JOURNAL OF ANTI-CORRUPTION LAW

2024 Volume 8 Pages 60 - 84

APPLYING THE ROUTINE ACTIVITIES THEORY TO CYBERCRIME: A CYBER-CAPABLE GUARDIAN*

Sagwadi Mabunda**

ABSTRACT

This paper presents the evolution of the Routine Activities Theory's third constitutive element, the capable guardian, as a strategic approach to combatting cybercrime. It advocates for a reconceptualisation of the capable guardian, transforming it into what is termed the cyber-capable guardian. Emphasising the imperative for countering cyber threats effectively, the integration of artificial intelligence technology, specifically machine learning, is championed to empower this Cyber-Capable Guardian, enabling it to undertake tasks beyond human capacity. Moreover, the discussion extends to section 52 of the South African Cybercrimes Act 19 of 2020, which delineates the provision for a Designated Point of Contact (DPoC). The argument posits that establishing the DPoC as an independent agency structured akin to a private company under The Companies Act 71 of 2008, could serve as a viable capable guardian. It contends that the DPoC, with its chief mandate focused on combatting cybercrime, should spearhead the development and operation of the Cyber-Capable Guardian.

1. INTRODUCTION

Cybercrime, compared to terrestrial crime, is a complex and relatively new threat to society. Over the past number of years, cybercrime has been a talking point as a menacing phenomenon which needs an immediate address, unfortunately, it is still inadequately understood.

---

* This paper is an adaption from ideas in my PhD thesis and as such it is fitting to give honour and dedicate it to my former supervisor Prof Raymond Koen (1955 – 2022) for his unwavering belief in me.

** LLB (Wits University), LLM (UWC/Humboldt University of Berlin), PhD (UWC). Senior Lecturer: Department of Criminal Justice and Procedure, Faculty of Law, University of the Western Cape smabunda@uwc.ac.za.

This paper is the second instalment in a series of discussions about how countries can conceptualise cybercrime in their domestic legislation. The series seeks to showcase a two-part framework which aims to guide States in this endeavour. The first leg of this framework is the application of the Routine Activities Theory as developed by Cohen & Felson. The second leg is the two-type classification of cybercrime as formulated by Gordon & Ford. The purpose of this framework is to provide a structure upon which law and policymakers can conceptualise cybercrime and construct measures to confront it.

The Routine Activities Theory (RAT) is a theory in criminology created by Cohen & Felson which posits that for a successful commission of a crime, a motivated offender and a suitable target need to converge in space and time where a capable guardian is absent.[1] The RAT thus presents the three constitutive elements to the successful commission of a crime: (1) the motivated offender; (2) the suitable target; and (3) the capable guardian. The second constitutive element, the suitable target, was discussed in the first instalment,[2] to lay the foundation for a better understanding of cybercrime. The justification for the adoption of this two-part framework was argued therein, providing the foundation upon which the present paper seeks to build. In this instance, the third constitutive element of the routine activities' theory, the capable guardian, is discussed.

This paper examines the concept of the capable guardian and discusses how it has developed since it was coined by the creators of RAT, Cohen & Felson, in 1979. It then discusses what a capable guardian in cyberspace would entail. It suggests a reimagining of the concept of the capable guardian from its original form to what this study calls the cyber-capable guardian. It argues that to combat cybercrime successfully, it is necessary to introduce artificial intelligence (AI) in the form of machine learning to create a cyber-capable guardian which will perform functions that a human-capable guardian would not be able to perform. Finally, it discusses section 52 of the Cybercrimes Act 19 of 2020 which provides for the Designated Point of Contact (DPoC). It argues that the DPoC would be a suitable capable guardian if it were to be created as an independent agency which is incorporated as a private

1 Cohen LE & Felson M (1979) "Social Change and Crime Rate Trends: A Routine Activity Approach" 44(4) *American Sociological Review* 588 – 608.

2 Mabunda S (2017) "Applying the Gordon & Ford Categorisation and the Routine Activities Theory to Cybercrime: A Suitable Target" in Cunningham P & Cunningham M (eds) *IST Africa Conference Proceedings* at 1.

company in terms of The Companies Act 71 of 2008. It further argues that the DPoC should be established with the chief mandate of combating cybercrime through the creation and operation of the Cyber-Capable Guardian.

Before turning to RAT, it is necessary to have a short discussion of Type I and Type II cybercrimes from the Gordon & Ford Categorisation of cybercrime as the second leg of the two-part framework. The distinction between Type I and Type II cybercrimes is essential to the viability of the cyber-capable guardian. If there is any hope of combatting cybercrime, it is important to understand it as a form of criminality in its own right. To tackle cybercrime successfully, it is important to locate where it fits into the criminal law and to do so in a way that does not comprehend its various forms simply as analogies of conventional crimes. While there are several characteristics present in cybercrime which are common with ordinary or terrestrial crimes – such as unlawful access to data or property – certain cybercrimes, by their very nature, do not have "real world" counterparts. The Gordan & Ford categorisation is thus quite instructive in how it differentiates between Type I and Type II cybercrimes.

1.1 Gordon & Ford categorisation

Gordon & Ford proposed the categorisation of cybercrime into two basic forms, Type I cybercrimes and Type II cybercrimes. Type I cybercrimes exist on one side of the cybercrime spectrum and Type II cybercrimes exist on the other. This dual categorisation of cybercrime is designed to provide a conceptual framework within which lawmakers can create legal definitions which are capable of being varied across different jurisdictions.

Type I cybercrimes are technical when compared to Type II cybercrimes. Type I cybercrimes, when viewed from the perspective of the victim, are generally singular or discrete events that often are facilitated by the introduction of crimeware programmes into the victim's computer system. Examples of Type I cybercrimes include phishing attempts, identity or data theft, and e-commerce or banking fraud facilitated by stolen credentials. Furthermore, these cybercrimes usually are facilitated through an exploitation of vulnerabilities in a computer system which allows for the introduction of crimeware.[3]

3 Gordon S & Ford R (2006) "On the Definition and Classification of Cybercrime" 2(1) *Journal in Computer Virology* 13 – 20 at 14; Mabunda (2017) at 5.

In Type I cybercrimes, crimeware is an important element because it is the malicious software (malware) that is introduced into a computer system to help the cybercriminal perpetrate the crimes. The installation of such crimeware is one of the elements that differentiates Type I cybercrimes from Type II cybercrimes. The crimeware can be used directly or indirectly in the commission of the crime. By definition, malware is undesirable from the perspective of the computer user. Gordon & Ford define crimeware as:

> [S]oftware that is used (directly or indirectly) in the commission of a criminal act; and not generally regarded as a desirable software application from the perspective of the computer user and is not involuntarily enabling the crime.[4]

This definition of crimeware is a common one for the most part, and Gordon & Ford found it appropriate to go a step further by including an element which provides that the software is "not involuntarily enabling the crime".[5] This last element, formulated as a double negative, plays an important role by distinguishing between software that is inherently malicious and software that is taken over to commit a malicious act. For example, the fact that a browser has a vulnerability which is exploited by an attacker does not mean that the browser software itself is crimeware. It may be simply flawed software.[6]

In terms of the discussion above regarding the definition of crimeware, the email client was used to deliver the spam email and, therefore, if interpreted strictly and contextually, the email facilitated the commission of a cybercrime, and could be incorrectly classified as crimeware. In this instance, the real crimeware is the software that created the clone of the bank's site and the keylogger software that stole the username and password.

Type II cybercrimes exist on the other end of the cybercrime spectrum. They have a more pronounced human element to them.[7] They include, but are not limited to, cyberstalking and harassment, extortion and blackmail, child predation, complex corporate espionage and planning or perpetrating online terrorism. Type II cybercrimes usually are facilitated by the use of programmes, such as social media platforms, which would not be classified normally as crimeware. Furthermore, these crimes generally are repeated events

4 Gordon & Ford (2006) at 17.
5 Gordon & Ford (2006) at 16.
6 Gordon & Ford (2006) at 16.
7 Gordon & Ford (2006) at 13.

from the perspective of the victim. The focus of the present paper is on Type I cybercrimes therefore it is unnecessary to delve into further detail on Type II cybercrimes.[8]

Cybercrimes should be considered in terms of a continuum ranging from those crimes which are almost entirely technological to those crimes which are in essence people-related. One will rarely find a complete attack that is purely Type I or purely Type II, as attackers readily deploy whichever means will achieve the desired and best result. In other words, an offender may use a combination of Type I and Type II cybercrimes to complete an attack such as using social engineering to instil fear into a suitable target to coerce her into downloading crimeware onto her computer system. Accepting that cybercrime exists on a continuum enables one to recognise the scale on which it exists and also to identify the areas in which different parties fall short. For example, police investigators are more likely to be able to handle people-centric cybercrimes whereas cybersecurity practitioners would be inclined to focus their efforts on technologically based crimes.[9] Recognition of this distinction allows for enhanced cooperation and would yield improved results.

The Gordon & Ford Categorisation of cybercrime has been adopted in this study because it helps evaluate the offences which are contained in cybercrime legislation in general and the South African Cybercrimes Act 19 of 2020, in particular.[10] The conceptual framework allows us to classify the offences into Type I or Type II cybercrimes by asking two questions:

1. Is the act a singular or isolated event?
2. Is the act facilitated by crimeware?[11]

If the answer to both these questions is yes, then the offence concerned most likely is classifiable as a Type I cybercrime; and if the answer is no to both questions, it is most likely a Type II cybercrime.

---

8 See Mabunda (2017).

9 Gordon & Ford (2006) at 16.

10 The Cybercrimes Act provides for both Type I and Type II cybercrimes in Chapter 2, Part I of the Act, titled "Cybercrimes". It identifies nine such offences which include unlawful access, unlawful interception of data, unlawful interference with computer data or computer programme and cyberfraud, to name a few. This author has argued elsewhere that the offence of cyberfraud is not a true cybercrime because it is only computer enabled rather than being computer dependent, and is not an attack on the confidentiality, integrity and availability (CIA) triad. See Mabunda S (2018) "Is it Cyberfraud or Good ol' Offline Fraud: A Look at Section of the South African Cybercrimes Bill" 2(1) *Journal of Anti-Corruption Law* 58 – 70.

11 Gordon & Ford (2006) 14, include a third consideration which accepts that the crimeware sometimes can be introduced into the computer through a vulnerability. However, it is not necessary to add a third question because this element can fit neatly into the considerations regarding crimeware.

1.2 Routine Activities Theory

The Routine Activities Theory has been developed and fused with other approaches to crime analysis.[12] However, it still rests upon the three factors which are essential for the commission of a predatory crime. The first is a supply of motivated offenders, the second is the availability of suitable targets or opportunities to commit the crime, and the third is the absence of capable guardians. It is emphasised that the lack of any of these elements is sufficient to prevent the perpetration of the offence.[13]

The theory specifies the kind of criminal violations with which it is concerned, namely, direct-contact predatory violations. Predatory violations are defined as "illegal acts in which someone definitely and intentionally takes or damages the person or the property of another".[14] Furthermore, the theory is confined to offences that involve direct physical contact between offender and target.[15] It should be noted that target, as opposed to victim, was chosen purposefully with a view to capturing also scenarios where a human is not directly and physically harmed by the offence. For example, a burglar could break into a home and steal objects while the occupiers are away. In this instance, the objects stolen cannot be referred to as victims but rather are targets. This distinction is vital because the word "target" is meant to highlight the fact that the theory focuses on offending from the viewpoint of the offender rather than that of the victim or society.[16] An example of a situation where a target could be referred to also as a victim would be a sexual assault or, to take a cyberspace example, cyberbullying in which victim and target converge.

One of the attractive features of RAT is that its approach can be contrasted to other theories[17] of crime which are focused on the criminal himself, that is, on the psychological,

12 See for example Miethe TD, Stafford MC & Long JS (1987) "Social Differentiation in Criminal Victimization: A Test of Routine Activities/Lifestyle Theories" 52(2) *American Sociological Review* 184 – 194; Pratt TC, Holtfreter K & Reisig MD (2010) "Routine Online Activity and Internet Fraud Targeting: Extending the Generality of Routine Activity Theory" 47(3) *Journal of Research in Crime and Delinquency* 267 – 296.

13 Felson M & Cohen LE (1980) "Human Ecology and Crime: A Routine Activity Approach" 8 *Human Ecology* 389 – 486.

14 Glaser D (1971) *Social Deviance* Ann Arbor: Markham Publishing Company: Ann Arbor at 4.

15 Cohen & Felson (1980) at 390.

16 Felson (2008) at 71.

17 For a discussion on the Rational Choice Theory see Paternoster R & Simpson S (2017) "A Rational Choice Theory of Corporate Crime" in Clarke RV & Felson M *Routine Activity and Rational Choice* London:

biological or social factors that motivate or drive him to offend.[18] RAT takes it as a given that individuals have criminal inclinations and, therefore, examines the social structures that allow people to commit offences.[19] To focus on the offender's viewpoint, as noted above, means to focus on why he regarded the target as suitable, rather than on what motivated him to take or harm the target.

It is prudent to make cursory comments on the first two constitutive elements of the Routine Activities Theory before zoning in on the last constitutive element which is the chief focus of this article, the capable guardian

.

1.2.1 The motivated offender

The first essential element of Routine Activities Theory is the presence of a motivated offender. A potential offender may be anyone who has the motive to commit an offence and the capacity to do so. Felson (among others),[20] in his later work on RAT, avoided using the term "motivated offender" because he considered it to be important to focus not on the disposition or the inclination of the offender but on the physical factors that made it possible for a person to become involved in the commission of an offence.[21] Focusing on the physical features does not mean necessarily that the perspective of the offender is dispensable, as the determination of what makes a target suitable is dependent on the offender's purposes and capacities in relation to the intrinsic characteristics of the target.[22]

1.2.2 The suitable target

The original aim of the theory that routine activities could explain crime trends was tested to confirm four hypotheses. The first was that the spreading of activities away from the home contributed to greater opportunities for offenders to gain access to suitable targets. The

Routledge at 37, for a discussion of the deterrence theory see Apel R & Nagin DS (2011) "General Deterrence: A Review of Recent Evidence" 4 *Crime and Public Policy* 411 – 436.

18 Miró F (2014) "Routine Activity Theory" in Miller JM (ed) *The Encyclopedia of Theoretical Criminology* John Wiley & Sons Ltd at 1.

19 Cohen & Felson (1980) 390.

20 Felson M & Boba RL (2010) "Crime and Everyday Life" (4ed) Sage Publications: London; and Cohen & Felson (1980) at 390.

21 Miró (2014) at 2.

22 Miró (2014) at 2.

second was that the victimisation rates could be increased by solitary life away from home. The fourth was that the structure of daily life was related to the rise in crime.[23] For purposes of the analysis of cybercrime, the second hypothesis pertaining to a suitable target is most relevant as it is one of the three core elements of RAT. The suitable target is discussed fully in the author's previous work.[24]

1.2.3 The human capable guardian

The capable guardian is defined in Felson's early works as someone who prevents a crime from occurring by being present at a particular place at a particular time. For one to be a capable guardian, one does not have to assume a particular position of authority, one can be a capable guardian by virtue of being in the right place at the right time.[25]

This means that although a police officer may be a capable guardian, a capable guardian need not be a police officer. The idea is that while police officers are able to prevent crimes, they are not always present at the scene of a potential crime to allow them to prevent it. It is more likely that a bystander is at the scene of a potential crime when the motivated offender is about to commit it.[26]

When Felson reconsidered the capable guardian in his later works, he defined the guardian as someone who keeps an eye on the potential target of a crime.[27] This suggests a level of awareness on the part of the capable guardian that she has some kind of responsibility towards the suitable target, albeit not necessarily a formalised responsibility. What underlined Felson's work was the idea of monitoring, which means that for the motivated offender to be deterred from committing an offence, there should exist an environment which suggests that there is someone who is watching and could detect untoward behaviour at any given moment.[28]

23 Miró (2014) at 2.
24 Mabunda (2017).
25 Felson M (1995) "Those who Discourage Crime" 4(3) *Crime and Place* 53 – 66.
26 Felson (1980) at 53.
27 Felson (1980) at 53.
28 Hollis-Peel ME et al (2011) "Guardianship for Crime Prevention: A Critical Review of the Literature" 56(1) *Crime, Law and Social Change* 53 – 70.

Felson developed the capable guardian by drawing a link between it and Hirschi's control theory as formulated in 1969.[29] The link to the control theory added another layer to the capable guardian by introducing the idea of adding "handles" to the motivated or would-be offender through established social bonds.[30] In short, the purpose of these handles is to create instruments of informal control over the motivated offender that would prevent her from offending against the suitable target in the first place. The introduction of the handles meant a shift of focus from the capable guardian as one who prevents the successful commission of a crime as a protector of the suitable target to the imposition of social responsibility over the motivated offender to not offend.[31] The control theory operates on many levels best illustrated by the socialisation of young people.

In short, informal social control requires attaching handles to youth on the primary level and then organising community life in such a way that handles may be grasped on the secondary level by the secondary handlers. A simplistic example of this would be a parent teaching a child the difference between right and wrong thus creating moral handles that can be grasped by her teachers to explain why being a bully is undesirable.

This control theory was further developed by Eck, who included the control or monitoring of places.[32] Eck argued that in the spatial structures of crime, there is a need for the inclusion of "managers" who are responsible for looking after places. Similarly, to the capable guardian, a manager is not so because she holds the formal title of manager, rather, a manager can be anyone who can serve to discourage the commission of a crime by looking after a place. A manager can be a homeowner, a receptionist, a concierge, a private security guard or a close neighbour.[33] Managers are thus a subset of capable guardians specifically responsible for looking after places.

Therefore, the successful commission of a crime will occur when an offender, who is free from her intimate handlers, finds a suitable target unprotected by a capable guardian in a place which does not have managers. Collectively, the capable guardian, the handlers and

29 Hirschi T (1969) *Causes of Deliquency* Berkley: University of California Press, Berkley.

30 Felson (2014) at 121.

31 Hollis-Peele et al (2001) at 55.

32 Eck JE (1994) "Drug Markets and Drug Places: A Case-Control Study of the Spatial Structure of Illicit Drug Dealing" Doctoral Dissertation, Research Directed by Institute of Criminal Justice and Criminology, University of Maryland at College Park.

33 Felson (1995) at 55.

the managers are referred to as controllers, and each of them may intervene to prevent the commission of a crime.[34]

1.2.4 The cyber-capable guardian

Despite the developments regarding the capable guardian, when it comes to cybercrimes there is a clear reluctance to re-imagine the capable guardian as something rather than someone. In other words, there is an insistence that the capable guardian has to be a human being who either is seen by the motivated offender or who implants the idea in the mind of the motivated offender that she is being watched at any given time. As a result, the trend has been towards what is referred to commonly as target hardening.[35]

Target hardening applies where the suitable target bears the responsibility for preventing perpetration of the offence.[36] Although some argue that the capable guardian is the most viable tenet of RAT to control the level of computer-crime victimisation,[37] the consensus is that target hardening significantly decreases risks of computer victimisation, such as cyberstalking and other forms of online harassment.[38] In computer security, the idea is to make it more difficult for the motivated offender to commit cybercrimes by updating and maintaining computer security, such as anti-virus software.[39] Some examples of target hardening against online victimisation include having adequate security software with appropriate filters, being proficient and skilled in computers and "online hygiene", and reducing risky information sharing.[40]

Target hardening is a very useful strategy against victimisation online. However, it must not be taken as a replacement for capable guardianship. This is because suitable targets are inherently vulnerable. If they always were able to protect themselves, there would be no need

34 Felson (1995) at 55.
35 Hollis-Peele et al (2001) at 54.
36 Johnson SD et al (2017) "Evaluation of Operation Swordfish: A-Near-Repeat Target-Hardening Strategy" 13 *Journal of Experimental Criminology* 505 – 525.
37 Choi KS (2008) "Computer Crime Victimization and Integrated Theory: An Empirical Assessment" 2(1) *International Journal of Cyber Criminology* 308 – 333 at 312.
38 Reyns BW, Henson B & Fisher BS (2016) "Guardians of the Cyber Galaxy: An Empirical and Theoretical Analysis of the Guardianship Concept from Routine Activity Theory as it Applies to Online Forms of Victimization" 32(2) *Journal of Contemporary Criminal Justice* 148 – 168.
39 Choi (2008) at 312.
40 Reyns et al (2016) at 154.

for this discussion. Furthermore, although some people may have the means to harden themselves as targets of terrestrial crime by installing alarm systems and high walls, this has not rendered police officers obsolete. Instead, what is needed is a reconceptualisation of the capable guardian, in line with technological advancement, to a kind of Cyber-Capable Guardian.

The Cyber-Capable Guardian needs to encapsulate the essence of the capable guardian but with a futuristic outlook. Traditionally, a capable guardian is someone who prevents the commission of a crime by converging in space and time with the motivated offender and the suitable target. Considering that only Type I cybercrimes may be considered true cybercrimes, as previously argued, the Cyber-Capable Guardian must be designed in such a way that it addresses the challenges which are specific to Type I cybercrimes.[41] The constitutive elements of a Cyber-Capable Guardian are considered below.

1.3 Big data

Big data has become a buzzword in recent years. It is a composite term which describes the emerging technological capabilities for solving complex problems. It is hailed as a new frontier for innovation, competition and productivity. Big data is multifaceted, multidimensional and multidisciplinary, finding application in various industries such as health, science, technology, transport and cybersecurity.[42] Each day billions of individual pieces of data are being amassed from various sources, including supplier data, delivery slips, employment records, health records, police criminal databases, DNA records, and user-generated content such as check-in locations, messages, photos and videos on social media platforms. The availability of big data contributes to motivating researchers from diverse fields — ranging across physics, computer science, genomics and economics — not only to investigate new methods and algorithms for detecting and sorting chunks of big data but also to invent them. Analysing more data faster can lead to better and more efficient decisions in areas such as finance, health and research.[43]

41 See Mabunda (2018).

42 Oussous et al (2018) 'Big Data Technologies: A Survey' 30(4) *Journal of King Saud University-Computer and Information Sciences* 431 – 448.

43 Oussous et al (2018) at 436.

There is still a lack of consensus on a single definition of big data, but it is clear, from the different areas where it is applicable, that there are common characteristics. The International Telecommunications Union (ITU) refers to the four Vs which often are used to characterise different aspects of big data. They are volume, velocity, variety and veracity.

Volume — the first V — refers to the issue of data anytime, anywhere, of anything and by anyone. One of the most attractive things about big data is volume. The statistics differ but it is estimated that more than 90 per cent of the data in the world was created in the past decade, with both machines and humans contributing to the data growth.[44] ITU estimates that there are almost seven billion mobile-cellular subscriptions worldwide which both consume and create data. The volumes of data created grow at a tremendous rate. The Exabyte era (1018 bytes or over 36 000 years' worth) of high definition (HD) videos has been superseded by the Zettabyte era (1021 bytes or a trillion gigabytes). We are told that it is merely a matter of time before we see the dawn of the Yottabyte (1024 bytes).[45]

Velocity — the second V — refers to the notion that every millisecond counts. A critical factor in the big data discussion is the speed of decision-making, that is, the time that elapses between the data input and the decision output. This involves the need for data processing capabilities for vast volumes of data which should occur in real- or near-real time.[46] Emerging technologies are gaining those capabilities, thereby aiding organisations to respond appropriately to changes in the field. Furthermore, not only should data systems have processing capabilities, they also should be able to handle data and draw links and patterns from it. An example of this may be gleaned from how data is used in the day trading practices of stock exchanges. The race for data velocity and tight feedback loops play a key part in giving an organisation a competitive advantage in several industries.[47]

Variety — the third V — refers to the reality that data is messy. Big data encompasses all kinds and structures of data, including text, sensor data, call records, maps, audio and visual information.[48] A vast majority (estimated at 80 per cent) of all data is said to be

44 International Telecommunications Union (ITU) (2013) 'Big Data: Big Today, Normal Tomorrow' *ITU-T Technology Watch Report*.
45 ITU (2013) at 11.
46 ITU (2013) at 8.
47 ITU (2013) at 8.
48 ITU (2013) at 8.

unstructured as it can present in many different forms, such as emails, call logger data and social media feeds. The term unstructured data refers to the fact that there is no latent meaning that is attached to the data in a way that a computer can understand. By contrast, structured data has a semantic meaning. For example, a database contains data that is represented by means of rows and columns which can be understood immediately by a computer system and may be transformed to be of interest to end-users, for example, Microsoft Excel spreadsheets.[49] An example of unstructured data would be a set of alphanumeric characters on a document which do not have a decipherable pattern and which do not have a programme which can discern their meaning, that is, an encrypted message without a decryption key. Structured data is easier to handle because there is more information available to a programme to enable it to determine what the data means.

Veracity — the final V — refers to the accuracy of the data upon which decisions will be based. Some data may be more reliable than others, for example, statistics from a research centre data as opposed to unverifiable statistics on a tweet. Veracity is influenced by the three other Vs and it can be uncertain because of many inconsistencies, incompleteness, ambiguities and latencies. Poor quality data is costly so it is essential that programmes are capable of distinguishing, evaluating and weighing different datasets so that veracity can be maintained.[50]

Veracity is linked closely to the acronym GIGO which stands for "garbage in, garbage out". GIGO is premised on the idea that, regardless of how good the system or programme is, if the input data is garbage, the output decision will be garbage also. In this case, there may even be instances where big data is simply not big enough to yield accurate results, as in a case where a sample size is simply too small to draw meaningful conclusions. Another challenge with big data is that the system or programme may not have the necessary mechanism in place to detect what data is garbage, which may lead to the perpetuation of garbage data outputs.[51] To gain insights and knowledge from big data, it is essential that the

49 ITU (2013) at 1.
50 ITU (2013) at 9.
51 Clegg B (2017) *Big Data: How the Information Revolution is Transforming our Lives* London: Icon Books Ltd.

dots are connected by aggregating and analysing the data so that patterns may be detected and accurate, comprehensive and actionable reports may be produced.[52]

Some of the challenges of big data relate to data protection, privacy and cybersecurity. Two main principles of data protection are data minimisation and data avoidance, which are in stark contrast to what big data represents. Big data can keep track of people's behavioural patterns with incredible accuracy and often without consent. This is of great concern for personal privacy when abused, but when done properly and legally it can be a great source of information for an entity such as the Cyber-Capable Guardian. For instance, large sets of phone call records or data traffic over the internet, even if anonymised and stripped of all personal information, can be used to identify individuals and create a highly accurate profile or "fingerprint" of users. Such a fingerprint, when used in combination with geo-location data, can be very helpful to an investigator.[53]

Internet services providers are typically good sources of big data, either as creators or as users. The relationship between big data, internet service providers and capable guardians is of cardinal significance because it is where some of the challenges of investigating and prosecuting cybercrimes begin to be solved.

1.4 Artificial intelligence and machine learning

Type I cybercrimes are technical in nature, which means that the spatio-temporal convergence will occur in cyberspace and not in the terrestrial realm. It may be extrapolated from this that a guardian who truly is capable of disrupting the commission of a cybercrime is one who is able to operate predominantly in cyberspace, that is, a Cyber-Capable Guardian.

It follows also that the proposed Cyber-Capable Guardian cannot be a human being. It will have to be a machine. Additionally, given the rate at which Type I cybercrimes develop, the machine-guardian must possess qualities which would allow it to adapt rapidly and without human intervention to any new and emerging threats. It should have a kind of intuition or intelligence that a human would have, while being endowed with the superior

52 ITU (2013) at 1.
53 ITU (2013) at 16.

technical capabilities of a machine. In other words, the proposed Cyber-Capable Guardian ought to be a machine based on the technology of artificial intelligence (AI).[54]

Artificial intelligence is still a relatively new field in the cyber world, with promising research being conducted in a host of areas. One of the aspects of AI which already has found many uses is machine learning. Machine learning is a result of the intersection between computer science and statistics.[55] While machine learning builds upon the issues raised by computer science and statistics, it pursues its own distinct concerns about how to enable computers to programme themselves from experience and initial structure.[56]

Machine learning is a field of knowledge which attempts to answer the question: "How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes?"[57] There is a broad range of learning tasks covered by this question. It includes the tasks of how to design autonomous mobile robots which would be able to navigate from their own experiences. It also includes learning tasks such as data mining, autonomous discovery, database updating, and programming by example.[58] Machine learning is special in that it can be applied to many different fields, from the study of human and animal learning in psychology and neuroscience to information security and data mining.[59] The flexibility of machine learning means that it may be applied to law in the fight against cybercrime, particularly because technology is the backbone of cybercrime. Machine learning uses data to learn programmes automatically. This is an attractive capability, because it reduces the burden of having to construct programmes manually.

Tasks related to cybercrime in which machine learning is used include web search optimisation, recommender systems, advertisement placement, credit scoring, fraud detection, stock trading and drug design. It is anticipated that machine learning will be the driver of the next wave of digital innovation. One of the areas in cybercrime where machine learning has been applied relatively successfully is in spam detection. Email clients and social

54 Russel S & Norvig P (2016) *Artificial Intelligence: A Modern Approach* (3ed) Essex: Pearson.
55 Mitchell TM (2006) "The Discipline of Machine Learning" Carnegie Mellon University: Pittsburgh.
56 Oussous (2018) at 435.
57 Mitchell (2006) at 1.
58 Mitchell (2006) at 1. Oussous (2018) at 434.
59 Mitchell (2006) at 1. Oussous (2018) at 434.

networking platforms experience many challenges with spam.[60] Spammers have used social networking platforms, such as Twitter, as a tool to post unsolicited messages to users, to spread malicious links and to hijack trending topics.[61] Twitter responded by employing target hardening techniques to combat the proliferation of spam. It did this by soliciting the users of the service to participate in detecting and identifying spam by adding a "report as spam" feature to the service. This was aimed at cleaning up accounts which were considered to be suspicious.[62]

It is apparent from the above discussion that great strides have been made in machine learning technology. Indeed, the extent of the capabilities of machine learning is being discovered still. Governments need to invest in research into machine learning and, it is submitted, must facilitate the construction of machines which will take on the role of a Cyber-Capable Guardian. It is here where big data would play its biggest role. It is essential that when developing the Cyber-Capable Guardian, sufficient safeguards are put in place to ensure that the data which is being fed to the AI is not garbage.

The elements identified in the preceding discussion are the essential building blocks of the Cyber-Capable Guardian which will be discussed shortly. As already noted, in order for the capable guardian to be effective, it needs to operate predominantly in cyberspace. This can be achieved only if it takes advantage of some of the resources that ISPs have, which include access to big data to create and run machine learning technology. Each of these elements plays a big and important role in the fight against cybercrime, but it must be borne in mind that in order to be a successful custodian of the Cyber-Capable Guardian, a multi-disciplinary and multi-party approach must be adopted. This means that the fight against cybercrime should not be viewed through a myopic lens. In other words, combating cybercrime cannot be seen only as a policing or criminal justice issue. Other disciplines, such as engineering and computer science, have to be factored in. Innovation is key.

60 Graham R & Triplett R (2017) "Capable Guardians in the Digital Environment: The Role of Digital Literacy in Reducing Phishing Victimization" 38(12) *Deviant Behavior* 1371 – 1382.

61 Wang AH (2010) "Detecting Spam Bots in Online Social Networking Sites: A Machine Learning Approach" *IFIP Annual Conference on Data and Applications Security and Privacy* at 335; El-Mawass N, Honeine P & Vercouter L (2018) "Supervised Classification of Social Spammers using a Similarity-Based Markov Random Field Approach" 14 *MISNC, Saint-Etienne, France* 1 – 8.

62 Wang (2010) at 1.

2. THE CAPABLE GUARDIAN AND THE CYBERCRIMES ACT

Article 35 of the Council of Europe Convention against Cybercrime (Budapest Convention) provides for the 24/7 network. It requires each Party to designate a point of contact, which must be available 24 hours a day and seven days a week, in order to ensure that there is immediate assistance available for the purpose of investigations or criminal proceedings related to cybercrime. Such assistance must include the facilitation or, if permitted by the domestic law, the direct provision of technical advice, the direct preservation of data in terms of articles 29 and 30, and the direct collection of evidence, supply of legal information and locating of suspects.[63]

The point of contact of one Party is required to have the capacity to carry out communications with the point of contact of another Party.[64] If the designated point of contact is not responsible itself for international mutual assistance or extradition, it must ensure that it is able to co-ordinate with the responsible authority on an expedited basis.[65] Having the Cyber-Capable Guardian function at optimal level is to the benefit not only of South Africa but also to the states which will seek mutual assistance from South Africa. One might say even that it is a country's implied duty to ensure the highest level of efficiency for the Cyber-Capable Guardian, even if the Convention itself does not impose this duty. The anti-cybercrime chain is only as strong as its weakest State.

The 24/7 network is based on the experience gained from the functioning of the network which already exists and operates under the auspices of the G8 group of nations. The G8 24/7 Cybercrime Network was created in December 1997. This Network consists of more than 70 countries, including South Africa. It exists primarily to preserve digital evidence which will be transferred subsequently through legal channels.[66] It was agreed amongst the members that the establishment of the Network is one of the most important means of

63 Art 35(1) of the Council of Europe Convention on Cybercrime, Budapest 23.XI (2001) (Budapest Convention).

64 Art 35(2)(a) of the Budapest Convention.

65 Art 35(2) of the Budapest Convention.

66 Ott C (2018) "What You Should Know about the 24/7 Cybercrime Network", available at https://www.dwt.com/files/uploads/documents/publications/What%20You%20Should%20Know%20About%20The%2024.pdf (visited 7 May 2024).

ensuring the effective and rapid response of law enforcement to cybercrime for which the Convention has provided.[67]

The Protocol statement of the G8 24/7 Cybercrime Network recognises that high tech crimes pose new challenges to law enforcement and hence it is imperative that technically literate investigators be capable of moving at unprecedented speeds to preserve electronic evidence and locate suspects during investigations, often by asking ISPs to preserve data. The G8 24/7 Cybercrime Network was created, therefore, to enhance and supplement (but not to replace) traditional methods of obtaining assistance. It was established as a new mechanism for the expedition of contacts between the participating Parties and any other autonomous law enforcement agencies.[68]

In order for a state to become party to this Network, it must have: (a) a 24/7 point of contact; (b) an English-speaking contact point; (c) a technically knowledgeable contact point; and (d) a contact point who is knowledgeable about domestic law.[69] Although the number of requests which have been made through this Network thus far have been relatively low, it has been shown to be beneficial as a crime-fighting route, as seen in the $10 million Norwegian bank robbery during which a police officer was shot dead. The Norwegian police used the Network to request the preservation of computer data from the K, an effort which led to Norway's most wanted suspect being traced to an internet café in Spain, where he was apprehended by the Spanish Police.[70]

2.1 The designated point of contact

Section 52 of the Cybercrimes Act is dedicated to the establishment of the Designated Point of Contact (DPoC), previously known as the 24/7 Point of Contact in version [B6 – 2017], which reflected the idea that the point of contact would have had to operate for 24 hours a day and seven days a week. This approach was informed by an acknowledgment that, because of the nature of the internet, cybercrime necessitates round-the-clock monitoring. Thus, despite the

67 Budapest Convention, Explanatory Report at 54.
68 Organisation of American States (undated) "The G8 24/7 Network of Contact Points Protocol Statement", available at http://www.oas.org/juridico/english/cyb_pry_G8_network.pdf (visited 8 May 2024).
69 Organisation of American States (undated) 1.
70 Organisation of American States (undated) 1.

change of the name, the DPoC retains the function of its predecessor, to provide non-stop vigilance against and assistance with cybercrime.

The South African National Police Commissioner is responsible for the administration and functioning of the DPoC, which must be designated to ensure that immediate and expedited assistance is available.[71] The services to be provided by the DPoC include: technical advice and assistance regarding cybercrime; anything which is authorised under Chapter 4 (powers to investigate, search and access or seize) and Chapter 5 (mutual assistance) of the Cybercrimes Act; legal assistance; the identification and location of an article;[72] the identification and location of a suspect; and co-operation with the appropriate authorities of foreign states.[73] The DPoC is meant to assist also in proceedings or investigations regarding the commission or intended commission of any of the offences provided for in the Cybercrimes Act. In addition to this, the DPoC will exercise power over any other crimes which may be similar to these or any crimes which may be facilitated by means of an article.[74] Further, section 52 provides for an international dimension by catering for offences committed in a foreign state but which are substantially similar to those described in the Cybercrimes Act.[75] The NDPP must avail members of the National Prosecuting Authority (NPA) to provide legal assistance to the DPoC, with a view to making it as effective as possible.[76] Such members of the NPA must have particular knowledge and skills in respect of any matter dealt with in the Cybercrimes Act. Also, they must possess the necessary security clearance as issued by the State Security Agency in terms of section 2A of the National Strategic Intelligence Act, 1994.[77]

71 Explanatory Note 63 (Version [B6 – 2017] of the Cybercrimes Act).

72 In section 1 of the Cybercrimes Act, the word "article" is defined as: "any data, computer programme, computer data storage medium or computer system which (a) is concerned with, connected with or is, on reasonable grounds, believed to be concerned with or connected with the commission of a crime; (b) may afford evidence of the commission or suspected commission; or (c) is intended to be used or is, on reasonable grounds, believed to be intended to be used in the commission of an offence in terms of chapter 2 or section 17, 18, or 19 or any other offence which may be committed by means of or facilitated through the use of such article, whether within the Republic or elsewhere".

73 Cybercrimes Bill [B6 – 2017], Explanatory Note 63.

74 Sec 52(3)(ii) of the Cybercrimes Act.

75 Sec 52(3)(iii)(bb) of the Cybercrimes Act.

76 Cybercrimes Bill [B6 – 2017], Explanatory Note 64.

77 Act 39 of 1994.

The Minister of Police will be entitled to make regulations for any matters related to the DPoC.[78] Said Minister will be required also to submit a report on the functions and activities of the DPoC to the Chairperson of the Joint Standing Committee on Intelligence at the end of each financial year. This report must include the number of matters in which technical advice and assistance were provided to a foreign state and the number of matters in which technical advice and assistance were received from a foreign state.[79]

Finally, the DPoC is to be established or designated by the National Police Commissioner within the existing structures of the South African Police Service (SAPS).[80] There well may be good reasons for housing the DPoC in the existing structures of SAPS, such as the Cybercrimes Unit. Unfortunately, the Cybercrimes Unit appears to be beset with its own challenges. In September 2018, it was reported that SAPS had ceased investigation of thousands of cases because certain software licences had lapsed due to non-payment. The deferred investigations included investigations into hacking, EFT scams and organised crime, which had to be halted because the software that is used to interpret cell-phone data had expired.[81] This is undoubtedly problematic.

Whereas it may appear feasible to locate the DPoC within SAPS, it may have more chance of success as a capable guardian against cybercrime were it to be established as an independent agency along the lines of the State Information Technology Agency (SITA), as established by the State Information Technology Agency Act (SITA Act).[82] SITA is the lead information technology agency for the South African state. It provides relevant ICT products, services and solutions to the government. It aims to: leverage economies of scale to cost effectively procure IT goods and services and to set standards for security and interoperability in government implementation of core ICT programmes.[83]

The SITA Act established a juristic person, called SITA, which is incorporated as a private company in terms of the Companies Act. The state is the sole shareholder in SITA, and it

78 Sec 52(4) of the Cybercrimes Act.
79 Sec 52(6) of the Cybercrimes Act.
80 Sec 52(1) of the Cybercrimes Act.
81 MyBroadband (2018) "SAPS Cybercrime Unit Crippled by Expired Software Licences", available at https://mybroadband.co.za/news/security/275703-saps-cybercrime-unitcrippled-by-expired-software-licences.html (visited 8 May 2024).
82 Act 18 of 1998.
83 *State Information Technology Agency*, available at http://www.sita.co.za/ (visited 8 May 2024).

resorts under the authority of the Minister of Public Services and Administration or anyone appointed by the President. When SITA was created, it comprised the Central Computer Services of the Department of State Expenditure, Infoplan, the sub-component Information Systems within the Department of Safety and Security, and any other department approved by the Minister of Public Services and Administration.[84] In general, the provisions of the Companies Act apply to SITA, unless excluded by said Minister.[85] SITA was created with the objective of providing information technology, information systems and related services in a maintained information systems security environment to, or on behalf of, participating departments and organs of state. It operates as an agent of the South African government.[86] It is governed and controlled by a Board of Directors which is appointed by the Minister of Public Services and Administration after consultation with Cabinet, and the Board approves the business and operational plans of SITA.[87] The Director of SITA must carry out her duties in accordance with the provisions of the SITA Act and the Companies Act.[88] SITA is funded mostly by monies it receives from the services it renders to participating departments and organs of state. However, the capital required to start SITA is obtained from funds that are agreed upon by the Minister of Public Services and Administration and the Minister of Finance, after consultation with relevant participating departments. Where there is a need for special funding or any other special financial agreements, such as government grants, these must be negotiated by SITA with the Minister of Public Services and Administration and the Minister of Finance and any other interested party. SITA is allowed also to receive donations, which must be approved by the Minister of Public Services and Administration and recorded in the annual report of SITA.[89]

There are a number of benefits to having an independent and privately incorporated agency operating as a capable guardian. With this in mind, it is submitted that the DPoC prescribed by the Cybercrimes Act ought to be incorporated as an agency, in the same way as SITA has been, in which capacity it will act as the overarching capable guardian charged with

84 Sec 3(4) of the SITA Act, 1998.
85 Sec 3(4) of the SITA Act, 1998.
86 Sec 6 of the SITA Act, 1998.
87 Sec 8 of the SITA Act, 1998.
88 Sec 9 of the SITA Act, 1998.
89 Sec 16 of the SITA Act, 1998.

protecting the country against cybercrime. This submission means that this incarnation of the capable guardian will deviate from the traditional conceptualisation by not being embodied in a single natural person charged with the responsibility for preventing attacks on the suitable target. The idea is that the DPoC will be in charge of the overall administrative functioning of the Cyber-Capable Guardian, which will be made up of the various machine learning instruments, networks and techniques.

To employ an imperfect analogy, imagine the DPoC as a version of SAPS and the machine learning technology (in whatever form) as a version of the police officers who actually would fight crime. The higher-ranking police officers, such as the commissioner, the captains and the detectives, may be likened to the human engineers, software developers, IT specialists and technology experts who would be in charge of creating and running the machine learning complex. There needs to be a close relationship between work that is done by the machines and the humans in the DPoC in order to obtain the best results. It is understood well that although artificial intelligence and machine learning can perform actions that human beings cannot, human beings are not dispensable (at least not for now). Soft skills such as leadership, emotional intelligence and creativity remain the forte of humans still. Therefore, it is necessary to utilise both machines and humans to produce the best solutions.

As to the benefits (referred to above) of having an independent and privately incorporated agency operate as the capable guardian, the first is that the money that is allocated by the Minister of Finance to establish the DPoC will be dedicated to combating cybercrime. If the DPoC were part of SAPS, the discretion as to how to fund the office of the DPoC would be left to the National Police Commissioner. It is foreseeable that, given the challenges which accompany the rampant crime in South Africa, the priority would be to tackle terrestrial crime, such as murder and robbery, instead of cybercrime, which is not understood fully. If, however, the DPoC were incorporated as a private company in the same way as SITA is, it could use the money allocated to it to lay the foundation of an agency that can generate its own income. Unfortunately, the agency would not be able to raise funds by offering shares on the stock market, as do other corporations (given that government would be the sole shareholder), but it can participate in income-generating activities, such as providing services to both the public sector and the private sector. It can be seen here how a mutually beneficial relationship between the DPoC and ISPs could be fostered.

The second benefit of an independent agency is that it would not be stifled by government bureaucracy. It would have the freedom to go about its business, such as entering into private contracts that generate income, unimpeded. This would mitigate the drawbacks of not issuing shares to corporations, while allowing the DPoC still to benefit from information and resource sharing. It would not be favourable for private companies to have shares in the agency because that could place it at the mercy of shareholders who want to further their own interests. It is important to remember that the ultimate goal of the DPoC would be to act as a Cyber-Capable Guardian for all of South Africa, and not to turn a profit.

The third benefit, similar to the second but from a different perspective, is that while the DPoC may be able to avoid government bureaucracy, it still would be subject to some government oversight and be involved in co-operation with other departments and entities, such as the NPA. Furthermore, it would be able to bridge the gap between the public and private sectors. One of the provisions in version [B6 – 2017] of the Cybercrimes Bill, that was not retained in the latest version [B 6B – 2017] and the Act, is section 55, titled "Nodal Points and Private Sector Computer Security Incident Response Teams". That section contained provisions which would have been useful in fostering co-operation between government and the private sector.

Section 55(1) mandated the Minister of Communications to declare that various sectors which provide electronic communications services are required to establish nodal points. These sectors would have been responsible for the establishment and operating costs of the nodal points. Where a designated sector failed to identify or establish a nodal point, then the Minister of Communications would have been allowed to identify and establish the nodal points as she deems fit, after engaging in consultations with the relevant sector.

Nodal points were intended to be structures which received and distributed information regarding cyber security incidents. The section in question also recognised private sector computer security response teams which would be expert groups that would handle cyber security incidents. Each nodal point would have been responsible for distributing information regarding cyber incidents to the other sectors. It would have been responsible also for receiving and distributing information. Version [B6 – 2017] also catered for referring cyber security incidents to the Cybersecurity Hub, which would have been created under the now deleted section 55(4). The Cybersecurity Hub was aimed at promoting cyber security in

the private sector. This body would have been an excellent supplement to the DPoC. Alternatively, it could have been subsumed under the DPoC in order to minimise costs relating to operating two bodies with similar mandates.

The fourth benefit of an independent agency is that when the Cyber-Capable Guardian is created, it may be customised from inception to cater for the needs of the country. The DPoC can engage in robust and dedicated research and development (R & D) in the same way that multinational corporations, like Google and Facebook, do. If the DPoC merely were incorporated into SAPS, there would be the temptation to apply existing crime fighting measures used by SAPS, which most likely would be outdated and ineffective in relation to cybercrime. The DPoC will have an opportunity to break the mould and ensure that the government invests in up-to-date and future-thinking technology which develops as fast as the technology that cybercriminals have at their disposal.

3. CONCLUSION

Whereas traditional ideas of controllers are useful for terrestrial crimes, cybercriminality requires innovation. RAT allows us to conceptualise clearly the life cycle of a crime. It tells us that the commission of a crime will occur when an offender, who is free from her intimate handlers, finds a suitable target unprotected by a capable guardian in a place which does not have managers. This understanding makes it relatively easy to figure out how to respond to cybercrime.

Hitherto, the responses to cybercrime have been dedicated, more or less, to target hardening, which shifts the focus from the motivated offender to the suitable target.

Unfortunately, due to the fact that suitable targets are inherently vulnerable to victimisation, target hardening is not enough. While target hardening has been shown to decrease risks of computer victimisation, a capable guardian remains indispensable. Furthermore, the technical nature of Type I cybercrimes necessitates the creation of a Cyber-Capable Guardian. While there are many instances of convergence between cyberspace and the "real world", each tends to face different challenges which require different responses.

A machine learning Cyber-Capable Guardian can achieve what a human being cannot, especially if it co-operates effectively with ISPs. The DPoC needs to be considered as an

independent agency taking the role of capable guardian against cybercrime. As overall capable guardian, it would be in charge of the overall administrative operation of the Cyber-Capable Guardian. A dedicated and independent agency will have the autonomy to dictate the strategic allocation of resources, keeping in mind cybercrime trends. Additionally, it would not suffer from the bureaucratic nightmares that have been observed to cripple many existing government departments. Administrative freedom is essential for running a functional and effective agency. However, the model proposed here also benefits from a kind of government oversight which provides necessary accountability measures. This is important because the task of fighting crime is ultimately a government responsibility. Co-operation with the NPA, for example, would ensure that the justice exercise is carried through to completion. Finally, an independent DPoC would have the space to break the mould by investing heavily in research and design: a Cyber-Capable Guardian cannot come into existence otherwise.