The Commonwealth

Commonwealth Cyber Journal \ 5
ISSN 2959-3018 (print), ISSN 2959-3026 (online)
thecommonwealth.org/cyber-journal

The Use and Legality of Honeypots, Tracers and Trackers in Active Cyber Defence

Brendan Walker-Munro[1], Andrew Cox[2], Grant Haroway[3], Joe Otway[4], Duncan Unwin[5] and Sascha Dov Bachmann[6]

Abstract

Australia, as an open market economy and democracy, is both dependent and reliant on the internet and online security for its prosperity, way of life and the functioning of our democracy. Cybersecurity, as a prerequisite for ever-increasing interconnectivity, is under assault from cyber-attacks and malicious cyber activity being conducted by states and 'hybrid actors', such as cybercriminals and syndicates.

Cyber-attacks pose a serious threat to the security and integrity of entities, especially when they involve trusted insiders who have access to sensitive data and systems. To counter this threat, this paper proposes that use of active cyber defence (ACD) – such as fake files and credentials that alert the security team when accessed by unauthorised users or tracking devices that report the network activity and location of genuine trading information – can deter and detect malicious actors, often more efficiently and effectively than other methods alone. By using these ACD techniques, organisations can increase their chances of preventing and identifying cyber-attacks, as well as of collecting evidence for potential legal action. However, this paper also acknowledges that there are some challenges and risks associated with the use of ACD, particularly in, though not limited to, the private sectors, such as ethical, privacy and regulatory issues. Therefore, this paper provides a legal analysis of the implications of using teasers and tracers in different jurisdictions, and highlights the following points:

1 Senior Lecturer (Law), Faculty of Business, Law and the Arts, Southern Cross University, Australia. Email: brendan.walker-munro@scu.edu.au
2 President, Active Cyber Defence Alliance Inc.; Principal Consultant, Avantgard
3 Managing Director, SiegeBrake Cyber Incident Readiness
4 Cyber Security Architect
5 Practice Manager, Business Aspect
6 Professor in Law, Canberra Law School, University of Canberra

- The use of ACD may constitute entrapment, deception or fraud depending on the legal definition and interpretation of these terms in different countries.
- The use of ACD may violate the privacy and data protection rights of the employees and customers of the financial institutions, as well as the third parties who may be affected by the cyber-attacks.
- The use of ACD may conflict with the contractual obligations and fiduciary duties of organisations, as well as industry standards and best practices.

This paper offers some recommendations and future directions for research, such as developing a clear and transparent policy for the use of teasers and tracers, obtaining the consent and co-operation of the relevant stakeholders, and conducting a risk assessment and evaluation of the effectiveness and impact of the techniques.

1. Introduction

> *'Ignorance of the law excuses no man; not that all men know the law, but because 'tis an excuse every man will plead, and no man can tell how to refute him.'*
>
> John Selden (1584–1654), English jurist, scholar and polymath

It is an unsurprising truism that in a field as thorny as cybersecurity, the use of inconsistent terminology and ill-defined concepts has created a storm of unnecessary complexity and confusion. Nowhere is that confusion more apparent than in the field of 'active cyber defence' (ACD). In one sense, the term ACD has arisen out of a desire for public policy to recognise that private actors 'are not allowed to retaliate or gather evidence beyond the perimeter of their own networks' (Broeders 2021: 1). Thus, a private actor can protect their own network but cannot – and never can be – authorised to 'hack back' a cyber-attacker, even in the direst of circumstances (Walker-Munro et al. 2022: 5). However, this has created a legal position across numerous jurisdictions that the cybercriminals (both organised and opportunistic) and foreign intelligence actors who engage in cyber-attacks are at less legal risk than the organisations who set out to protect against those attacks (Walker-Munro and Dov Bachmann 2024).

In another sense, the term ACD has arisen to provide a dichotomy between acts which identify and/or expose a cyber-attacker from all other forms of cyber defence, which are necessarily 'passive'. Indeed, it is on this basis that numerous analyses have suggested that firewalls, antivirus software, network resilience and good 'cyber hygiene' would occupy this latter category (Curry 2012; McGee et al. 2013; Dewar 2014; Creado and Ramteke 2020).

The US Department of Defense is widely regarded as having coined the term ACD, defining it as using military systems or capabilities 'to discover, detect, analyze, and mitigate threats and vulnerabilities' (United States Department of Defense 2011). Rosenzweig (2013: 2) then added that ACD must also 'operate at network speed using sensors, software and intelligence to detect and stop malicious activity ideally before it can affect networks and systems'. Dewar went further, proposing that ACD was in fact 'an approach to achieving cyber security predicated upon the deployment of measures to detect, analyse, identify and mitigate threats to and from communications systems and networks in real-time, combined with the capability and resources to take proactive or offensive action against threats and threat entities including action in those entities' home networks' (Dewar 2013: 10).

Then, in 2017, the National Institute for Science and Technology (NIST) defined ACD holistically to describe *any* capability, tool or technique which offers '[s]ynchronized, real-time capability to discover, detect, analyze, and mitigate threats and vulnerabilities' (NIST 2024), with a particular emphasis on the production of cyber-threat intelligence. ACD was then provided with some additional nuance through the work of Lin, Harknett and Smeets, who suggested a delineation between ACD which involved an 'attack' on a system (with the intention of damage or destruction) and 'exploitation' (with the intention of gathering information or intelligence about the attacker/s or their network/s and system/s) (Lin 2010; Lin 2016; Harknett and Smeets 2022).

This paper proposes a synthesised definition which subsumes that nuanced approach to suggest ACD involves the use of 'sensors, software and intelligence to actively – rather than passively – detect, expose and potentially disrupt a cyber-attacker during a cybersecurity incident'. Therefore, one aim of this paper is to socialise that definition of ACD in the literature on cybersecurity.

A second aim is to provide a brief overview of the evolution and development of technologies and related concepts in ACD, often referred to as the 'honey world' because of references to technologies like honey tokens and honeypots (because they appear to be legitimate system resources, and are thus attractive to potential cyber-attackers like honey is attractive to bears (Juels and Ristenpart 2014)). These are techniques that intentionally expose vulnerabilities or false information (also referred to as faux data) to lure, deceive or confuse potential attackers, while monitoring their activities and collecting evidence.

The third aim is to examine and discuss the legal implications and challenges (using Australia as a contextual case study, but with international examples where appropriate) when private companies seek to use honey tokens, honeypots and related techniques. These issues can involve the possible violation of privacy rights, data protection laws, computer misuse statutes or entrapment doctrines. The paper concludes by identifying the areas where the use of ACD-related techniques remains controversial and which require further legal and regulatory clarity.

2. Context

At the Commonwealth Heads of Government Meeting in London in 2018, the Commonwealth Cyber Declaration was signed, which emphasised a strong commitment to common standards, harmonised legal approaches and improved interoperability, including 'through the use of Commonwealth model laws' (The Commonwealth 2018). However specific legal approaches to ACD across the Commonwealth are fragmented and nascent, to the extent that they exist at all.

In Canada, criminal laws still prohibit the use of 'hacking' as a defensive capacity, even despite recognition of ACD in Canadian cybersecurity policy since at least 2017 (Government of Canada 2017). Indeed, even the laws supporting both Canadian intelligence and military forces to engage in ACD have been criticised for undermining interoperability with NATO and US allies, as well as the 'lethargic' pace in establishing the necessary military and civilian capabilities (Rudolph 2021). Malaysia, Singapore and New Zealand permit their national intelligence and police forces to engage in limited forms of ACD, usually under the imprimatur of broad enabling legal authorities (Walters 2023). A study by Thinvane and Christine (2020: 8) that included numerous Commonwealth nations found:

> *'Almost half of Asia-Pacific states have no national cybersecurity strategies yet. Some instead have master plans that cover aspects of cybersecurity in the form of national digital policy (e.g., Pakistan, Brunei, Lao People's Democratic Republic), ICT masterplans (e.g., Cambodia, Solomon Island, Micronesia), and e-governance masterplans (e.g., Myanmar). Countries such as Indonesia, Mongolia, and Pakistan have laws and programmes related to cybersecurity, including cybersecurity centres and national computer emergency/incident response team (CERT/CIRT), but do not have a cybersecurity strategy yet. Meanwhile, some countries, such as Nepal and Fiji, are still in the process of drafting their strategies.'*

African nations fare even more poorly. A study by Ajayi (2016) showed that of the 54 nations making up the continent, only four have laws that explicitly deal with countering cybercrime. Thus, much of the emerging debate on ACD in the private sector in the Commonwealth has come by way of research in the United Kingdom (which in turn has been informed by debate in the US) so the paper reviews those jurisdictions before examining the contextual case study of Australia.

In the US, ACD and 'hacking back' have been interwoven into public debate for many years. Private companies – indeed, many of the major technology companies like Google/Alphabet, Meta and Apple – are headquartered in the continental US, and are prohibited from hack-backs by title 18 of the US code (section 1030), also known as the 1986 Computer Fraud and Abuse Act or CFAA. In addition, numerous US federal statutes positively prohibit the use of certain ACD tools and techniques (Cook 2018). Attempts by Congress in 2017 and again in 2019 to resolve this dilemma using the appropriately titled Active Cyber Defence Certainty Act failed to receive support and has never been

resurrected (Broeders 2021: 2). However, it seems widely regarded that the use of 'tarpits and honeypots, denial and deception, and beaconing on your own network' are acceptable forms of ACD conduct by US entities (Center for Cyber and Homeland Security 2016).

The UK, on the other hand, has taken more of an enablement approach towards ACD, in that the private sector would provide such capabilities to government under the close inspection and supervision of the Crown. Former Prime Minister David Cameron made clear in the 2015 *National Security Strategy* that ACD tools – including a full spectrum of capabilities to detect, analyse and track cyber-attacks pre-emptively – would be considered as 'national capabilities, developed and operated by the private sector' (Cameron 2015). However, like many other jurisdictions, the UK seems mired in policy limbo, lacking a specific definition of the term and clear operational guidance as to how ACD will actually be done. Sexton, for example, wrote that relying on the private sector to provide any form of security for government was not only paradoxical but risked legitimising the use of cyberweapons for corporate interests (Sexton 2016).

In Australia, a concise policy position remains largely elusive. There is no federal (Commonwealth) policy on ACD, having failed to address it in successive defence White Papers (Ball and Waters 2013) as well as the recent 2023–2030 *Cyber Security Strategy* (Department of Home Affairs 2023) and *Cyber Security Bill* (Department of Home Affairs 2024). This is not without industry consistently raising ACD as a matter for policy and legal clarification. In the absence of that position, numerous actors in the private sphere have attempted to delineate a position that ACD involves 'offer[ing] a diversion tactic that makes the cyber-criminal think they're on to something of value' (Powell and Dolan 2021) or 'the use of countermeasures by businesses and corporations to identify, slow down or hinder hackers in executing cyber attacks and malicious cyber activities' (Walker-Munro and Dov Bachmann 2024). At the same time, others in industry have constructed positions which attempt to align ACD either with existing Australian law (Shackelford et al. 2019), or with industry guidance produced or provided by the government (Powell 2021). Yet it appears from the literature that the Australian government has no formalised public position on the activities of ACD (other than those undertaken under the imprimatur of its intelligence services (Hanson and Uren 2018)).

Taking a broad and generalist approach to the various jurisdictions, one could, therefore, surmise that the law struggles to recognise the concept of ACD. From that above examination of the literature, it is generally considered that the internal network of a trading corporation is that corporation's *domaine réservé*, such that engagement in limited forms of deception in the pursuance of ACD is legally acceptable. There are several edge cases which the next section deals with: specifically, where the potential exists for internal networks utilising deception methodologies to still be customer-facing in some aspect and thus potentially incurring a claim against misleading trade practices or prohibitions against 'passing off'. Depending on the jurisdiction, such claims can be offset

by a broad statement in corporate documentation – such as a privacy policy or terms-of-use statement – that discloses to customers that their use of a particular network or system may involve the use of ACD.

3. Case for using ACD

The various methodologies or practices of ACD can be difficult to precisely define or formulate. However, the following is a non-exhaustive list of common ACD techniques or tools.

- **Tracers:** Cookies or similar programs attached to genuine trading information, which periodically transmit back to the incident response team network transmission and movement information, including IP addresses and/or physical locations of potential attackers' systems or networks, if the data are exfiltrated from the host system (Zhang and Thing 2021).
- **Honey objects, honey tokens, honeywords and honey encryption:** Falsified files, databases and user credentials which appear genuine to an external actor but alert the incident response team when accessed. As the files and/or credentials themselves are falsified, there is no genuine need for those files to be accessed (Juels and Rivest 2013; Juels and Ristenpart 2014).
- **Honeypot:** A broader term referring to an entire system (e.g. a web server) or system resource (e.g. a network) that is designed to be attractive to potential intruders, which is configured to alert when an attempt is made to access it. Typically, these are deployed within the boundaries of an organisation (Han et al. 2018).
- **Deception networks/systems/operations:** While nominally aligned with honey resources, deception activities generally rely upon the creation and maintenance of false (but realistic) networks and system resources intended to permit real-time monitoring of cyber-attackers. Given there is no legitimate need for users to be in a deception network, generally any activity in such locations will be unauthorised (Bushby 2019; Steingartner et al. 2021).

As a broad generalisation, the use of deception and 'honey' objects in ACD is considered legally permissible. This is because there is no legitimate need for a customer or other authorised user to ever need to access a 'honey' object – as the object is known by the business to be false, and managed accordingly, there is no way a customer or authorised user would ever find their way to that object. The result is that, by exclusion, any access of that file must be with malicious intent. Equally, the use of 'tracing' technology, such as a piece of code or software which captures an IP address or physical address of the person accessing the system, is legally fraught. In almost every case, consent must be provided, even if that consent is constructive in the sense that it forms part of the standard contractual obligations of persons accessing a given network or resource.

Thought must also be given to internal staff, i.e. employees and contractors. These individuals must be notified (either through employment contracts and/or organisational policies) that the corporation employs ACD, and given sufficient details about those programs to ensure the employee or contractor gives their informed consent to operate in a monitored environment. Generally, specific details are unnecessary. But there must be sufficient and cogent reasoning governing the recording of employee activity to avoid lawsuits later on, i.e. that network monitoring occurs 'for security and acceptable use'.

Even though honeypots can be useful for detecting and analysing cyber-attacks, they also pose some legal challenges that cybersecurity practitioners need to be aware of, in part due to the lack of policy and legal clarity. Depending on the jurisdiction, the use of honeypots may violate laws related to privacy, data protection, computer misuse, entrapment or unauthorised access. Moreover, the interaction between the honeypot operator and the attacker may create liabilities or obligations that are not clearly defined or regulated.

4. Legal issues

There are several interconnected but discrete issues associated with the use of ACD. This paper separates these issues into two categories: those related to international law, and those arising under domestic legal restrictions.

4.1 International law

This paper deals briefly with limitations arising under international law as these are largely beyond its scope. Numerous scholars have done an excellent job of enumerating these international legal problems, including that attribution during a cyber event can be nearly impossible even months afterwards (Berghel 2017; Tran 2018) and that attribution is a sovereign political decision that private companies are usually not authorised to make (Rid and Buchanan 2015; Wanner and Ghernaouti 2019). Such responses might be ruled 'use of force' under international law (Waxman 2011; Halberstam 2013; Corn and Jensen 2018) or even encourage endless cycles of 'Tom and Jerry' retaliatory actions (Gallagher 2022). Although scholars are generally in agreement that the level of permissiveness for ACD will not specifically cross a use of force threshold at international law, this is not always the case with some of the higher-end uses of ACD, and may not actually implicate states that seek to respond even to the mildest provocation (Van Dine 2019). Broeders (2021: 3) described it this way:

> *'If a company follows an attacker down the rabbit hole of the global internet there is no* a priori *telling in which country and jurisdiction it is going to resurface. If private parties conduct disruptive [ACD] operations on foreign, perhaps even state operated or affiliated networks, this can easily have an escalatory effect as foreign actors are likely, and may*

> *even be keen, to take offense. Especially in the current times of heightened geopolitical tensions some states will not look kindly on private companies that are legally licensed by the American government to conduct intrusive and disruptive cyber operations.'*

Looking at domestic legal restrictions, there appear to be several concise domains where ACD can cause private companies a degree of legitimate caution.

4.2 Criminal law

The first and most prominent domain generally relates to the common criminal prohibitions against unauthorised access to computer resources, i.e. hacking. Given that numerous states are signatories to the *Budapest Convention on Cybercrime*, most jurisdictions will have criminalised computer infiltration, data theft and similar acts as part of their ratification processes, which can include some techniques of ACD (Basu and Hickok 2020).

Many western states do have prohibitions against actions by private corporations which would allow unauthorised access to *any* system, even that of a cyber-attacker during a live incident. Canada (Gerke 2021), Japan (Jun 2023), Singapore (Housen-Couriel 2021), the US (Dewar 2014; Broeders 2021) and the UK (Sexton 2016; Montasari 2023) stand out as exemplars of legislative regimes where gaining unauthorised access to any form of computer resource is a crime, irrespective of the nature of the actor and/or the nature of any provocation such an actor may be facing.

However, the specifics of each jurisdiction are patchy and largely unexplored. As the Global Commission on the Stability of Cyberspace (2019: 45) warned in 2019, '[s]ome states do not control or may actively ignore these practices... However, in many states such practices would be unlawful, if not criminalized, while in other states they appear to be neither prohibited nor explicitly authorized'.

Australia's position is likewise that access to any computer that is unauthorised will be unlawful (Walker-Munro et al. 2022). The *Criminal Code* (Commonwealth of Australia) creates, for example, an offence (section 477.3(1)) where a person causes any unauthorised 'impairment' of communication to or from a computer, and that impairment is unauthorised. The use of tracers, deception networks or honey objects could conceivably cause modification in an attacker's data which impairs their ability to access the network; after all, one of the purposes of ACD is to preclude the attacker from ongoing access to corporate information. Another section of the *Criminal Code* (Commonwealth of Australia) creates an offence (section 478.1(1)) for 'unauthorised access to, or modification of, restricted data', where such data are restricted by an access control system. If a private corporation employing ACD were to gain intelligence about a cyber-attacker, behind, for example, a firewall or password-protected file, this could also lead to the commission of an offence for engaging in ACD. Without an immunity (which government agencies enjoy), private corporations could face the very real prospect of criminal charges and conviction.

Further, these legal regimes usually do not extend traditional defences into the cyber domain, i.e. self-defence to 'allow for the use of force against an attacker and thus render an otherwise illegal act lawful, provided it was necessary to defend one's own interests' (Stevens 2020: 320). In some cases, the failure to recognise self-defence flows from a legal disconnect where 'data' or 'information' are not recognised as a tangible form of property, the likes of which can be protected by a party engaging in otherwise unlawful conduct (Rosenzweig 2014; Hoffman and Nyikos 2018). In others, the limitations for the application of self-defence doctrines arise because, while data and information can be considered special forms of property, the scope and purview of rights which vest in that digital property (and hence how those rights can be defended) differ markedly from tangible real-world items (Lawrence 2007; Boerding et al. 2018; Grimmelman and Mulligan 2023). Focusing on Australia, computer data and information are not defined as 'property' under the *Criminal Code* (Commonwealth of Australia) or the *Corporations Act 2001* (Commonwealth of Australia),[7] and so cannot be subject to the doctrine of self-defence.

There are other defences worth examining: namely state of emergency, necessity and provocation. 'State of emergency' is a legal doctrine which can excuse certain illegal conduct where it is in response to an 'emergency' in which the illegal conduct was the only reasonable way of escaping or de-escalating that emergency (Ackerman 2003; Jakab 2006; Crusto 2015). 'Necessity' involves a response to a particular threat, described generally as where 'he or she carries out the conduct constituting the offence in response to an emergency which forced him to ward off immediately an immediate peril against himself or his property or against another person or his property' (Al Qudat 2009). 'Provocation' as a doctrine generally operates in English law systems to reduce a charge of murder to manslaughter because 'the accused killed during a sudden loss of self-control caused by provocation which was enough to make a reasonable man do as he did' (Ashworth 2009; Gruber 2015). All three defences have potential applications to using ACD, dependent on the law of the jurisdiction in question.

For example, the use of ACD could be excused if the ACD is a response that is reasonable and proportionate to the circumstances of a cyber incident – where that ACD involves potentially unlawful conduct such as accessing another person's computer or network or modifying or impairing data transmission. If an incident is in real-time or ongoing or involves serious compromise of sensitive or personal information and/or real-world damage or destruction or requires the use of ACD as the only available option of intervention (i.e. all other 'passive' measures have failed), the use of ACD under doctrines of emergency are more likely to be legally excusable. Necessity could also excuse such conduct if a three-element test is met:

7 Where property includes any 'legal or equitable estate or interest (whether present or future and whether vested or contingent) in real or personal property of any description', but does not cover computer data, information or intangible property of that sense: *Corporations Act 2001* (Commonwealth of Australia), section 9.

> *'First, the criminal act or acts must have been done only in order to avoid certain consequences which would have inflicted irreparable evil upon the accused or upon others whom he was bound to protect…The [second] element…[is]…that the accused must honestly believe on reasonable grounds that he was placed in a situation of imminent peril…thus if there is an interval of time between the threat and its expected execution it will be very rarely if ever that a defence of necessity can succeed. The [third] element of proportion simply means that the acts done to avoid the imminent peril must not be out of proportion to the peril to be avoided. Put in another way, the test is: would a reasonable man in the position of the accused have considered that he had any alternative to doing what he did to avoid the peril?'*
>
> R V Loughnan [1981] VR 443: 448

Provocation, as a partial defence to murder only, will likely never apply. But from a legal philosophy perspective, it does have some attraction in ACD because it could be used as a defence to excuse otherwise criminal conduct in response to the 'provocation' of a cyber-attack being conducted on a private organisation.

In Australia, the *Criminal Code* (Commonwealth of Australia) provides for defences of 'duress' (section 10.2) and 'sudden or extraordinary emergency' (section 10.3), but not provocation. In cases of both duress and emergency under Australian law, a person is not criminally responsible for conduct in response to situations of emergency or threats if the illegal conduct is the only reasonable response. Clearly the facts of a given ACD incident will matter. An argument could be easily made for example if a cyber-attack is occurring in real-time, and an ACD technique is immediately required to prevent sensitive data exfiltration or real-world impacts, i.e. the Colonial Pipeline incident which paralysed fuel supplies across the US eastern seaboard (Easterly and Fanning 2023). If the cyber-attackers have had access to the network or servers of a private entity for a period, such as several months, ACD is less likely to be an excusable response. Provocation on the other hand, as a partial defence to murder, is dealt with as a matter of state law in Australia and is slowly being eroded by a greater recognition that it has excused violent conduct in the past (Ramsey 2010). Its application to ACD is, therefore, unlikely to ever arise without a substantial body of law reform.

4.3 Privacy law

For many western states, there are also implications for privacy. Under the General Data Protection Regulation of the European Union (EU) for example, cyber-threat intelligence generated by ACD activities may contain confidential or protected information, which in turn has implications for how that information can be captured, analysed or shared with other entities and bodies to protect against cyber-attacks (Albakri et al. 2019). The collection of data or information, such as name, physical address or IP address, which could lead to identifying a person, may also be treated as 'personal information' subject to privacy statutes, and may be prohibited from collection, use or disclosure unless the

person to whom that information relates gives free and informed consent (Irving 2013; Miraglia and Casanove 2016). These provisions are increasingly being tailored to counter the emergence of behaviour referred to as 'doxing' (a person's real-life identity, address or contact details are made publicly available (Karimi et al. 2022)), a practice that is now *prima facie* illegal under most privacy legislation (Kukul 2023). Of course, the generalisation that privacy will always be an issue in the conduct of ACD can be countered by specific jurisdictional idiosyncrasies, e.g. in the United States, the Supreme Court has ruled that a person does not have a right to privacy when engaging in illegal activities, like hacking into a network or system (Simonato 2014).

In Australia, the *Privacy Act 1988* (Commonwealth of Australia) prohibits 'serious and repeated infringements with privacy' (s 13G), including repeated or sustained conduct which breaches the Australian Privacy Principles (APP) (section 13). These principles include, for example, the need to communicate clearly and transparently the reasons and mechanisms of data collection, processing, use, disclosure and storage (APP1.3 and 1.4) as well as dealing with 'unsolicited' information (APP4) which may arise from the use of ACD. Identification of a cyber-attacker's personal identity, physical or virtual location in cyber-threat intelligence shared with other bodies or entities may also breach APPs if the sharing is not with law enforcement agencies (APP6.1 and 6.3).

4.4 Consumer protection law

The third and final domain within which ACD may cause thorny legal challenges relates to commercial and consumer protection laws. This is because many jurisdictions adopt legislative standards to protect consumers of various goods and services from the predations of those making false or misleading claims as to efficacy, standard, utility or numerous other benefits of their products (Pengilley 2007; Cooper and Shepherd 2016; Willis 2020). Therefore, a private corporation which, while otherwise providing its goods or services as an entity 'in trade or commerce', deploys a deception network or honey objects, may fall foul of those misleading conduct provisions.

The devil is clearly in the detail, and the whole domain itself is woefully under-explored. For example, under the United Kingdom's *Consumer Protection from Unfair Trading Regulations 2008* (sections 5(2) and 5(3)) a business will behave unlawfully if it engages in a 'commercial practice' that 'contains false information' about a product such that it would 'cause the average consumer to take a transactional decision he would not have taken otherwise'. Clearly, an 'average consumer' is never likely to be exposed to an ACD operation unless that consumer is themselves doing something illegal, i.e. breaking into the corporation's network. That position can be contrasted with the US, where section 5(a) of the *Federal Trade Commission Act* (15 USC section 45) prohibits 'unfair or deceptive acts or practices in or affecting commerce'. As the US legislation does not rely on an 'average consumer' standard, there is arguably more room for it to potentially apply to private operators employing ACD.

In Australia, the Australian Consumer Law (ACL) operates as a schedule to the *Competition and Consumer Act 2010* (Commonwealth of Australia). Under the ACL, a bald prohibition exists (section 18) to a person engaging in conduct during 'trade or commerce' that is deceptive or misleading. These provisions clearly apply to 'computer software' and 'any component part of, or accessory to' that software. Conduct is deceptive or misleading if it 'induces or is capable of inducing error' (van Wyk 2015). As in the continental US, a private corporation that engages in ACD – even one entirely ancillary to, and protecting the underlying rationale for, its core business – may breach these provisions if its ACD program incorporates elements of deception.

5. Recommendations

To overcome legal ambiguities and facilitate responsible ACD adoption, the following amendments to legislation are recommended (an Annex is attached that demonstrates how these changes could be achieved in the context of Australian law):

5.1 Recommendations for the Commonwealth

- Ensure that corporations law, privacy law and crimes/criminal code legislation is amended to clarify that digital information, digitally stored data and information on a computer or network are considered an intangible form of property. This will ensure that private entities can use ACD while availing themselves of the 'self-defence' doctrines present in numerous Commonwealth jurisdictions.
- Ensure that, if private parties are authorised to conduct ACD, they do so within strict boundaries only and that they adhere to all guidelines issued by the national government. This will ensure Commonwealth nations have strong control over where and how private entities can engage in ACD.
- Closely monitor international developments in cybercrime and ACD legality to ensure that national laws do not breach international legal obligations.

5.2 Clarify self-defence, emergency and provocation defences

- The criminal statutes relating to criminal responsibility, excuses and defences (especially computer crimes) could otherwise be amended such that private organisations can take action (i.e. ACD) to protect their proprietary or commercial data, or the data of third parties for which they owe a statutory or common law duty of care, i.e. the personal information of customers.
- Alternately, specific computer crime offences (such as those enacted by parties to the Budapest Convention) need to be redrafted to exclude conduct of ACD by private organisations in response to a cyber-attack or cyber incident. This could

be by the creation of an exclusionary provision or a broader defence of 'acting in good faith'. In any event, private organisations should have the confidence that their position in utilising ACD is legally defensible.

- Alternately, computer crime offences could be subject to a limited defence of either emergency or provocation. In such situations, offence provisions should not apply to the conduct of ACD by persons in a private organisation either because of emergency, i.e. a reasonable and appropriate response to protect information for which the entity has responsibility; or because they are responding to a 'provocation', i.e. a cyber-attack or cyber incident against some asset for which those persons have some responsibility.

5.3 Safeguards

- **Privacy:** The use of ACD (especially tracing) will have serious consequences for privacy law, as it may enable the collection, use and dissemination of information on persons who are not cyber-attackers. There should be strong safeguards that limit the use of information gathered during ACD practices to protect individual privacy rights under both international and domestic law.
- **Anti-consumer practices:** Companies that deploy ACD (especially tracing) should be prohibited from using that information to engage in behaviour which offends consumer protection laws, i.e. by using tracing technology for advertising or marketing. Consumer laws should clearly define prohibited uses of information collected during ACD to prevent collection of excessive data or conduct invasive surveillance that infringes consumer privacy.
- **Oversight and transparency:** If these measures are adopted, judges could be permitted to immunise conduct of ACD by private entities. Relevant safeguards need to be included such as regular audits and public reporting by a competent and independent authority on the issuance and outcomes of such activities.
- **Unintended effects on legitimate users and scams:** While these provisions could empower businesses to combat cyber threats more effectively, there is a risk of collateral damage to innocent parties if compromised systems are mistakenly targeted. Clear safeguards and limitations on the scope of actions permissible under these amendments could help mitigate this risk.
- **Clarity on data usage limitations:** To prevent excessive or unjustified data collection, guidelines should clearly define the limits of what constitutes 'relevant' data that can be collected by ACD actions and how the data should be managed.

6. Conclusion

This paper has examined current ACD measures and the legal defences that may be invoked by individuals or organisations using ACD measures to protect their networks and data from cyber-attacks. It has explained the concepts of intervening conduct or event, sudden or extraordinary emergency, and duress, and how they relate to the use of ACD measures. It has also discussed the limitations and challenges of applying these defences in the context of ACD, such as the uncertainty of the law, the proportionality of the response, the attribution of the attacker and the potential harm to third parties. The paper concludes that the use of ACD measures requires careful assessment of the legal risks and consequences, and that more clarity and guidance from government authorities and the courts are needed to ensure the legitimacy and effectiveness of such measures.

Implementing and receiving the full value of ACD requires legal clarity. The amendments suggested in this paper would remove the current legal ambiguity to businesses, providing legal certainty so they can build and defend their organisations while operating within the bounds of the law. Without such clarity, businesses may inadvertently operate in legal grey areas, compromising their ability to protect themselves and their clients effectively.

In the Australian context, to meet its vision of being a world leader in cybersecurity by 2030, the Australian government needs to promote the use of ACD technologies and techniques to improve understanding of the actions and intent of cyber-attackers and, therefore, of the threats to Australian organisations and citizens. Globally, before we can encourage the use of ACD technologies and techniques, there needs to be legal clarity about the use of ACD. Therefore, this paper should serve as a call to action for legislators, especially in Australia, to make the changes that will remove the legal grey areas and allow private organisations to contribute to their respective jurisdictions' constructions of cybersecurity as part of the Australian government's call for a whole-of-nation effort of shared responsibility across the wider community.

Annex: Proposed Amendments to Commonwealth Criminal Laws

1 Insert:

Self-defence of data, data security or information systems

(1) A person is not criminally responsible for an offence if:

(a) the person believes that the conduct constituting the offence is necessary to defend data, data security or information systems that belong to the person or another person from unlawful injury; and

(b) the conduct is a reasonable response in the circumstances as the person perceives them.

(2) In determining whether the conduct is a reasonable response, regard must be had to:

(a) the nature and extent of the injury to data, data security or information systems that is threatened or inflicted;

(b) the potential or actual consequences of the injury to data, data security or information systems for the person, another person, or the public interest;

(c) the proportionality of the force used to the injury to data, data security or information systems that is threatened or inflicted;

(d) the availability and feasibility of any alternative means of preventing or mitigating the injury to data, data security or information systems;

(e) any relevant laws, policies, standards or codes of conduct that regulate or govern the use, protection or management of data, data security or information systems;

(f) any other relevant factors.

(3) For the purposes of this section:

(a) data means any information that is stored, processed, transmitted, or communicated by any means, whether electronically, digitally, optically, magnetically or otherwise;

(b) data security means the protection of data from unauthorised access, use, disclosure, modification, deletion or destruction;

(c) information system means any system, device, network or infrastructure that is used for the creation, storage, processing, transmission or communication of data;

(d) injury to data, data security or information systems means any act or omission that causes or is likely to cause damage, loss, impairment, disruption, interference or degradation to data, data security or information systems;

(e) unlawful injury means injury to data, data security or information systems that is contrary to law, or that exceeds or violates any lawful authority, consent or permission.

2 Insert:

Liability for certain acts – tracing and information gathering

A person is not subject to civil or criminal liability inside or outside [STATE] if the person causes any unauthorised access to data held in a computer or any compromise of the security system protecting the data held on a computer, if:

(a) the person is, or acts on behalf of, the owner of a computer system that has been subject to unauthorised access or exfiltration of data by another person;

(b) the person deploys software or hardware on the computer system of the other person for the purpose of gathering information about the unauthorised access or exfiltration of data;

(c) the person does not use or disclose any information collected by the software or hardware that is not reasonably considered relevant to identifying the person or system responsible for the unauthorised access or the defence of the person's system; and

(d) the person does not intentionally cause any damage, loss, or harm to the computer system or data of the other person, or any other person, as a result of the deployment of the software or hardware.

Liability for certain acts – damage or impairment to the computer of an attacker

A person is not subject to any civil or criminal liability for engaging in conduct inside or outside [STATE] that causes or is intended to cause computerrelated act, event, circumstance or result on the computer of another person (target computer) if:

(a) the person is, or acts on behalf of, the owner of a computer system that has been subject to unauthorised access or exfiltration of data by another person; and

(b) the person is reasonably satisfied that the target computer is the source of the attack; and

(c) the person is reasonably satisfied that the owner or operator of the target computer caused or permitted the attack to take place; and

(d) the conduct is likely to:

  (i) delete, damage or erase data present on the target computer without authorisation; and/or

  (ii) prevent or disrupt cybercrime; and

(e) the computer related act, event, circumstance or result is authorised by an active defence authority.

Active defence authorities

(1) A person may apply to a judge for an active defence authority.

(2) The application must include the evidence of each of the matters required by above.

(3) An application for an active defence authority must be dealt with in the absence of the public but is otherwise to be dealt with in such manner as is decided by the judge to whom the application is made.

(4) A judge must not issue an active defence authority unless the judge is satisfied that the application for the authority shows that reasonable grounds exist to justify its issue.

(5) When determining whether there are reasonable grounds to issue an active defence authority, a judge must have regard to the seriousness of the unlawful activity with which the application is concerned and the potential benefits of the conduct that would be authorised.

References

Ackerman, B (2003), 'The emergency constitution', *Yale Law Journal,* 113(5), 1029-1092.

Albakri A, Boiten EA and Lemos, RD (2019), 'Sharing cyber threat intelligence under the General Data Protection Regulation', in Naldi, M, Italiano, GF, Rannenberg, K, Medina, M and Bourka, A (Eds.), *Privacy technologies and policy*, Springer, Verlag, 28–41.

Al Qudat, MM (2009), 'Corporate criminal liability under the criminal laws of Jordan and Australia: A comparative analysis', *Journal Sharia and Law,* 37(8), 27-88.

Ashworth, AJ (2009), 'The doctrine of provocation', *Cambridge Law Journal,* 35(2), 292-320.

Ajayi, EFG (2016), 'Challenges to enforcement of cyber-crimes laws and policy', *Journal of Internet and Information Systems,* 6(1), 1-12.

Ball, D and Waters, G (2013), 'Cyber defence and warfare', *Security Challenges,* 9(2), 91-98.

Basu, A and Hickok, E (2020), 'Conceptualizing an international framework for active private cyber defence', *Indian Journal of Law & Technology,* 16(1), 16-47.

Berghel, H (2017), 'On the problem of (cyber) attribution', *Computer,* 50(3), 84-89.

Boerding, A, Culik, N, Doepke, C, Hoeren, T, Juelicher, T, Roettgen, C and Schoenfeld, MV (2018), 'Data ownership: A property rights approach from a European perspective', *Journal of Civil Law Studies,* 11(2), 323-370.

Broeders, D (2021), 'Private active cyber defense and (international) cyber security: Pushing the line?', *Journal of Cybersecurity,* 7(1), tyab010.

Bushby, A (2019), 'How deception can change cyber security defences', *Computer Fraud & Security,* 2019(1), 12-14.

Cameron, D (2015), *National security strategy and strategic defence and security review 2015: A secure and prosperous United Kingdom*, available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/478936/52309_Cm_9161_NSS_SD_Review_PRINT_only.pdf

Center for Cyber and Homeland Security (2016), *Into the gray zone: The private sector and active defense against cyber threats*, Washington, George Washington University.

Cook, C (2018), 'Cross-border data access and active cyber defense: Assessing legislative options for a new international cyber security rulebook', *Stanford Law & Policy Review,* 29, 205-236.

Cooper, JC and Shepherd, J (2016), 'State unfair and deceptive trade practices laws: An economic and empirical analysis', *Antitrust Law Journal,* 81(3), 947-980.

Corn, G and Jensen, E (2018), 'The use of force and cyber countermeasures', *Temple International & Comparative Law Journal,* 32(2), 127-134.

Creado, Y and Ramteke, V 2020, 'Active cyber defence strategies and techniques for banks and financial institutions', *Journal of Financial Crime,* 27(3), 771-780.

Crusto, MF (2015), 'State of emergency: An emergency constitution revisited', *Loyola Law Review,* 61(3), 471-524.

Curry, J (2012), 'Active defence', *ITNOW,* 54(4), 26–27, https://doi.org/10.1093/itnow/bws103

Department of Home Affairs (2023), *2023–2030 Australian cyber security strategy,* available at: https://www.homeaffairs.gov.au/cyber-security-subsite/files/2023-cyber-security-strategy.pdf

Department of Home Affairs (2024), *Cyber security legislative reforms engagement*, available at: https://www.homeaffairs.gov.au/help-and-support/how-to-engage-us/consultations/cyber-security-legislative-reforms

Dewar, RS (2014), 'The triptych of cyber security: A classification of active cyber defence', in Brangetto, P, Maybaum, M and Stinissen, J (Eds.), *Proceedings of the 2014 6th international conference on cyber conflict*, Tallinn, CCDCOE, 7-21.

Easterly, J and Fanning, T (2023), *The attack on colonial pipeline: What we've learned & what we've done over the past two Years*, available at: https://www.cisa.gov/news-events/news/attack-colonial-pipeline-what-weve-learned-what-weve-done-over-past-two-years

Gallagher, H (2022), 'Recognising a right to hack back: Tom and Jerry in cyberspace?', *Trinity College Law Review,* 25, 56-82.

Gerke, K (2021), 'Canadian hack-back?: A consideration of the Canadian legal framework for private-sector active cyber defence', *Alberta Law Review,* 59(1), 171-200.

Global Commission on the Stability of Cyberspace (2019), *Advancing cyberstability*, available at: https://cyberstability.org/assets/images/report/GCSC-Advancing-Cyberstability.pdf

Government of Canada (2017), *Defence policy: Strong, secure, engaged*, available at: https://www.canada.ca/en/department-national-defence/corporate/reports-publications/transition-materials/transition-assoc-dm/defence-policy-sse.html

Grimmelman, J and Mulligan, C (2023), 'Data property', *American University Law Review,* 72(3), 829-884.

Gruber, A (2015), 'A provocative defense', *California Law Review,* 103(2), 273-334.

Halberstam, M (2013), 'Hacking back: Re-evaluating the legality of retaliatory cyberattacks', *George Washington International Law Review*, 46(1), 199-238.

Han, X, Kheir, N and Balzarotti, D (2018), 'Deception techniques in computer security: A research perspective', *ACM Computing Surveys (CSUR),* 51(4), 1-36.

Hanson, F and Uren, T (2018), *Australia's offensive cyber capability*, available at: https://www.aspi.org.au/report/australias-offensive-cyber-capability

Harknett, RJ and Smeets, M (2022), 'Cyber campaigns and strategic outcomes', *Journal of Strategic Studies,* 45(4), 534-567.

Hoffman, W and Nyikos, S (2018), *Governing private sector self-help in cyberspace: Analogies from the physical world*, available at: https://carnegieendowment.org/research/2018/12/governing-private-sector-self-help-in-cyberspace-analogies-from-the-physical-world?lang=en

Housen-Couriel, D (2021), 'Hacking back under international law: Toward effective remedies against cyberattacks for non-state actors', in Siboni, G and Ezioni, L (Eds.), *Cybersecurity and Legal-Regulatory Aspects*, World Scientific, New York, 103-133.

Irving, L (2013), *Active cyber defense: A framework for policymakers*, available at: https://www.cnas.org/publications/reports/active-cyber-defense-a-framework-for-policymakers

Jakab, A 2006, 'German constitutional law and doctrine on state of emergency: Paradigms and dilemmas of a traditional (continental) discourse', *German Law Journal,* 7(5), 453-477.

Juels, A and Rivest, RL (2013), 'Honeywords: Making password-cracking detectable', *Proceedings of the 2013 ACM SIGSAC conference on computer & communications security.*

Juels, A and Ristenpart, T (2014), 'Honey encryption: Security beyond the brute-force bound', *Annual international conference on the theory and applications of cryptographic techniques*.

Jun, O (2023), 'Direction of Japan's new cybersecurity policy', *Asia-Pacific Review,* 30(3), 63-78, https://doi.org/10.1080/13439006.2023.2295707

Karimi, Y, Squicciarini, A and Wilson, S (2022), 'Automated detection of doxing on Twitter', *Association for Computing Machinery,* 6(3).

Kukul, B (2023), 'Personal data and personal safety: Re-examining the limits of public data in the context of doxing', *International Data Privacy Law,* 13(3), 182-192.

Lawrence, DE (2007), 'It really is just a game: The impracticability of common law property rights in virtual property', *Washburn Law Journal,* 47(1), 505-550.

Lin, HS (2010), 'Offensive cyber operations and the use of force', *Journal of National Security Law & Policy,* 4, 63-86.

Lin, P (2016), *Ethics of hacking back*, U.S. National Science Foundation, San Luis Obispo, CA.

McGee, S, Sabett, RV and Shah, A (2013), 'Adequate attribution: A framework for developing a national policy for private sector use of active defense', *Journal of Business & Technology Law,* 8(1), 206.

Miraglia, A and Casenove, M (2016), 'Fight fire with fire: The ultimate active defence', *Information & Computer Security,* 24(3), 288-296.

Montasari, R (2023), 'Cyber threats and the security risks they pose to national security: An assessment of cybersecurity policy in the United Kingdom', in Montasari, R (Ed.), *Countering cyberterrorism: The confluence of artificial intelligence, cyber forensics and digital policing in US and UK national cybersecurity,* Springer, Verlag, 7-25.

National Institute of Standards and Technology [NIST], (2024), *Computer Security Resource Center, Glossary*, available at: https://csrc.nist.gov/glossary/term/active_cyber_defense

Pengilley, W (2007), 'Fair trading, misleading or deceptive conduct', *University of Queensland Law Journal,* 26(1), 215-218.

Powell, J and Dolan, A (2021), *Active cyber defence tips the scales back in favour of the enterprise*, available at: https://purple.telstra.com.au/insights/thought-leadership/active-cyber-defence-tips

Powell, J (2021), *Deception in the essential eight*, available at: **https://acda.group/articles/**

Ramsey, CB (2010), 'Provoking change: Comparative insights on feminist homicide law reform', *Journal of Criminal Law & Criminology,* 100(1), 33-108.

Rid, T and Buchanan, B (2015), 'Attributing cyber attacks', *Journal of Strategic Studies,* 38(1–2), 4–37.

Rosenzweig, P (2013), 'International law and private actor active cyber defensive measures', *Stanford Journal of International Law,* 47(1), 1-15.

Rosenzweig, P (2014), 'International law and private actor active cyber defensive measures'. *Stanford Journal of International Law,* 50(1), 103-118.

Rudolph, A (2021), *Canada's active cyber defence is anything but active*, available at: https://www.cgai.ca/canadas_active_cyber_defence_is_anything_but_active

Sexton, M (2016), 'U.K. cybersecurity strategy and active cyber defence: Issues and risks', *Journal of Cyber Policy,* 1(2), 222-242.

Shackelford, SJ, Charoen, D, Waite, T and Zhang, N (2019), 'Rethinking active defense: Comparative analysis of proactive cybersecurity policymaking', *University of Pennsylvania Journal of International Law,* 41(2), 377-428.

Simonato, M, (2014), 'Defence rights and the use of information technology in criminal procedure', *Revue Internationale de Droit Penal,* 85(1), 261-310.

Steingartner, W, Galinec, D and Kozina, A (2021), 'Threat defense: Cyber deception approach and education for resilience in hybrid threats model', *Symmetry,* 13(4), 597-622.

Stevens, S (2020), 'A framework for ethical cyber-defence for companies', in Christen, M, Gordijn, B and Loi, M (Eds.), *The ethics of cybersecurity*, Springer, Cham, 317-330.

The Commonwealth (2018), *Commonwealth cyber declaration*, available at: https://thecommonwealth.org/commonwealth-cyber-declaration-2018

Thinyane, M and Christine, D (2020), *Cyber-resilience in the Asia Pacific: A review of national cybersecurity strategies*, available at: https://collections.unu.edu/eserv/UNU:7760/n2020_Cyber_Resilience_in_Asia-Pacific.pdf

Tran, D (2018), 'The law of attribution: Rules for attribution the source of a cyber-attack', *Yale Journal of Law & Technology,* 20, 376-441.

United States Department of Defense (2011), *Department of Defense strategy for operating in cyberspace*, available at: https://csrc.nist.gov/CSRC/media/Projects/ISPAB/documents/DOD-Strategy-for-Operating-in-Cyberspace.pdf

Van Dine, A (2019), 'When is cyber defense a crime? Evaluating active cyber defense measures under the Budapest Convention', *Chicago Journal of International Law,* 20(2), 530-564.

van Wyk, W (2015), 'Deception by omission: A warning about lawyers' duty to the court', *Bulletin (Law Society of South Australia),* 37(5), 37.

Walker-Munro, B, Mount, D and Ioannou, R (2022), 'The hacker strikes back: Examining the lawfulness of "offensive cyber" under the laws of Australia', *Computers & Law,* 94, 5.

Walker-Munro, B and Dov Bachmann, S-D (2024), *When cyber defence involves attack: Issues for Australia*, available at: https://www.lowyinstitute.org/the-interpreter/when-cyber-defence-involves-attack-issues-australia

Walters, R (2023), *Cybersecurity and data laws of the Commonwealth: International trade, investment and arbitration*, Springer, Cham.

Wanner, B and Ghernaouti, S (2019), 'Conceptualizing active cyber defence in cyber operations', *St Antony's International Law Review,* 15(1), 58-82.

Waxman, MC (2011), 'Cyber-attacks and the use of force: Back to the future of article 2(4)', *Yale Journal of International Law,* 36, 421-460.

Willis, LE (2020), 'Deception by design', *Harvard Journal of Law & Technology,* 34(1), 115-190.

Zhang, L and Thing, VL (2021), 'Three decades of deception techniques in active cyber defense: Retrospect and outlook', *Computers & Security* 106, 102288-102307.

About the authors

Dr Brendan Walker-Munro (lead author) is Senior Lecturer (Law) with the Faculty of Business, Law & the Arts at Australia's Southern Cross University. Brendan's focus is on "research security", the use of law and policy to protect university research from national security threats such as espionage, foreign interference, hacking, and technology transfer.

Andrew Cox is President of Australia's Active Cyber Defence Alliance Inc. and Principal Consultant at Avantgard, where he guides organisations in advancing from static to active cyber defence.

Grant Haroway is Managing Director, SiegeBrake Cyber Incident Readiness

Joe Otway is Cyber Security Architect, Australian Competition and Consumer Commission. Introduced to IT security in 1992, Joe has since operated in a range of information and cyber security roles in both Australia and overseas. He works with organisations that are critically dependent on secure information and IT.

Duncan Unwin is Practice Manager for Business Aspect, the consulting business of Data3. His cyber security work focuses on strategic threat intelligence in high risk and critical infrastructure environments.

Sascha Dov Bachmann is a Professor in Law at Canberra Law School, University of Canberra; Extraordinary Reader (Docent) in War Studies, Swedish Defence University (FHS), Stockholm; Fellow at NATO SHAPE, Hybrid War, and Lawfare Pacific and is working on various aspects of cyber operations and human security.