



BIG DATA CONTRA O CRIME: EFEITO MINORITY REPORT

Renan Saisse¹

INTRODUÇÃO

A tecnologia avança sem cessar em uma velocidade imensurável capaz de transformar significativamente a vida em sociedade em pouco tempo. Vale citar a própria *Internet* que em poucos anos revolucionou a forma de comunicação ao redor do mundo. A ficção nunca esteve mais próxima da realidade e a cada dia vemos algo, antes inimaginável, se tornando comum e corriqueiro. Temos recentemente estudos sobre o holograma *wifi* que permitirá a visão através de estruturas sólidas, os colisores de partículas que permitem estudar a origem do universo, carros autônomos, óculos de correção que permitem aos daltônicos enxergar as cores como de fato elas são.

Diante de uma nova realidade, de uma nova Sociedade, que muitos intitulam como Sociedade da Informação, temos uma quantia imensurável de dados sobre quase tudo que existe. Esses dados podem ser obtidos de diversas formas por intermédio de diferentes tecnologias, o que permite a formação do conhecimento para estratégias de negócios, mapeamentos de perfis de clientes para direcionamento de publicidades, previsões com certa precisão de eventos da natureza por meio da captação de dados meteorológicos, climáticos ou sísmicos, estudos geopolíticos por captação de dados em pesquisas de campo ou online, etc. Vivemos em um verdadeiro emaranhado colossal de dados prontos para serem coletados, estudados e interpretados.

1 BIG DATA

Esta nova realidade fez surgir em meados da década de 90 o termo “Big Data” cunhado pela NASA ao descrever a quantidade massiva de dados disponíveis, além das capacidades computacionais e analíticas por processos tradicionais, ou seja, a geração de conhecimento se tornava cada vez mais complexa devido à própria complexidade das informações em quantidades inimagináveis para a época. Trazendo para os dias atuais, um projeto de Big Data, por exemplo, pode explorar as redes sociais a partir da mineração, processamento e análise de dados, gerando conhecimento suficiente para tomadas de decisão. Rodrigo Arrigoni¹, afirma que:

¹ Analista de Sistemas atuante nas áreas de Faturamento, Fraude, Revenue Assurance e Roaming Internacional na Oi S.A. Graduado em Análise e Desenvolvimento de Sistemas (UniverCidade); Pós-Graduado em CyberCrime, CyberSecurity e Perícia Forense Aplicada à Informática (Faculdade Unyleya); Acadêmico do curso de Bacharel em Direito (UFRJ); Possui cursos de Extensão em Ciência Política (USP), Direito Digital (EMERJ) e Linguagem de Programação Python (Harvard); Certificado em ITILV3 e Cobit V4.1; E-mail: rcaisse@gmail.com.



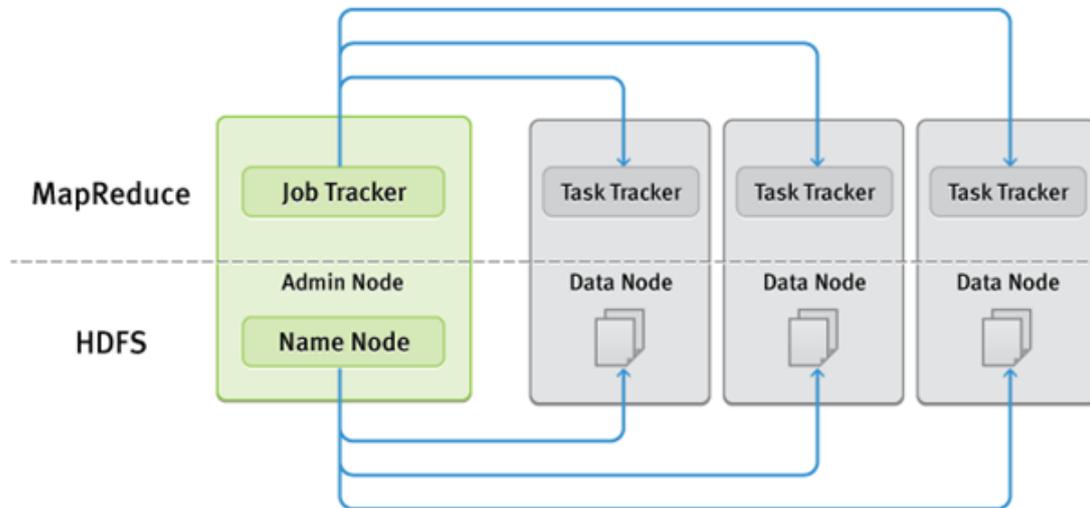
Big Data pode ser considerado como uma região de conhecimento transversal que abrange diversas áreas produtivas e pesquisas científicas. É uma disciplina focada em gerar conhecimento e inteligência a partir de grandes quantidades de dados complexos.

Somente em meados dos anos 2000 que o termo Big Data ganhou a mídia de forma abrangente devido à veiculação de Doug Laney, membro da Gartner, com sua definição em 3 Vs: **Volume**, dados massivos de inúmeras fontes, **Velocidade**, capacidade e tempo de processamento sustentável ao objetivo e a necessidade de processamento em tempo real para diversos segmentos, e **Variedade**, diversos formatos entre vídeos, fotos, *hashs*, transações financeiras, etc. Há ainda a empresa SAS que possui uma família de software SAS (Statistical Analysis System) em que, além dos 3Vs de Laney, utilizam mais duas dimensões para descrever o Big Data que são a **Variabilidade**, inconsistência de fluxos de dados variados com picos sazonais, e a **Complexidade** no qual dispensa definição. Porém, é adotado majoritariamente Volume, Variedade, Velocidade, Veracidade e Valor como os cinco “Vs” do Big Data.

Toda essa estrutura que se tem hoje para coleta de dados se tornou possível com a sofisticação tecnológica, podemos citar a plataforma *Hadoop*, desenvolvida em linguagem java, de caráter distribuído para processamento de dados em *clusters*. Seu desenvolvimento é constante por uma comunidade aberta de contribuidores iniciada pela Apache. O *Hadoop*, brevemente falando, possui dois focos principais: Armazenamento e Processamento. O Armazenamento está no componente HDFS (*Hadoop Distributed File System*), responsável por manipular e armazenar os dados coletados para todas as máquinas do *cluster* onde o *Hadoop* esteja em execução. O HDFS pode ser definido como um sistema de arquivos escalonável e distribuído baseado no GFS (*Google File System*). Já o Processamento é de responsabilidade do componente *Map-Reduce*, que é basicamente um paradigma de programação baseado em conceitos pautados em processamento de grandes massas de dados. A figura abaixo



demonstra graficamente o que venha a ser a estrutura dos componentes no *framework Hadoop*:



Fonte: Apache - <https://www.apache.org/>

O advento do *Big Data* trouxe consigo também a proliferação de novos termos ou *buzzwords*, que apesar de não ser o escopo aqui, vale citar alguns como *fast data*, *slow data*, *small data*, *medium data*, *dark data* e *dirty data*. Esses exemplos ainda não são muito bem explicados, porém vale conferir seus significados para o *Big Data*.

2 TIPOS DE ANÁLISES EM *BIG DATA* (*Big Data Analytics*)

Inicialmente se faz necessário entender que toda e qualquer informação precisa ser coletada, estudada e associada entre si para se obter conhecimento, ou seja, a analítica de dados é primordial. Essa pode ser classificada em Prescritiva, Diagnóstica, Descritiva e Preditiva.

2.1 – Análise Prescritiva

É a análise baseada em estudos de casos ou fatos traçando possíveis consequências sobre ações tomadas. É o tipo de análise utilizada para identificar em cada possível ação quais seriam seus possíveis resultados.

2.2 – Análise Diagnóstica

As análises diagnósticas traçam uma história dos eventos ocorridos para viabilização de tomadas de decisões baseadas em fatos. É a análise que responde as seguintes perguntas: Quem? Quando? Onde? Por que?



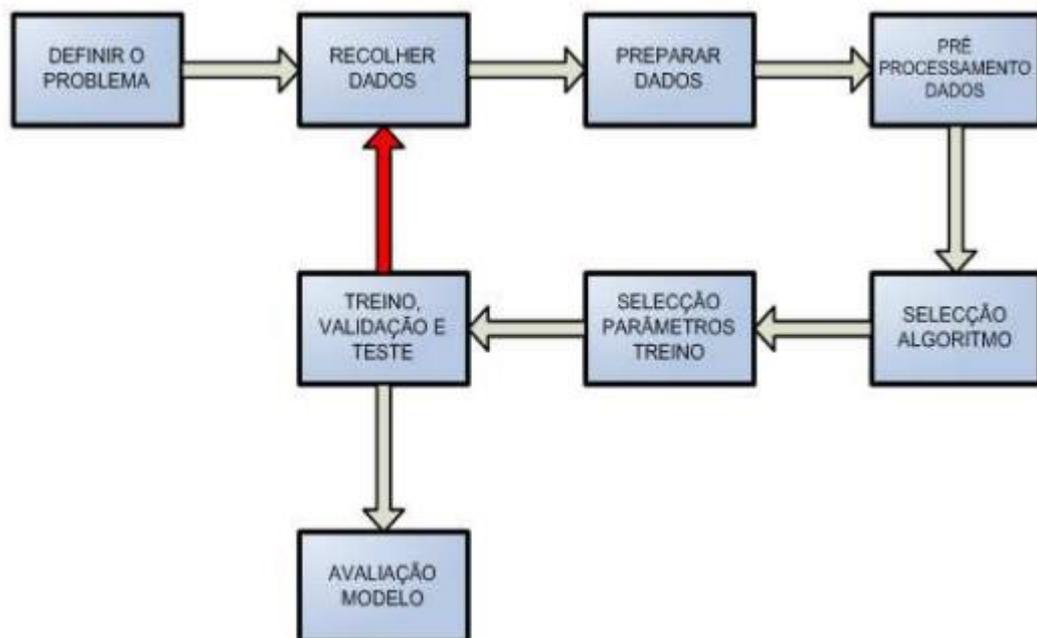
2.3 – Análise Descritiva

É a análise de dados em tempo real. É a principal forma de resposta ao presente, ou seja, é a mineração dos dados captados para entendimento do seu significado atribuindo respostas para o presente momento.

2.4 – Análise Preditiva

É a famosa previsão do futuro. Utiliza a mineração de dados históricos para traçar tendências futuras ou possibilidades futuras. Para se implantar modelos preditivos de análise primeiramente é preciso entender que um modelo corresponde a uma função matemática capaz de aprender e mapear uma gama de variáveis de entrada de dados, sejam os dados agrupados ou não, e uma variável de saída.

Para entender melhor é preciso citar que há três tipos de modelos para a Ciência tradicional, são eles: modelos determinísticos, pautados no conhecimento; modelos paramétricos, pautados em pressupostos e os modelos não-paramétricos que são pautados em dados. Os modelos não-paramétricos que deram origem às atuais técnicas de mineração de dados em grandes quantidades e conforme Bação (2006:47) pode ser representado da seguinte forma:



Fonte: Paulo João (2009:58)



Partindo da representação acima, podemos classificar os modelos não-paramétricos em **modelos de aprendizagem supervisionada e não supervisionada**.

Em modelos preditivos por aprendizagem supervisionada os dados são manipulados ou “ensinados” e introduzidos repetidamente para se obter entradas e saídas desejadas até que se desenvolva o mapeamento automático e reconheça os perfis programados exibindo resultados esperados. Este tipo de modelo utiliza árvores de decisão, redes neurais de retorno de propagação ou máquinas de vetor.

Já em modelos preditivos por aprendizagem não supervisionada há um aprendizado “autodidata”, reconhecendo padrões e categorias nos dados apresentados, interpretando e codificando estes em uma saída.

3 BIG DADA CONTRA O CRIME: EFEITO *MINORITY REPORT*

Os aficionados em tecnologia e o universo criminológico ao seu redor, certamente recordam do filme *Minority Report: A Nova Lei*, dirigido por Steven Spielberg, lançado em 2002 e estrelado por Tom Cruise. Esse filme foi baseado no conto de Philip K. Dick, escritor americano de ficção científica e tem por enredo a cidade de Washington e Northern Virgínia e se passa no ano de 2054. No contexto do filme há um departamento de polícia especializada que promove a paz com a prisão de “criminosos” com base no conhecimento prévio de fatos trazidos à tona por três videntes chamados “Precogs” evitando assim que estes venham a ocorrer. Fazendo uma analogia aos dias de hoje estes “videntes” podem ser comparados às análises preditivas em *Big Data*.

Diante dos acontecimentos recentes como a criação do Observatório Nacional de *Big Data* pelo Governo Argentino que pretende, dentre outros objetivos, realizar análises de dados para previsão de eventos futuros de qualquer natureza ou mesmo o exemplo de algumas cidades dos EUA como Memphis e Nova York que usam, já há alguns anos, modelos matemáticos em uma grande infraestrutura de TI para processamento de dados na prevenção de crimes, ainda longe da realidade do filme citado, porém pode-se dizer que não está tão distante como parece.

A melhor maneira de entender análise criminal com big data é compará-los com a previsão do tempo. Os especialistas podem prever os padrões climáticos futuros, mas ainda são previsões, geralmente limitadas a dez dias, que muitas vezes são alteradas ou estão simplesmente erradas. Richard Janikowski, professor do departamento de criminologia da Universidade de Memphis e um dos responsáveis por implementar a análise de big data na cidade.

3.1 – Polícia Preditiva

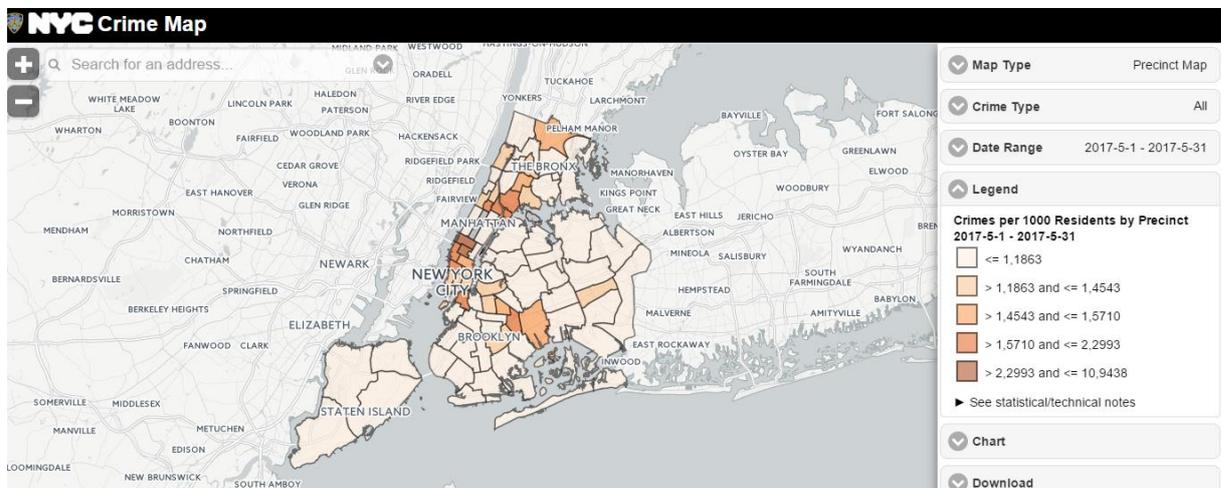
A iniciativa *Blue Crush (Criminal Reduction Utilizing Statistical History)* na Cidade de Memphis onde foram coletadas todas as ocorrências policiais analisando padrões e locais de. Os dados são registrados por PDAs que cada policial carrega, após são coletados pelo SPSS (*Statistical Package*



for the Social Sciences - pacote estatístico para as ciências sociais), um software preditivo da IBM, em conjunto com o GIS (*Geographic Information System*), serviço de mapeamento da empresa ESRI, tornando possível a criação de um mapa com os possíveis locais de novos crimes. A partir desses dados a polícia direciona o patrulhamento para esses locais, trazendo maior precisão e poupando recursos financeiros. O sistema *Crush* também é utilizado na Inglaterra.

Já a polícia de Chicago conta com uma unidade inteira para predição criminal, realizando diligências preventivas a possíveis alvos para conscientização ou averiguação. Chicago utiliza a tecnologia *ShotSpotter* capaz de registrar sons de tiros e sua localização aproximada, porém recentemente integraram esse sistema à mapas computadorizados. Essa tecnologia utiliza algoritmos que identificam padrões e tendências a partir dos dados do *ShotSpotter* de acordo com o relatório da BBC Click. Conta ainda com a integração a um sistema chamado *HunchLab* que ajuda a polícia a tomar decisões de acordo com a análise massiva das informações detenções, chamadas 911, atividades de gangues e outros dados relevantes. Chicago também é um exemplo a ser seguido, pois, de acordo com um relatório em *Chicago Tonight*, a criminalidade está reduzindo na cidade em 13%, porém os dois distritos onde a tecnologia foi implantada viram 49% e 66% de redução nos tiroteios em fevereiro e março de 2017. Steven Sesso, capitão de polícia em Chicago, afirma que “isso nos aproxima desse crime, e nos dá a probabilidade e a possibilidade de contrair infratores e prestar ajuda de maneira mais rápida”.ⁱⁱ

Com a ajuda da Microsoft, Nova York também lançou em dezembro de 2012 o mapa interativo de crimes que possibilita filtrar bairros e ruas, além de verificar quais tipos de crimes possuem maior incidência em cada região. O Mapa é largamente utilizado e pode ser acessado através do endereço <https://maps.nyc.gov/crime/>. O sistema por trás do mapa é o DAS (*Domain Awareness System*) capaz de processar os dados telefônicos das chamadas de emergência (911), imagens de câmeras de segurança, reconhecendo rostos, placas, radiação e até mesmo localizar o policial mais próximo no mapa. Claro que isto gera um largo debate sobre a constitucionalidade da constante vigilância sobre a privacidade que merece atenção. Ray Kelly, comissário de polícia em NYC afirma “O mapa ajuda a melhorar a compreensão dos nova-iorquinos e dos investigadores sobre onde os delitos e os crimes violentos persistem”.ⁱⁱⁱ



Fonte: [NYC Crime Map](#)

3.2 – PredPol (*Predictive Policing Software*)

PredPol é uma iniciativa concreta desenvolvida pelo Departamento de Polícia de Los Angeles em parceria com a UCLA (Universidade da Califórnia em Los Angeles), com o objetivo de descobrir como usar dados do CompStat, que é um sistema/“modelo” que auxilia as forças policiais na identificação dos problemas relacionados à criminalidade e suas soluções, fornecendo ajudas na gestão de recursos e escolha de estratégias, para além de fins históricos, focando em prevenção de fatos futuros. Basicamente o PredPol é utilizado para geoposicionar oficiais de polícia em local e hora corretos para prevenir um possível crime. Toda sua estrutura é baseada em nuvem (arquitetura *cloud*).

A criação do PredPol uniu matemáticos e cientistas da UCLA e da Universidade de Santa Clara para avaliar a extrema variedade de dados e modelos comportamentais e preditivos. Eles contaram com a ajuda de analistas criminais, oficiais da LAPD (Los Angeles Police Department) e o Departamento de Polícia de Santa Cruz. Após as pesquisas realizadas, foi possível antecipar locais e horários de possíveis crimes respondendo três perguntas: Tipo do Crime? Localização? Data e Hora? Segundo informações dos mantenedores do PredPol, a atual plataforma PredPol representa um investimento significativo de mais de 70 anos de pesquisa de análise, modelagem e desenvolvimento de nível de doutorado. Ele foi testado e comprovado em dezenas de departamentos de todos os tamanhos em vários países ao redor do mundo usando apenas três pontos de dados para previsões: o que, onde e quando - não quem.

Pode-se concluir, segundo informações, que o PredPol não utiliza informações pessoais sobre indivíduos ou grupos de indivíduos, eliminando quaisquer ameaças à liberdade pessoal e questões de perfil, assim, há um encorajamento para a estratégia policial preditiva viabilizando a proximidade com parceiros e a comunidade em locais prováveis de crimes.

Eu não vou conseguir mais dinheiro. Eu não vou conseguir mais policiais. Eu tenho que ser melhor em usar o que eu tenho, e é disso que o policiamento preditivo fala. Se esse antigo policial de rua pode mudar a maneira como ele pensa sobre essas coisas,



então sei que meus [oficiais] podem fazer o mesmo. Charlie Beck, Chefe da Polícia de Los Angeles.

Com o uso de *analytics* (análises em big data), milhares de *petabytes* podem ser destrinchados por investigadores para trabalharem de forma preditiva identificando possíveis crimes e ataques. Dados isolados são meros dados, mas quando há uma interligação entre eles com ferramentas de mineração e visualização cria-se valiosos *insights* e respostas.

Big data é, sem dúvida, algo que não sai da nossa cabeça. É fácil perceber a grande relevância para investigação, mas não é fácil implementar. Estamos em um projeto cuja ideia é possibilitar a procura de dados em redes sociais, executar algoritmos de reconhecimento facial sob imagens postadas, ou identificar eventos que podem acontecer baseados no volume de discussão a respeito de um tema. Silvino Schlikmann Júnior. Diretor-assistente de Pesquisa e Inovação da Interpol.

3.3 – Unisys Digital Investigator

A Empresa de Tecnologia Unisys lançou o software *Digital Investigator*, como parte do projeto Cidades Seguras (*Safe Cities*). A Ferramenta possui um portal público para colaboração de cidadãos, acompanhamento de ocorrências por *smartphones* e outros dispositivos. O Sistema pode ser integrado a ferramentas de análise de mídias sociais auxiliando na coleta de evidências digitais. Permite ainda o compartilhamento de dados entre unidades investigativas e de inteligência.

Mark Forman, líder global para o Setor Público da Unisys, cita que a iniciativa 'Cidades Seguras' da Unisys foi criada para oferecer às organizações de Justiça e segurança pública as soluções necessárias para que possam utilizar os dados de fontes como dispositivos móveis e Internet das Coisas da melhor maneira, possibilitando um ambiente público seguro. O *Digital Investigator* é um componente crucial dessa iniciativa. O produto oferece aos departamentos de polícia um conjunto de ferramentas que aumenta a produtividade por meio da capacidade de operar diversos formatos de dados, em múltiplas linhas de investigação, além de interligar diferentes informações relacionadas.

Segundo a Unisys, o *Digital Investigator* permite a coleta de informações além dos limites tecnológicos e geográficos, pois possui os recursos necessários para coletar e analisar a crescente variação e complexidade do conjunto de dados, inclusive de redes sociais, necessários ao trabalho de captura, interceptação de grupos criminosos e terroristas.

3.4 – Banco de Dados Genéticos e a Previsão Comportamental rumo ao Pré-Crime

A primeira etapa de uma visão preditiva baseada em perfis genéticos e comportamentais inicia-se por pesquisas que buscam garantias psíquicas semelhantes através de perfis genéticos. Essa questão nos deixa um passo a frente na luta contra o crime e assim aumentando exponencialmente a justiça



criminal em relação a natureza punitiva. Os estudos realizados buscam trazer à realidade o pré-crime e são tidos como relevantes ao futuro apesar de parecerem assustadores.

Nós já fazemos o que é chamado de polícia preditiva ou polícia de pontos de acesso, que se baseia em estatísticas de onde certos crimes acontecem. Agora, estamos a investigar os comportamentos dos indivíduos para prever se eles podem cometer um ato violento, e a genética é parte disso. Está muito longe do pré-crime, mas eles são definitivamente passos nessa direção. Colleen Berryessa, gerente de programa do Centro de Integração de Pesquisa em Genética e Ética de Stanford.

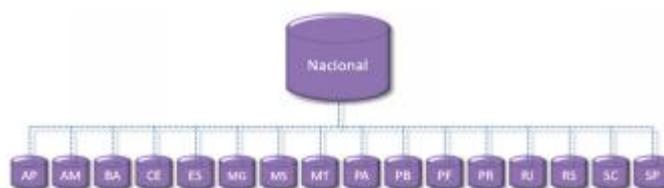
Os bancos de dados de DNA é o cerne da análise preditiva rumo ao pré-crime e seu aumento é uma realidade inevitável. De acordo com a Interpol, 54 países investiram em bancos de dados de DNA somente em 2009. O número cresceu muito desde então, assim como o alcance do que essas bases de dados podem coletar.

Um dos empreendimentos mais ambiciosos está acontecendo nos Emirados Árabes Unidos, que iniciou a criação de um banco de dados central de DNA de seus 5,5 milhões de cidadãos, incluindo os beduínos nativos.^{iv} Iniciado em 2006, e pode-se dizer que seria uma das maiores bases de dados de DNA do planeta.

Em 2010, o Reino Unido criou o Crime and Security Act, ^vpermitindo que os policiais limpassem as faces de criminosos condenados e adicionassem as amostras a um banco de dados de DNA contendo milhões de perfis. As autoridades do Reino Unido inicialmente permitiram a retenção dos dados genéticos por tempo indeterminado, porém duas pessoas absolvidas de crimes foram até o Tribunal Europeu de Direitos Humanos que autorizou a retirada dos seus dados genéticos.

O maior banco de dados de DNA atualmente é o CODIS(Combined DNA Index System - Sistema Combinado de Índices de DNA)^{vi} do FBI nos EUA, que possui informações genéticas de algo em torno de 11 milhões de pessoas suspeitas ou condenadas por crimes, de acordo com a AP. No ano passado, o Supremo Tribunal Americano julgou procedente a decisão que permite que as agências de segurança extraíam DNA de qualquer pessoa que tenha sido presa sem a necessidade de uma autorização judicial.

No Brasil foi implantado o CODIS em 2010, porém alguns Estados não possuíam laboratório de DNA na época para participar da Rede Integrada de Perfis Genéticos. A Rede de Dados integrada de Laboratórios Genéticos possui como membros os seguintes Estados:



Fonte: Instituto Nacional de Criminalística, Brasília/DF2



A meta do Ministério da Justiça é incluir todos os Estados nesta rede de compartilhamento de dados genéticos. Não é o intuito aqui discorrer sobre os dispositivos legais aplicados ao banco de dados de DNA.

3.6 – Big Data e a Segurança no Brasil

Em 2014 o governo do Estado de São Paulo firmou uma parceria com a Microsoft, para utilização da ferramenta desenvolvida pela parceria entre a Microsoft e a Prefeitura de Nova York. O sistema intitulado Detecta possui funções de análise de dados em *big data* e BI indexando e associando dados policiais de fácil consulta e possui o maior cinturão eletrônico de radares e câmeras de vigilância interligado da América Latina e está em constante expansão.

CINTURÃO ELETRÔNICO DO DETECTA (2661 CÂMERAS)	
Câmeras OCR	
Vale do Ribeira	12 câmeras
Baixada Santista e ABC	208 câmeras
Litoral Norte	49 câmeras
Alto Tietê	67 câmeras
Vale do Paraíba	144 câmeras
TOTAL	480 câmeras OCR
Câmeras de videomonitoramento	
Praia Grande	1.500 câmeras
Santos	517 câmeras
Guarujá	72 câmeras
Cubatão	32 câmeras
Itanhaém	28 câmeras
Campos do Jordão	32 câmeras
TOTAL	2181 câmeras

Fonte: Secretaria de Segurança Pública de São Paulo

Temos 2.661 câmeras, não apenas ligadas ao Copo, mas também nos batalhões, nas delegacias seccionais e em todos os tablets nas viaturas. Alexandre de Moraes. Ex-Secretário de Segurança de São Paulo. Ex-Ministro da Justiça. Ministro do STF.

Esse sistema pode ser associado ao sistema Omega^{vii} da Polícia Civil de SP desenvolvido para investigar, mapear e combater a criminalidade no estado. O Omega acessa informações de cadastros dos registros civil, criminal, de armas, veículos roubados e furtados, Junta Comercial, Disque-Denúncia, Delegacia Eletrônica (boletins de ocorrência e inquéritos policiais), e Detran, incluindo Certificado de Registro de Veículo, Carteira Nacional de Habilitação e multas, dados do Infocrim (informações criminais) e mapas, além do sistema de identificação biométrica *Phoenix*, que armazena dados como impressões digitais, gravação de voz e fotos de suspeitos em várias dimensões, sendo possível a



investigação preditiva de suspeitos. O projeto do sistema venceu o Prêmio Mário Covas^{viii} na categoria "Eficiência no Uso dos Recursos Públicos e Desburocratização".

Há também um convênio firmado também em 2014 entre o Governo do Paraná e a Universidade de Chicago para implementar softwares para se ter algo como a já citada unidade de polícia preditiva de Chicago. Na época houve a seguinte colocação do secretário de segurança em exercício:

A intenção é aprimorar o intercâmbio técnico-operacional-científico entre as instituições. É o estilo de uma polícia moderna e que tenta se antecipar aos crimes. E planejamos colocar em prática. Leon Grupenmacher. Secretário de Segurança Pública

Conforme informações veiculadas a parceria é acadêmica e não envolve recursos financeiros.

É uma transferência de tecnologia acadêmica para as instituições, para uma parceria que se pretende ser duradoura. Não é uma interferência, e sim fornecimento de elementos para que a própria Secretaria da Segurança Pública possa desenvolver seus parâmetros e novas metas. Fabio Losso. Pós-doutor em Políticas Públicas e Presidente do Conselho Internacional do Reitor da Universidade de Chicago na época.

As Polícias Judiciárias Estaduais de modo geral estão se modernizando e coletando dados úteis além de desenvolver técnicas de análises diagnósticas e até descritivas para investigações, porém poucas possuem processos de análise preditiva como fonte investigativa para estratégia de policiamento e prevenção de crimes

Já a Polícia Federal está bem a frente nesse quesito, tendo uma gama de softwares que permitem o correlacionamento de dados de diversas fontes. Apesar de se utilizar análises descritivas e diagnósticas, a PF possui a capacidade de implementar também unidades preditivas. Ao longo do tempo vimos ferramentas inovadoras utilizadas pela PF como o Nudetective^{ix}, apresentado em diversos congressos internacionais e desenvolvido pelos peritos criminais Pedro Eleutério e Matheus Polastro. Possui funcionalidades focadas na detecção de pornografia infantil por meio de análises de imagens, nomes, *hashs* e vídeos. O software foi originado a partir da alteração do ECA (Estatuto da Criança e do Adolescente) que tipificou o crime de posse de arquivos de pornografia infanto-juvenil. O *Nudetective* possui princípios que podem ser explorados para expansão conceitual e técnica instituindo uma análise forense direcionada para uma coleta massiva de dados em tempo real.

Outro que merece ser citado é o EspiaMule. ^xSoftware de destaque na PF, realizava pesquisas e coletas de informações de usuários da rede Emule, catalogando endereços e criando um mapa da distribuição de imagens de pornografia infantil, o que atualmente é realizado também na rede P2P *Utorrent* e técnicas são aplicadas também na *Deep Web*. A PF possui também um sistema de informações integradas desde 2016 que permite identificar criminosos estrangeiros, procurados pela



Organização Internacional de Polícia Criminal (Interpol) sempre que eles comprarem passagens aéreas ou embarcarem em um voo com destino ao Brasil. Isso permite identificar onde e quando quaisquer criminosos internacionais estarão caso optem por utilizar as vias aéreas do país. José Eduardo Cardozo, ex ministro da justiça, afirma “Esse sistema de cruzamento de dados possibilita estarmos em pé de igualdade com os países mais desenvolvidos, em termos de segurança, e tem se mostrado exitoso”.^{xi}

Porém a PF mergulhou na era *Big Data* com o advento da Operação Lava Jato, criando um banco de dados unificado e procedimentos analíticos próprios, sendo intitulado como o *big data* para o combate à corrupção.

Como poderíamos desenvolver 38 fases, uma quantidade de material imensa sem um banco de dados unificado? Isso permite uma análise mais rápida e técnica de todo o serviço. Criamos procedimentos para facilitar a vida de quem está investigando. Maurício Moscardi Grillo. Delegado Federal. Coordenador da Operação Lava jato em Bauru.

A partir do modelo criado é possível adotar futuramente uma análise preditiva, caso os dados sejam alimentados constantemente, e assim traçar perfis e tendências relacionadas a transações fraudulentas.

4 FILOSOFIA PRÉ-CRIME E O ORDENAMENTO JURÍDICO BRASILEIRO: BREVES CONSIDERAÇÕES

Primeiramente, apesar do título deste artigo ser pautado na teoria criada pela obra de ficção *Minority Report*, a real aplicação das análises massivas de dados pauta-se sobre a mineração de dados históricos ou em tempo com avaliações preditivas de tendências criminais, acontecimentos prováveis, regiões e/ou horários possíveis e assim provendo subsídios para a criação de estratégias de policiamento preventivo, intervenção social em determinadas áreas ou direcionada a determinados indivíduos para se evitar a escalada criminal.

Difere-se totalmente da realidade criada na ficção baseada em premonições de crimes e aplicação do *jus puniendi*^{xii} estatal aos supostos autores apenas por serem os autores de atos ilícitos em previsões, não necessitando que estes pratiquem o ato para tal ou mesmo realizem atos executórios.

Para se falar sobre *jus puniendi* é preciso descrever brevemente a teoria do crime que majoritariamente define o crime como um fato típico, ilícito e culpável, sendo o *iter criminis*^{xiii} dividido basicamente em cogitação, atos preparatórios, atos executórios e consumação.

Pode-se concluir que é preciso atender aos três elementos do crime para se punir, não sendo por exemplo um fato típico a mera cogitação de um crime, uma mera tendência sobre a possibilidade de alguém cometer um crime ou atos preparatórios não tipificados. O Estado apenas poderá agir punitivamente a partir dos atos executórios, havendo a única alternativa agir preventiva a partir das



análises preditivas. Conforme cita Masson: A Preparação ou atos preparatórios correspondem aos atos indispensáveis à prática da infração penal, municiando-se o agente dos elementos necessários para a concretização da sua conduta ilícita. Precisa ir além do simples projeto interno (mínimo), sem que se deva, contudo, iniciar a imediata realização tipicamente relevante da vontade delitiva (máximo). Os atos preparatórios, geralmente, não são puníveis, nem na forma tentada, uma vez que não se iniciou a realização do núcleo do tipo penal. De fato, o art. 14, II, do CP vinculou a tentativa à prática de atos executórios. Em casos excepcionais, é possível a punição de atos preparatórios nas hipóteses em que a lei optou por incriminá-los de forma autônoma. São os chamados crimes obstáculo. (Masson, 2014).

Os casos excepcionais citados por Masson são crimes autônomos como a associação para o tráfico conforme art. 288 do CP. Porém, com o advento dos jogos olímpicos, foi promulgada a Lei Antiterrorismo 13.260/2016 que regulamentou o crime de terrorismo anteriormente descrito na Constituição Federal de 88 em seu art. 5º, inciso XLIII. Apesar do mesmo já ter sido tipificado vagamente na Lei de Segurança Nacional 7170/83, a referida lei antiterrorismo criminalizou atos preparatórios em seu art. 5º que dispõe sobre a punição da realização de atos preparatórios de terrorismo com o propósito inequívoco de consumir tal ato. O referido dispositivo transforma este tipo de ato preparatório em crime autônomo e aproxima a filosofia pré-crime existente no enredo do filme de ficção à atualidade que nos encontramos, possibilitando a punição através de análises que demonstrem a possibilidade de terrorismo desde que se confirme tais atos.

Vale ainda citar que não se deve confundir a persecução penal preditiva com flagrante forjado. Uma vez que o flagrante forjado é nulo por viabilizar prisão a delito inexistente, além de criminalizar o autor da forja por denúncia caluniosa ou por abuso de autoridade, se funcionário público. Não é intuito aqui esgotar a lei antiterrorismo ou todas as teorias referentes ao policiamento preditivo.

Visando ainda a preservação dos direitos fundamentais do cidadão, segundo a Ph.D. em psicologia Dr. Collen McCue, é possível prever onde delitos irão acontecer sem a necessidade de coleta de dados pessoais.

Nosso trabalho não envolve informações pessoais sobre os indivíduos, mas lida com informações geoespaciais sobre ações. Colleen McCue, Ph.D. em psicologia pelo Dartmouth College e diretora da companhia de imagens por satélite Digital Globe Analytics.

A título de curiosidade, Collen foi responsável direta pela prisão do atirador Yonathan Melaku em 2011, no cemitério militar da Virgínia, nos Estados Unidos, após avaliar todos os dados de ataques anteriores de Melaku e prever o próximo local.

Vale citar também que a realidade legislativa brasileira sobre tal tema pode ser resumida como afirma Marc Goodman, ex-policial, consultor na área de Cibercrime, fundador do Future Crimes



Institute “As leis não acompanham o avanço tecnológico, e as forças policiais, sobretudo no Terceiro Mundo, não estão preparadas”.^{xiv}

CONCLUSÃO

O *Big Data* é o diferencial para excelência em decisões, sejam elas em quaisquer ramos ou segmentos e atualmente é imprescindível para as forças de segurança pública na prevenção criminal, principalmente no momento em que se encontra a segurança em nosso país. Muitas vezes os esforços em análises preditivas esbarram na questão legal que precisa ser constantemente atualizada e esclarecida para implementar limites e garantir direitos.

A questão da coleta de dados no Brasil é uma realidade e possibilita sim a análise preditiva para questões de segurança pública e nacional, porém este acesso a dados massivos, no que tange dados pessoais, é extremamente sensível e merece atenção. Ainda não há uma legislação específica em relação à coleta de metadados, como a geolocalização no *Twitter* ou origem/destino de chamadas telefônicas por exemplo, pois há uma indefinição sobre a questão dos metadados serem informações pessoais ou não, apesar da existência de algumas posições doutrinárias nesse sentido, porém o Projeto de Lei 5276/2016 que dispõe sobre o tratamento de dados pessoais para garantia do livre desenvolvimento da personalidade e da dignidade da pessoa natural, conforme descrito no site da câmara, encontra-se em tramitação na Câmara dos Deputados. A criação de uma lei para a proteção dos dados pessoais se faz necessária para garantia de nossa intimidade e privacidade.

Estamos diante de uma realidade inegável e diversos debates sobre as novas tecnologias estão cada vez mais frequentes, principalmente no que tange a regulação destas, pois, como citado anteriormente, atividades de vigilância podem violar direitos fundamentais dos cidadãos, porém aplicadas ao policiamento preditivo podem salvar vidas.

REFERÊNCIAS

MASSON, Cleber. **Direito penal esquematizado – parte geral – vol 1**. Rio de Janeiro: Forense; São Paulo: Método. 2014.

EAGLE, Nathan. GREENE, Kate. **Reality Mining: Using Big Data To Engineer a better World**. The MIT Press, 2014.

RUDGER, Christian (autor). AZEVEDO, Patricia (tradutor). **Dataclisma**. Rio de Janeiro: Best Seller, 2015.
SCHÖNBERGER, Viktor Mayer. CUKIER, Kenneth. **Big Data: A Revolution That Will Transform How We Live, Work and Think**. Hardcover, 2013.

SIEGEL, Eric. **Análise Preditiva: O poder de prever quem vai clicar, comprar, mentir, morrer**. Alta Books, 2017.

TANG, Chunlei. **The Data Industry: The Business and Economics of Information and Big Data**. John Wiley & Sons, Inc, Hoboken, New Jersey, 2016.



New York Crime Map. Disponível em: <<https://maps.nyc.gov/crime/>>. Acesso em: 19 jun. 2017.

How big data fights crime. Disponível em: <<https://fcw.com/articles/2012/11/02/big-data-memphis.aspx>>. Acesso em: 16 jun. 2017.

Polícia de SP usará sistema baseado em big data para combater crime. Disponível em: <<http://exame.abril.com.br/tecnologia/policia-de-sp-usara-sistema-baseado-em-big-data-para-combater-crime/>>. Acesso em: 15 jun. 2017.

Da Cibersegurança à Ciberguerra: O Desenvolvimento de Políticas de Vigilância no Brasil. Disponível em: <<http://artigo19.org/wp-content/blogs.dir/24/files/2016/03/Da-Ciberseguranc%CC%A7a-a-%CC%80-Ciberguerra-WEB.pdf>>. Acesso em: 10 jun 2017.

Predicting crime, LAPD-style. Disponível em: <<https://www.theguardian.com/cities/2014/jun/25/predicting-crime-lapd-los-angeles-police-data-analysis-algorithm-minority-report>>. Acesso em: 09 jun 2017.

COMPSTAT: mudou a forma como a polícia reage à criminalidade. Disponível em: <<http://segurancaedefesa.blogs.sapo.pt/ele-mudou-a-forma-como-a-policia-reage-158822>>. Acesso em: 08 jun 2017.

Big Data e as análises na predição de crimes e fraudes. Disponível em: <<http://www.datacloudbrasil.com.br/big-data-e-as-analises-na-predicao-de-crimes-e-fraudes/>>. Acesso em: 08 jun 2017.

Secretaria da Segurança Pública discute análise preditiva de crimes com representantes da Universidade de Chicago. Disponível em: <<http://www.seguranca.pr.gov.br/modules/noticias/article.php?storyid=7658>>. Acesso em: 07 jun 2017.

A ferramenta que permite saber quando e onde acontecerá um crime. Disponível em: <http://brasil.elpais.com/brasil/2017/03/09/tecnologia/1489078250_691655.html>. Acesso em 10 jun 2017.

Unisys lança sistema de apoio a investigações policiais. Disponível em: <<http://computerworld.com.br/unisys-lanca-sistema-de-apoio-investigacoes-policiais>>. Acesso em: 16 jun 2017.

Big Data no combate ao terrorismo. Disponível em: <<http://www.bigdatabusiness.com.br/big-data-no-combate-ao-terrorismo/>>. Acesso em: 10 jun 2017.

Genetic Profiling and Predictive Policing Are Taking Us to the Pre-Crime Future. Disponível em: <https://motherboard.vice.com/en_us/article/genetic-profiling-and-predictive-policing-are-taking-us-to-the-pre-crime-future>. Acesso em: 17 jun 2017.

Sistema de prevenção e investigação. Disponível em: <<http://www.ssp.sp.gov.br/acoes/leAcoes.aspx?id=36288>>. Acesso em: 15 jun 2017.

ⁱ Rodrigo Arrigoni É Especialista em Comunicação Estratégica, Mestre em Semiótica pela PUC-SP, pesquisador em comunicação e lógica com foco em hábito versus performance e processos sociais criativos, mineração de dados sociais, análise de big data e machine learning (Natural Language Processing and Social Networks Analysis), desenvolvedor de softwares, SaaS, intranets 2.0, plataformas de venda de ingressos, aplicativos mobile e jogos para treinamentos.

ⁱⁱ <https://nakedsecurity.sophos.com/2017/05/10/minority-report-in-chicago-as-police-aim-to-stop-crime-before-it-happens/>

ⁱⁱⁱ <http://exame.abril.com.br/tecnologia/big-data-contra-o-crime/>

^{iv} <http://www.genewatch.org/sub-566821>

^v <http://www.legislation.gov.uk/ukpga/2010/17/contents>

^{vi} <https://www.fbi.gov/services/laboratory/biometric-analysis/codis>



-
- vii <http://www.saopaulo.sp.gov.br/spnoticias/na-imprensa/sistema-secreto-da-policia-pode-rastrear-qualquer-um/>
- viii <http://www.premiomariocovas.sp.gov.br/conteudo/apresentacaoTipoUmConteudo.aspx?9uaVfyQKNukCj/Aj9nwWWe8/aNZ2UC1i>
- ix <http://www.eleuterio.com/nudetective.html>
- x <http://apcf.org.br/AgênciaAPCF/tabid/65/ctl/Details/mid/397/ItemID/835/Default.aspx>
- xi <http://agenciabrasil.ebc.com.br/geral/noticia/2014-06/sistema-vai-permitir-que-pf-identifique-criminosos-antes-que-cheguem-ao-brasil>
- xii Para Mir Puig, o Direito objetivo equivale ao conjunto das normas penais. Por sua vez, o Direito subjetivo (também chamado jus (ou ius) puniendi ou direito de punir) é o direito que corresponde ao estado de criar e aplicar o Direito Penal objetivo.
- xiii Iter Criminis o caminho percorrido até a consumação do crime.
- xiv <http://exame.abril.com.br/tecnologia/big-data-contra-o-crime/>