido criado um ponto central de armazenamento de informação, toda a informação produzida no decorrer da atividade operacional da PSP constitui um dos pontos fortes deste sistema de informação. Combinando esse fator com a possibilidade de acesso imediato durante o serviço a toda essa informação, via centro de comando e controlo, o SEI torna-se uma ferramenta importante no apoio à atividade operacional. O Subcomissário Nuno Lopes, Chefe do Núcleo de Sistemas de Informação e Comunicações do COMETPOR (comunicação pessoal, 16 de abril 2019) corrobora esta ideia dizendo que uma das grandes mais-valias do SEI é “a disponibilidade imediata da informação”. O facto de estar integrado na RNSI possibilita que em qualquer lugar, desde que se tenha acesso à rede, seja possível consultar a informação nele contida. Esta característica não se prende só com o acesso à informação, mas também com a atualidade da informação. Significa isto que mal os dados são introduzidos no sistema ficam logo disponíveis, permitindo, segundo Nuno Lopes (comunicação pessoal, 16 de abril 2019) “extrair relatórios de criminalidade por zona ou tipo de crime, (...) baseado na informação diária inserida pelos elementos policiais no âmbito da sua atividade operacional”. Neste sentido, o Comissário Jorge Teixeira, Chefe do Núcleo de Informações Policiais do COMETPOR (comunicação pessoal, 04 de abril de 2019) afirma que é importante usarmos todas as ferramentas ao nosso dispor no SEI, em particular no que concerne ao planeamento de operações, pois isso impede que a informação seja perdida em registos pessoais.

Por outro lado, o SEI tem um sistema de credenciações que permite pôr em prática o princípio da necessidade de conhecer, possibilitando, segundo Nuno Lopes (comunicação pessoal, 16 de abril 2019), “que cada elemento só tenha acesso ao que realmente necessita (...) impedindo a inserção de dados por quem não o pode fazer e não permitindo a leitura a quem não deve ter acesso à informação”.

Um dos pontos fracos do SEI é a qualidade da informação. Nuno Lopes (comunicação pessoal, 16 de abril 2019) afirma que ainda existe uma grande duplicação de registos. Quer isto dizer que são feitos vários registos para o mesmo item, o que tem um impacto na qualidade dos resultados do tratamento da informação. Teixeira (comunicação pessoal, 04 de abril de 2019) adianta uma explicação para este problema afirmando que, embora a perspetiva da instituição esteja certa quanto à valorização que dá aos dados contidos no SEI, existe muito trabalho a fazer quanto à mentalidade dos elementos policiais, no que concerne ao rigor na introdução dos dados e à própria perceção da sua importância. Ele verifica que, de um modo geral, qualquer elemento policial “olha para o SEI como uma ferramenta informática de produção de expediente e não como uma ferramenta informática na qual vai colocar informação, que vai ser usada futuramente.” Deste modo, também ele afirma que ainda existe um grande problema no que concerne à qualidade dos dados. Este problema tem origem, por um lado, na falta de rigor dos elementos policiais aquando da introdução dos dados no sistema e, por outro lado, a disposição pouco intuitiva e *user friendly* do próprio sistema (Lemos, 2011).

Como conseguimos perceber ao longo da nossa pesquisa, não só é fundamental que o sistema esteja preparado do ponto de vista da informação, tecnologia e métodos, mas também é preciso ter em conta o impacto que estas mudanças vão ter na instituição. Começando pelos utilizadores, deve ser tido em conta que, tal como aconteceu com a introdução do SEI na organização em 2004, pode haver uma certa resistência à mudança e dificuldades de adaptação, o que pode atrasar os efeitos da implementação do novo sistema. De facto, levou ainda algum tempo até que os objetivos de formação SEI fossem atingidos visto que em 2011, 7 anos depois da sua implementação, segundo o inquérito feito por Lemos (2011), a formação que era fornecida aos elementos não era suficiente, o que permitia que houvesse um mau manuseamento deste sistema de informação. Neste momento, o RFSEI impõe no seu art.º 12º, nº2 e nº3 que a formação SEI deve ser feita de 3 em 3 anos, podendo este prazo ser extensível para 5 anos. Os elementos são sujeitos a avaliação aquando desta formação.

Por outro lado, outro ponto fraco apresentado por Teixeira (comunicação pessoal, 04 de abril de 2019) tem a ver com a interoperabilidade dos sistemas. Sendo esta uma das diretrizes traçadas para a elaboração deste sistema de informação, ele afirma que todas as aplicações “que se relacionam com o SEI (...) têm de estar preparadas para cruzar dados com o SEI”. Este revela uma dificuldade acrescida na pesquisa pelo facto dos diferentes sistemas terem critérios de pesquisa não compatíveis com o SEI.

Segundo Nuno Lopes (comunicação pessoal, 16 de abril, 2019), apesar de o SEI ser uma ferramenta essencial no suporte à atividade operacional e, definitivamente ser uma mais valia para a PSP em geral, “a sua evolução encontra-se a um ritmo mais lento”, no que concerne ao desenvolvimento das ferramentas informáticas, do que as “restantes empresas em Portugal e as restantes polícias europeias e até mesmo que a própria sociedade”. Na sua ótica, todos os elementos deveriam dispor de acesso ao SEI de forma imediata, independentemente do serviço, nomeadamente, através de uma aplicação móvel, acessível através de *smartphones* ou *tablets* distribuídos pela própria PSP. Deste modo, para Nuno Lopes (comunicação pessoal, 16 de abril, 2019) a alteração prioritária a fazer ao SEI é a “portabilidade do sistema”. Trata-se de desenvolver o sistema no sentido de este poder ser portátil, ou seja, acessível em tempo oportuno, em qualquer lugar, porque “nos dias atuais [isso] significa uma fração de segundos”.

Para Jorge Teixeira (comunicação pessoal, 4 de abril, 2019) uma das alterações a fazer tendo em vista esse objetivo seria introduzir mecanismos de responsabilização dos elementos que introduzem as informações no sistema. Estes, combinados com uma melhor formação, trariam indubitavelmente benéficos para a qualidade geral dos dados, tal como já referido. Atualmente, as equipas que estão responsáveis pela supervisão da informação introduzida no SEI ainda não dispõem de nenhuma ferramenta informática que os apoie nesta tarefa, ou que o faça por eles. Desta feita, toda a informação que é introduzida no SEI diariamente tem de ser verificada manualmente, o que, dada a quantidade de informação existente, não é suficiente para assegurar a qualidade de informação necessária ao seu tratamento. Contudo, já existem métodos de tratamento que permitem tratar a vastas quantidades de informação melhorando, assim a sua qualidade. Segundo Jorge Teixeira (comunicação pessoal, 4 de abril, 2019) estas alterações trariam melhorias a nível do “trabalho quer das informações policiais, quer das informações criminais e até para a própria dinâmica de um comandante de esquadra e para a próprio elemento que está no CP”.

Capítulo III – Método

1. Abordagem Qualitativa

A presente investigação tem por base um estudo exploratório, uma vez que é nosso intento descrever e compreender o processo de desenvolvimento do SEI segundo uma

perspetiva de *Big Data*. Esta opção justifica-se pelo facto de pretendermos definir os limites do objeto de estudo e, por esta razão, os resultados têm um carácter não definitivo, tal como esclarecem Ramos e Coutinho (2011), caracterizando a investigação exploratória como uma investigação “de carácter provisório, na medida que se realiza para obter um primeiro conhecimento da situação que se quer estudar” (p. 37).

Neste sentido, são utilizadas técnicas de recolha de dados “sem medição numérica para descobrir ou aperfeiçoar questões de pesquisa, que podem ou não provar hipóteses [no] seu processo de interpretação” (Sampieri, Collado & Lucio, 2006, p. 5), ou seja, métodos qualitativos. Estes métodos são os indicados para o nosso objeto de estudo, uma vez que são ‘abertos’ o suficiente para se adaptarem à sua complexidade (Flick, 2005).

Os métodos qualitativos são “centrados na investigação em profundidade, conduzida de acordo com procedimentos regulares e repetidos” (Santo, 2015, p.27), tais como, “observação não estruturada, entrevistas abertas, revisão de documentos, discussão em grupo, avaliação de experiências pessoais, inspeção de histórias de vida, análise semântica e de discursos quotidianos [...]” (Sampieri, Collado & Lucio, 2006, p. 10). Dentro deste rol de técnicas, optámos pela revisão de documentos, isto é, pela análise documental (Bardin, 2018), que nos permitirá “passar de um documento primário (em bruto) para um documento secundário (representação do primeiro)” (Bardin, 2018, p.47), podendo esta representação ser a conjugação de diversos documentos, como é o nosso caso. É a partir daqui que, posteriormente, extraímos as nossas conclusões, pois, tal como refere Flick (2005, p.4), uma das partes essenciais da investigação qualitativa, é “a reflexão do investigador sobre a investigação, como parte do processo de produção do saber”.

Com este estudo qualitativo não é pretendida uma generalização dos resultados para realidades mais amplas, almejamos sim que este proporcione uma visão mais simplificada e completa, tendo em conta a recolha, categorização e análise dos dados efetuada. Desta feita, este será um estudo útil para futuras investigações nesta área, possibilitando assim, “a sua réplica por outros investigadores” (Pais, 2004, p.251).

2. Corpus

O *corpus*, segundo Bardin (1977, p. 122), consiste no “conjunto dos documentos tidos em conta para serem submetidos a procedimentos analíticos”. A documentação selecionada para constituir o *corpus* foi o Plano Estratégico de Sistemas de Informação (PESI), o Estudo SEI+ - Estudo de Utilização e Evolução do SEI (Sistema Estratégico de

Informação, Gestão e Controlo Operacional da PSP), Visão Global de Operacionalização da Estratégia para as TIC na PSP 2013-2016 e Grandes Opções Estratégicas da PSP para 2017-2020. Na recolha destes documentos foi tida em conta o valor estratégico dos mesmos, bem como a temporalidade da sua elaboração e dos seus efeitos de modo a permitir o estabelecimento da progressão da perspetiva estratégica institucional sobre os sistemas de informação, desde um período prévio à implementação do SEI até à atualidade.

3. Instrumento: Análise de Conteúdo

Inicialmente foi feita uma análise documental ao *corpus*. Este tipo de análise é definido por Chaumier (cit. in Bardin, 2018, p.47) como “uma operação ou conjunto de operações visando representar o conteúdo de um documento sob uma forma diferente da original, a fim de facilitar, num estado ulterior, a sua consulta e referência”. Desta feita, a interpretação dos textos pode ser elaborada com dois objetivos opostos: “um é revelar (...) ou contextualizar as afirmações feitas no texto, o que conduz normalmente à ampliação do material (...); o outro visa reduzir o material textual, parafrazeando-o, resumindo, categorizando-o” (Flick, 2005, pp. 179-180). Nesta senda, e tal como refere Flick (2005), estas duas estratégias podem ser aplicadas sucessivamente, que é o que se verifica no presente estudo, onde, primeiramente, categorizamos o material para, posteriormente, reconstituirmos o texto com os acrescentos que consideramos adequados.

Sendo assim, após a leitura dos documentos, seguimos com a análise categorial, pois, tal como é referido por Bardin (2018, p.145), “a maioria dos procedimentos de análise organiza-se (...) em redor de um processo de categorização”. Deste modo, os dados constantes do *corpus* são “reduzidos (codificados, formatados) para serem, em seguida, tratados” (Lessard-Hébert, Goyette & Boutin, 2008, p. 163). Isto pressupõe que seja feito um “desmembramento do texto em unidades, [isto é], em categorias, segundo reagrupamentos analógicos” (Bardin, 2018, p.199). Por sua vez, as categorias “são rubricas ou classes (...) [que] reúnem um grupo de elementos (unidades de registo no caso da análise de conteúdo) sob um título genérico, agrupamento esse efetuado em razão das características comuns desses elementos” (Bardin, 2018, p.145). Este processo de categorização deve ser feito com rigor, tornando as “categorias e unidades de registo claras” (Serrão & Pais, 2006, p.153), assentando na “crença de que a categorização (passagem de dados em bruto a dados organizados), não introduz desvios (por excesso ou por recusa) no material” (Bardin, 2018, p.147). Para que tal seja cumprido, é necessário que uma série de ‘qualidades’ (Bardin,

2018), das quais fazemos uma sùmula, sejam tidas em conta. Temos a ‘exclusão mútua’, que se baseia na não repetição de um elemento em duas ou mais categorias; a ‘pertinência’ que determina que a categoria deve ser adaptada ao contexto definido; a ‘objetividade e a fidelidade’, que estabelecem que o analista deve definir com clareza as variáveis e os critérios que determinam a entrada de um elemento numa categoria (Bardin, 2018, pp. 147-148).

Tendo em consideração o exposto, as categorias do presente estudo foram feitas de acordo com os critérios, estabelecidos por Vala (1999), da exaustividade e exclusividade na sua elaboração, que permitem caracterizá-las como fiáveis e válidas. Para o efeito anteriormente referido, definimos cinco categorias combinando o conteúdo dos documentos constitutivos do *corpus* com as ilações retiradas com enquadramento teórico realizado. As categorias “informação”, “tecnologias”, “métodos” e “impacto” foram determinadas diretamente através da divisão feita no enquadramento teórico, pelo facto de constituírem os núcleos conceptuais mais revelantes na perspetiva do *Big Data* (cfr. Anexo 2). A categoria “polícia” surgiu na sequência de uma síntese dos temas mais importantes no contexto do uso do *Big Data* nas polícias (cfr. Anexo 2).

4. Procedimento

Tal como mencionado, o *corpus* do nosso trabalho é constituído por quatro documentos, dos quais, tendo em conta os princípios estruturantes da análise de conteúdo, nos propusemos a analisar e, por conseguinte, de acordo com os objetivos da presente dissertação, extrair o que se afigurava de mais proveitoso para retirarmos conclusões.

Em consonância, a recolha do (referido) material que integrado no *corpus* do nosso estudo obedeceu aos seguintes critérios: impacto estratégico dos documentos, isto é, a preponderância que estes documentos representam ao nível da tomada de decisão estratégica relativa às novas tecnologias e inovações referentes ao SEI; o momento da sua elaboração, sendo que, neste ponto, é necessário ressaltar que não existe documentação recente que nos permita fazer uma análise mais atual, pelo que, os documentos por nós utilizados remontam a um período compreendido entre 2002 e 2016¹¹, procurando assim, percorrer as diferentes fases de implementação e desenvolvimento do SEI.

¹¹ Note-se que o documento, “Grandes Opções Estratégicas para a PSP 2017-2020” é, deveras, um documento atual, no entanto, o mesmo não versa só sobre os sistemas de informação especificamente.

Para prosseguir os nossos fins, e tal como supramencionado, recorreremos à análise categorial, uma vez que esta vem por meio das suas especificidades, dar uma forma simplificada e conveniente que nos permita tirar ilações perante os resultados apresentados (Bardin, 2018). Para o efeito, definimos cinco categorias, as quais enformam as subcategorias que, ao longo da leitura e análise documental, se foram mostrando relevantes, tendo as mesmas sido enquadradas, logicamente, na categoria correspondente. Importa então mencionar quais as categorias por nós estipuladas: Informações (onde se enquadram as ideias relacionadas com os atributos dos dados e com as operações tendentes à *digitization* e *datafication*); Tecnologias (engloba a informação referente às exigências tecnológicas associadas ao *Big Data*); Métodos (abarca as noções relacionadas com o tratamento e manuseamento dos dados nos sistemas); Impacto (encerra as unidades de registo relativas às influências das alterações proporcionadas pelo *Big Data*); Polícia (como seria de esperar, inserem-se aqui as referências com particular enfoque na atividade das instituições policiais).

Desta feita, assumimos que o preenchimento ou não das categorias apresentadas é indicativo da relevância atribuída a essas ideias pela PSP na implementação e no desenvolvidamente do SEI, tendo sido realizadas inferências e interpretações dos dados com base nesse pressuposto.

De mencionar ainda que, com o intuito de complementar a descrição do objeto de estudo, foram realizadas entrevistas a dois oficiais de polícia da PSP¹², uma delas no sentido de compreender a perspetiva do Núcleo de Sistemas e Informação e Comunicações e a outra para integrar na nossa descrição a perspetiva do Núcleo de Informações Policiais. Neste âmbito, procedeu-se ao pedido de autorização (cfr. Anexo 1) para a realização das mesmas, tendo esta sido concedida (cfr. Anexo 1).

¹² Referenciadas como comunicação pessoal ao longo do texto.

Capítulo IV – Apresentação, Análise e Discussão de Resultados

Na presente investigação foi realizada uma análise de conteúdo, tendo sido utilizada uma grelha categorial para codificar as ideias contidas no *corpus*. Assim, apresentamos os resultados da codificação que foram distribuídos pelas diferentes categorias da seguinte forma:

Tabela 1 - Distribuição das u.r. pelas categorias

| <i>Categoria</i> | <i>Distribuição (u.r.)</i> | <i>Percentagem (%)</i> |
|-----------------------|----------------------------|------------------------|
| <i>A. Informação</i> | 17 | 12.5 |
| <i>B. Tecnologias</i> | 26 | 19.1 |
| <i>C. Métodos</i> | 30 | 22.1 |
| <i>D. Impacto</i> | 32 | 23.5 |
| <i>E. Polícia</i> | 31 | 22.8 |
| <i>Total</i> | 136 | 100 |

Analisando a distribuição apresentada na tabela 1 verificamos que todas as categorias foram preenchidas. Este facto não suscitou qualquer surpresa porque como já foi referido as categorias foram teoricamente derivadas, com exceção da categoria E. A distribuição das u.r. pelas diferentes categorias foi praticamente uniforme ressalvando o caso da categoria A.

No que concerne à distribuição das u.r. pelas subcategorias apresentamos a tabela 2.

Tabela 2 Distribuição das u.r. pelas subcategorias

| <i>Subcategoria</i> | <i>Distribuição (u.r.)</i> | <i>Percentagem (%)</i> |
|---------------------|----------------------------|------------------------|
| <i>A.1</i> | 8 | 5.9 |
| <i>A.2</i> | 7 | 5.1 |
| <i>A.3</i> | 2 | 1.5 |
| <i>B.1</i> | 1 | 0.7 |
| <i>B.2</i> | 2 | 1.5 |
| <i>B.3</i> | 7 | 5.1 |
| <i>B.4</i> | 6 | 4.4 |
| <i>B.5</i> | 10 | 7.4 |
| <i>B.6</i> | 0 | 0 |

| | | |
|--------------|-----|------|
| <i>C.1</i> | 14 | 10.8 |
| <i>C.2</i> | 7 | 5.1 |
| <i>C.3</i> | 0 | 0 |
| <i>C.4</i> | 8 | 5.9 |
| <i>C.5</i> | 1 | 0.7 |
| <i>D.1</i> | 13 | 9.6 |
| <i>D.2</i> | 1 | 0.7 |
| <i>D.3</i> | 12 | 8.8 |
| <i>D.4</i> | 6 | 4.4 |
| <i>E.1</i> | 4 | 3.0 |
| <i>E.2</i> | 6 | 4.4 |
| <i>E.3</i> | 3 | 2.2 |
| <i>E.4</i> | 8 | 5.9 |
| <i>E.5</i> | 10 | 7.4 |
| <i>Total</i> | 136 | 100 |

Considerando a distribuição das u.r. pelas subcategorias, apresentada na tabela 2, verifica-se que todas elas foram preenchidas, excetuando as subcategorias B.6. (Avaliação das tecnologias) e C.3 (Objetividade da análise). Podemos então refletir sobre o não preenchimento destas subcategorias que, no nosso entender, revela, à partida, que as estratégias traçadas para o SEI são omissas quanto a avaliação do desempenho dos sistemas (*benchmarking*), o que, segundo Xiong *et al.* (2013), é uma vertente essencial da implementação destas tecnologias. Deste modo, a omissão desta vertente nos documentos estratégicos é algo surpreendente. Portanto, podemos considerar que a avaliação do desempenho dos sistemas através de *benchmarking* não foi um aspeto central na conceção e desenvolvimento do SEI. A ausência de u.r. na subcategoria C.3 (objetividade da análise) também não era expectável, uma vez que o que está em causa é o impacto das sucessivas intervenções humanas (Ferguson, 2017) que podem influenciar a isenção e a precisão da análise dos dados. Esta carência leva-nos a considerar uma possível ausência de estratégia para a fiabilidade da análise, estando em causa possíveis distorções da realidade.

No quadro teórico verificámos que as componentes dos sistemas de *Big Data* são informação, tecnologias, métodos e impacto (De Mauro *et al.*, 2016). Chan e Moses (2016) afirmaram que este fenómeno pode ser uma mais valia para as polícias, pelo que,

foi criada uma categoria denominada “Polícia”. Analisada a distribuição das u.r. nas diferentes categorias não foram verificadas grandes assimetrias, destacando-se apenas a categoria A (informação) com menor prevalência. Este equilíbrio é de fácil compreensão, uma vez que, considerando a natureza/dimensão estratégica dos documentos que constituem o *corpus*, estas vertentes são as componentes constitutivas da arquitetura do SEI.

1. Impacto

No que diz respeito ao impacto, a subcategoria mais prevalente foi a D.1. (Suporte à atividade operacional). O apoio à tomada de decisão e à atividade operacional foi um desígnio estratégico para o SEI (*e.g.* u.r. 131 “Considera-se como um vetor primordial de desenvolvimento neste âmbito a mobilidade do acesso aos sistemas de informação de processamento e apoio da atividade operacional bem como a necessidade da organiza”), bem como melhorar o tratamento e organização de dados (*e.g.* u.r. 134 “... bem como dota-los de maiores capacidades de *business intelligence* para análise e apoio à decisão”), e ainda dotar a polícia de ferramentas que permita o acesso aos sistemas independentemente do tempo e lugar (*e.g.* u.r. 67 “Para um correcto controlo e planeamento da sua actividade operacional diária, as forças policiais, neste caso a PSP, tem absoluta necessidade de saber quais os locais com maior índice de criminalidade, quais os locais com maior número de acidentes de viação, etc. Mais importante do que isso, é a necessidade de dispor dessa informação em tempo útil, para que as diligências necessárias sejam tomadas no mais curto espaço de tempo possível”). Vimos anteriormente nesta dissertação que Evans (2012) corrobora esta ideia, dizendo que esta é condição de aproveitamento da análise. Espera-se, assim, que estes sistemas de informação permitam melhorar a qualidade e a quantidade de informação disponibilizando-a para que os polícias possam tomar decisões operacionais baseadas nessa informação. Portanto, percebe-se o alcance estratégico que os documentos constantes do *corpus* pretenderam dar ao apoio à atividade operacional.

De igual modo prevalente nesta categoria, a subcategoria D.3. (reestruturação e planeamento) foi uma das vertentes tidas em conta na implementação e desenvolvimento do SEI, tal como era expectável. Nesta subcategoria são mencionadas as ideias relativas às mudanças do ponto de vista estrutural que ocorrem quando um sistema de informação desta dimensão e importância é introduzido no seio de uma instituição. De um ponto de

vista geral, houve alterações respeitantes à multiplicação e comercialização destas tecnologias (*e.g.* u.r. 52 “ao nível dos meios computacionais com massificação do uso da Internet, a explosão da indústria das telecomunicações, os fenómenos multimédia e, em particular, o acesso cada vez mais facilitado à informação (novos canais e dispositivos de acesso disponíveis a custos cada vez menores), introduziram alterações estruturais na forma como as pessoas e as organizações se podem relacionar”), o que, à partida, fará a PSP alterar não só os seus objetivos como também a sua estrutura (*e.g.* u.r. 80 “Os objectivos pretendidos para a PSP são complexos e ambiciosos, sendo necessário operar algum tipo de mudança a vários níveis: organizativo, processual e tecnológico, através do lançamento de diversas iniciativas em cada uma das vertentes enunciadas”). Neste sentido, estas alterações devem ser planeadas e implementadas (*e.g.* u.r. 102 “devendo garantir o alinhamento da estratégia de SI com os objectivos estratégicos da organização, constituindo-se assim como um elemento fundamental na sua modernização”), pois, atualmente, este pode ser o critério de sucesso das organizações na prossecução dos seus objetivos, tal como referiu McAfee e Brynjolfsson (2012). De facto, estas ideias foram mencionadas várias vezes ao longo das u.r. o que revelador da consciência da instituição para com estas necessidades.

O mesmo não se poderá dizer quanto à subcategoria D.2 (Privacidade e proteção de dados). Verifica-se que esta apenas é mencionada uma vez de acordo com a nossa análise (*e.g.* u.r. 113 “garantir a privacidade e integridade da informação, permitir utilização dos sistemas para suportar realização dos processos policiais”) numa vertente de segurança da informação (subcategoria E.3.) e não numa vertente do desenvolvimento de políticas de proteção de dados pessoais. Segundo Manyika *et al.* (2011) qualquer instituição que lide com este tipo de dados deve desenvolver políticas da gestão dos mesmos de modo não só a garantir a sua proteção contra o acesso e divulgação indevida, mas também reforçando a confiança dos cidadãos na instituição responsável pela detenção e manipulação dos seus dados pessoais. Deste modo, perante a quase total ausência da menção desta vertente, constatamos que esta não tem sido uma preocupação primordial na implementação e desenvolvimento do SEI.

2. Polícia

A segunda categoria com mais prevalência foi baseada nas componentes mais relevantes para o *Big Data* no contexto securitário e, mais concretamente, nas instituições

policiais. A subcategoria mais prevalente foi a E.5. (acessibilidade e partilha de informação), com 10 u.r. Por conseguinte, constata-se que a visão estratégica do desenvolvimento do SEI integrou, por um lado a acessibilidade à informação nas suas diretrizes, tanto numa perspetiva interna (*e.g.* u.r. 6 “Criação de um repositório único e integrado, transversal e partilhado por todas as unidades da PSP”), como de uma perspetiva externa (*e.g.* u.r. 12 “Consultas a bases de dados internas ou externas para pesquisa de informação relevante para prossecução da atividade operacional e conferir maior nível de segurança dos agentes no terreno”). Por outro lado, também a partilha de informação pode ser vista da perspetiva interna (*e.g.* u.r. 103 “o modelo promove uma completa integração, coerência e partilha da informação, garantindo a sua disponibilização onde e quando necessário”) e da perspetiva externa (*e.g.* u.r. 127 “estabelecer protocolos e acordos em relação a trocas de informação entre a PSP e outras entidades”). Segundo Manyika *et al.* (2011) cada vez mais as empresas têm apostado na integração de informação de fontes externas e, de facto, Wyllie (2013a) refere que esse é um dos fatores preponderantes na obtenção de resultados concretos na atividade policial.

Não se tendo verificado assimetrias evidentemente vincadas no que concerne à distribuição das u.r., destaca-se, devido à relevância dada por alguns autores (Manyika *et al.*, 2011; Scott & Kirby, 2012) no contexto desta investigação, a subcategoria E.1. (responsabilidades), na qual foram enquadradas ideias quanto ao estabelecimento de responsabilidades no manuseamento destas tecnologias e na gestão do desenvolvimento destes sistemas de acordo com as necessidades da instituição (*e.g.* u.r. 118 “A definição das funções e responsabilidades da organização de SI é um aspecto particularmente relevante por forma a garantir o alinhamento entre a orientações estratégicas da PSP e a estrutura organizacional desejada”). Contudo, tendo em conta o referido no primeiro capítulo, existem outras dimensões da responsabilidade que não são abrangidas, tais como, a da responsabilidade por quem introduz a informação nos sistemas (Ferguson, 2017) e por quem desenvolve os algoritmos analíticos, o que pode levar a quebra de rigor nestes processos (Ferguson, 2017; Speranza, 2017).

3. Métodos

Nesta categoria destaca-se a subcategoria C.1. (Análise e produção de conhecimento). Esta compreende as ideias relativas aos métodos de análise de informação com vista à produção de conhecimento, constatando-se que, no quadro da perspetiva

estratégica da implementação e desenvolvimento do SEI, se destaca a produção de conhecimento a partir da informação disponível na PSP (*e.g.* u.r. 19 “Aplicação de métodos analíticos e estatísticos a dados de histórico para previsão de tendências e padrões de comportamento futuros (recurso sistemático a entidades estatísticas)”), a disponibilização dessa informação de forma útil ao utilizador final (*e.g.* u.r. 17 “Capacidades de exploração de informação, de uma forma gráfica e muito “*user-friendly*”, com base na apresentação de redes de ligação/interdependência entre elementos informacionais relevantes para o contexto de decisão em causa (por exemplo, ligações entre diferentes indivíduos relacionados num processo de investigação criminal) - a aplicar aos sistemas operacionais da PSP ou às capacidades analíticas a implementar”) e a integração de fontes não só internas mas também externas (*e.g.* u.r. 48 “Integração de ferramenta de análise de dados e identificação de padrões de comportamento, associações e emissão de alertas (cruzamento de dados do SEI e de outras fontes - sites, bases documentais, sistemas de outras entidades, etc.)”). Esta frequência de menções no *corpus* analisado, não se assume como imprevisível, na medida em que, tal como Fernandes (2014) integralmente referiu, estes métodos vão permitir uma previsão das tendências criminais, e, desse modo, potenciar uma aplicação dos recursos mais eficiente e eficaz, garantindo melhores resultados na atividade policial.

De destacar, embora com menor prevalência, as subcategorias C.4. (formação) e C.2. (qualidade dos dados), com elevada importância nos documentos estratégicos analisados. Estando estas relacionadas, estabelecemos que a qualidade dos dados deve ser garantida (*e.g.* u.r. 77 “Não deverá assim existir multiplicidade de registos da mesma informação, evitando, dessa forma, redundâncias, incoerências e multiplicações do esforço associado à sua recolha e tratamento”) por via não só da responsabilização, tal como já foi referido anteriormente, mas também através da formação dos elementos policiais (*e.g.* u.r. 110 “A existência de meios humanos qualificados e a utilização das tecnologias adequadas permitirão a obtenção de ganhos significativos ao nível da simplificação, da redução dos custos e do aumento da qualidade associada aos processos de trabalho da organização”), que inicialmente tiveram dificuldades na adaptação a estas novas tecnologias (*e.g.* u.r. 26 “apontadas deficiências ao nível da formação inicial e contínua de utilizadores”). Noutra perspetiva, relevam-se ainda as necessidades de formação de elementos técnicos especializados (*e.g.* u.r. 91 “Haverá uma grande necessidade de formação e adaptação dos recursos humanos técnicos da PSP, por via das novas tecnologias a adotar, não sendo de excluir o eventual reforço das capacidades

internas”). Não nos surpreende esta preocupação demonstrada com a formação pois, em coincidência com Babuta (2017), esse é um dos pilares fundamentais no correto uso destas tecnologias. Quando estas são usadas deficientemente, podendo, como vimos, a formação estar na sua origem, pode ser comprometida a sua própria eficácia, nomeadamente, por via da inviabilização da análise, sendo uma das causas a baixa de qualidade da informação (Jagadish, 2015).

4. Tecnologias

A subcategoria B.5. (inovação) foi uma constante ao longo dos vários documentos analisados. A consciência de que as tecnologias se encontram num ciclo de rápida evolução, que deve ser acompanhada pela PSP (*e.g.* u.r. 2” esforços na modernização desta área, procurando, de forma contínua e estratégica, alinhar as suas ferramentas e sistemas aos requisitos da sua atividade operacional e à rápida e transversal evolução da Tecnologia a nível mundial”), levou a que esta componente fosse mencionada frequentemente, não só pelas próprias exigências do cidadão (*e.g.* u.r. 59 “Estes são alguns dos desafios que se apresentam a uma organização como a PSP, que terá que continuar o seu processo de constante adaptação e modernização, por forma a não defraudar as expectativas daqueles que são os seus “clientes”: os cidadãos”) mas também pela possibilidade de atualização face às tecnologias disponíveis (*e.g.* u.r. 79 “Alinhamento com as tendências de mercado em matéria tecnológica, tirando partido da simplicidade e das potencialidades oferecidas pela micro-informática, aproveitando a abundância de recursos disponíveis e, conseqüentemente, beneficiando dos factores concorrenciais que permitirão a obtenção de mais capacidade com menores custos”), o que implica a implementação de um sistema preparado para sofrer atualizações quase constantes (*e.g.* u.r. 109 “a continuidade e escalabilidade da solução tecnológica, de forma a garantir que não se atingem os pontos de saturação da infra-estrutura que se verificam hoje em dia”). Esta aproximação dos documentos que constituem o *corpus* às soluções tecnológicas para o SEI é expectável pois, tendo em conta a lei de Moore (2006) e a pressão que o constante aumento da frequência de produção de dados (Dhanklad, 2019) exercem no fabrico de novas tecnologias e, conseqüentemente, novas formas de estas servirem os objetivos traçados pelas organizações. Como tal, existe base teórica que sustente o desenvolvimento de uma estratégia que tenha em conta este fator.

A par da subcategoria B.6. (avaliação das tecnologias), mencionada anteriormente, verificam-se duas subcategorias residuais no quadro da categoria B. Por

um lado, na subcategoria B.1. (processamento) apenas é referido a necessidade de existência (e.g. u.r. 87 “novas infra-estruturas de processamento, constituindo um ambiente de produção para o novo Sistema de Informação, bem como ambientes de desenvolvimento e de testes de qualidade, para suportar a implementação das novas aplicações”), o que não é condizente com a elevada importância dada pela Microsoft Research (2013) ao superior poder de processamento necessário ao tratamento de *Big Data*. Por outro lado, a subcategoria B.2. (armazenamento) apenas apresenta duas u.r. fazendo menção à necessidade de uma possível renovação do centro de armazenamento de dados (e.g. u.r. 88 “O *Data Center* da PSP deverá ser alvo de renovação ou de eventual substituição, em função de uma avaliação mais detalhada das futuras necessidades em termos de instalações físicas”), sendo omissa, contudo, no que concerne à capacidade de armazenamento em si. Neste aspeto e pelo exposto, podemos considerar que, tal como NBD-PWG (2014) referiu, a escalabilidade do sistema deve ser um fator preponderante aquando da implementação de um sistema de informação, da perspetiva do *Big Data*.

5. Informação

Esta categoria foi a menos prevalente no quadro das u.r. codificadas. Apuramos que do ponto de vista estratégico as subcategorias A.1. (organização e estruturação) e A.2. (introdução) têm relevância na implementação e desenvolvimento de um sistema de informação para a PSP. Assim, a capacidade de o sistema organizar e estruturar dados materializa-se na (e.g. u.r. 13 “criação de um repositório integrado de dados que centralize a informação dos diferentes sistemas operacionais e de suporte da PSP num modelo único e otimizado, e implementação dos respetivos processos de carregamento [extração, transformação e carregamento]”), pois este permitirá que a informação, que se apresenta em variados formatos (e.g. u.r. 76 “De facto, nos dias de hoje a informação não assume apenas a sua forma mais tradicional baseada em dados estruturados, havendo possibilidade e necessidade de recorrer a formatos complementares que cada vez mais se revelam fundamentais para as organizações policiais”) seja organizada segundo os mesmos critérios. Este, segundo Mayer-Schönberger e Cukier (2013), é um passo essencial na preparação da informação para que possa ser posteriormente analisada. Consequentemente, é perceptível a razão que levou à adoção de medidas estratégicas neste sentido.

Na perspetiva da introdução de dados, esta passou, numa fase inicial pela conversão dos documentos em suporte físico para digital (*e.g.* u.r. 50 “Carregamento, no repositório digital, de todos os documentos em suporte papel que tenham entrado na PSP. Utilização de scanners para digitalização ou importação de ficheiros electrónicos”), evoluindo no sentido de fornecer algum suporte nos registos diretos no sistema (*e.g.* u.r. 82 “Além destas entidades e analisando em pormenor alguns dos macro-processos mais importantes, encontramos outro tipo de necessidades de informação que importa registar. Por exemplo, aquando de uma ocorrência, e além da informação relativa às entidades já referidas, há necessidade de registar informação sobre a própria ocorrência [entidade Ocorrência], ou sobre a detenção de determinado indivíduo [entidade Detenção], podendo a mesma ter ou não associado um processo [contra-ordenação, contravenção, crime, etc.] cuja tramitação interessa gerir e acompanhar [entidade Processo]”). Esta perspetiva revela-se adequada pois a obtenção dos elementos essenciais de informação permitirá reduzir o “barulho”, permitindo direccionar corretamente o foco da análise (Ratcliffe, 2016).

No entanto, verificamos um descrédito da subcategoria B.3. (atributos dos dados). Nesta subcategoria há apenas uma menção direta aos desafios que estes tipos de dados geram (*e.g.* u.r. 135 “A quantidade e qualidade de dados existentes nos sistemas de informação atuais da PSP motivam acrescidas preocupações de gestão da informação (*Big Data*)”), o que nos permite concluir que os cuidados com a superação destes desafios não foram transversais a todo o processo de desenvolvimento do SEI. Na nossa perspetiva, perante o exposto por De Mauro *et al.* (2016), o reconhecimento destes atributos será essencial para a adoção de medidas futuras que permitirão o aproveitamento do valor inerente aos *Big Data*.

Capítulo V - Conclusões

O *Big Data* apresenta-se como um conceito vasto e complexo que gera por si só “um fenómeno cultural, tecnológico e académico” (Boyd & Crawford, 2012, p. 663). Este fenómeno abrange cada vez mais áreas da sociedade, sendo uma delas a segurança. Assim, o *Big Data*, enquanto fenómeno, gera potencialidades de suporte e melhoria do desempenho das organizações, de entre as quais as policiais, acarretando, contudo, alguns desafios identificados ao longo desta dissertação, que devem ser tidos em conta aquando da implementação das soluções de *Big Data*.

A presente dissertação pode ser considerada como inovadora, uma vez que nunca haviam sido desenvolvidos estudos acerca do *Big Data* na PSP. Desta feita, desenvolvemos o presente trabalho em torno da problemática de investigação apresentada na introdução, procurando perceber como é que o SEI foi desenvolvido numa perspetiva de *Big Data*. Para esse efeito, recorreremos a métodos qualitativos utilizando a análise de conteúdo, nomeadamente a categorial, que nos permitiu retirar do *corpus* aquilo que considerámos relevante, expondo-o de forma relacionada com o enquadramento teórico previamente feito, o que permitiu comparações entre o que está definido no seio da PSP, e aquilo que devem ser as características do *Big Data*.

Após a análise realizada, e tendo sido apresentados e discutidos os resultados, podemos afirmar que é possível identificar que a maior parte das diretrizes definidas pela PSP para a implementação e desenvolvimento do SEI coincidiam com vertentes fundamentais do *Big Data*. Neste sentido, verificámos que somente a avaliação das tecnologias e a objetividade da análise não constam nos documentos analisados. Isto expõe a necessidade de que estes segmentos precisam de ser desenvolvidos pela PSP, uma vez que são ambos de uma grande relevância. O primeiro, (avaliação das tecnologias), implica o uso de técnicas de *benchmarking*, ou seja, o estabelecimento de sistemas de referência que permitam perceber que melhorias há a fazer, e em que sentido. Isto impede a cristalização dos sistemas de informação em uso, pois, uma vez que o avanço tecnológico é contínuo, é necessário que o SEI tenha a capacidade de o acompanhar. Quanto à objetividade da análise, esta implica que a análise seja feita com isenção, sem influência das convicções pessoais e ideologias que a possam enviesar. Estas duas dimensões constitutivas do *Big Data*, têm de constar dos documentos estratégicos de implementação e desenvolvimento do SEI, uma vez que são cruciais para que o mesmo seja uma ferramenta séria da atividade policial, e um verdadeiro sistema de informação preparado para o *Big Data*.

Desta feita, cabe-nos agora tecer considerações quanto à aplicabilidade dos resultados. De facto, os resultados obtidos na sequência da análise de conteúdo dos documentos estratégicos referidos ao longo desta dissertação, permitem-nos afirmar que é necessário enquadrar estas duas vertentes em futuras orientações estratégicas, de modo a que também estes possam ter uma expressão estratégica da sua implementação no SEI e restantes sistemas de informação da PSP.

Para além destas considerações, propomo-nos agora refletir sobre toda a investigação explorando eventuais soluções para uma aproximação da PSP ao fenómeno do *Big Data*.

Assim, tal como explanado no enquadramento teórico, os modelos de policiamento existentes nunca são aplicados de uma forma pura. Aliás, nem sequer é necessário que a sua aplicação seja feita em regime de exclusividade, o que significa que cada polícia pode extrair aquilo que achar mais relevante desses modelos de policiamento para construir as suas próprias estratégias e táticas. Desta forma, aproximamo-nos mais da realidade de cada Polícia e, se assim for decidido, de cada esquadra. Tenhamos como exemplo a potencial conjugação do POP com o POI já referido neste trabalho. Podemos e devemos aproveitar lições extraídas do modelo de implementação do POP concebido por Scott e Kirby (2012), conjugando-o com a metodologia analítica do POI. Deste modo, garantimos um modelo completo de implementação de alterações profundas na instituição, proporcionando uma mudança no paradigma policial no seio da PSP.

Outra das soluções para a PSP que se nos afigura plausível perante este estudo é a implementação e desenvolvimento de uma plataforma, tal como as apresentadas ao longo desta dissertação, que consiga responder às necessidades da PSP e melhorar os seus resultados operacionais. Deste modo, essa plataforma deverá ter uma arquitetura escalável, isto é, que seja expansível para quando seja necessário fazer melhoramentos ao sistema, deixando espaço para que atualizações e inovações sejam feitas aos sistemas, seja em que vertente for (recolha, armazenamento, tratamento ou processamento). Esta plataforma tem de estar preparada para apresentar a informação que lhe é pedida numa variedade de formatos. Esta exigência fundamenta-se não só na utilidade que ela poderá ter para o apoio à análise da criminalidade, mas também à própria tomada de decisão. Tomada de decisão esta que pode ser do patrulheiro que se depara com um suspeito ou uma viatura e da informação fundamental para a resolução da ocorrência ou do dirigente máximo que precisa de delinear estratégias fundamentadamente e com base em informações o mais aproximadas da realidade possível. No mesmo sentido, tal como foi referido ao longo desta dissertação, as tecnologias e métodos de *Big Data* podem servir de auxílio à execução das missões atribuídas às Polícias e aos processos inerentes a cada modelo de policiamento. Remetendo novamente para a situação comum do patrulheiro, poderão surgir situações onde ter a informação celeremente pode determinar o sucesso da resolução da ocorrência.

Deste modo, a introdução de uma plataforma de gestão e tratamento da informação com base nos exemplos de soluções de *Big Data* constitui-se como um potencial aliado à prossecução das missões atribuídas à PSP, desde que estas sejam implementadas holisticamente segundo modelos planeados e adaptados à instituição. Neste âmbito, podemos referir a solução apresentada pela Google, que combina a base do *Hadoop* com outras aplicações facilmente reprogramáveis e, portanto, com um elevado grau de adaptabilidade às condicionantes de cada organização. A grande vantagem desta solução, sendo esta implementada na PSP seria o aproveitamento dos recursos tecnológicos atualmente existentes, pois cada computador passaria a ser um *node*, isto é, passaria a contribuir para o armazenamento e processamento da informação. Uma outra vantagem derivada desta solução seria o facto de não ser necessário recorrer a entidades externas para estes processos, o que significa que a informação ficaria permanentemente na posse da PSP. Por outro lado, esta plataforma deveria ser compatível com os restantes sistemas e deveria integrar recursos de avaliação do desempenho do sistema (*benchmarking*), dos métodos e da satisfação dos utilizadores finais com a mesma. A plataforma em questão deveria não só possibilitar a integração de informação dos sistemas já existentes na PSP, mas também informação proveniente de fontes externas que se revelassem relevantes à atividade da PSP.

Do ponto de vista do utilizador, surgiria a necessidade de formação adicional dos elementos da PSP para a adaptação a estas alterações. Em linha com este pensamento, a formação deveria ser a chave para o aumento dos níveis de confiança nestas tecnologias, para aumentar o desempenho geral da instituição e para aumentar a qualidade dos dados. Quanto a este último ponto, poderiam ainda ser criados mecanismos de responsabilização, que vimos ser importantes ao longo desta dissertação.

Na atualidade, a existência de equipas de supervisão dos dados na PSP seria uma mais valia neste processo na medida em que poderíamos aproveitar estes recursos profissionais e dotá-los de mais conhecimentos nesta matéria. É evidente, no entanto, que uma equipa única não será suficiente para tratar todos os dados que são inseridos no SEI diariamente, garantindo, assim, a qualidade desejada e necessária para o seu posterior tratamento. Como tal, estas poderiam servir de complemento a novos métodos de tratamento de dados que esta plataforma integraria.

Estas alterações iriam implicar mudanças do ponto de vista cultural e estrutural. Deste modo, seria primordial desenvolver medidas para a gestão e acompanhamento desta mudança segundo essas duas perspetivas, sob pena, tal como foi demonstrado no

enquadramento teórico, da ineficácia das mesmas. Um dos mecanismos passíveis de ser utilizados tanto na fase de implementação, como na fase de desenvolvimento destes potenciais sistemas, seria a realização do diagnóstico periódico da perspetiva e das necessidades dos utilizadores.

Apesar das várias conclusões retiradas deste estudo, houveram algumas dificuldades preponderantes na sua elaboração. Em primeiro lugar, o tempo disponível para a realização da dissertação foi manifestamente curto, o que dificultou o desenvolvimento que se expectava das várias fases da investigação. Em segundo lugar, a dificuldade em obter respostas às entrevistas planeadas para esta dissertação, que foram enviadas atempadamente, por questões alheias à nossa compreensão. Em terceiro lugar, encontrar relatórios atualizados do SEI, por inexistência ou por indisponibilidade dos mesmos, que permitam fazer uma caracterização atualizada do sistema.

Este estudo, pelo seu carácter exploratório, teve como principal intuito uma condensação deste conceito abrangente que é o *Big Data*, procurando uma primeira aproximação à PSP e, mais concretamente, ao SEI. Neste sentido, foi possível concluir que existem potencialidades na implementação deste tipo de soluções, desde que esse processo seja devidamente planeado, gerido e avaliado. Nesta senda, podemos apenas ressaltar as inúmeras perspetivas sob as quais investigações futuras poderão ser desenvolvidas com apoio nas sugestões supramencionadas. Para além destas, perante os resultados obtidos na presente dissertação sugerimos que outros estudos sejam feitos no sentido de perceber em que ponto se encontra a implementação concreta do que foi analisado no enquadramento teórico, ou seja, perceber em que medida as componentes basilares *Big Data* são concretizados no contexto atual da PSP e do SEI.

Bibliografia

- Accenture / Polícia de Segurança Pública. (2002). Plano Estratégico de Sistemas de Informação. S.I.: Accenture / Polícia de Segurança Pública
- Accenture. (2009). Estudo SEI+ - Estudo de utilização e evolução do SEI – Sistema estratégico de informação gestão e controlo operacional da PSP. S.I.: Accenture
- Agnellutti, C. (Eds.). (2014). *Big Data: An exploration of opportunities, values, and privacy issues*. Retrieved from: <https://eds.a.ebscohost.com/eds/detail/detail?vid=1&sid=8923adf5-56a8-4970-88ec-496864751989%40sdc-v-sessmgr01&bdata=JnNpdGU9ZWRzLWxpdmU%3d#AN=811106&db=nlebk>
- Babuta, A. (2017). *Big Data and policing: An assessment of law enforcement requirements, expectations and priorities*. London: The Royal United Services Institute
- Baraças, B. (2009). *Teoria e Modelos de Polícia e Ética Policial - Do Policiamento de Proximidade ao Policiamento Orientado pelas Informações*. Lisboa: Instituto Superior de Ciências Sociais e Políticas.
- Bardin, L. (1977). *Análise de Conteúdo*. Lisboa: Edições 70.
- Bardin, L. (2018). *Análise de Conteúdo* (4.^a ed.). Lisboa: Edições 70.
- Barrett, M. (2006). The Need for Intelligence-Led Policing. *Domestic Preparedness Journal*.
- Bayley, D. H. (2016). The complexities of 21st century policing. *Policing: A Journal of Policy and Practice*, 10(3), 163–170. doi:10.1093/police/paw019
- Beyer, M.A. & Laney, D. (2012). *The importance of “Big Data”*: A definition (Report No. G00235055) Retrieved from Gartner: <https://www.gartner.com/en>
- Bovens, M. A. P., Schillemans, T. & Goodin, R.E. (2014). Public accountability in MAP. In M. A. P. Bovens, R. E. Goodin & T. Schillemans (eds.) *The oxford handbook of public accountability* (1–18). doi:10.1093/oxfordhb/9780199641253.001.0001
- Boyd, D. & Crawford, K. (2012). Critical questions for *Big Data*: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15 (5), 662-679. doi:10.1080/1369118X.2012.678878

- Brown, E. (2016). 21 Open Source Projects for IoT. *Linux.com*. Retrieved from: <https://www.linux.com/NEWS/21-OPEN-SOURCE-PROJECTS-IOT>
- Buhl, H.U., Röglinger, M., Moser, F. & Heidemann, J. (2013), *Big Data. Business & Information Systems Engineering*, 5(2), 65-69. doi:10.1007/s11576-013-0350-x
- Candela, L., Castelli, D., Ferro, Y., Koutrika, G., Meghini, C., Pagano, P., Ross, S., Soergel, D., ... Schuldt, H. (2007). The DELOS digital library reference model. Foundations for digital libraries (Report No. 507618) Retrieved from: https://www.researchgate.net/publication/225035550_The_DELOS_Digital_Library_Reference_Model_Foundations_for_Digital_Libraries_Version_096
- Candela, L., Manghi, P. & Ioannidis, Y. (2012), Fourth workshop on very large digital libraries: on the marriage between very large digital libraries and very large data archives, *ACM SIGMOD Record*, 40(4), 61-64. doi:10.1145/2094114.2094130
- Caragliu, A., Del Bo, C. & Nijkamp, P. (2011). Smart cities in Europe. *Journal of Urban Technology*, 18(0048). doi: 10.1080/10630732.2011.601117
- Chan, J. & Moses, L. B. (2016). Making sense of *Big Data* for security. *The British Journal of Criminology*, 57(2), 299–319. doi:10.1093/bjc/azw059
- Chen, H., Chiang, R. & Storey, V. (2012). Business intelligence and analytics: From *Big Data* to big impact. *MIS Quarterly*, 36(4), 1165-1188. doi:10.2307/41703503
- Chen, H. & Yan, Z. (2016). Security and privacy in *Big Data* lifetime: A review. doi:10.1007/978-3-319-49145-5_1
- Clemente, P. J. L. (2010). Polícia e Segurança – breves notas. *Lusíada: Política Internacional e Segurança*, 4, 141-171. Disponível em: <http://hdl.handle.net/11067/1005>
- Coyle, K. (2006). Mass digitization of books, *Journal of Academic Librarianship*, 32(6), 641-645. doi:10.1016/j.acalib.2006.08.002
- Crawford, K. (2014). The Anxieties of *Big Data*, *The New Inquiry*. Retrieved from: <https://thenewinquiry.com/the-anxieties-of-big-data/>

- De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of *Big Data* based on its essential features. *Library Review*, 65(3), 122–135. doi:10.1108/LR-06-2015-0061
- Dhanklad, S. (2019). A Brief Summary of Apache Hadoop: A Solution of *Big Data* Problem and Hint comes from Google, *Towards Data Science*. Retrieved from: <https://towardsdatascience.com/a-brief-summary-of-apache-hadoop-a-solution-of-big-data-problem-and-hint-comes-from-google-95fd63b83623?fbclid=IwAR2qGnPlGHsgtoW4Za0IgJsD2uFMB5hv4xMLZaouy1OgdNhlPUJdSJyW0b0>
- Dijcks, J. (2013). *Oracle: Big Data for the enterprise*, Oracle White Paper, Oracle Corporation, Redwood Shores, CA.
- Dumbill, E. (2013). Making sense of *Big Data*. *Big Data*, 1(1), 1-2. doi:10.1089/big.2012.1503
- Elias, L. (2018). *Ciências Policiais e Segurança Interna. Desafios e Prospetiva*. Lisboa: Instituto Superior de Ciências Policiais e Segurança Interna
- Evans, R. M. (2012). The Diamond Matrix: A Science-Driven Approach to Policing with Crime Intelligence. *Policing: A journal of police and practice*, 6(2), pp. 133-143. doi: 10.1093/police/par063
- Ferguson, A. G. (2017). *The rise of Big Data policing: Surveillance, race, and the future of law enforcement*. New York: New York University Press.
- Fernandes, L. F. (2014), *Intelligence e Segurança Interna*. Lisboa: Instituto Superior de Ciências Policiais e Segurança Interna
- Fisher, D., DeLine, R., Czerwinski, M. & Drucker, S. (2012). Interactions with *Big Data* analytics. *Interactions*, 19(3), 50-59. doi:10.1145/2168931.2168943
- Flick, U. (2005). *Métodos qualitativos na investigação científica*. Lisboa: Monitor.
- Goldstein, H. (1979) Improving policing: A problem-oriented approach. *Crime & Delinquency*, pp. 236-258. doi: 10.1177/001112877902500207
- Goldstein, H. (1990). *Problem-oriented policng*. New York: McGraw-Hill
- Goldstein, H. (1995). Problem-oriented policing: The rationale, the concept and reflections on its implementation (p.13). Presented at the conference on crime

control in theory and practise, the robinson college, cambridge university,
england: home office

Google. (2019, Fevereiro 6). Google books library project – an enhanced card catalog of the world’s books. Retrieved from: www.google.com/googlebooks/library

Intel IT Center. (2012). *Big Data analytics: Intel’s IT manager survey on how organizations are using Big Data*. Intel IT Center, Intel Corporation, Santa Clara, CA.

Jagadish, H. V. (2015) *Big Data and Science: Myths and Reality*. *Big Data Research*, 2(2), pp. 49-52. doi: 10.1016/j.bdr.2015.01.005

Jifa, G. (2013). Data, information, knowledge, wisdom and meta-synthesis of wisdom-comment on wisdom global and wisdom cities. *Procedia Computer Science*, 17(2013), 713-719. doi:10.1016/j.procs.2013.05.092

Joh, E. E. (2014). Policing by numbers: *Big Data* and the fourth amendment. *Washington law review*, 89(35), pp.35-68.

Kalron, A. (2018, Fevereiro 12). *How do Hadoop and Spark Stack Up?*. Retrieved from: <https://logz.io/blog/hadoop-vs-spark/>

Lacinák, M., & Ristvej, J. (2017). Smart City, Safety and Security. *Procedia Engineering*, 192, pp.522-527. doi: 10.1016/j.proeng.2017.06.090

Laney, D. (2001). 3-D data management: Controlling data volume, velocity and variety. *META Group*. Retrieved from: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>

Lemos, J. M. L. (2011). *Sistemas de informação e qualidade de dados: O caso do sistema estratégico de informação, gestão e controlo operacional da polícia de segurança pública* (Dissertação de Mestrado, não publicada). Instituto Superior de Ciências Policiais e Segurança Interna, Lisboa.

Lessard-Hébert, M., Goyette, G. & Boutin, G. (2008). *Investigação qualitativa: Fundamentos e práticas*. (3ª Ed.). Lisboa: Instituto Piaget.

- Lohr, (2013). The Origins of 'Big Data': An Etymological Detective Story, [*Bits*], Retrieved from: <https://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/>
- Manovich, L. (2012). Trending: the promises and the challenges of big social data. Retrieved from: <http://manovich.net/content/04-projects/067-trending-the-promises-and-the-challenges-of-big-social-data/64-article-2011.pdf>
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. & Byers, A.H. (2011). *Big Data: The next frontier for innovation, competition, and productivity*. New York: McKinsey Global Institute
- Martins, D. F. (2013). *A Aplicabilidade do Policiamento Orientado pelas Informações na GNR*. (Dissertação de Mestrado, não publicada). Academia Militar, Lisboa.
- Mayer-Schönberger, V. & Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray.
- McAfee, A. & Brynjolfsson, E. (2012). *Big Data: the management revolution*. *Harvard Business Review*, 90(10), 61-67. Retrieved from: <https://hbr.org/2012/10/big-data-the-management-revolution>
- Michel, J. B., Shen, Y. K., Aiden, A., Veres, A., Gray, M. K., Pickett, J. P., ... Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, 331(6014), 176-182. doi:10.1126/science.1199644
- Microsoft Research (2013), *The big bang: how the Big Data explosion is changing the world*. Retrieved from: <https://news.microsoft.com/2013/02/11/the-big-bang-how-the-big-data-explosion-is-changing-the-world/>
- Mohanty, S. P., Choppali, U. & Kougianos, E. (2016) *Everything You Wanted to Know About Smart Cities*. doi: 10.1109/MCE.2016.2556879
- Moore, G.E. (2006). Cramming more components onto integrated circuits, reprinted from *Electronics*, volume 38, number 8, April 19, 1965, pp.114 ff. *IEEE Solid-State Circuits Newsletter*, 11(5), 33-35. doi:10.1109/N-SSC.2006.4785860
- Nagin, D. S., Solow, R. M., & Lum, C. (2015). Deterrence, criminal Opportunities, and police. *Criminology*, 53(1), 74–100. doi:10.1111/1745-9125.12057

- NBD-PWG (2014). NIST *Big Data* public working group: Draft of *Big Data* Definition. Retrieved from: www.bain.com/images/bain_brief_big_data_the_organizational_challenge.pdf.
- Olesker, A. (2012). White paper: *Big Data* solutions for law enforcement. *CTOlabs*. Retrieved from: <https://core.ac.uk/download/pdf/30678906.pdf>
- Pais, L. G. (2004). *Uma história das ligações entre a psicologia e o direito em Portugal: Perícias psiquiátricas médico-legais e perícias sobre a personalidade como analisadores* (Tese de doutoramento não publicada). Faculdade de Psicologia e de Ciências da Educação da Universidade do Porto, Porto.
- Perry, W. L., McInnis, B., Price, C. C., Smith, S. C. Hollywood, J. S. (2013). *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. Santa Monica: RAND Corporation
- Peterson, M. B., Smith, A., & Fahlman, R. C. (Eds.). (1997). *Intelligence led policing. International perspectives on policing in the 21st century*. New Jersey: International Association of Law Enforcement Intelligence Analysts.
- Polícia de Segurança Pública. (2013). *Visão global de operacionalização da estratégia para as TIC na PSP 2013-2016*. Lisboa: Direção Nacional da Polícia de Segurança Pública.
- Polícia de Segurança Pública. (2016). *Grandes Opções Estratégicas para a PSP 2017-2020*. Lisboa: Direção Nacional da Polícia de Segurança Pública.
- Ramos, S., Coutinho, C. P. (2011). Metodologia da Investigação em Ciências Sociais e Humanas: Teoria e Prática. *Interações: Sociedade e as novas modernidades*, 11(20).
- Ratcliffe, J. (2002). Intelligence-led policing and the problems of turning rhetoric into practice. *Policing and Society: An International Journal of Research and Policy*, 12(1), 53-66. doi: 10.1080/10439460290006673
- Ratcliffe, J. (2016). *Intelligence-led policing*. New York: Routledge.
- Regulamento de formação do sistema estratégico de informação (SEI), de 1 de fevereiro de 2011.

- Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163-180. doi: 10.1177/0165551506070706
- Sampieri, R. H., Collado, C. H., & Lucio, P. B. (2006). *Metodologia de pesquisa* (3ª Ed.). São Paulo: McGraw-Hill.
- Santo, P. E. (2015). *Introdução à Metodologia das Ciências Sociais – Génese, Fundamentos e Problemas*. (2.ª Ed.). Lisboa: Edições Sílabo.
- Serrão, J. C., & Pais, L. G. (2006). A delinquência juvenil segundo os seus protagonistas: Uma reflexão sobre o género. In C. Machado, L. Almeida. M. A. Guisande, M. Gonçalves, & V. Ramalho (Coords.), *Actas da XI Conferência Internacional de Avaliação Psicológica: Formas e Contextos* (pp.149-158). Braga: Psiquilíbrios.
- Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D. and Tufano, P. (2012). Analytics: The Real-World Use of *Big Data*. Retrieved from: http://www.informationweek.com/pdf_whitepapers/approved/1372892704_analytics_the_real_world_use_of_big_data.pdf
- Scott, M. S., Kirby, S. (2012). *Implementing Pop: Leading, Structuring, and Managing a Problem-Oriented Police Agency*. Washington D.C: United States. Department of Justice. Office of Community Oriented Policing Services.
- Skogan, W. G., & Fryll, K. (Eds.) (2004) *Fairness and effectiveness in policing: The evidence* (National Research Council.). Washington, D.C: The National Academies Press
- Somers, J. (2017) Torching the Modern-Day Library of Alexandria, *The Atlantic*. Retrieved from: <https://www.theatlantic.com/technology/archive/2017/04/the-tragedy-of-google-books/523320/>
- Smartcities. (2019). Smart cities – Cidades sustentáveis. Retrieved from: <http://smart-cities.pt/>
- Speranza, A. (2017). *Big Data* and Law Enforcement: On Predictive Policing. *Medium*. Retrieved from: <https://medium.com/@alishope/big-data-and-law-enforcement-on-predictive-policing-a16afd882dd2>

- Suthaharan, S. (2014). *Big Data* classification: Problems & challenges in network intrusion prediction with machine learning. *Performance Evaluation Review*, 41(4), pp. 70-73. doi: 10.1145/2627534.2627557
- Tapscott, D. & Tapscott, A. (2016) *Blockchain revolution: Ho the technology behind bitcoin is changing money, business and the world*. London: Penguin books.
- Thuma, J. (2018). What is Apache Hadoop?, *Dataseries*. Retrieved from: <https://medium.com/dataseries/fnu-for-non-unicorns-what-is-apache-hadoop-991230b704c?fbclid=IwAR0wdjV18i2mOwZ8AxogmYBBOeJKI4pMF3d5UaEZawdjHtPkcCBtdkXphA>
- Treverton, G. F., Wollman, M., Wilke, E., & Lai, D. (2011). *Moving toward the future of policing*. Santa Monica: RAND Corporation.
- Vala, J. (1999). A análise de conteúdo. In A. Silva, & J. Pinto (Eds.), *Metodologia das ciências sociais* (10ª Ed.), pp. 101-128. Lisboa: Edições Afrontamento
- Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, 12(4), pp. 5–33. Retrieved from: <https://doi.org/10.2307/40398176>.
- Wang, R. Y., Ziad, M., & Lee, Y. W. (2002). *Data Quality. Advances in Database Systems*. Dordrecht: Kluwer Academic Publishers.
- Ward, J.S. & Barker, A. (2013). *Undefined by data: a survey of Big Data definitions*. Retrieved from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.705.9909&rep=rep1&type=pdf>
- Wyllie, D. (2013a). How ‘*Big Data*’ is helping law enforcement, *PoliceOne*. Retrieved from: <https://www.policeone.com/police-products/software/Data-Information-Sharing-Software/articles/6396543-How-Big-Data-is-helping-law-enforcement/>
- Wyllie, D. (2013b). Rise of the crime analyst, *PoliceOne*. Retrieved from: <https://www.policeone.com/police-products/software/Data-Information-Sharing-Software/articles/6396540-Rise-of-the-crime-analyst/>
- Wyllie, D. (2011). IBM supercomputers help law enforcement gather, analyze, and manage crime data, *PoliceOne*. Retrieved from: <https://www.policeone.com/police-products/software/Data-Information-Sharing-Software/articles/4241913IB>

M-supercomputers-help-law-enforcement-gather-analyze-and-manage-crime-data/

Xiong, W., Yu, Z., Bei, Z., Zhao, J., Zhang, F., Zou, Y. & Xu, C. (2013). A characterization of *Big Data* benchmarks. In *Proceedings – 2013 IEEE International Conference on Big Data, Big Data* (pp. 118-125). Retrieved from: <http://doi.org/10.1109/BigData.,2013.6691707>

Zaslavsky, A., Perera, C. & Georgakopoulos, D. (2013). *Sensing as a service and Big Data*. Retrieved from: https://www.researchgate.net/publication/234017925_Sensing_as_a_Service_and_Big_Data

Zikopoulos, P., Eaton, C., deRoos, D., Deutsch, T. & Lapis, G. (2011), *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. New York: McGraw

ANEXOS

Anexo 1 – Autorização para a realização de entrevistas

POLÍCIA SEGURANÇA PÚBLICA

INSTITUTO SUPERIOR DE CIÊNCIAS POLICIAIS E SEGURANÇA INTERNA
DIRECÇÃO DE ENSINO
SECRETARIA ESCOLAR



Exmo. Senhor
Diretor Nacional Adjunto/Unidade Orgânica de Recursos
Humanos
(Departamento de Formação)
DN/PSP Largo da Penha de França, N.1
1199-010 LISBOA

Sua Referência:
Sua Comunicação:

Nossa Referência: 97/SECDE/2019

Classificador: 080.01.10

Processo: SECDE201800001ASP

Data: 2019-03-19

Assunto: PEDIDO DE COLABORAÇÃO EM TRABALHO DE DISSERTAÇÃO DE Mestrado Integrado em Ciências Policiais

1. O Curso de Mestrado Integrado em Ciências Policiais (CMICP), no 5.º ano - Estágio, compreende a elaboração de uma dissertação/trabalho de projeto que deverá, obrigatoriamente, incidir sobre um tema das áreas científicas de ciências policiais, ciências jurídicas, ciências sociais e humanas e/ou ciências de desenvolvimento e adaptação motora.
2. O Aspirante a Oficial de Polícia Miguel Rodrigues Pereira irá realizar o seu estudo numa daquelas áreas científicas, subordinado ao tema "Big data: O caso do Sistema Estratégico de Informação, Gestão e Controlo Operacional da Polícia de Segurança Pública", do qual é orientador o Sr. Superintendente Sérgio Felgueiras.
3. Deste modo, solicita-se a V.ª Ex.ª autorização para realização de entrevista às seguintes individualidades da PSP:
 - 3.1 Exmo. Sr. Diretor do Departamento de Sistemas de Informação e Comunicações, Superintendente João Ribeiro;
 - 3.2 Exmo. Sr. Chefe do Núcleo de Sistemas de Informação e Comunicações do COMETLIS, Comissário Diogo Paulo;
 - 3.3 Exmo. Sr. Chefe do Núcleo de Informações Policiais do COMETPOR, Comissário Jorge Teixeira;
 - 3.4 Exmo. Sr. Chefe do Núcleo Sistemas de Informação e Comunicações do COMETPOR, Subcomissário Nuno Lopes.
4. Atendendo à natureza do estudo as entrevistas revelam-se pertinentes na recolha de dados, visando constituir um capítulo da dissertação que sustente todo o trabalho de investigação realizado.
5. Mais se informa V.ª Ex.ª de que o Aspirante a Oficial de Polícia Miguel Pereira se compromete a obter de todos os participantes um termo de consentimento informado, bem como a utilizar as respostas obtidas apenas no âmbito da dissertação em questão.

O Diretor em substituição

Paulo Jorge Gonçalves Sampaio
Superintendente



R. 1.º de Maio, nº3 1349-040 Lisboa Tel.: 213613900 Fax: 213610535 www.iscpsi.pt |
iscpsi@psp.pt

147458
Página 1/1

Anexo 2 - Grelha de análise de conteúdo (com exemplos)

Categoria A. – Informação – Nesta categoria inserem-se as ideias relacionadas com os atributos dos dados e com as operações tendentes à *digitization* e *datafication*.

Subcategoria A.1. – “**Organização e estruturação**” – Codificam-se nesta subcategoria as u.r referentes à capacidade de o sistema organizar e estruturar dados.

Exemplo: “a existência de um repositório estruturado permitirá a exploração da informação segundo diferentes perspetivas”

Subcategoria A.2. – “**Introdução**” – Codificam-se nesta subcategoria as u.r. relacionadas com a introdução de dados no sistema.

Exemplo: “O futuro sistema deverá assim valorizar e registar todos os tipos de informação disponíveis”

Subcategoria A.3. – “**Atributos**” – Codificam-se nesta subcategoria as ideias relacionadas com os atributos dos dados em *Big Data* (volume, velocidade e variedade).

Exemplo: “A quantidade e qualidade de dados existentes nos sistemas de informação atuais da PSP motivam acrescidas preocupações de gestão da informação (*Big Data*)”

Categoria B. – Tecnologias – Nesta categoria insere-se a informação referente aos requisitos tecnológicos associados ao *Big Data*.

Subcategoria B.1. – “**Processamento**” – Codificam-se nesta subcategoria as ideias relativas ao poder de processamento necessário perante dados com tais atributos (volume, velocidade e variedade).

Exemplo: “Ao nível central, será necessário criar novas infra-estruturas de processamento, constituindo um ambiente de produção para o novo Sistema de Informação”

Subcategoria B.2. – “**Armazenamento**” – Codificam-se nesta subcategoria as ideias relativas à capacidade de armazenamento disponível nos sistemas da PSP.

Exemplo: *“O Data Center da PSP deverá ser alvo de renovação ou de eventual substituição, em função de uma avaliação mais detalhada das futuras necessidades em termos de instalações físicas.”*

Subcategoria B.3. – **“Transmissão”** – Codificam-se nesta subcategoria as ideias relacionadas com a transmissão de dados na rede da PSP.

Exemplo: *“Outro aspecto a ter em conta são as dificuldades evidenciadas pela rede de comunicações que não disponibiliza o acesso às aplicações nas condições em que o deveria fazer”*

Subcategoria B.4. – **“Interoperabilidade”** – Codificam-se nesta subcategoria as u.r. relacionadas com a compatibilidade e interoperabilidade dos vários sistemas.

Exemplo: *“Integração do SEI com outros sistemas da PSP (por exemplo, Gestão Documental) e de entidades externas (por exemplo, Tribunais)”*

Subcategoria B.5. – **“Inovação”** – Codificam-se nesta subcategoria as ideias relacionadas com a perspectiva da instituição em relação às necessidades inovação nos seus sistemas.

Exemplo: *“esforços na modernização desta área, procurando, de forma contínua e estratégica, alinhar as suas ferramentas e sistemas aos requisitos da sua atividade operacional e à rápida e transversal evolução da Tecnologia a nível mundial.”*

Subcategoria B.6. – **“Avaliação das tecnologias”** – Codificam-se nesta subcategoria as u.r. relativas a mecanismo de avaliação de desempenho dos sistemas (*benchmarking*).

Categoria C. – Métodos – Nesta categoria inserem-se as ideias relacionadas com o tratamento e manuseamento dos dados nos sistemas.

Subcategoria C.1. – **“Análise e Produção de conhecimento”** – Codificam-se nesta subcategoria as ideias relativas aos métodos de análise de informação com vista à produção de conhecimento.

Exemplo: *“implementação de ferramentas e processos que permitam a criação, exploração e partilha do conhecimento.”*

Subcategoria C.2. – “**Qualidade dos dados**” – Codifica-se nesta subcategoria as ideias relativas à veracidade dos dados, bem como aos processos de preparação para o seu tratamento.

Exemplo: “Módulo de qualidade de dados que permita gerir a informação capturada, visualizá-la, tratá-la para que o produto final seja um conjunto de dados codificados, coerentes e com qualidade.”

Subcategoria C.3. – “**Objetividade da análise**” – Codifica-se nesta subcategoria o grau de isenção e precisão da análise.

Subcategoria C.4. – “**Formação**” – Codificam-se nesta subcategoria as u.r. referentes às necessidades de formação para o manuseamento do SEI e restantes sistemas.

Exemplo: “um aumento quantitativo e qualitativo da componente formação é essencial para responder às cada vez maiores exigências a que estão sujeitas as organizações policiais”

Subcategoria C.5. – “**Avaliação de métodos**” – Codificam-se nesta subcategoria as ideias relativas à avaliação da implementação e do desempenho dos métodos.

Exemplo: “Avaliação de alternativas de operacionalização dos processos (revisão das atividades contempladas, sistemas de informação para suporte às funções, externalização de tarefas repetitivas e facilmente “industrializáveis”, ...)”

Categoria D. – Impacto – Nesta categoria inserem-se as u.r. referentes aos impactos que as alterações subjacentes ao *Big Data* têm na instituição policial e na sociedade.

Subcategoria D.1. – “**Suporte à atividade operacional**” – Codifica-se nesta subcategoria o grau de suporte que o sistema fornece à tomada de decisão e à atividade operacional.

Exemplo: “Exploração das TIC, a qual determina a aposta contínua nas TIC como suporte à capacitação operacional da PSP”

Subcategoria D.2. – “**Privacidade e proteção de dados pessoais**” – Codificam-se nesta subcategoria as ideias relativas às medidas tendentes à proteção da privacidade e dos dados pessoais.

Exemplo: “*garantir a privacidade e integridade da informação, permitir utilização dos sistemas para suportar realização dos processos policiais*”

Subcategoria D.3. – “**Reestruturação e planeamento**” – Codificam-se nesta subcategoria as u.r. que digam respeito ao planeamento e implementação ao nível institucional de novas soluções tecnológicas.

Exemplo: “*Os objectivos pretendidos para a PSP são complexos e ambiciosos, sendo necessário operar algum tipo de mudança a vários níveis: organizativo, processual e tecnológico, através do lançamento de diversas iniciativas em cada uma das vertentes enunciadas*”

Subcategoria D.4. – “**Mudança cultural**” – Codificam-se nesta subcategoria as medidas tendentes ao acompanhamento e gestão de mudanças na cultura da PSP derivadas à introdução de novas tecnologias.

Exemplo: “*Claramente, um dos pontos chave para a mudança reside nas pessoas. Elas, em primeiro lugar, terão que ser o motor da mudança, sem o qual tecnologia alguma poderá ser bem sucedida.*”

Categoria E. – Polícia – Nesta categoria inserem-se as u.r. que representam uma particular relevância no que concerne à atividade das instituições policiais.

Subcategoria E.1. – “**Responsabilidades**” – Codificam-se nesta subcategoria as ideias relativas à distribuição das responsabilidades no manuseamento e organização dos sistemas.

Exemplo: “*A definição das funções e responsabilidades da organização de SI é um aspecto particularmente relevante por forma a garantir o alinhamento entre as orientações estratégicas da PSP e a estrutura organizacional desejada.*”

Subcategoria E.2. – “**Resultados**” – Codifica-se nesta subcategoria a avaliação final dos resultados das operações e decisões apoiadas no conhecimento extraído.

Exemplo: *“Disponibilização de um conjunto de capacidades de suporte à atividade da PSP no terreno, que permitam aumentar os níveis de performance, segurança e serviço ao Cidadão.”*

Subcategoria E.3. – **“Segurança da informação”** – Codificam-se nesta subcategoria as u.r. relativas às medidas de segurança dos sistemas e de acesso à informação.

Exemplo: *“É essencial a criação e aplicação de políticas globais de segurança da informação, abrangendo de forma coerente as vertentes de procedimentos, responsabilidades e tecnologia, apoiadas ao mais alto nível da organização.”*

Subcategoria E.4. – **“Suporte Administrativo”** – Codificam-se nesta subcategoria as u.r. referentes ao apoio e aos meios dispostos pela organização para a utilização dos sistemas.

Exemplo: *“outra área transversal que deverá contar com um help-desk local de 1ª linha ao nível das unidades orgânicas, mais vocacionado para a vertente de suporte técnico, e um help-desk central de 2ª linha que, embora mais vocacionado para as vertentes funcionais”*

Subcategoria E.5. – **“Acessibilidade e partilha de informação”** – Codificam-se nesta subcategoria as ideias relativas aos meios de acesso e partilha de informação por parte da instituição policial.

Exemplo: *“o modelo promove uma completa integração, coerência e partilha da informação, garantindo a sua disponibilização onde e quando necessário”*