LA TROBE UNIVERSITY

Australia's 2020 Cyber Security Strategy – A call for views

1 November 2019

A personal view on:

whether URLs in general, and specifically, suspected and known malicious URLs should legally be classified as 'the contents of a communication', under Australian law

To: Cyber Security Policy Division
Department of Home Affairs
4 National Circuit Barton ACT 2600

ENQUIRIES
Dr. Stanley Shanapinda
Research Fellow
La Trobe University
Victoria 3086

T [redacted]
E
latrobe.edu.au

Contents

I. Executive summary

This submission is in response to a call for views regarding Australia's 2020 Cyber Security Strategy development process.

Malicious Uniform-Resource Locator (URL)-based cyber-attacks are rising since the end of 2018 at the rate of 26%. To effectively detect and prevent URL-based attacks, law enforcement such as the Australian Federal Police (AFP), the Australian Securities and Investments Commission (ASIC), security agencies such as the Australian Security Intelligence Organisation (ASIO) (jointly referred to as the Agencies), and other agencies such as the Australian Signals Directorate (ASD), with the Australian Cyber Security Centre (ACSC) that is part of the ASD, may need to collect, access and use malicious URLs to investigate, inquire into, research, prevent and propose cybersecurity countermeasures. This is so given that backlisting of domains and URLs is ineffective. Machine learning techniques can be used to improve this situation. The problem is this however - currently it is unclear whether URLs are regarded as the contents or substance of a communication under Australian law and policy. This creates uncertainty in the regulatory environment and hampers the sharing of cybersecurity data because the privacy protection of individuals is not clearly articulated in the law. Any future cybersecurity strategy, policy and law, must seriously address this concern if cyber risks are to be effectively addressed.

This submission therefore argues that URLs must legally be classified as the contents or substance of a communication under the *Telecommunications (Interception and Access) Act 1979* (Cth) (*TIA Act 1979*). URLs can be argued to be both content and non-content – URLs are non-binary. URLs can be said to be metadata because of the way they relate to web pages as the content. URLs are web address that links the user with the content when clicked. However, the inquiry does not end here. URLs can also be said to be the contents of a communication based on the following:

i. the URL contains the path and query portions which is content generated by the users web browsing activities;
ii. when one IP address is used for many websites the authority part of the URL must be inspected by the web server, which means the web server is the recipient of a message that it inspects as opposing to retrieving the website without examining the authority part, which would otherwise be metadata;
iii. a malicious URL is sent in a phishing email or an SMS, as its contents, addressed to the individual Australian citizen and resident to click on, who is directed to a fake website with the intention to defraud the recipient, whether personally targeted or not; and
iv. the legitimate URL is hi-jacked and injected with inputs so that the individual receives and clicks on the URL as a message sent to them, to be redirected to a malware site or to download malware to steal their personal information.

Preference must however be given to the content nature of URLs as the contents or substance of a communication. Therefore, a preservation notice, and a warrant process should be followed, as provided for in *TIA Act 1979,* to collect and access URLs for the investigation and detection of attack types. This scenario is no different to people conspiring to commit a crime and messaging each other over SMS or email. To gain access to the content of these messages, law enforcement agencies must apply for a warrant to request same.

Denying the content nature of URLs and allowing URLs to be collected under the authorisation and notification process under the *TIA Act 1979* will not ensure public trust and confidence in the cybersecurity strategy, policies and laws. Individuals should be allowed to request access to their URLs and to donate same for cybersecurity research purposes. The Telco should be allowed to de-anonymise URLs and use the information to research cybersecurity tools to better combat malicious URLs and to share the data with academia to collaborate on such projects. Individuals should be allowed to object to the collection and use of the URLs they accessed, in open court. This framework should be extended beyond Australian telecommunications companies (Telco(s)), Australian Internet Service Providers (ISPs) and multinational Social Media Platform Companies (SMPCs) such as Google and Facebook, to require private businesses to share malicious URLs under preservation notices and judicial warrants, issued by an independent court and judges, that are competently qualified, subject to robust appeal and judicial review processes. There should also be a general requirement, as is the case in respect of Telco's and ISP's under section 313 of the *Telecommunications Act 1997* (Cth) (*TA Act 1997*), for private Australian business as well, to do their best to ensure the confidentiality, integrity and availability of information and communications. In this manner private businesses will be obliged to roll out best practice malicious URL detection and introduce the best countermeasures.

However, private companies run and operate propriety networks that are their private domain. Access to that private domain can only be allowed through their consent and permission and through other lawful means by the Agencies, as is the norm in a democratic society. These lawful means are the judicial warrants, the preservation notices process under the *TIA 1979* and the assistance and access regime under the *Telecommunications and Other Legislation Amendment (Assistance and Access) Act 2018* (Cth) (*Assistance and Access Act 2018*). However, these processes only apply to Telco's, SMCPs and ISPs. As regards entities such as critical infrastructure companies, private Australian businesses, private Australian citizens and residents, judicial warrants and the preservation notices processes under the *TIA Act 1979* do not apply to them. Under a new cybersecurity law, a similar preservation notice process and judicial warrant process should be considered to enable the collection of malicious URLs by the Agencies and ASD. The private business and individuals should be informed in advance and be allowed the right to object on open court to such requests, followed by review and appeals processes, in a robustly open, transparent and democratic legal system.

II. About the author

My name is Dr. Stanley Shanapinda. I am a Research Fellow at the Optus La Trobe University Cyber Security Research Hub in Melbourne, Australia.

I am a graduate of the Australian Centre for Cyber Security (ACCS), based in Canberra at the Australian Defence Force Academy (ADFA), of the University of New South Wales (UNSW).

I would like to thank the Department of Home Affairs for the opportunity to present my personal views on the envisaged 2020 Cyber Security Strategy. The views are expressed herein are solely for this express purpose and are not meant to offend or cause harm.

A. Experience and Research

I have over 16 years of experience in the telecommunications regulatory space as in-house legal counsel of a telecommunications corporation, inaugural CEO of an ICT regulator and researcher.

I continue to research the dynamic relationships between:

- the powers of law enforcement and national security agencies, telecommunications companies, ISPs and social media tech companies, to access and use telecommunications content and metadata of individuals for their investigatory, commercial and other functions;
- the role of the Agencies, Telco's, ISPs and Social Media Platforms, to protect privacy, ensure cybersecurity resilience, digital services, digital transformation, the digital economy and digital disruption;
- the latest developments in information and communications technologies, web applications and cybersecurity trends; and
- the role of independent oversight, accountability, governance and ethics.

I have researched and published in the areas of telecommunications, cybersecurity, privacy and governance, and have offered expert advice and opinions in the media, such as ABC News 24, The Conversation, SBS and the ABCs RNDrive radio program.

I have lectured on ICT Regulation in South Africa and Brazil; developed teaching material on for the Cybersecurity Governance master's program at La Trobe University; and is busy developing and will teach the inaugural Cyber Policy and Regulation course to cybersecurity students at La Trobe University in 2020. I am also currently developing a Digital Strategy roadmap for the Ministry of ICT in Namibia, with a focus on cybersecurity. I am also affiliated with the Learning Information Networking Knowledge (LINK) Centre at the University of the Witwatersrand (Wits) Tshimologong Digital Innovation Precinct in Johannesburg, South Africa.

My contact details and professional profiles can be found at:

- https://scholars.latrobe.edu.au/display/sshanapinda;
- https://www.linkedin.com/in/stanley-shanapinda-phd-b4909b21/; and
- https://twitter.com/stanamor.

B. Personal views and no conflict of interest

The views expressed herein are my personal and independent research and views, and do not represent the views or opinions of the Optus La Trobe Cyber Security Research Hub or the University of La Trobe Melbourne, or any of our associates or sponsors. I have no conflict of interest to declare, and no relationships with any other third party that may have an interest in the development or outcome of the envisaged strategy.

I am available to discuss the views expressed herein at the convenience of the Department of Home Affairs.

III. Introduction and Background

The digital age has ushered in major cyber risks. All types of smart technology devices, from computers to phablets to tablets, are connected to the Internet. It is reported that 95 per cent of breaches are the result of human error.[1] The individual uses the internet to browse and search for information and to communicate using emails and social media applications. Cyber threats have become more sophisticated and dynamic. Vulnerabilities are self-replicating and are hard to detect, taking longer to discover and to recover from. Real time monitoring, with a cyber situational approach may be what is required to effectively protect against cyber threats. Devices and equipment belonging to private individuals and business are used as vectors to pose cyber risks to national interests, such as electricity systems, water management, the financial services sector and the health sector. Governments are not able to address such national interests risks effectively given the private ownership of the devices and the personal nature of the information and communications stored on and shared via these devices and technologies. Access can only be granted via lawful means, as is the norm in a democratic government system and with the consent of the private business and private individual.

URL links are sent in emails and social media messages and accessed on personal and business devices. URL links can in automated fashion, or when clicked on infect whole networks with cyber threats. Google search results can also return a malicious URL that the individual then clicks on. So, there are various types of URLs that are generated at the initiative of the individual, when browsing the Internet; when presented in a phishing email and then clicked on; or when presented with automatically downloaded malware without the individual clicking on the link, by simply visiting the malicious website. In all these instances, the URL forms part of the activity history of the individual using the device. This can be referred to as the 'web browsing history'. As such, URLs need to be collected, stored, analysed and used to inquire into and investigate cybercrimes, cyber incidents and cyber-attacks more effectively.

A. Scope of the Submission

The Australian government called for views regarding Australia's 2020 Cyber Security Strategy development.[2] The government expressed its desire to be a world leader in cyber security by stating:

> For nationally significant systems, such as those that control our power and water, Australia must position itself as a world leader in cyber threat detection, prevention and response.[3]

Known malicious URLs are increasingly used as a threat vector, and suspected malicious URLs, that individual Australian citizens and residents, in their personal and private capacity or as employees of a public body or private business, may need to be collected, accessed and analysed to detect and prevent cyber incidents and cyber-attacks.

[1] Byron Connolly (CIO). 06 March, 2019 11:46. Malicious URLs now rampant: study. https://www.cio.com.au/article/658501/malicious-urls-now-rampant-study/
[2] Commonwealth of Australia 2019. Australia's 2020 Cyber Security Strategy. A call for views. <https://www.homeaffairs.gov.au/reports-and-pubs/files/cyber-security-strategy-2020-discussion-paper.pdf>
[3] Ibid Pg. 4

Malicious URLs can be used to target power and water systems. It however seems that it is not clear how access to URLs must be obtained legally by the relevant authorities. This depends on the legal classification of URLs as either telecommunications data or as the contents or substance of a communication, an issue that is technologically complex and seems legally and policy-wise unresolved.

This issue is crucial to resolve in order to assess the applicable governance system when it comes to the level privacy protection to be afforded to Australian citizens and residents. If URLs are legally 'content' then preservation notices and warrants must be issued to collect URLs. On the other hand, if URLs are legally 'telecommunications data', or commonly 'metadata', URLs may be accessed by the Agencies issuing their own notifications and authorisations to access and use URLs, with not reasonable suspicion standard required to be met, as discussed in Part VIII.

Not following the correct legal process may have serious implications for the investigations and inquiries conducted by the Agencies and the evidence collected and presented in court to attribute a cybercrime to a particular individual or organisation that may have aided in, created and distributed the malware. It is therefore crucial that this policy issue is addressed and settled, to create a predictable and clear legal environment for all private and public businesses and the ordinary citizens and residents that might be impacted by it.

As such, the submission will try and address the following:

1. How URLS are used in the commission of cybercrimes, cyber incidents and cyber-attacks;
2. The architectural structure of URLs in general and known malicious URLs;
3. Whether URLs are telecommunications data or the contents or substance of a communication:
    a. How the collection, storage, disclosure and use of URLs (as telecommunications data or as the contents or substance of a communication) are addressed to investigate and inquire into cyber security risks, cybercrimes, cyber incidents and cyber-attacks under existing Australian laws such as the *TIA Act 1979*, *Telecommunications Act 1997*, and *Assistance and Access Act 2018,* and in relation to Telco's, ISPs and Social Media Platform Companies (SMPCs) such as Google and Facebook, referred to as Designated Service Providers (DSPs). These are referred to as Category A Entities;
    b. How the collection, storage, disclosure and use of URLs (as telecommunications data or as the contents or substance of a communication) are addressed to investigate and inquire into cyber security risks, cybercrimes, cyber incidents and cyber-attacks are addressed under existing Australian laws, and in relation to privately owned privately owned and operated Australian companies, partnerships, businesses, corporations, charities, institutions, pubic bodies, statutory bodies, civil society organisations etc. These are referred to as Category B Entities;
4. Recommendations:
    a. How suspected malicious URLs should be legally treated in respect of both Category A and B Entities to better prevent cyber risks and better protect privacy.
5. Recommended Strategies:
    a. To introduce an evidenced-based regulatory framework, for private and public entities, that enables access to and the use of suspected and known malicious URLs

to investigate cyber threats, in a manner that protects privacy better and instils public trust.

B. Questions Answered from the Cyber Security Discussion Paper

The key question for any cyber security strategy, as paraphrased from Question 10 of 'Australia's 2020 Cyber Security Strategy - discussion paper', is this:

> is the current regulatory framework appropriate for the Agencies to collect, access and use of URLs, whether clicked on by the individual or not, for the purpose of inquiring into and investigating the URLs, to detect malicious threats?

The discussion about the existing framework will touch on the existing role the Agencies play and how the Agencies maintain trust. The proposals on changes to the existing framework will address the revised role of the Agencies and how the Agencies can continue to maintain trust when it comes to the handling of personal information and the contents of communication.

The submission broadly relates to the following ten questions from the discussion paper: Questions 1, 2, 3, 5, 10, 11, 14, 18, 21 and 26. The answers to the questions, are found in Part XIII. A, and are preceded by the discussions below.

C. Out of Scope

The submission will not directly cover the following.

1. How the collection, storage, disclosure and use of URLs (as telecommunications data or the contents or substance of a communication) are addressed as it relates to privacy and personal information under the *TIA Act 1979*, the *Telecommunications Act 1997* and the *Privacy Act 1988.*

The only statements that may be made in relation to privacy are these:

- If URLs are regarded as the contents of a communication then, the privacy question may be settled by confirming that access to URLs must be obtained with a judicial warrant, like other types of content, such as voice communications, which are also personal information.
- Even if URLs are considered personal information, it may still be disclosed to the Agencies under the notification and authorisation process set out in the *CAC Determination* 2015, under the *TIA Act 1979*. It is only if the URL is legally classified as the contents of a communication that it may not be legal to disclose under the notification and authorisation process set out in the CAC Determination 2015, under the *TIA Act 1979*. The URL must then be disclosed under the preservation notice and warrant system that regulates the disclosure of the contents of a communication. Telecommunications data that is required to be retained is deemed to be personal information under the *TIA Act 1979.* Even if the telecommunications data is not deemed to be personal information, the telecommunications data may still be shown to be personal information if it meets the definition of personal information under the *Privacy Act 1988.* In other words, the Agencies may issue notifications and authorisations the *CAC Determination* 2015, under the *TIA Act 1979* to collect personal information, unless the said personal information is also the

contents of a communication. If it is the contents of a communication, the notifications and authorisations under the *CAC Determination* 2015 would be inappropriate to use. The preservation notices and the warrants under the *TIA Act 1979* would be the appropriate tools to use to collect the retained contents of a communication, whether it is personal information or non-personal information. This is illustrated by Table 2.

- If URLs are legally considered as the contents of a communication, then a more acceptable legal process of interfering with the privacy of the individual is followed. This process is the use of judicial warrants and preservation notices. A. In this manner the privacy of the individual Australian citizen and resident is better protected.

IV. The architecture of the URL

A Uniform Resource Locator (URL) is commonly referred to as a ***web address***, or the address of a World Wide Web page (WWW):

> '"A URL is an identifier, such as a webpage reference, used to locate a resource on the Internet… The URL is analogous to the name used when addressing a postal envelope."'.[4]

The hyperlink below is an example of a URL:
<https://cybersecuritystrategy.homeaffairs.gov.au/AssetLibrary/dist/assets/images/PMC-Cyber-Strategy.pdf>

The URL consist of various parts: the scheme, host, port, the optional path and query.[5]

i. It includes the **scheme**. In this case the scheme is the part indicated by the 'https'. It is also the protocol;
ii. It consists of the **host**. The part 'homeaffairs.gov.au' is the hostname. It refers to the server that is connected to the Internet. The host is the domain name, appearing in human readable form. It is also represented by a 32-bit IPv4 address or a 64-bit IPv6 address. The host is referred to as network address.[6] It is also referred to as the **authority.** This submission will inquire into whether the **authority** of the URL should legally be classified as the contents or substance of a communication.;
iii. It also includes the **port**. The port number for the 'http' scheme is 80 and the port number for the 'https' scheme is 443, which are the addresses generally identified with accessing the internet;[7]
iv. The **path** is generally separated by slashes. For example, the part: '/AssetLibrary/dist/assets/images/PMC-Cyber-Strategy.pd' is the path. This submission will inquire into whether this part of the URL should legally be classified as the contents or

[4] Telstra Corporation Limited and Privacy Commissioner [2015] AATA 991 (18 December 2015) [50.]; URL. Living Standard — Last Updated 21 October 2019 [4.1.] https://url.spec.whatwg.org/; IETF. RFC 1738. https://tools.ietf.org/html/rfc1738; IETF. RFC 3986. https://tools.ietf.org/html/rfc3986; IETF. RFC 2396. https://www.ietf.org/rfc/rfc2396.txt

[5] Bellovin, Steven M., Matt Blaze, Susan Landau, Stephanie K. Pell, 'It's Too Complicated: The Technological Implications of IP-Based Communications on Content/Non-Content Distinctions and the Third Party Doctrine,' (2016) 30(1) Harvard Journal of Law & Technology

[6] URL. Living Standard — Last Updated 21 October 2019 [3.1.] https://url.spec.whatwg.org/

[7] Ibid [4.2.]

substance of a communication. The submission will also critically analyse whether the whole of the URL should be regarded as the contents or substance of a communication.

v. The **query** part is optional, and is the part preceded by a question mark, in the URL below. That part includes the words typed in to the web browser by the individual. For example, if I type in the words 'home affairs cyber strategy' the browser suggest to me by Google may appear as:
<https://www.google.com/search?q=home+affairs+cyber+strategy&rlz=1C1CHBF_en-GBAU820AU820&oq=home+affairs+cyber+strategy&aqs=chrome..69i57j0j69i60l3.1837j0j7&sourceid=chrome&ie=UTF-8>
It includes the content typed into the search bar by the individual. This part is a communication from the individual Australian citizen or resident, to the web server, via the web browser. This URL would qualify as a record of the web browsing history. The question is given this, whether the URL should not be classified as the contents or substance of a communication. With malicious URLs however, this part may not be created by the individual, but instead manipulated by the attacker, as discussed in Part V. C.

When clicking on this URL, the Google server will re-direct the request for the web page back to the actual Home Affairs URL:
<https://cybersecuritystrategy.homeaffairs.gov.au/AssetLibrary/dist/assets/images/PMC-Cyber-Strategy.pdf>

A web browser such as Google, is a software application that makes an http or https (Hypertext Transfer Protocol Secure) request for the information the individual is looking for from the web server that is hosting the information and return options of what is available to the individual. If the individual clicks on the URL link that is returned the individual is then presented with the web page.

Regarding malicious URLs, an example is shown below. The malicious URL was sent in a phishing email, a very common form of distribution for malicious URLs. This issue is one that raises key pointers about why a URL, communicated in an email, must legally be classified as the contents of a communication:

Figure 1. Example content-less email.

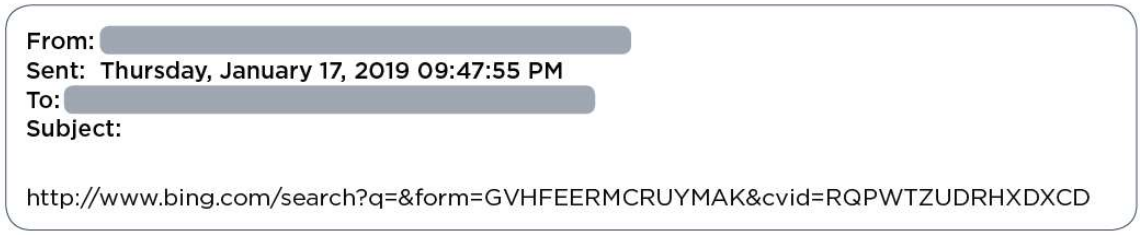

Figure IV.1 Example of a malicious URL sent in an email[8]

[8] Fireeye. Email Threat Report – Q1 2019. Pg. 3. https://www.fireeye.com/offers/rpt-email-threat.html

All parts that are common across all URLs, as discussed above, are evident from the above example of a malicious URL. The architectural structure is described in Part V. C.

V. The Cybersecurity Risks Inherent in URLs

URLs are prone to spoofing. The URL link is fraudulent. It looks similar to the legitimate URL. The website that is set up is also fake and is a replica of the legitimate website. The aim is to deceive the user and steal their personal and sensitive information.[9] It is an attack whereby the user can confuse one host or URL for another when looking at the alphanumeric structure of the URL. Misspelled words may not be easy to pick up. URLs may be shortened and that can be misused for spreading malicious URLs. These URLs come from untrusted sources and can be injected with inputs that allow for the leakage of personal and sensitive information, thereby harming the user, and perhaps leading to major cyber incidents and attacks that can potentially target water and power systems. [10] A typical example of an attack that injects inputs is a Cross-Site Scripting (XSS) attack.

Other ways in which malicious URLs present cyber risks include: link manipulation, fast-flux hosting, URL obfuscation, JavaScript obfuscation, semantic URL attacks, encoded URL attacks and reflected URL attacks.[11]

As a precautionary measure, the URL may be modified to prevent URL spoofing. As such, the domain name or the scheme may be left out and not be displayed.[12]

[9] NORDVPN. URL Spoofing: Definition and explanation. Jan 29, 2019. https://nordvpn.com/blog/url-spoofing/
[10] URL. Living Standard — Last Updated 21 October 2019 [2.] https://url.spec.whatwg.org/
[11] ShymalaGowri Selvaganapathy, Mathappan Nivaashini & HemaPriya Natarajan (2018) Deep belief network based detection and categorization of malicious URLs, Information Security Journal: A Global Perspective, 27:3, 145-161, Pg. 159, DOI: 10.1080/19393555.2018.1456577.
[12] Fn 10 [4.2.], [4.8.1.]

A. Recent Trends in URL-based cyber attacks

One key vector used for such cyber threats are Uniform Resource Locators (URLs). Since the last quarter of 2018, there has been a 26% increase in malicious URLs using HTTPS to launch cyber threats. Globally, URL-based attacks distributed via phishing emails were the main instrument for the delivery of malicious content between January and March 2019. This trend started in 2018 and continued into 2019. HTTPS was considered more secure than HTTP, but it seems this new trend is threatening the trust users have in HTTPS. Given how new this phenomenon is, it is a challenge to detect and identity URL-based attacks. A more urgent and dynamic method of detection is therefore required.[13]

This new trend is dynamic in the following ways:

i. the emails contain only the malicious URL as its content which makes it easier to bypass email filters;
ii. non-clickable URLs are used, which means the link is not live. The link is activated when copied and pasted into the browser. The link is not made live because in this way it is able to bypass security filters.[14]

URL-based attacks have become increasingly popular since the end of 2018.

B. Malicious URLs as #FakeWebsites

A malicious URL can also be referred to as a malicious website.[15] A malicious URL is a URL that has been created with the aim of data theft, money theft or the theft of personal information through the use of fake website.[16] Malicious URLs play host to unwelcome content. These include spam, phishing and drive-by downloads. The individual is then presented with the infected website. Malicious software will then download to the computer of the individual to spy on the activities on

[13] Fireeye. Email Threat Report – Q1 2019. Pg. 2-3. https://www.fireeye.com/offers/rpt-email-threat.html
[14] ibid
[15] Sahoo, D., Liu, C., & Hoi, S. C. H. (2019). Malicious URL detection using machine learning: A survey. Ithaca: Cornell University Library, Pg. 1, arXiv.org. Retrieved from http://ez.library.latrobe.edu.au/login?url=https://search-proquest-com.ez.library.latrobe.edu.au/docview/2075353706?accountid=12001; Sirageldin A., Baharudin B.B., Jung L.T. (2014) Malicious Web Page Detection: A Machine Learning Approach. In: Jeong H., S. Obaidat M., Yen N., Park J. (eds) Advances in Computer Science and its Applications. Lecture Notes in Electrical Engineering, vol 279. Springer, Berlin, Heidelberg; RongWangYanZhuJiefanTanBinbinZhou. Detection of malicious web pages based on hybrid analysis. Journal of Information Security and Applications Volume 35, August 2017, Pages 68-74; ShymalaGowri Selvaganapathy, Mathappan Nivaashini & HemaPriya Natarajan (2018) Deep belief network based detection and categorization of malicious URLs, Information Security Journal: A Global Perspective, 27:3, 145-161, DOI: 10.1080/19393555.2018.1456577 .
[16] ShymalaGowri Selvaganapathy, Mathappan Nivaashini & HemaPriya, Natarajan (2018), Pg. 145

the computer and to steal personal and sensitive information. Innocent Australian citizens and residents become victims of scams such as monetary loss, theft of private information and malware installation because of malicious URLs. Generally, malicious URLs cause losses of billions of dollars every year.[17] This theft could lead to the loss of credentials, used to access Supervisory Control and Data Acquisition (SCADA) systems that control utility services such as water and electricity, if stolen from the Australian citizen or resident that has authorised access to it. The person can be targeted with a phishing email to click on an infected link that monitors their activities on the device and steal the passwords to gain access to these critical infrastructures. In all these instances the individual is navigating to these fake websites and interfaces or can be said to be browsing the websites, as they have enough time to enter their personal details, under the false impression that it was the legitimate URL they clicked on and therefore the legitimate website they visited. They may not know that they were duped until they have lost money, the control systems were accessed, or a fraud report is made about an account opened with their details, that they have no idea about. At all times the person may be under the impression they visited a legitimate website and may consider that access a personal and private matter:

- that no Australian business of which they are an employer,
- that no Telco or ISP that provided the internet service should disclose without their knowledge and prior informed consent,
- that none of the Agencies can simply access without a judicial warrant,
- and that access to the URLs should only be with their free and informed consent, and voluntary cooperation.

Under section 289 of the *Telecommunications Act 1997*, personal information may be disclosed with the knowledge or consent of the person concerned. In the 2018 financial year 1,802,706 disclosures were made with the knowledge or consent of the individual.[18] A similar process can be undertaken for the cyber security strategy to detect and investigate malicious URLs.

A system that allows for access to URLs under these circumstances is probably a regime the Australian citizen and resident can trust and have confidence in - their privacy is well protected throughout. People's fears that they may have done something wrong and acted illegally, for which they may be in trouble, may be allayed and their cooperation be obtained if such guidance is clear and unquestionable as to its ambiguity. The individual would want to know what their rights are and how they can ensure the uncomplicated and easy enforcement of their privacy rights, and still be

[17] Sahoo, D., Liu, C., & Hoi, S. C. H. (2019). Malicious URL detection using machine learning: A survey. Ithaca: Cornell University Library, Pg. 1, arXiv.org. Retrieved from http://ez.library.latrobe.edu.au/login?url=https://search-proquest-com.ez.library.latrobe.edu.au/docview/2075353706?accountid=12001; Sirageldin A., Baharudin B.B., Jung L.T. (2014) Malicious Web Page Detection: A Machine Learning Approach. In: Jeong H., S. Obaidat M., Yen N., Park J. (eds) Advances in Computer Science and its Applications. Lecture Notes in Electrical Engineering, vol 279. Springer, Berlin, Heidelberg; RongWangYanZhuJiefanTanBinbinZhou Detection of malicious web pages based on hybrid analysis, Journal of Information Security and Applications Volume 35, August 2017, Pages 68-74; ShymalaGowri Selvaganapathy, Mathappan Nivaashini & HemaPriya Natarajan (2018) Deep belief network based detection and categorization of malicious URLs, Information Security Journal: A Global Perspective, 27:3, 145-161, DOI: 10.1080/19393555.2018.1456577 .

[18] ACMA. Australian Communications and Media Authority annual report 2018–19, pg. 128

able to voluntarily participate in cybersecurity investigations, without potentially exposing themselves legally.

C. Three Types of Malicious URLs

There can be said to be three types of malicious URLs. These are: i) landing URLs; ii) distribution URLs; and iii) URLs that are generated by malware after drive-by downloads.

Figures V.1 and V.2 below illustrate some of the examples.

Table 1 – Feature sets and features.

Feature Set	Feature Name	Malicious URL Examples
Host-based features	similar domain	wsad004.asia, wsad066.asia, wsad123.asia
	similar subdomain	webhosting50.1blu.de, webhosting52.1blu.de
	similar IP prefix	104.149.195.115, 104.149.195.116
	similar port	174.139.26.238:801, 174.139.30.222:802
	same domain	couponmoeum.com/PEG/css/1.js, couponmoeum.com/PEG/css/ad.html
	same subdomain	aa.sswangima.com, aa.wangma1q.com
	same IP	175.126.74.101/files/2.js, 175.126.74.101/files/1.js
	same port	2288.org:8832, 6600.org:8832, 8800.org:8832, 8866.org:8832
Path-based features	similar pathname	sbsjob.co.kr/PEG/js/check_38746.js, **dowmi.net/PEG/ad/check_38746.js**
	same pathname	braxico.com/images/m2dcbAnA.php, kappen-orth.de/images/iRsGKfv2.php
Filename-based features	similar query type	rcjhvqmtkmp?bdhkuyjqwmpi=6621548, rckjyihjpoggki?bxvuwwsceqt=6621548
	similar filename	199.188.107.109/yy.html, 199.188.107.109/zz.html
	same query type	175.45.4.158/index.php?id=kim031, 175.45.4.158/index.php?id=kim032
	same filename	199.188.107.112/xiaoyu.html, 198.2.221.203/xiaoyu.html

Figure V.1 Examples of malicious URLs[19]

[19] Sungjin Kim, Jinkook Kim, Brent ByungHoon Kang, pg. 791.

Table A1 – Malicious URL examples.

Malicious IP Examples	Malicious Domain Examples
110.34.196.117/cake.php	www.zilphotography.com/.errordocs/FOZ4fFXK.php
110.34.196.125/cake.php	ropeholders.info/.errordocs/sQodNA2k.php
103.251.36.92/index.html	iambrucehan.com/.errordocs/75yYodGX.php
103.251.38.103/index.html	www.wigsislandstudio.com/.errordocs/GQsvQJzH.php
103.251.37.211/index.html	www.cleanenergyhi.com/.errordocs/6UhfUL3J.php
103.251.37.213/sb.html	ozactivity.com/.errordocs/wEXfiNFD.php
103.251.37.214/sb.html	perivaleproductions.com/.errordocs/zB0ypBk9.php
103.251.37.213/main.html	live-counter.net/?click=13950265
103.251.37.214/main.html	hosttracker.net/?click=6621593
103.251.37.213/index.html	hostverify.net/?click=1110828
103.251.37.214/index.html	webexperience13.com/?click=85009921
1.226.83.40/ts/index.html	thedeadpit.com/?click=341881
1.226.83.40/ts/dy.html	internetcountercheck.com/?click=13218787
113.10.187.41/oacs19/29.html	google-ana1yticz.com/?click=172078
113.10.187.42/oacs19/29.html	coaipr.org/aqgy.html?i=1958545
113.10.187.41/oacs19/man.html	petalconsultancy.info/aqgy.html?i=1958545
113.10.187.42/oacs19/man.html	lindsethcpas.com/aqgy.html?i=1958545
126.19.87.31/2222/tiancai.html	innerbath.com.au/crpy.html?j=1958545
126.19.87.31/2222/index.html	stevebeam.com/wrpy.html?i=1958545
116.81.235.128/3333/tiancai.html	ptsolutionsgroup.com/crgt.html?i=1958545
116.81.235.128/3333/index.html	morehead-motorsports.com/eqgy.html?i=1958545
126.114.226.40/3333/shifu.html	petalconsultancy.info/aqgy.html?i=1958545
126.114.226.40/3333/index.html	cronicadelcorrugado.com/mqpt.html?i=1958545
103.240.197.28/b.html	tvpasiones.com/arpy.html?i=1958545
103.240.197.30/c.html	abinnetsol.ca/eqgy.html?i=1958545
103.240.197.33/c.html	eastmead1.ipower.com/hrpt.html?i=1958545
103.240.197.35/k.html	wheresweems.com/argt.html?i=1958545
103.240.197.36/j.html	bestdeckshoes.com/oqpt.html?i=1958545
103.240.197.37/v.html	lindsethcpas.com/aqgy.html?i=1958545
113.10.187.41/live3/qq.html	burtcasey.net/wrpt.html?i=1958545
113.10.187.42/live3/qq.html	3diporn.com/eqgy.html?i=1958545
113.10.187.41/live2/qq.html	kirtidan.com/mqpt.html?i=1958545
113.10.187.42/live2/qq.html	budgetcancun.com/oqpy.html?i=1958545
113.10.187.41/code0002/qq.html	prmd.biz/wrpy.html?i=1958545

Figure V.2 Examples of malicious URLs[20]

i. Landing URLs are very similar to a benign URL with the aim of concealing the identity of an attack. They resemble benign URLs in terms of the URL length and lexical format. This malicious URL will look like a typical URL discussed in Part IV. above, for example: <https://cybersecuritystrategy.homeaffairs.gov.au/AssetLibrary/dist/assets/images/PMC-Cyber-Strategy.pdf>.

   Figure V.3 below illustrates the example.

[20] Sungjin Kim, Jinkook Kim, Brent ByungHoon Kang, pg. 804.

Table A2 – Landing URLs used in Alexa top 100.	
Landing URL	Distribution URL
http://l.facebook.com/lsr.php?u=http%3A%2F%2Fwww.duzonbiz.co.kr%2Fkey%2Fv3k.html&ext=1424074132&hash=AcnLdlvGrTbcTzTqzyVNC7QqyiAk4ctqMDAzdmxsPHkZEjv-	(2015-02-16) http://www.duzonbiz.co.kr/key/v3k.html
http://www.google.com/url?url=http://www.korailtour.com/UserFiles/gg/index.html&rct=j&frm=2&q=&esrc=s&sa=U&ei=VcHCVO2KDYf8ygOZ0YCICQ&ved=0CBQQFjAA&usg=AFQjCNE_7SOlEFqdS9l_cxgNPcFU6Et1yg	(2016-02-01) http://www.korailtour.com/UserFiles/gg/index.html
http://webcache.googleusercontent.com/search?q=cache:f3h3Cq-Aer4J:www.hyunjinsn.com/+&cd=1&hl=ko&ct=clnk&gl=kr	(2015-12-14) http://www.wonartschool.com/xe/libs/PEAR/view.html
http://yandex.ru/clck/jsredir?from=yandex.ru%3Bsearch%3Bweb%3B%3B&text=&etext=908.1vNg2-G_cw8lJCLRBthLM0kjY94_4GOVdgllaUV2Iyd3SZMpI_s09gbsMXpaQObG.bfe5ebfa9ebea90c7c0fe4daacc03ed318203d59&uuid=&state=AiuY0DBWFJ4ePaEse6rgeAjgs2pI3DW9J0KiE5XNXd0dp0ZMwFHoviUoYa6nzP7MFsomsouu4qcHbQqcq9usxGGO7RUCBA3CQOuv8Jg-Hj9QrokjqARXAhk_ZBegv2NHoKEopnuLoVMWYZiPqM4fPRV81es3G38m59_Blx_owikL3-IrlWDWd7PQ5JPN3hN4-uArkW6PSOiSO-qJhCUR3Q&data=UlNrNmk5WktYejR0eWJFYk1LdmtxdERJZnBDcW5pYTVPTjE4Mm42NHdyZmZPWDV4cnByVVNBTVFpS29iUkFhaE9NYkp4TFQ1WXZ2RzZCSXFlOU95emw3VnVPMFNvckxUY3Zzczl0Ukp1TEZiNlJvcm5rVVBLUQ&b64e=2&sign=cdaa450e91e8a20642b38a47b1c1c638&keyno=0&l10n=ru&cts=1450587773490&mc=5.24638133345	(2015-12-20) http://infobank.kit.ac.kr/yy/1.html

Figure V.3 Examples of Landing and Distribution URLs[21]

The malicious landing URLs in Figure V.3, when compared to the legitimate URL example in Part IV. v., shows no apparent distinction about the malicious nature: 'https://www.google.com/search?q=home+affairs+cyber+strategy&rlz=1C1CHBF_en-GBAU820AU820&oq=home+affairs+cyber+strategy&aqs=chrome..69i57j0j69i60l3.1837j0j7&sourceid=chrome&ie=UTF-8'.

The alphanumeric structure of the two will make little difference to the ordinary Australian citizen and resident. An expert cybersecurity analyst will have to conduct a comprehensive evaluation of the unknown URL to characterise it as malicious.

ii. Distribution URLs contain an exploit toolkit. These URLs differ from landing URLs in terms of the length and lexical format. An exploit toolkit is the second most common type of attack used.[22]

Another tactic used by attackers with distribution URLs, is that distribution URLs create random queries, incomprehensibly changing the pathname. It also renames the file-name in ways that closely resemble the original file. Specific URL segments are also replaced.[23]

Distribution URLs runs on a web service. The kit holds an assortment of already known exploits. Visitors to the website, when they are browsing the internet are then infected with malware. The individual may receive a phishing email with a link to the malicious website. A false advert may also appear on a website visited by the individual. This advertisement is referred to as 'malvertisement'. The individual may also be re-directed from the legitimate website they were visiting to a fake website. An events website could be used to infect visitors to the site.[24]

[21] Sungjin Kim, Jinkook Kim, Brent ByungHoon Kang, pg. 805.
[22] Australian Cyber Security Centre (ACSC) Threat Report 2017. Pg. 27. https://www.cyber.gov.au/sites/default/files/2019-03/ACSC_Threat_Report_2017.pdf
[23] Fn 21 Pg. 794
[24] Australian Cyber Security Centre (ACSC) Threat Report 2017. Pg. 58.

The dynamic nature of this exploit is that the person does not need to download a file to install the malware. A simple visit to the fake website will do the job of infecting the computer of the individual to start harvesting their personal and sensitive information,.[25] This was the case with the Mazar BOT malware that spread through unsolicited SMSs and pop-up downloads on some websites.[26]

The following was an example of an SMS send to an Android phone with a malicious link:

> You have received a multimedia message from + [country code] [sender number] Follow the link http://www.mmsforyou.net/mms.apk to view the message.[27]

Once installed, the phone was in the total control of the BOT, to make calls and send messages. This is done by means of MITM (man-in-the-middle) attack. This tactic can also be used to collect other information related to their job, and to try and gain access to confidential servers the individual has access to. Several hidden capabilities of Mazar are still being discovered.[28] These capabilities may pose threats to national security interests which may not yet be known. Imagine a senior public official's Android device being attacked in this manner to spread fake news during political campaigns, leading to reputational damage. Once the truth is discovered the damage may have been done and public trust would have eroded in the political system. Senior US government and military officials were already targeted by the May 2019 hack of WhatsApp using the Pegasus tool, and used malware delivery methods such as spearphishing messages containing links to malicious code.[29] This was revealed in the lawsuit filed by WhatsApp and Facebook on the 29th of October 2019.[30] There is therefore no telling what the full capability of an exploit is until it is researched properly in a safe environment. Academia can help in this regard. Democratic systems and infrastructure are all at risk. It is for the envisaged strategy to prioritise and rank the risks in order of national interests.

iii. URLs that are generated by malware after drive-by downloads.[31]
Drive-by downloads are downloads of malware by the user, without the intention or the knowledge to download malware. The person may be clicking on a malicious URL without

[25] Australian Cyber Security Centre (ACSC) Threat Report 2017. Pg. 27
[26] Ibid Pg. 29
[27] Symantec Corporation NORTON. 2019. Mazar BOT malware invades and erases Android devices. https://us.norton.com/internetsecurity-emerging-threats-mazar-bot-malware-invades-and-erases-android-devices.html
[28] Ibid
[29] Reuters. Christopher Bing, Raphael Satter NOVEMBER 1, 2019 / 3:04 AM / UPDATED 4 HOURS AGO Exclusive: Government officials around the globe targeted for hacking through WhatsApp – sources. https://www.reuters.com/article/us-facebook-cyber-whatsapp-nsogroup-idUSKBN1XA27H
[30] WHATSAPP INC., and FACEBOOK, INC., v. NSO GROUP TECHNOLOGIES LIMITED and Q CYBER TECHNOLOGIES LIMITED, Case 3:19-cv-07123. https://context- cdn.washingtonpost.com/notes/prod/default/documents/bf5edf35-5672-49fa-aca1-edefadff683f/note/8ef25c0d-fee9-416a-b7f9-e0a4dedc66f2.pdf#page=1
[31] Fn 21 Pg. 792

knowing that it contains malware, which is then downloaded, as a result of clicking on the hyperlink. and include downloads that are automated, without the user having to click on the link.

D. Public Education: Creating a Cyber-aware Community

As recognised by the Australia's 2020 Cyber Security Strategy discussion document, human behaviour is taken advantage of in committing cybercrimes. Phishing emails preys on the ignorance of the victim.[32] This is true when it comes to discerning a legitimate URL from a manipulated one.

Given the complex structure and features of URLs, and the ways in which URLs are manipulated, the general public may not be able to competently distinguish malicious URLs from legitimate URLs. This gap in knowledge is what attackers rely on. It is only with cybersecurity experts learning about these features and analysing the habitual and tactical behaviour of attackers to manipulate or to re-write the URLs, as was done by the researchers from the 'Graduate School of Information Security, School of Computing, Korea Advanced Institute of Science and Technology'.[33] The researchers then developed a similarity matching technique to better detect malicious URLs. The tool can be used at pre-processing and as a web filter.[34]

The general public can be educated about the tree types of malicious URLs, their architectural structure, how to try and identify same and to avoid them altogether. The public can be educated about the various features of malicious URLs as illustrated in Figure V.1.

The general message should always be not to click on any links and not to copy and paste URL links into browsers. This is however easier said than done.

E. Chicken or egg dilemma

The challenge in the timely and effective detection and prevention of the threats posed by malicious URLs may rest on whether URLs are legally classified as content or as non-content, under Australian government policy and law. Malicious URLs are by their very nature not created by the web queries of ordinary residents and citizens, but by malicious actors. A malicious URL may look innocent, especially a landing URL. As such it is difficult to tell from just the format, the features and length that it has the properties of a malicious URL and can therefore be suspected of being malicious. To investigate the URL link, it must be clicked on, copied and pasted into a search bar and the malware installed or downloaded to obtain the evidence about the malicious nature of the URL. The similarity matching technique is a tool that can be used at pre-processing and as a web filter.[35]

These malicious properties can be classified as follows:

[32] Commonwealth of Australia 2019. Australia's 2020 Cyber Security Strategy. A call for views. Pg.16 <https://www.homeaffairs.gov.au/reports-and-pubs/files/cyber-security-strategy-2020-discussion-paper.pdf>
[33] Fn 21 Pg. 792
[34] Ibid
[35] ibid

Table 2 – Properties of malicious URLs.		
Features	Description	Section
Alexa-based properties	Adversaries mainly use subdomains in Alexa top-ranked domains attack. High-ranked domains for landing URLs and relatively lower-ranked sites for distribution URLs are used.	4.1.1
Geographical trends	Between landing URLs and distribution URLs show high geolocation difference.	4.1.2
Code changes in EKs	The type of some EKs can be predicted via the frequency of code changes.	4.2
Landing URL	Landing URLs in terms of length and path depth resemble benign URLs.	4.3.1
Serial IP zone	New malicious IPs are often found in /16 and /24 prefixes around malicious IPs.	4.3.2
Pathname	Attackers reuse pathnames.	4.3.3
Filename	Typical filenames: very short, default and its variants, foreign-language-based, and random. Distribution URLs' size is shorter in general than that of landing URLs. Alphanumeric filenames are more popular than numeric filenames.	4.3.4

Figure V.4 Properties of Malicious URLs[36]

However, the unknown or suspected URL must first be collected from the Telco or web browsing history of the individual Australian citizen or resident, especially given their benign looks. It is only once the URL is collected and analysed that it may be revealed that it was legitimate or malicious. By that time the improper legal procedure may have been used to collect, access and use the URL for a cyber incident inquiry or investigation. The privacy rights of the individual may already have been breached by then. The question then is whether, given how deceiving URLs can be, and given that URLs may contain content in the authority, path and query parts, whether the URL should not be legally classified as content and be subject to the preservation and warrant processes under the *TIA Act 1979*. The URLs must therefore first be collected and analysed for their malicious properties. Moreover, malicious URLs are sent, addressed to the target, as messages inside phishing emails and SMS, making them the contents of a communication.

F. The methodologies for investigating, detecting and blocking suspected malicious URLs

Malicious URLs are considered a common and serious threat to cybersecurity. As such, it is urgently important to detect malicious URLs and to identify their attack type in order to prevent such attacks and to implement the required and effective countermeasures.[37] It is therefore advisable that a future cyber security strategy tackles this challenge head-on by specifically mentioning the threat posed by malicious URLs and propose policy, legal and technological strategies to address same.

To investigate and inquire into suspected malicious URLs, researchers should collect unknown URLs including spamming, phishing, malware and advance persistent threat (APT) URLs from various

[36] Ibid pg. 794

[37] ShymalaGowri Selvaganapathy, Mathappan Nivaashini & HemaPriya Natarajan (2018) Deep belief network based detection and categorization of malicious URLs, Information Security Journal: A Global Perspective, 27:3, 145-161, Pg. 145, DOI: 10.1080/19393555.2018.1456577; Sahoo, D., Liu, C., & Hoi, S. C. H. (2019). Malicious URL detection using machine learning: A survey. Ithaca: Cornell University Library, Pg. 1, arXiv.org. Retrieved from http://ez.library.latrobe.edu.au/login?url=https://search-proquest-com.ez.library.latrobe.edu.au/docview/2075353706?accountid=12001; Sirageldin A., Baharudin B.B., Jung L.T. (2014) Malicious Web Page Detection: A Machine Learning Approach. In: Jeong H., S. Obaidat M., Yen N., Park J. (eds) Advances in Computer Science and its Applications. Lecture Notes in Electrical Engineering, vol 279. Springer, Berlin, Heidelberg; RongWangYanZhuJiefanTanBinbinZhou Detection of malicious web pages based on hybrid analysis. Journal of Information Security and Applications Volume 35, August 2017, Pages 68-74;

repositories and datasets. URLs must be fed into a deep belief network (DBN) to abstract benign or malicious features to detect malicious URLs and identify attack types.[38] Effective countermeasures can then be developed.

Machine learning techniques and blacklist have all been used to detect and prevent malicious URLs.[39] Researchers from the 'Graduate School of Information Security, School of Computing, Korea Advanced Institute of Science and Technology' described the failure of domain blacklisting - it leads to the blocking of whole domains and access to benign websites are also accidentally blocked.[40]

In 2014 ASIC issued section 313(3) requests under the *Telecommunications Act 1997* for an IP-based block. The result was that it wrongly led to simultaneously blocking 250,000 unrelated websites, that all used the same IP address.[41]

Also, relying on published blacklist leaves organisations vulnerable, as new domains are constantly being registered.[42] As such a better way of detecting and blocking URLs that are suspected as being malicious is required. The similarity matching technique can also be used at pre-processing and as a web filter in this regard.[43]

G. Cooperation between the Agencies and ASD

The AFP cooperates with ASIO on investigations.[44] ASIO in turn co-operates with intelligence and law enforcement agencies such as the ASD in connection with the performance of their functions.[45] Hence, in order to comply with cybersecurity requirements under the laws, and in order to inquire into and investigate cybercrimes, under existing laws and any proposed national strategy and Federal laws, the Agencies may in future adopt research methods similar to the ones described in Part V. F., above, to detect malicious URLs and identify attack types, in the real world, if they have not done so already.

This submission proposes that the Agencies continue with their inquiry and investigative powers, under the *TIA Act 1979*, under the cybersecurity strategy as well. The current framework allows for the Agencies to cooperate with the ASD as the lead cybersecurity organisation. In this cooperation the Agencies can be the ones interfacing with the public and private businesses and individuals, using existing processes, such as warrants and preservation notices to collect suspected malicious URLs. The Agencies are currently legally empowered to share such collected information with the ASD to analyse and provide their assistance to the Agencies. In this manner the ASD will continue to

[38] Fn 37 ShymalaGowri Selvaganapathy, Mathappan Nivaashini & HemaPriya Natarajan (2018).
[39] Sahoo, D., Liu, C., & Hoi, S. C. H. (2019). Malicious URL detection using machine learning: A survey. Ithaca: Cornell University Library, arXiv.org. Retrieved from http://ez.library.latrobe.edu.au/login?url=https://search-proquest-com.ez.library.latrobe.edu.au/docview/2075353706?accountid=12001
[40] Fn 21 Pg. 792
[41] Rohan Pearce (Computerworld). New ASIC guidelines for web blocking awaiting legal sign-off. 30 May, 2018 11:45. https://www.computerworld.com.au/article/641732/new-asic-guidelines-web-blocking-awaiting-legal-sign-off/
[42] Fireeye. Email Threat Report – Q1 2019. Pg. 14. https://www.fireeye.com/offers/rpt-email-threat.html
[43] Fn 40
[44] *AFP Act 1979* (Cth) s 8(1) (bf)(ii).
[45] *ASIO Act 1979* (Cth) ss 17, 19A

serve as the central hub for national cyber threat intelligence. The ASD can use the information for national cyber situational awareness, and to advise private business, citizens and residents.

VI. AFP Accessing URLs

The *Cybercrime Act 2001* (Cth) and the *Crimes Legislation Amendment (Telecommunications Offences and Other Measures) Act* 2004 (Cth) amend the *Criminal Code Act 1995* (Cth) to criminalise cybercrimes. In order to inquire into and investigate cyber-crimes, we can reasonably assume the AFP collected and analysed URLs obtained from the Telco. The Agencies may employ strategies similar to the deep belief network (DBN) to collect and analyse benign-looking URLs to detect the attack types, and to inquire into and investigate cyber-crimes and cyber incidents. The resulting effect may be the blocking of the malicious domain and website.

In 2014, the AFP issued notices under section 313 of the *Telecommunications Act 1997* to collect:

> …to prevent the distribution of peer-to-peer malicious software (malware) which was designed to steal personal banking and financial credentials from the platforms of Australian computer users.[46]

The power was used to block known malicious domains. The AFP commented on its use of this power as follows:

> Section 313 provides the AFP with an effective means to disrupt illegal online activity where other mechanisms to prevent the activity have been or are unlikely to be successful. The AFP considers its use of Section 313 to block internet content has been reasonable and proportionate to the threat of the criminal activity.[47]

Between 2015 and 2018 the AFP made 3821 requests for historical data to investigate and inquire into cybercrime and telecommunications offences.[48] During the same period, 3984 requests were made for fraud, deception and related offences. These requests would have been made for both 'retained data' under the mandatory data retention regime and for the collection of information, such as URLs that are not required to be retained.

Given that URLs are not required to be retained under the mandatory data retention regime of 2015, and the indications that URLs may be treated and disclosed as metadata, as telecommunications data, as non-content data, section 313 of the of the *Telecommunications Act 1997* may still be used to access URLs, but along with notifications and authorisations issued under the *CAC Determination* 2015 and the *TIA Act 1979*.

The Communications and Media Authority (ACMA) reported no such requests were made by either ASIC or the AFP during the 2017 nor the 2018 financial years to request assistance to block malicious online content.[49]

---

[46] AFP. Inquiry into the use of subsection 313(3) of the Telecommunications Act 1997 by government agencies to disrupt the operation of illegal online services Submission 20. Pg. 3

[47] Ibid Pg. 4

[48] Australian Federal Police. Review of the mandatory data retention regime Submission 15. https://www.aph.gov.au/DocumentStore.ashx?id=5e9dc811-bb16-4f71-bb1d-2e124354c530&subId=668079

[49] ACMA and the Office annual reports 2017–18, pg. 190; Australian Communications and Media Authority annual report 2018–19. Pg. 130. https://www.acma.gov.au/-/media/mediacomms/Report/pdf/ACMA-and-eSafety-annual-reports-2018-19-pdf.pdf?la=en

A. URLs disclosed on rare instances only to the Agencies

In the court case: *Privacy Commissioner v Telstra Corporation Limited [2017] FCAFC 4 (19 January 2017),* on the 15th of June 2013, Mr. Grub requested access to URLs retained by Telstra that were disclosed to the Agencies. He claimed the URLs were personal information, and were metadata or telecommunications data:

> The metadata would likely include… URLs [Uniform Resource Locators] of websites I have visited…[50]

Mr. Grubb did not claim that the URLs were the contents or substance of a communication. In both 2013 and in 2015, Telstra confirmed that it discloses URLs to the AFP, although this is only done in extremely rare instances:

> (2) Except in extremely rare instances, the Law Enforcement Liaison group does not give law enforcement agencies with … or Uniform Resource Locators (URLs) involved in mobile data communications.[51]

These rare instances were not described. It does not appear as if the disclosures referred to were under a preservation notice or a warrant process. Mr. Grubb's case was about the disclosure of metadata without a warrant process being used. The point is that the Telco does in fact disclose URLs to the Agencies, even though it may be only in rare instances. The question is: what legal process was used to make the disclosures that were made.

Given the widespread nature of malicious URLs, even in 2013, one can reasonably assume that some of the URLs Mr. Grubb requested may have been malicious. Given the rise in URL-based cyber risks since the end of 2018, one can assume that such requests for URL disclosures may increase, beyond the rare instances referred to by Telstra and the few instances referred to by the AFP in 2014.

B. Department of Home Affairs may have implied web browsing history is not content

The Department of Home Affairs confirmed to the PJCIS that content is not required to be retained under the mandatory data retention regime:

> The legislation requires providers to retain the details of a communication, without capturing its content. In addition to content data, other datasets are explicitly ruled out of the regime, such as a subscriber's web browsing history.[52]

The Home Affairs Department did not address the issue of the disclosure of URLs to the Agencies in this statement to the PJCIS. This may be because the review by the PJCIS was focussed on the data retention regime and did not cover the disclosure of the voluntarily retained or the disclosure of the mandatorily retained data.

---

[50] [8.]
[51] Telstra Corporation Limited and Privacy Commissioner [2015] AATA 991 (18 December 2015) [63.] – [66.]
[52] Parliamentary Joint Committee on Intelligence and Security Review of the mandatory data retention regime Home Affairs Portfolio submission. Review of the mandatory data retention regime Submission 21 [39.]

With the above quote, the Department of Home Affairs may have implied that web browsing history is not the contents of a communication.[53] The questions that are raised and that may need to be addressed in any future cybersecurity strategy, that hopes to effectively address malicious URLs as a threat vector, are these:

i. Even if web browsing history is not required to be retained under the mandatory data retention regime, is web browsing history not required to be retained because it is the contents or substance of a communication?
ii. Is web browsing history required to be disclosed to the Agencies under the authorisation and notification process of the *CAC Determination* 2015, and not under the preservation notice and warrant process of the *TIA Act 1979*?
iii. Even if the web browsing history is not required to be retained, if the Telco has the web browsing history in its possession, is the Telco required by law to disclose the web browsing history to the Agencies under the authorisation and notification process of the *CAC Determination* 2015, and not under the preservation notice and warrant process of the *TIA Act 1979*?

C. The Commonwealth Ombudsman concerns about the AFP accessing URLs as metadata

In its submission to the PJCIS, the Commonwealth Ombudsman referred to web browsing history as '… ***the Uniform Resource Locators (URLs) they have searched.***' The Commonwealth Ombudsman raised concern that it is not clear whether URLs collected by the AFP is classified as content under the *TIA Act 1979*. Despite the above statement by the Department of Home Affairs that content is not required to be retained and the implication that web browsing history is non-content, but is not telecommunications data, the Commonwealth Ombudsman appears to be contradicting the Department of Home Affairs by implying that web browsing history may be the contents of a communication. The Commonwealth Ombudsman implied that the Telco may be disclosing content simultaneously with the URLs it discloses to the Agencies.

Despite the statements from Telstra that it does not disclose the content parts of the URLs, it seems that other Telco's may be disclosing the content parts of the URL as well. If different Telco's are complying to different standards, there needs to be clarity on the exact compliance standard. This needs to be addressed in any future cybersecurity strategy that is envisaged.

The Commonwealth Ombudsman requested clarity in this regard - whether URLs, which the Commonwealth Ombudsman refers to as '… ***the Uniform Resource Locators (URLs) they have searched.***' should be legally defined as the contents of a communication. If URLs are classified as content, then the process set out in Part VIII must be strictly followed. The submission did not specifically address malicious URLs. It can however be reasonably assumed the AFP does request access to URLs, some which may be malicious, to investigate and inquire into cybercrimes. In its submission to the PJCIS the Commonwealth Ombudsman implied the AFP may not always follow the requirements under section 172 of the *TIA Act 1979.* Section 172 of the *TIA Act 1979* prohibits the disclosure of content to the AFP under the authorisation and notification process of the *CAC Determination* 2015. This

[53] ibid

means content must be disclosed as described in Part VIII, with preservation notices and with warrants. The Commonwealth Ombudsman stated:

> 'Determining what constitutes 'content'
>
> Under s 172 of the TIA Act, an authorisation does not permit the disclosure to an agency of information that is the content or substance of a communication, or a document that contains the content or substance of a communication. The term 'content or substance of a communication' is not defined in the TIA Act.
>
> For the majority of telecommunications data we inspect, our Office is able to determine whether the disclosed information breaches the restriction in s 172. For example, the telecommunications data of a phone call can include the date, time and location(s) of the call but cannot include the substance of what the parties said.
>
> ***For other data types, it has been difficult for our Office to make this determination. For example, we have identified instances where carriers have returned telecommunications data that has included a person's Internet Protocol (IP) address and the Uniform Resource Locators (URLs) they have searched.***
>
> …
>
> It is unclear whether such information breaches the restriction under s 172 and depends on a case by-case assessment. Where we consider that information does amount to content, we notify the agency and suggest it:
>
> - quarantine that information from any further use, and
> - where relevant, seek legal advice on any use of the information to date.
>
> We follow up at our next inspection to confirm what, if any, remedial action the agency has taken. Clarity on what constitutes 'content' would likely assist:
>
> - carriers in determining what information they can provide to agencies under an authorisation
> - agencies in identifying when they may have received content from a carrier so that they can take immediate steps to limit any use of that information, and
> - our Office in assessing whether the telecommunications data accessed by an agency complies with the restriction under s 172.
>
> **The Committee may wish to consider whether the Act should be amended to include a definition of the term 'content or substance of a communication or document**' (emphasis added).[54]

The quote from the Commonwealth Ombudsman seem to agree with the indication from Telstra that parts of the URL may be considered as content in on a case by case basis, and that those parts

[54] Submission by the Commonwealth Ombudsman, Michael Manthorpe July 2019. Review of the mandatory data retention regime Submission 20. https://www.aph.gov.au/DocumentStore.ashx?id=4e65d856-df71-462c-a807-5c54d8a13da8&subId=668159

may be severed from the non-content parts of the URL. As argued by Bellovin et al. (2106) in Part VII. B. 4. and C., the URL may contain both content parts and non-content parts.[55]

The phrase '… ***the Uniform Resource Locators (URLs) they have searched***' may refer to the web browsing history. This would be the websites, or the URLs links clicked by the person. The person may have navigated to malicious URLs that may be responsible for a cyber incident or cyber-attack that the Agencies are inquiring into or investigating.

Despite the legal requirements listed in Part VIII, from the concerns expressed by the Commonwealth Ombudsman, it would seem as if in some instances, some web browsing histories may have been treated as non-content. The Commonwealth Ombudsman implied that URLs were not disclosed to the AFP under a Stored Communications Warrant (SCW), an Interception Warrant (IW), or a Computer Access Warrant (CAW). Instead, URLs appear to have been disclosed under authorisations and notifications and not under the section 172 of the *TIA Act 1972* and therefore under the *CAC Determination* 2015.[56]

Given the additional safeguards the Commonwealth Ombudsman introduced to protect the URLs it considered as content, it appears as if the Commonwealth Ombudsman would want URLs to be treated as content.

VII. Are URLs content or non-content under Australian law?

Australian law, similarly to American and British law, categorises evidence related to information and communications in two main parts. The evidence is either the contents of a communication or the non-contents of a communication. This distinction was simple to apply to traditional fixed-line telephony. The time of the call is non-content but the conversation between the two people are the contents of a communication. As such, if law enforcement wanted to listen in on a conversation or obtain a recording of the conversation, a warrant, or a wiretap order was obtained from a judge. To collect the time of the call the telecommunications company was simply required to hand that over when the police requested for it. This general legal process was first implemented in 1975 in respect of ASIO.[57]

As it now stands, under the *TIA Act 1975*, the same legal distinction between content versus non-content is upheld. There seems to be no legally recognised third category. However, with the introduction of the assistance and access framework, under the *Assistance and Access Act 2018*, the Telco and SMPCs are required to disclose 'technical information' to the Agencies.[58] The Agencies will still apply for search warrants, computer access warrants and interception warrants to access the content of communications. This still fits in with the traditional distinction between content and no-content. However, regarding the legal duty to disclose 'technical information', there is no legal

[55] Bellovin, Steven M., Matt Blaze, Susan Landau, Stephanie K. Pell, 'It's Too Complicated: The Technological Implications of IP-Based Communications on Content/Non-Content Distinctions and the Third Party Doctrine,' (2016) 30(1) Harvard Journal of Law & Technology, pg. 69.
[56] Fn 54
[57] (Shanapinda, S.: Advance metadata fair: The retention and disclosure of location information as metadata for law enforcement and national security, and the impact on privacy – An Australian story. Dissertation, UNSW Sydney (2018) (unpublished)) Pg. 97
[58] Section 317E(1)(b)

definition for the term 'technical information'. It is therefore not clear what 'technical information' entails. As such, no analysis can be made whether 'technical information' may be content or non-content. This may have the impact of bypassing the content versus non-content dichotomy.

A. Legal Definitions: Content versus Non-content

The legal phrase used under the *TIA Act 1975* to define the term 'content', for the purposes of this submission is: 'information that is the contents or substance of a communication'.[59]

The legal phrase used to define 'non-content', the metadata or telecommunications data is: '... any information or document that ... relates to ... the contents or substance of a communication'.[60]

For the purposes of this submission the phrase 'technical information' is placed in a category of its own.

As illustrated in Figure VII.1, content may not be collected by using any of the other two processes. The preservation notice and warrant must be used to collect the content or substance of a communication. The strategic question posed by this submission is in which category URLs fall, it seems that it is both.

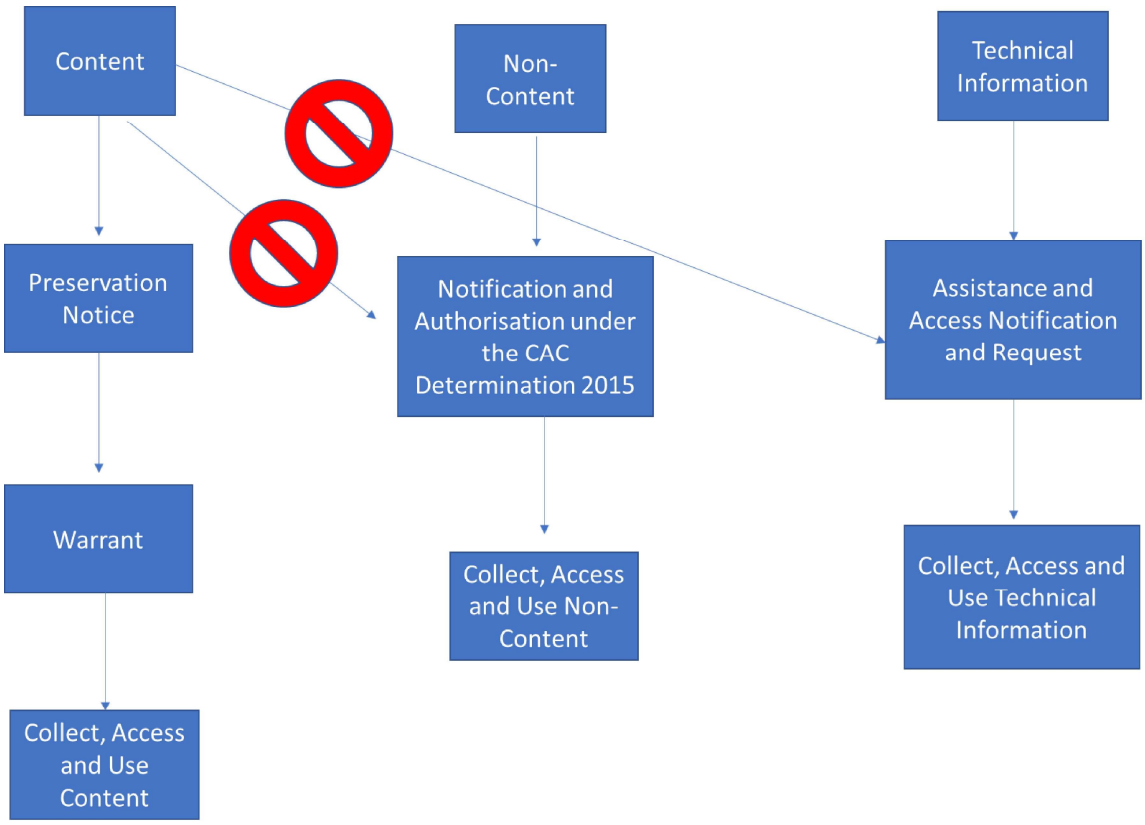

Figure VII.1 The distinction between content and non-content

[59] *TIA Act 1979* s 172(a); Explanatory Memorandum, Telecommunications (Interception and Access) Amendment (Data Retention) Bill 2015, xxxvi.
[60] *TA 1997* s 276, 280, 313(7)(d), (e) (TA 1997).

The 'content' of a communication is not able to be disclosed under the *CAC Determination* 2015.'[61]

B. The definition of 'content' under Australian law

Explanatory memoranda from proposed Bills may be used as extrinsic material to try and interpret the laws. In 2007, Chapter 4 of the *TIA Act 1979* titled 'Access to telecommunications data' was introduced, by transferring the relevant parts of the *Telecommunications Act 1997* to the *TIA Act 1979.* The aim was to empower the Agencies to authorise the access and use of telecommunications data.[62]

The phrase 'content or substance of a communication' were not defined at the time and has remained undefined since. The word 'communication' was however defined, and it may shine some light on what the word 'content' may mean. Table 1 below refers to two similar definitions of communication used under current laws.

Table 1 The definition of communication(s)

TIA Act 1979	***Telecommunications Act 1997***
communication includes ***conversation*** and a ***message***, and any part of a conversation or message, whether: (a) in the form of: (i) speech, music or other sounds; (ii) data; (iii) text; (iv) visual images, whether or not animated; or (v) signals; or (b) in any other form or in any combination of forms (emphasis added).[63]	***communications*** includes any communication: (a) whether between persons and persons, things and things or persons and things; and (b) whether in the form of speech, music or other sounds; and (c) whether in the form of data; and (d) whether in the form of text; and (e) whether in the form of visual images (animated or otherwise); and (f) whether in the form of signals; and (g) whether in any other form; and (h) whether in any combination of forms.[64]

The words 'conversation' and 'message' require further reading.

The Macquarie Dictionary gives 'conversation' its semantic or ordinary grammatical meaning as follows:

[61] *TIA Act* 1979 s 172; Communications Access Coordinator's (CAC) Telecommunications (Interception and Access) (Requirements for Authorisations, Notifications and Revocations) Determination 2015 (Cth) (at 9 October 2015) (CAC Determination 2015); Shanapinda (Dissertation) Pg. 216
[62] Explanatory Memorandum, Telecommunications (Interception and Access) Amendment Bill 2007 [Provisions] (Cth), 3.
[63] *TIA Act 1979* s 5(1) (definition of 'communication').
[64] *TA 1997* s 7 (definition of 'communication').

1. informal interchange of thoughts by spoken words; a talk or colloquy.

2. an instance of this.

3. …[65]

The Macquarie Dictionary gives 'message' its semantic or ordinary grammatical meaning as follows:

1. a communication, as of information, advice, direction, or the like, transmitted through a messenger or other agency.

2. …'[66]

The Macquarie Dictionary gives content its ordinary grammatical meaning, and incidentally refers to the example of a web page as being 'content':

'1. …

4. the information which is in a communication, as opposed to the format, design, etc.: the content of a web page.'[67]

These definitions will be incorporated in the discussions below.

1. Is the URL the content of a message?

When an Internet bot sends the malicious URL in an email or SMS message, this is machine to machine communication, using text, and the browser is fetching or retrieving the available information resource or content from the web server, whether malicious or not, and displays the resulting web page on the screen of the laptop or smart phone to be seen, heard or interacted with by the individual. So, is the retrieval of the web page and its delivery and presentation on the screen of the user's device a message with the visuals, sounds and images contained therein, the content or substance of a communication? It would appear so. As per the ordinary definition of the word 'content' from the Macquarie Dictionary, a web page is content. A web page contains a combination of speech, music, other sounds, data, text and visual images, whether they are animated or not.

The question is whether a URL or website address can be said to be conversation or a message, or any part of a conversation or a message, communicated as content to the individual? With URL-based HTTPS phishing, the individual is sent an emailed malicious URL link. The user interacts with the URL by clicking on it, copying and pasting it and being directed to the website it intends to visit using the emailed URL link, or is re-directed to another website, whether legitimate or malicious. In this manner the URL is also the contents of a communication sent as a message, as illustrated in Figure IV.1, intended to be opened by the user as its intended recipient.

---

[65] Macquarie Dictionary. https://www.macquariedictionary.com.au/features/word/search/?search_word_type=Dictionary&word=conversation

[66] ibid https://www.macquariedictionary.com.au/features/word/search/?search_word_type=Dictionary&word=message

[67] ibid https://www.macquariedictionary.com.au/features/word/search/?search_word_type=Dictionary&word=content

The message can be transmitted using an instrument, which is denoted by the use of the words 'other agency'. The Internet bot that is generating the unsolicited emails and unsolicited SMS messages is the 'other agency' is sending these messages, which messages include the web address as its content, on its way to be delivered to the individuals device, so that the individual receives, as the contents of the message, the malicious URL and click on it, to be defrauded. This is the aim of malicious websites, sending fake URLs as messages, intended to be received and read and clicked on to download malware and steal the personal details of the individual. Without the individual receiving this fake URL and clicking on it malicious websites will not operate. The web page is delivered in that message to individual who sees it and reacts to it by clicking on it.

The email is evidence of the alleged cybercrime. As such, the rules that apply to collecting evidence to investigate a crime would apply. In this instance, under the *TIA 1979*, if the email is considered as content, or the parts of it that are, a warrant would be required, which is preceded by a preservation notice. If the email is regarded as telecommunications data a notification and authorisation under the *TIA 1979* would be issued to collect the email. To intercept email communications, in real-time or near-real time, from private companies and private citizens, on the basis of national security interests, existing laws about how content is accessed will apply, under the *TIA Act 1979*. Given that email is the contents of a communication, even if it may originate from an Internet bot, a non-Australian resident or citizen, but is messaged to an Australian citizen or resident as the intended recipient of the message, with the URL contained as its contents, the expected warrant process under the *TIA Act 1979* must also apply. According to Bellovin, Blaze, Landau and Pell (2016), according to US law, emails may be regarded as content and a warrant would be needed to access same.[68]

2. Details of Internet sessions as content

In the 2007 Explanatory memorandum, regarding web browsing, the term 'content' was taken to refer to '… the details of Internet sessions'. The term 'content' however does not refer to '… the sender's and recipient/s' Internet addresses…' such as the IP address of the users device or the IP address of the domain:

> For ***Internet based*** telecommunications, such as …, ***web browsing***, … ***[telecommunications] data*** includes the *sender's* and *recipient*/s' ***Internet addresses***,... The information *does not include content* such as … the ***details of Internet sessions*** (emphasis added).[69]

The quote above implies two things:

i. As regards web browsing, 'Internet addresses' are regarded as telecommunications data, or as non-content; and
ii. Also as regards web browsing, the 'details of Internet sessions' are regarded as its content.

Given that the 'details of Internet sessions' are regarded as content the obvious question is what 'details of Internet sessions' are. The details would typically include messages exchanged between the client and the web server. The client is the recipient, which is the web browser like Google

[68] Bellovin, et al. (2016) 30(1).
[69] Explanatory Memorandum, Telecommunications (Interception and Access) Amendment Bill 2007 [Provisions] (Cth), 8.

Chrome, installed on the persons laptop, for example. The web server is the sender. The client is given an ID for the session to know its identity, and for a period of time it is still connected and authenticated to access the web server. The web server is able to remember earlier request from the same client by using cookies to serve as its memory. This cookie is sent between the client and the server, and contains details of its earlier messages, using the ID issued to identify it when the connection between the two were established. A good example of this is when shopping online. All the items selected and put in the online shopping basket are stored using cookies, for the duration of the online shopping. When the shopper closes the connection and returns to continue shopping, they can start where they have left off.[70]

Internet sessions may be referred to as HTTP sessions. To retrieve the web page, two common methods are used. They are referred to as GET and POST methods. The difference between the two are distinct. With the GET method, the query data sent is part of the URL. With the POST method the parameters are not stored in the browser history or in the web server logs. Google for example uses the GET method. The POST method is popular with email usage, where the query data is kept in the body of the message.[71] The GET command includes query information in the URL, so if the 'details of Internet sessions' are regarded as the contents or substance of a communication, it would seem that query information in the URL may only be legally collected, accessed and used as described in Part VIII, with preservation notices and warrants.[72]

3. Web Browsing history as 'other datasets'

However, in 2019, the Home Affairs Department implied that web browsing history may not be the contents of a communication:

> The legislation requires providers to retain the details of a communication, without capturing its content. In addition to content data, other datasets are explicitly ruled out of the regime, such as a subscriber's web browsing history.[73]

Web browsing history is not required to be retained. However, as indicated by the Commonwealth Ombudsman it appears, if the information is available to the Telco, the AFP does request for access to the information and the Telco does disclose it to the AFP, sometimes with content and at other times without. Given the aforementioned, the above quote raises several questions. The more important one is: is web browsing history classified as 'other datasets', as opposed to content or non-content? Does that mean that the traditional distinction between content and non-content as illustrated in Figure VII.1 is further expanded with a fourth category to be known as 'other datasets'?

Does web browsing history fall into a separate and third category of 'other datasets' which category is not found in the *TIA Act 1979*? According to the *TIA Act 1979* there are only two categories, so under which of the two categories would web browsing history fall?

[70] Young B. Choi, Yin L. Loo, Kenneth LaCroix. Cookies and Sessions: A Study of what they are, how they can be Stolen and a Discussion on Security (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 1, 2019, pp. 32-36

[71] Bellovin et al. (2016), pg. 67-68

[72] This submission will not delve into the details of this issue, but simply focus on the issue of malicious URLs.

[73] Parliamentary Joint Committee on Intelligence and Security Review of the mandatory data retention regime Home Affairs Portfolio submission. Review of the mandatory data retention regime Submission 21 [39.]

The alternative question is, 'other datasets' are differentiated from 'content data', so does this mean that web browsing history is considered as non-content data, as telecommunications data, as metadata, by official policy? In other words, web browsing history is not excluded from the mandatory data retention regime because it is legally considered as the contents or substance of a communication? It seems, the fact that the web browsing history is excluded from the legal obligation to retain does not mean web browsing history is the contents or substance of a communication.

4. Analysing the architectural structure of the URL to determine its nature as content or non-content

The URL is commonly referred to as a web address, that refers to a web resource. The address indicates where the resource is located. The web page is the resource that is located on a given web server, identified by a domain. However, despite this, the URL has a dual nature. The query part of the URL, as indicated in Part IV, may be describing the some of the contents of the communication, i.e. the web page. In Figure VII.2, the malicious URL also contains the path and query as its content.

Table 1 – Feature sets and features.

Feature Set	Feature Name	Malicious URL Examples
Host-based features	similar domain	wsad004.asia, wsad066.asia, wsad123.asia
	similar subdomain	webhosting50.1blu.de, webhosting52.1blu.de
	similar IP prefix	104.149.195.115, 104.149.195.116
	similar port	174.139.26.238:801, 174.139.30.222:802
	same domain	couponmoeum.com/PEG/css/1.js, couponmoeum.com/PEG/css/ad.html
	same subdomain	aa.sswangima.com, aa.wangma1q.com
	same IP	175.126.74.101/files/2.js, 175.126.74.101/files/1.js
	same port	2288.org:8832, 6600.org:8832, 8800.org:8832, 8866.org:8832
Path-based features	similar pathname	sbsjob.co.kr/PEG/js/check_38746.js, dowmi.net/PEG/ad/check_38746.js
	same pathname	braxico.com/images/m2dcbAnA.php, kappen-orth.de/images/iRsGKfv2.php
Filename-based features	similar query type	rcjhvqmtkmp?bdhkuyjqwmpi=6621548, rckjyihjpoggki?bxvuwwsceqt=6621548
	similar filename	199.188.107.109/yy.html, 199.188.107.109/zz.html
	same query type	175.45.4.158/index.php?id=kim031, 175.45.4.158/index.php?id=kim032
	same filename	199.188.107.112/xiaoyu.html, 198.2.221.203/xiaoyu.html

Figure VII.2 Examples of malicious URLs[74]

The URLs in Figure VII.2 and Part IV, contain query parts. According to Bellovin et al. (2016), the query part would generally be content. The malicious actor must have created the URL string and forwarded it via a phishing email to the individual Australian citizen or resident. The malicious URL is evidence of the wrongdoing, once that is confirmed by the detection methods used, such as the similarity matching technique, and the theft of data confirmed, or the fraud committed. Prior to that it may appear benign. In any other ordinary criminal investigation, the written content in an email, is the contents of a communication, that is evidence of fraud. This content evidence, whether it is known beforehand, or suspected, or unknown, may only be collected with a preservation notice and a warrant issued to the Telco to preserve and then disclose the email from the email server. If a suspected URL is content or the substance of a communication, the same procedures would need to be followed. However, as discussed in this submission, in various parts, it is not clear whether URLs or web browsing history is content, non-content or some other third type of data. The view of this

[74] Sungjin Kim, Jinkook Kim, Brent ByungHoon Kang pg. 791

submission is that URLs in general, and suspected malicious URLs are the contents of a communication.

Table A1 – Malicious URL examples.

Malicious IP Examples	Malicious Domain Examples
110.34.196.117/cake.php	www.zilphotography.com/.errordocs/FOZ4fFXK.php
110.34.196.125/cake.php	ropeholders.info/.errordocs/sQodNA2k.php
103.251.36.92/index.html	iambrucehan.com/.errordocs/75yYodGX.php
103.251.38.103/index.html	www.wigsislandstudio.com/.errordocs/GQsvQJzH.php
103.251.37.211/index.html	www.cleanenergyhi.com/.errordocs/6UhfUL3J.php
103.251.37.213/sb.html	ozactivity.com/.errordocs/wEXfiNFD.php
103.251.37.214/sb.html	perivaleproductions.com/.errordocs/zB0ypBk9.php
103.251.37.213/main.html	live-counter.net/?click=13950265
103.251.37.214/main.html	hosttracker.net/?click=6621593
103.251.37.213/index.html	hostverify.net/?click=1110828
103.251.37.214/index.html	webexperience13.com/?click=85009921
1.226.83.40/ts/index.html	thedeadpit.com/?click=341881
1.226.83.40/ts/dy.html	internetcountercheck.com/?click=13218787
113.10.187.41/oacs19/29.html	google-ana1yticz.com/?click=172078
113.10.187.42/oacs19/29.html	coaipr.org/aqgy.html?i=1958545
113.10.187.41/oacs19/man.html	petalconsultancy.info/aqgy.html?i=1958545
113.10.187.42/oacs19/man.html	lindsethcpas.com/aqgy.html?i=1958545
126.19.87.31/2222/tiancai.html	innerbath.com.au/crpy.html?j=1958545
126.19.87.31/2222/index.html	stevebeam.com/wrpy.html?i=1958545
116.81.235.128/3333/tiancai.html	ptsolutionsgroup.com/crgt.html?i=1958545
116.81.235.128/3333/index.html	morehead-motorsports.com/eqgy.html?i=1958545
126.114.226.40/3333/shifu.html	petalconsultancy.info/aqgy.html?i=1958545
126.114.226.40/3333/index.html	cronicadelcorrugado.com/mqpt.html?i=1958545
103.240.197.28/b.html	tvpasiones.com/arpy.html?i=1958545
103.240.197.30/c.html	abinnetsol.ca/eqgy.html?i=1958545
103.240.197.33/c.html	eastmead1.ipower.com/hrpt.html?i=1958545
103.240.197.35/k.html	wheresweems.com/argt.html?i=1958545
103.240.197.36/j.html	bestdeckshoes.com/oqpt.html?i=1958545
103.240.197.37/v.html	lindsethcpas.com/aqgy.html?i=1958545
113.10.187.41/live3/qq.html	burtcasey.net/wrpt.html?i=1958545
113.10.187.42/live3/qq.html	3diporn.com/eqgy.html?i=1958545
113.10.187.41/live2/qq.html	kirtidan.com/mqpt.html?i=1958545
113.10.187.42/live2/qq.html	budgetcancun.com/oqpy.html?i=1958545
113.10.187.41/code0002/qq.html	prmd.biz/wrpy.html?i=1958545

Figure VII.3 Examples of malicious URLs[75]

The question is whether any of the above examples of malicious URLs can be said to be the contents or substance of a communication. It may be challenging to make out what the content is by just looking at the URLs. It is only when one clicks on the URLs that the web page, as its content or substance of the communication is revealed. The URLs display query types, as displayed in the second last line, that are alphanumerical and reveals little about what the web page the individual may have clicked on contains. According to Kerr[76] and Bellovin et al. (2016), everything after the domain name, which is for example, '.com', which is 'path' portion of the URL, should be considered as content, under American law:

> '… path portion of the URL … it functions much like a file name on a web server. It therefore reveals communicative content because it describes what the user is requesting from a website'.[77]

The important note from the quote above is that based on how the internet works, irrespective of Australian or American law, is that the path portion of the URL describes what the user is requesting from a website. If the individual typed in the ordinary words into a search bar and these words

[75] Sungjin Kim, Jinkook Kim, Brent ByungHoon Kang pg. 804
[76] In re Google Cookie Placement Consumer Privacy Litigation, 806 F.3d 125, 139 (3d Cir. 2015)
[77] Bellovin et al. (2016), pg. 69.

appear in the URL then the URL contains communicative content in the query part. As such, under American law, it would be communicative content. This is discussed further in Part VIII B. 4. As there is no legal definition for content under Australian law, the in the ordinary meaning of the word, the content is what is in the communication. The malicious URL is what is in the communication, the communication being the phishing email and SMS. It may not be the words typed into the search bar, but it may be copied from the email and pasted into the search bar by the individual. Whether it is a malicious URL, when the individual clicks on the link, they are making a request, whether they have been duped into making that request for that web resource and did not intend to visit the malicious web page. Clicking on the link is the making of the request for the resource. That is in itself generally a communication of content, that is then transmitted to via the web browser to the web server and the response is the malicious web page.

Bellovin et al. (2016) consider the authority portion of the URL to be non-content under American law.[78] However, it gets complicated, because in some instances this portion may even be considered as content, based on the how the technology works:

> The authority portion of a URL, while generally non-content DRAS, can become architectural content in certain web hosting arrangements. If a single physical server hosts multiple web servers for different domains, the server uses the authority field that is sent to it as part of the HTTP request to determine which of its web servers should process the request. As we previously noted, in this hosting arrangement the authority acts both as non-content when it is translated to the server's IP address and used to establish network communication, and as architectural content when the original host name from the URL string is sent to the web server. When a single web server exclusively provides services to web sites owned by a single entity, there is no third party involved in serving the web page. In the case where a single web server is shared by different entities (as can be the case in commercial services), however, the operator of the server program must route the HTTP request to the appropriate web page. The particular hosting arrangement that determines whether a third party receives the authority portion of the URL is a decision made and implemented by the hosting service operator. The user does not make a voluntary conveyance of information to a third party, as the user cannot control or know if or when a third party will receive the information. Accordingly, in a web hosting arrangement where a single server provides services to web sites owned by multiple entities, a court cannot rely upon the third-party doctrine to determine the appropriate access standard when law enforcement compels the authority portion of a URL from a third party. The court would need to conduct a reasonable expectation of privacy analysis without the benefit of the third-party doctrine.[79]

The URL may be redirected by attackers for phishing attacks or malware distribution.[80] In that instance the person did not intend on landing at the malicious website. However, the person reacted to the contents of a message, which is the malicious URL link and clicked on. The person was clearly the intended recipient of the malicious link and requested the retrieved page, even if they were not aware of its malicious nature.

Bellovin et al. (2016), cautions as follows:

> (2) The Current Rules Distinguishing Content and Non-Content are Too Difficult to Apply.
> Understanding where the boundary is between metadata and content is specific to the situation and

[78] Bellovin et al. (2016), Pg. 71

[79] Pg. 71-72

[80] MitsuakiAkiyamaaTakeshiYagiaTakeshiYadaaTatsuyaMoribYoukiKadobayashic. 2017. Analyzing the ecosystem of malicious URL redirection through longitudinal observation from honeypots. Computers & Security Volume 69, August 2017, Pages 155-173 Computers & Security. https://doi.org/10.1016/j.cose.2017.01.003

> the communications protocol used. Simple guidelines such as "email addresses are metadata" are often misleading. A detailed understanding of the technical minutiae of Internet protocols is therefore required to begin the analysis. As we have seen in many cases (for example, URLs and service location ambiguity), it is necessary to do a deep analysis of the specific fact pattern of each desired interception to determine where the boundary may lie.[81]

It must be made clear that Kerr[82] and Bellovin et al. (2016), were only discussing URLs in general and under American privacy law under the Fourth Amendment of its Constitution, the Wiretap Act and the Pen/Trap statute, and not malicious URLs, the architecture and display of which differ and would be considered differently under Australian law. The URL, web address or web browsing history is not required to be retained under the *TIA Act 1979*. As such, it is not legally regarded as 'retained data' and is not deemed to be personal information. This submission does not delve into the question whether URLs are personal information, only whether URLs are the contents or substance of a communication.

Whereas for Bellovin et al. (2016), the discussion about content was tightly linked to the privacy question under the US Constitution, for the purposes of this submission a look at privacy and personal information protection is not as relevant to make a difference in how the information is accessed and used, as discussed in Part III. C. Notwithstanding the aforementioned, the discussion by Bellovin et al. (2016) about the technical nature of URLs and how the internet works is still instructive, and the legal aspects are good benchmarks for Australia.

a) The URL may contain parts that are considered content

Section 172 of the *TIA Act 1979* prohibits the disclosure of content under a notification and authorisation issued by the Agencies to themselves. Section 172 states:

> **172 No disclosure of the contents or substance of a communication**
>
> Divisions 3, 4 and 4A do not permit the disclosure of:
>
> (a) information that is the contents or substance of a communication; or
>
> (b) a document to the extent that the document contains the contents or substance of a communication.

The URLs were only disclosed by Telstra to the extent that they did not reveal the content of communications:

> "Any telecommunications data or meta data but not the content or substance of a communication. It may include:
>
> ...
>
> Internet Protocol (IP) addresses and **Uniform Resource Locators (URLs) to the extent that they do not identify the content of a communication**, and
>
> ..." (emphasis added)[83]

[81] Pg. 92
[82] In re Google Cookie Placement Consumer Privacy Litigation, 806 F.3d 125, 139 (3d Cir. 2015)
[83] Telstra Corporation Limited and Privacy Commissioner [2015] AATA 991 (18 December 2015) [64.]

The Telco retains the URL or web address and discloses the URL to the Agencies. It appears Telstra is implying that portions of the URL may contain the contents of a communication. Telstra indicated that it does not disclose web browsing history to the extent that it reveals content and its only in rare instances that web browsing history is disclosed. Those parts of the URL that reveal the content, such as the query portion must be omitted from any disclosures that are made to the Agencies. Although the URL is a web address that refers to a web page, which web page is the content, the URL link still may include traces of content which are indications of the contents of the web page. In other words, to the extent that the URL reveals content, the URL should not be disclosed to the Agencies without a preservation notice and a warrant.

However, matters may not be as simple as Telstra explained, given the explanation by Bellovin et al. (2016), and the introduction of the term 'architectural content' in Part VII. C below. In regard to malicious URLs, and noting that Telstra was not discussing malicious URLs, even if the content parts of a URL, such as path and query are removed, because they are content, if the URL was contained in a phishing email, in a SMS message on an Android phone, as was the case with the Mazar BOT, the URL link are the contents of that SMS message and the email. Only disclosing the authority part, which is the domain such as 'homeaffairs.gov.au', as illustrated in Part IV, as long as it was contained in an SMS message or a phishing email, the whole malicious URL is the contents or substance of a communication. Under these circumstances disclosing the web browsing history even under rare instances would require following the preservation notice process and the warrant process. If an individual send an email message to another, to commit fraud, which may be investigated by ASIC, ASIC would be looking to collect the contents of the email. The legal way to collect the contents of the email that proofs the fraud conspiracy would be to serve a preservation notice on the Telco and ISP, and to obtain a warrant, for the disclosure of the email contents. The same principle applies to a known or unknown malicious URLs contained in an SMS message and a phishing email, sent to the Australian citizen or resident.

C. A view on the definition of 'contents' under American law

Under American law, the contents of an email message or a web page, is regarded as content.[84] The challenge arises when it comes to URLs. Bellovin et al. (2016) argue that the Department of Justice may also be accessing content when collecting URLs, even though it may not intend to do so. Similar circumstances, in addition to other factors, may be at play in the brief scenario sketched by the Commonwealth Ombudsman. With greater transparency about how URLs are accessed and used by the AFP, we can obtain the direct evidence and have a fact-based technical and legal analysis. In the meantime, we can hypothesise. Bellovin et al. (2016), state that the authority part of the URL, as described in Part IV, is generally regarded as metadata. However, the authority part may sometimes, based on how the internet works, become the contents of a communication. This may happen as follows:

> While the DOJ may be trying to prevent the collection of content with a Pen/Trap order, this exemption from the "phone home to Main Justice" policy may actually lead to the collection of content with a trap and trace device. Specifically, content may be improperly collected in the following example: Since some web servers host multiple web sites sharing a single IP address, the specific web site that is being accessed is not itself derivable solely from the server's IP address; thus, the server must inspect the authority field

[84] Bellovin et al. (2016), pg. 45; Wiretap Act 18 U.S.C. § 2510(8) (2012)

> of the URL to determine what web page to serve. That information is transferred as part of the HTTP session. In that case, the authority field is architectural content, not metadata, to the network, although it may be metadata to a server run by a third party (i.e., one that is not the owner of the hosted web sites).[85]

Bellovin et al. (2016), used the Wiretap Act as the basis to define what architectural content was. Architectural content is not communicative content, it is not the substance of a communication. Architectural content is metadata or non-content, but it is contained in a message that is sent within the telecommunications network, from machine-to-machine. For example, it is sent from the web browser to the web server, based on programmatic interfaces and protocols, that may not be visible to the individual user. Bellovin et al. (2016), state:

> We formally define "architectural content" to mean information that — from a given point in the network and network stack — is simply transported, unexamined, even if it is not "information concerning the substance, purport, or meaning of that communication. We define its complement, "architectural metadata," as information intended for the potential use of a particular layer in the stack."[86]

Content is defined under the Wiretap Act as:

> "contents", when used with respect to any wire, oral, or electronic communication, includes any information concerning the substance, purport, or meaning of that communication.[87]

The Wiretap Act refers to "communicative content" and is based on the semantic meaning of a communication.[88]

The definition of 'electronic communication' under American law is:

> "electronic communication" means any transfer of signs, signals, writing, images, sounds, data, or intelligence of any nature transmitted in whole or in part by a wire, radio, electromagnetic, …, but does not include—
>
> (A)any wire or oral communication;
>
> (B)…;' [89]

The definition of 'wire communication' under American law is:

> (1)"wire communication" means any *aural transfer*[90] made in whole or in part through the use of facilities for the transmission of communications by the aid of wire, cable, or other like connection between the point of origin and the point of reception (including the use of such connection in a switching station) furnished or operated by any person engaged in providing or operating such facilities for the transmission of interstate or foreign communications or communications affecting interstate or foreign commerce;[91]

Under American law, Bellovin et al. (2016), makes the distinction between two types of content: communicative content and architectural content, when looking at URLs. This is based partly on the distinction between 'wire communication' and 'electronic communication'. Under Australian law, there is no such distinction. There is only the general reference to 'communication'. Under

[85] Pg. 72
[86] Pg. 45
[87] 18 U.S.C. § 2510(8) (2012) https://www.law.cornell.edu/uscode/text/18/2510
[88] Iobid; Bellovin et al., (2016) Pg. 6
[89] Fn 87
[90] (18)"aural transfer" means a transfer containing the human voice at any point between and including the point of origin and the point of reception;
[91] Fn 87

Australian law, the legal definition of the term 'communication' under the *TIA Act 1979*, is somewhat a combined version of the American definitions.

The content and metadata can be distinguished in that the content is unexamined and not used when it is being transported within the layers of Transmission Control Protocol/Internet Protocol (TCP/IP) stack. So, the authority part of the URL may be unexamined and not used by the web server to retrieve the web page. However, if a web server is using one IP address for several web pages, it may have to examine and use the authority part of the URL turning the authority part into content, albeit architectural content, as opposed to the substance or meaning of the communication, under the Wiretap Act of the US.

Under Australian law, the URL can either be the content or substance of a communication. You may have noticed that I have stayed clear from attempting a discussion about whether URLs could possibly be the substance of a communication. For the reasons Bellovin et al. (2016), do not describe URLs as the substance of a communication, I have also opted to stay clear of that discussion, because URLs may not convey semantic meaning. URLs however can be regarded as content under Australian law, for the same technical explanation offered by Bellovin et al. (2016) above. The authority part that is inspected by the web server is information which is in a communication, addressed to the web server to inspect and retrieve the requested web page. The query part is the information which is in a communication sent to the web browser which the web browser relays to the web server and returns the web page. Clicking on a malicious URL link received in a phishing email is information which is in a communication sent to the web browser which the web browser relays to the web server and returns the malicious web page.

VIII. Treating URLs as the contents of a communication

As discussed in above, this submission argues URLs should legally be regarded as the contents of a communication. Based on this thesis, Part VIII discusses how URLs may need to be treated under the existing legal framework and the potential changes that may need to be considered in any future cybersecurity strategy, policy and law, to fully cater for URLs as content.

A. The process to retain the contents of a communication to the Agencies treated

Table 2 demonstrates the differences between how content is treated under the *TIA Act 1979* versus how non-content is treated. Content is only accessible under stricter requirements, whereas non-content is easily accessible. Based on the categorisation URLs are given, they may fall under one of the two categories.

Table 2 How the contents of communication are treated to non-content under the TIA Act 1979[92]

Agency	Non-content/Telecommunications Data	Content
ASIO	No investigation is required for access to historical or prospective telecommunications data. This means that suspicion of an offence, based on reasonable grounds is not required to access the telecommunications data.[93] If URLs are regarded as telecommunications data, the URL does not need to be suspected as being malicious to collect the URL.	No investigation is required, which means suspicion of an offence, based on reasonable grounds is not required.[94] If URLs are regarded as telecommunications data, the URL does not need to be suspected as being malicious to collect the URL.
	Historical telecommunications data may be accessed and used for non-serious offences.[95]	
	The person does not have to be named in the authorisation and the notice issued to the Telco.[96]	The individual person must be named and accessing the contents of a group of persons from a single preservation notice is not allowed.[97]
	Store the LI for a period of two years after the LI came into existence.[98]	Telco to store the contents after the preservation notice has been issued.[99]
	Make a statement that the LI will be used in connection with the functions of ASIO.[100]	Consider whether there are reasonable grounds that the contents might assist with the investigation.[101]
AFP	No investigation is required for access to HLI which means suspicion of an offence, based on reasonable grounds, is not required.[102]	An investigation is conducted, which means there must be suspicion of an offence, based on reasonable grounds.[103]

[92] Shanapinda, Dissertation, Pg. 221- 224
[93] *CAC Determination* 2015 Schedule 1, Part 1 sections 1.01.
[94] ibid s 107J (1).
[95] ibid ss 175, 176.
[96] ibid ss 175, 176.
[97] ibid s 107H (3).
[98] ibid s 187C.
[99] ibid s 107H.
[100] CAC Determination 2015 Schedule 1, Part 1 sections 1.01.
[101] TIA Act 1979 s 107J (1).
[102] CAC Determination 2015 Part 3 s 3.01 (1) Items 3 b (vii) and (viii), (ix), Part 2 s 2.01 (1) Items 1–10, Part 2 Section 2.01(1) Items 8–9; TIA Act 1979 s 178(2).
[103] TIA Act 1979 s 107J (1).

	Access and use data for non-serious offences.[104]	Access and use content for serious contraventions. [105]
	The individual and the telecommunications service do not have to be named in the authorisation and the notice issued to the Telco.[106]	The individual and the telecommunications service must be named in the notice and the warrant. Accessing the contents of a group of persons from a single preservation notice is not allowed.[107]
	Store the data for a period of two years after the LI came into existence.[108]	Telco to store the contents after the preservation notice has been issued.[109]
	Make a statement about the likely relevance and usefulness of the data.[110]	Consider whether there are reasonable grounds that the contents might assist with the investigation.[111]
	Make a statement that the AFP is satisfied that the disclosure of the data is reasonably necessary for the enforcement of the criminal law.[112]	Consider whether there are reasonable grounds that the stored communications might assist in connection with the investigation.[113]
	The authorised officer must make a statement that he or she is satisfied, on reasonable grounds, that any interference with the privacy of any person or persons that may result from the disclosure or use is justifiable and proportionate, having regard to the likely relevance and usefulness of the data.[114]	The Judge, magistrate or AAT member must make a statement that he or she is satisfied on reasonable grounds that any interference with the privacy of any person or persons that may result from the disclosure or use is justifiable and proportionate, having regard to the likely relevance and usefulness of the data.[115]

[104] TIA Act 1979 ss 6A, 6B.
[105] ibid s 107J (1).
[106] CAC Determination 2015 Schedule 1, Part 1 sections 1.01.
[107] TIA Act 1979 ss 107H (3), 107J (1).
[108] Ibid s 187C.
[109] Ibid s 107H.
[110] CAC Determination 2015 Schedule 1, Part 2 section 2.01 (1) items 1–10.
[111] TIA Act 1979 s 107J (1).
[112] CAC Determination 2015 Schedule 1, Part 2 section 2.01 (1) items 1–10.
[113] TIA Act 1979 s 107J (1).
[114] CAC Determination 2015 Schedule 1 Part 2 section 2.01 (1) items 1–10.
[115] ibid Item 9.

	The data may be accessed after receipt of the notice from the Agencies.[116] The warrant must be issued in addition to the notice to access the data.[117]	The contents may only be accessed after the warrant has been approved.[118]
Both Agencies	The individual does not have to be informed of the access to the telecommunications data.[119]	The individual does not have to be informed about the access to the content.[120]

B. The process to mandatorily disclose the contents or substance of a communication to the Agencies

After the Telco is issued with the preservation notice, the Telco must store the contents the Agencies. Once stored, the Agencies and the Telco are prohibited to access the stored communications, or to authorise or permit another person to access the stored communications.[121] The Agencies and the Telco must not do any act or thing that will enable another person to access the stored communications.[122] The stored communications may now only be accessed under a Stored Communications Warrant (SCW), an Interception Warrant (IW), or a Computer Access Warrant (CAW).[123]

The Agencies are prohibited from accessing the stored communications without the knowledge of the intended recipient of the stored communication or the person who sent the stored communication.[124] There are however exceptions to this prohibition - the Agencies are not prohibited from accessing the stored communications without the knowledge of the individual under a preservation notice.[125] As discussed in Part XII. A and B., these rights may need to be transposed to any future cybersecurity strategy, policy and law that seeks access to web browsing histories or URLs.

IX. Treating URLs as telecommunications data

Part VIII. argued that URLs should be treated as the contents of a communication. In Part IX it may be fair to also assess the alternative argument – that URLs may be metadata, telecommunications data, the non-contents of a communication.

[116] CAC Determination 2015 Schedule 1, part 2 section 2.01 (1) items 1–10.
[117] TIA Act 1979 s 6DC; section 6DC was added by the Telecommunications (Interception and Access) Amendment (Data Retention) Act 2015 (Cth) on 13 April 2015.
[118] TIA Act 1979 s 9(3).
[119] CAC Determination 2015 Schedule 1, Part 3.
[120] TIA Act 1979 s 108(1).
[121] ibid
[122] ibid
[123] Ibid; IGIS Act s 25A; Attorney-General's Department, 'Telecommunications (Interception and Access) Act 1979 Annual Report 2015–16', (2016), vii; Shanapinda, Dissertation, Pg. 227
[124] ibid
[125] Ibid

A. The non-binary nature of URLs: both content and no-content

If URLs meet the criteria of being '… information or document that … relates to … the contents or substance of a communication',[126] then URLs may be telecommunications data or metadata. URLs are web addresses. The web page is the contents of a communication. The alternative legal question is whether URLs are '…information or document that … relates to … the contents or substance of a communication'.[127]

Given there is no legal definition for the word 'content', to assist with this inquiry, the Macquarie Dictionary defines the word 'relate' and the phrase 'relate to' as follows:

> 1. to tell.
>
> 2. to bring into or *establish association, connection, or relation*.
>
> –verb (i) 3. to have reference (to).
>
> 4. *to have some relation (to).*
>
> –phrase 5. *relate to, to understand and often identify with*: parents and teenagers often find it hard to relate to each other; he seemed to relate to that character in the film (emphasis added).[128]

The simple argument can be said to be this: URLs are the addresses of the web page, which web page is the content, as such, the URL relates to the contents. Clicking on the URL takes you to the contents. In this manner the URL and the web page have an established association, a connection or a relation. It can therefore be said that the URL relates to the contents that is the web page. This argument seems an acceptable one.

However, does the inquiry stop there? Can it be that URLs simply and only relate to the web page full stop? Can the URL have a dual or multiple nature, which simply cannot be ignored? Can it be that URLs relate to web pages as the content, but that URLs themselves are also the contents of a communication, as discussed in Part VII. Part VII stated URLs are sent in phishing emails and SMS as the contents of a communication, addressed to the targeted victim, or the random person who then receives the message, opens the message and interacts with, clicks on it and falls victim to a scam or personal data theft. The URL relates to the web page in that it re-directs the individual to the fake web page, but that malicious URL is distributed as a landing URL, as the contents of a message. As such, the URL does not simply stop with its association to a web page as the content that it relates to, the malicious URL itself is also transported in a message as contents, delivered through SMSs and phishing emails, as its contents to the victim to act upon.

Furthermore, the dual nature of URLs are well explained by Bellovin et al. (2016), by referring to URLs as architectural content when the authority part is inspected by the web server. These features of the URL can simply not be ignored or wished away but must be recognised and accounted for.

In the court case: *Privacy Commissioner v Telstra Corporation Limited [2017] FCAFC 4 (19 January 2017)*, and although speaking *obiter dictum* and in the context of personal information, the Federal

---

[126] *TA 1997* s 276, 280, 313(7)(d), (e) (TA 1997).
[127] ibid
[128] Macquarie Dictionary https://www.macquariedictionary.com.au/features/word/search/?search_word_type=Dictionary&word=relates

Court of Australia (FCA) underscored a key principle. In contrast to the Administrative Appeals Tribunal (AAT) earlier decision in the Mr. Ben Grubb case, that telecommunications data may only be about one thing, i.e. the location information was solely for the purpose of delivering the communications to Mr. Grubbs mobile phone, the FCA highlighted in general the principle that telecommunications data can be about various things, and are not confined, but are instead non-binary, when it stated:[129]

> Information and opinions can have multiple subject matters.[130]

This submission is about the meaning of the word content and not about the definition of personal information. Future research can examine whether URLs are personal information given these cases. The FCA's opinion is however cited for its persuasive value, and not for its authoritative value. Like other metadata that can have multiple subject matters, URLs too can have multiple subject matters. Whether URLs can have multiple subject matters so that can pass the test of being personal information, will not be discussed in this submission. However, the principle is that the multiple relationships URLs have with web pages as content and as being content itself, carried in addressed messages and inspected by web servers, should not be disregarded and a simplistic argument be made that URLs relate to web pages as content and therefore URLs in and of themselves being contents of a communication is not legally and technologically true. Making this binary distinction would classify URLs as non-content whereas, based on how URLs are distributed in phishing emails and SMSs, their dual nature as content would be denied, to maintain the status quo – an unclear legal framework, that has the potential of being exploited, by disclosing URLs under self-certified notices as opposed to obtaining warrants issued by the independent judiciary. It is therefore best to lean towards recognising the content nature of URLs as opposed to denying the content nature of URLs, because they simply cannot be one or the other, that is not how technology in question functions.

Ignoring the content nature of URLs ushers in with it restrictions to individual privacy rights that are protected under the warrant processes of the *TIA Act 1979*. The reverse is also true – denying the metadata nature of URLs is to give preference to its content nature and to encumber the Agencies with greater accountability, in carrying out their cybersecurity functions. Would it be an unreasonable burden and restriction on the Agencies carrying out their functions to require warrants to collect known and unknown malicious URLs, in relation to the protecting the privacy of the individual? Which is the lesser evil, so to speak? Given the discussion in Part VII. and Part VIII., the content nature of URLs must surely outweigh the non-content nature of URLs and preserve the privacy rights of individual Australian citizens and residents. We must always choose to err on the side of caution and choose the option that protects and preserves rights as opposed to the alternative, when faced with such complex and contrasting technological and legal arguments, that have a direct impact on individual human, civil and political rights. These must be the values to respect and protect, as this will surely meet the legitimate expectation of the general populace.

[129] Telstra Corporation Limited and Privacy Commissioner [2015] AATA 991 (18 December 2015) 8 [9.]
[130] *Privacy Commissioner v Telstra Corporation Limited [2017] FCAFC 4 (19 January 2017)* [63]

B. Treating URLs as telecommunications data

Given the discussion in Part VII that URLs must be regarded as the contents of a communication, and should be treated as such under current laws, the discussion below contrasts how URLs, is potentially being treated as metadata under the current legal framework. This contrast will highlight a legal process that will not instil public confidence and trust in a cybersecurity strategy that will classify URLs as metadata, to be collected by means of a self-certification notification and authorisation process under the *CAC Determination* 2015 as supported by the *TIA Act 1979*.

The discussion below starts with how web browsing history is not required to be retained, but not because web browsing history is regarded as content. Web browsing history is also not clearly classified as the contents of a communication. This gives rise to an ambiguous and confusing current legal framework, that if continued to exist, will hamper any future cybersecurity strategy.

1. The process to retain telecommunications data to the Agencies

Under the telecommunications data retention regime, web browsing history is not required to be retained. Web browsing history is not required to be retained, so web browsing history is not 'retained data'. Only telecommunications data that is required to be retained is regarded as 'retained data'. Section 187A(4) describes the type of information the Telco is ***not required to retain*** as follows:

> (4) This section does not require a service provider to keep, or cause to be kept:
>
> (a) information that is the contents or substance of a communication; or
>
> Note: This paragraph puts beyond doubt that service providers are not required to keep information about telecommunications content.[131]
>
> (b) information that:
>
> (i) states an address to which a communication was sent on the internet, from a telecommunications device, using an internet access service provided by the service provider; and
>
> (ii) was obtained by the service provider only as a result of providing the service; or
>
> Note: This paragraph puts beyond doubt that service providers are not required to keep information about subscribers' web browsing history.[132]

According to Section 187A(4) above, there are two things the Telco is not required to do under the mandatory data retention scheme. The **first** is, the Telco and ISP is not legally required to retain '*information that is the contents or substance of a communication*'. As discussed in Part VII. B. 4. a), the Telco is not required under section 172 to disclose these contents or substance of a communication under a notification or authorisation under the *CAC Determination* 2015. According to the first note above, '*information that is the contents or substance of a communication*' is the same as '*information about telecommunications content'*. The two phrases can therefore be interpreted as being given the same meaning.

The **second** is, information that is described as '*information about subscribers' web browsing history*' is not required to be retained by the Telco under the mandatory data retention scheme.

[131] It must however be stated that the notes in the above sections are not part of the law. They are simply notes that can be used to help with interpreting the section in the law.
[132] ibid

The **second** note also does likens:

> *information that states an address to which a communication was sent on the internet, from a telecommunications device, using an internet access service …*
>
> to,
>
> *information about subscribers' web browsing history*.

The two phrases can therefore be interpreted as being given the same meaning. The individual Australian citizen or resident would be using a telecommunications device, like a smart phone, to access email and click on a web address that was contained in the email or use a web browser such as Google Chrome and type in a search query into the search bar and be presented with the search results. This process takes place under the HTTP/S session between the web browser and the web server, as discussed in Part VII. B. 2.

The Telco is not required to retain the web browsing history, which history is represented by past URLs visited from the device of the individual Australian citizen or resident. The web page itself, which is the content, or the web URL, which is the web address from where the web page itself was retrieved by the web browser, would not be required to be retained. The web address or the URL is therefore not the information that states an address to which a communication was sent on the internet, from the web server, as the telecommunications device.

What is not legally clear is whether these two exceptions are made because information that states an address to which a communication was sent on the internet, from a telecommunications device, URLs, or the web browsing history, is legally classified as the contents or substance of a communication. Section 187A(4)(a) above only refers to telecommunications content, that is not required to be retained but does not refer to web browsing history in the accompanying note. Section 187A(4)(b) refers to web browsing history but to contrast it against the IP address. In other words, the IP address of the users device may be retained and disclosed to the Agencies, but the web browsing history may not be. So, is this because web browsing history is legally classified as content?

The fact that Section 187A(4)(a) and Section 187A(4)(b) are separated and there is no reference between the web browsing history and content, makes it difficult to make an interpretation that web browsing history and content is equated. The separation between the two sub-sections may be made because Parliament wanted to avoid the interpretation that web browsing history is or likely to be the contents or substance of a communication. As such, what is content is left vague. But what is clear is that web browsing history is not required to be legally retained, but is likely not content, and so it may be voluntarily retained for maintenance, business, commercial and cybersecurity purposes of the Telco.

Despite the two exceptions, the Telco is however not prohibited to voluntarily retain any of the following types of information:

i. *information that is the contents or substance of a communication*;
ii. *information about telecommunications content;*
iii. *information that states an address to which a communication was sent on the internet, from a telecommunications device, using an internet access service provided by the service provider; and was obtained by the service provider only as a result of providing the service*; or

iv. *information about subscribers' web browsing history*.

Even if web browsing history was legally classified as content, the Telco may still retain same for its own purposes. The preservation notices and the warrant processes under the *TIA Act 1979* operates as follows: The Telco may already have the content in its possession, based on its internal storage, network maintenance, research and development strategies and cybersecurity policies etc. The Telco is only then issued with a notice to preserve the content it already has stored and possesses, and to continue to preserve the said content. This is followed up with a stored communications warrant, to access the content that has been stored under the preservation notice.[133]

The Telco may voluntarily retain any of the above types of information for its own commercial, network maintenance, quality of service and cybersecurity purposes. The *TIA Act 1979* does not specifically state that the retention of this information is voluntary. However, the cybersecurity purpose may also in addition to being voluntary under the *TIA Act 1979*, be an implied obligation under section 313 of the *Telecommunications Act 1997.*

2. The process to mandatorily disclose telecommunications data to the Agencies

Under the *TIA Act 1979,* the Agencies may request any of the telecommunications data to be disclosed to it. Telco is allowed to disclose telecommunications data if the disclosure is authorised by law, or if the disclosure is required under the law. The Telco is required to disclose telecommunications data it is legally required to retain.[134] The Telco is also legally required to disclose telecommunications data the Telco is not required to retain but chooses to retain.[135] The Telco may also choose to voluntarily disclose any of the telecommunications data, even if it has no legal duty to do so.[136] The disclosure of the information to the Agencies must be 'reasonably necessary' to safeguard national security.[137] The Telco is required to give such help as is 'reasonably necessary' to enforce the criminal laws and to safeguard national security.[138]Giving such help as is reasonably necessary means allowing the Agencies to intercept the contents of a communication.[139] It also means to disclose information or documents under a stored communications warrant.[140] Disclosing information or documents as requested under an authorisation and notification issued by the Agencies under the *CAC Determination* 2015 is another way of giving help to the Agencies, that is reasonably necessary.[141] The Agencies may issue authorisations and notifications in terms of the *CAC Determination 2015* to request access to telecommunications data.[142]

---

[133] Chapter 3 *TIA Act 1979* (Cth).
[134] *TA 1997* s 276.
[135] Shanapinda, Dissertation, Chapter 4.
[136] *TA Act* 1997 ss 275A, 276, 280, 313(3), (4), (7); TIA Act 1979 ss 174, 177.
[137] *ibid* s 313(3).
[138] Ibid s 313(7)(d), (e).
[139] Ibid s 313(7)(d), (e).
[140] Ibid s 313(7)(d), (e).
[141] *TIA Act 1979* s 183; CAC Determination 2015 Schedule 1, Part 1-4. Part 2 section 2.01 item 5; Communications Access Coordinator's (CAC) Telecommunications (Interception and Access) (Requirements for Authorisations, Notifications and Revocations) Determination 2015 (Cth) (at 9 October 2015).
[142] Shanapinda, Dissertation, Pg. 193-194

The Telco is therefore required to disclose any of the telecommunications data it has in its possession, whether it chose to retain the telecommunications data voluntarily or whether the Telco was required to retain the telecommunications data under the mandatory data retention scheme – it makes no difference to the obligation to disclose the types of information requested.

a) The process under section 313 of the Telecommunications Act 1997

(1) The process to retain threat intelligence voluntarily or under implied legal obligations under section 313 of the *Telecommunications Act 1997*

The *Telecommunications Act 1997* does not specifically require the Telco to retain information and data, such as suspected malicious URLs, for the purposes of complying with their legal obligations to detect and prevent cyber incidents and attacks under section 313(1) – (2). This obligation can however be said to be an implied obligation. Alternatively, it can be said to be the voluntary retention of the relevant threat intelligence.

As discussed in Part V, malicious URLs pose cyber security risks to private Australian and to Australia's national security. As such, the Telco is impliedly required, in compliance with section 313 (1)-(2) to employ the cyber security controls, similar to the types of controls discussed in Part V F., in order to detect and prevent malicious URLs from causing cyber incidents and cyber-attacks. In doing so Telco's may voluntarily store and therefore have in their possession malicious URLs the Agencies may have an interest in collecting and using to inquire into and investigate cybercrimes, as discussed in Part V. G. and Part VI.

(2) The process to disclose threat intelligence under legal obligations under section 313 of the *Telecommunications Act 1997*

The Telco may also be requested to disclose URLs they voluntarily retained to comply with cyber security obligations under the law.[143] Under section 313(1)-(3) the Telco is required to comply with the cybersecurity legal requirement of doing their best to ensure the confidentiality, integrity and availability of the information and communication in their possession, when transported and at rest. When URLs are transported in the servers, they may manage ISPs would be required to introduce anti-spoofing countermeasures as best it can, as best practice, especially given the unprecedented nature of recent attacks.

The Telco may only disclose the suspected malicious URLs lawfully. The lawful process is to disclose the suspected malicious URLs as either the contents of a communication and using the relevant process, or as telecommunications data, and using the relevant process. As such, according to section 313(3) – (7) of the *Telecommunications Act 1997* read with the disclosure obligations under the *TIA Act 1979*, the Telco is required to disclose telecommunications data and the contents of a communication to enforce criminal laws and to safeguard the national interest.

The disclosure process under section 313(3) – (7) of the *Telecommunications Act 1997* can work separately or work in tandem and be complementary, with the legal duties of the Telco to disclose

[143] *TA 1997 s* 280

telecommunications data to the Agencies under the notification and authorisation scheme of the *CAC Determination* 2015, as permitted by sections 175 to 180 of the *TIA Act 1979*. The two disclosure processes work together and are not opposing each other. It allows the Agencies and non-law enforcement agencies alternative legal options by which to collect, access and use other types of information, the Telco has in its possession, but did not store for the specific purposes of the mandatory data retention regime under the *TIA Act 1979* as amended by the *Data Retention Act of 2015*. In other words, even if URLs or web browsing history are not required to be retained, if the Telco has the URLs in its possession the URLs must be disclosed under section 313(3) – (7) of the *Telecommunications Act 1997* read with sections 175 to 180 of the *TIA Act 1979*. Alternatively, separate request may also be issued under section 313(3) – (7) of the *Telecommunications Act 1997* as was done by the AFP in 2014 prior to the introduction of mandatory data retention scheme. According to the ACMA it however seems as if none of the Agencies have used section 313(3) – (7) of the *Telecommunications Act 1997* separately for online content blocking, which bocking may stem from individual complaints of being defrauded via fake websites and/or using web browsing histories of the individual. According to the Commonwealth Ombudsman, even if section 313(3) – (7) of the *Telecommunications Act 1997* may not have been used to collect any information relevant to blocking fake websites, the AFP did collect URLs to investigate crimes, which crimes may be fraud, whether online or not, and cybercrimes. The latter information may have been collected under the notification and authorisation scheme of the *CAC Determination* 2015, as permitted by sections 175 to 180 of the *TIA Act 1979*.

3. The process to voluntarily disclose telecommunications data to the Agencies

The Telco may voluntarily disclose telecommunications data to ASIO.[144] The voluntary disclosure must be about ASIO performing its functions.[145] The Telco may voluntarily disclose telecommunications data to the AFP.[146] The disclosure must be reasonably necessary to enforce the criminal laws.[147] If the AFP initiates the request, it is not considered a voluntary disclosure.[148]

However, the Telco may not voluntarily disclose the contents of a communication. The Telco may also not voluntarily disclose the contents of a communication as if it were telecommunications data under the *CAC Determination* 2015. The Telco may have been under the impression that URLs are legally classified as telecommunications data and that it may be disclosed as telecommunications data under the *CAC Determination* 2015. This may have happened under the circumstances described by the Commonwealth Ombudsman. This uncertainty may therefore lead to an exploitation of the circumstances, were URLs although considered as content by the oversight body, may be disclosed as telecommunications data instead. Given that the law has no definition of what is content, and the Explanatory memorandum are inconclusive, the cyber security strategy must clearly state that URLs are the contents or substance of a communication.

[144] *TIA Act 1979* s 174.
[145] ibid
[146] *TIA Act 1979* s 177.
[147] ibid
[148] Shanapinda, Dissertation, Pg. 210; *TIA Act 1979* s 174.

X. The processes under section 280 of the *Telecommunications Act 1997*

In in the 2018–19 financial year, a total of 8,432 disclosures were made to non-law enforcement agencies under section 280 because the disclosure or use was required or authorised by or under a given law.[149] It is not clear to what extent the disclosures under section 280 fallow the content versus non-content traditional distinction.

In its submission to the Parliamentary Joint Committee On Intelligence and Security (PJCIS) Telstra indicated non-law enforcement agencies request huge quantities of data they are not always able to properly interpret.[150] It is not publicly known whether such disclosures included web browsing histories or URLs. In the envisaged cybersecurity strategy, section 280 may need to be clarified as to how it would operate in relation to accessing the contents of a communication, such as URLs. If URLs are content, section 280 should not be used to access URLs. The provisions under the *TIA Act 1979* should be harmonised with section 280 so that non-law enforcement agencies are also subject to similar governance and accountability measures. This recommendation was also made by Telstra in its submission to the PJCIS. [151]

A. The processes to disclose threat intelligence under section 280 of the *Telecommunications Act 1997*

Threat intelligence is information and data that leads an organisation to take decisive action and to change its behaviour in terms of how it has always managed cybersecurity risks. Known and unknown malicious URLs may constitute such threat intelligence, in one form or another.

Web browsing histories refer to a past record of the websites visited. The other name for URLS are websites. As such, URLs or the web browsing history is not required to be retained to be retained under the mandatory data retention scheme. And any data that is not required to be retained for the specific purpose of the mandatory data retention scheme may be retained voluntarily or under an implied legal obligation and be disclosed to non-law enforcement bodies under section 280 of the *Telecommunications Act 1997*, when the disclosure or use is required or authorised by or under a given law.[152]

[149] Australian Communications and Media Authority annual report 2018–19, pg. 128
[150] Telstra. Review of the mandatory data retention regime Submission 35. Parliamentary Joint Committee On Intelligence and Security. Review Of The Mandatory Data Retention Regime. Pg. 2-4.
[151] ibid. Pg. 3.
[152] Section 280(1B)(b)

XI. The process under the *Assistance and Access Act 2018*

A. The process to disclose 'technical information' under the legal obligations of the *Assistance and Access Act 2018*

Under the *Assistance and Access Act 2018*, the Telco, ISPs and Social Media Platform companies like Google and Facebook are required to assist the Agencies by disclosing 'technical information' to the Agencies. The word 'technical information' is not defined in the law. It is therefore not clear whether a URL is considered to be 'technical information' under the *Assistance and Access Act 2018*. The question is whether suspected and known malicious URLs would qualify as 'technical information'. This issue needs to be clarified under any future cybersecurity strategy.

If the telco has this information available, as part of its cybersecurity checks, the Telco may volunteer the information or comply with an assistance and access request under the *Assistance and Access Act 2018* if URLs are classified as 'technical information' by the Agencies, or under the authorisation and notification process of the *CAC Determination* 2015, if the URL is described as telecommunications data. However, given that the URL is a web address, and even though it may not be required to be retained, it may not be required to be disclosed to the Agencies as telecommunications data, as non-content, because it is contained in a phishing email addressed to the individual Australian citizen or resident to activate, as its intended and targeted recipient. URLs may therefore not qualify as 'technical information' under the *Assistance and Access Act 2018* because known and suspected URLs are already the contents of a communication under the *TIA Act 1979*.

B. The process to retain 'technical information' voluntarily or under the implied legal obligations of the *Assistance and Access Act 2018*

Subsection 317ZGA(3) of the *Assistance and Access Act 2018* prevents Technical capability Notices (TCNs) from being used to build or extend data retention requirements to designated service providers (DSPs) or Social Media Platform companies (SMPCs) like Google and Facebook.

Subsection 317ZGA(4) of the *Assistance and Access Act 2018* prevents TCNs from being used to create technical capabilities to store the browsing history of internet users.

However, as stated in Part XII A., the Telco, ISP and SMPCs may store URLs as part of their cybersecurity compliance and governance framework. These entities are also prohibited from doing anything to prevent compliance with the *Assistance and Access Act 2018*. The impact of the combination of the disclosure obligations may be that these entities choose to hold URLs as a result. Under these circumstances there would be no need for the *Assistance and Access Act 2018* to require the retention of URLs logs. Additional implied retention obligations may also stem from the obligations under section 313 of the *Telecommunications Act 1997* that requires the Telco follow best practice in preventing cyber-attacks and ensure the confidentiality, integrity and availability of information and communications.

C. Overlapping Interests?

Moreover, Google and Facebook do retain web browsing histories anonymously as part of their legitimate businesses operations to sell advertising and in this manner monetise URLs. Social psychologist Shoshana Zuboff refers to this practice as 'surveillance capitalism'. The resulting impact is this: in this manner law enforcement and major tech companies support each other's interests in a sort of a symbiotic relationship, willing or unwilling.

This statement is not meant to be a criticism of either party, but a reflection of the resulting dynamic. The statement is not contradictory to proposals that SMPCs and private businesses should also be legally required to protect cyber space and in doing so hold URLs. The cybersecurity threat does inadvertently require many organisations to continue to hold data to research and investigate unknown threats, following the best practice of cyber situational awareness. While doing so, the organisations may as well try and monetise the data, if this was not their traditional business strategy, hence the continued rise of Big Data analytics. Any direct or implied legal or moral obligation, or business imperative to secure systems, given unprecedented levels of dynamic attacks, may eventually result in these types of effects. It is best if these eventualities are anticipated and properly addressed, in advance.

XII. Standardised Framework for Category A and B Entities

This section discusses how the collection, storage, disclosure and use of URLs (as the contents or substance of a communication) may need to be addressed to investigate and inquire into cyber security risks, cyber-crimes, cyber incidents and cyber-attacks, based on existing Australian laws, and in relation to privately owned and privately operated Australian companies, partnerships, businesses, corporations, charities, institutions, pubic bodies, statutory bodies and civil society organisations. These entities may be referred to as Category B Entities.

A. Standardised and harmonised legal framework to collect and share threat data

In the same way that the *TIA Act 1979,* the *Telecommunications Act 1997* and the *Assistance and Access Act 2018* puts legal obligations on Category A Entities to directly and indirectly prevent cyber incidents, store threat intelligence, analyse and share required details with the relevant Agencies, when required to do so, Category B Entities could share similar legal obligations.

One such obligation may include the sharing of suspected and known malicious URLs, which increasingly poses greater and dynamic cybersecurity threats. Section 313 requires Category A Entities to assist the Agencies by sharing information. The *Assistance and Access Act 2018* requires Category A Entities to share technical information when requested to do so. Only the *Assistance and Access Act 2018* allows Category A Entities to be informed of their rights to object to such requests. The scope of these laws could be extended to all private and public business; public officials and private citizens and residents, but with stronger rights to lodge objections, have decisions reviewed and appeals lodged in an independent and open court, with a competent judge with relevant knowledge and experience in relevant matters.

B. An open, transparent and democratic governance framework to collect and share data

The sharing of threat data must however only be done under a governance framework, that poses the least threats to the privacy of the individual, human rights, civil and political rights and the least interference with private business undertakings. The framework must first address the strategically imperative question that URLs, malicious or otherwise, are the contents and substance of a communication. This will give the framework public trust and confidence. The current scenario where this critical issue is unclear will not instil any public confidence, if tomorrow the Agencies would want to access URLs or web browsing history to investigate a cyber-attack that caused harm to life and property. The public will better trust a governance framework that outlines that URLs may only be collected with a preservation notice issued, based on reasonable grounds, and the disclosure of the URL is authorised by an independent member of the judiciary, with relevant experience. The public would also want to be informed if a warrant is being applied for to access URLs, and to object to such warrants in open court. Some of these elements are provided for in the laws listed above and can be transferred and improved on in any future strategy, policy and law aimed at strengthening cybersecurity in the national interest.

C. Data philanthropy

The AATA denied Mr, Grubb access to URLs about his web browsing history, arguing Mr. Grubb was not entitled to access the same information as the Agencies:

> Mr Grubb has asked why he cannot have the same information as that available to law enforcement agencies. The answer is that the entitlements of Mr Grubb and those of law enforcement agencies are the subject of different legislative regimes. Each regime seeks to achieve a balance of policy considerations and desirable outcomes. Those policy considerations include protection of an individual's privacy, search and rescue, security and law enforcement issues and public safety.[153]

Any successful cybersecurity strategy and any regime created under it must harmonise the rights of the individual to be able to request and access URLs held by their service providers, in a usable format, if they so wish. Google already does this. Individuals can run tests and research cybersecurity challenges and countermeasures, in collaboration with others like Universities and with the service providers themselves.

In May 2019 WhatsApp was hacked with the Pegasus malware tool. WhatsApp users and WhatsApp itself shared the evidence with the University of Toronto's Citizen Lab. It was as part of this investigation that it was found that this tool was used to target journalist and human rights lawyers. The same tool can be used to impact Australia's national interests. As a result, we can learn more about the vulnerability in WhatsApp and was addressed. Such collaboration should be encouraged, and an enabling environment fostered between citizens whose tools are used to launch attacks, social media platforms whose applications are used as vectors, law enforcement and academia.[154]

[153] [114.]

[154] NSO Group / Q Cyber Technologies. Over One Hundred New Abuse Cases. October 29, 2019. https://citizenlab.ca/2019/10/nso-q-cyber-technologies-100-new-abuse-cases/

The envisaged governance framework can allow individuals and private businesses to 'donate' and 'trade' URLs, or generally threat intelligence of their choosing, for research and the development of detection countermeasures, to universities, the ASD and the Agencies. In this manner cybersecurity tools can be developed based on real-world experiences. This philanthropic effort can be incentivised by offering cybersecurity advice and assistance from the ASD, especially to small and medium size businesses. This will help make the threat landscape tougher for would-be attackers, and the security tools more relevant, effective and designed for purposes, based on the latest tactical behaviours of attackers.

D. Same difference

The *TIA Act 1979* , *Telecommunications Act 1997* , and the *Access and Assistance Act 2018* complement each other. As such, the same 'details of Internet sessions' that may be legally classified as 'content' under the *TIA Act 1979* may be legally classified as 'technical information' under the *Access and Assistance Act 2018* or be requested under section 280 of the *Telecommunications Act 1997* and be disclosed by the Telcos, ISPs and SMPCs under these legal procedures, to the Agencies and to non-law enforcement bodies. Each party may request and collect the information that is materially the same but under a different law and with a different legal description. The Agencies and non-law enforcement bodies would be forum shopping, using the more flexible legal procedure. The Agencies and the non-law enforcement bodies may do so under current legal provisions. The same content or telecommunications data may be used to inquire into and investigate cybercrimes, ensure cybersecurity resilience as well as minor and serious offences.[155]

If a new overarching cybersecurity law is envisaged under the Cybersecurity Strategy, the various pieces of legislation must be harmonised. This harmonisation must address the types of prospective and historical information and data that is similar to 'technical information', 'content' and 'telecommunications data' the Agencies and non-law enforcement bodies would collect, access and use to inquire into, investigate, detect and prevent cyber incidents and cyber-attacks. One such piece of information would include suspected malicious URLs. A URL can easily be legally classified as 'technical information', 'content' or as 'telecommunications data'. Any envisaged law must clearly technically describe URLs, describe the nature of URLs as per the three types of URLs as described herein, and whether all types of URLs would be required to be accessed and used by relevant bodies and Agencies, as the contents or substance of a communication, as opposed to non-content, under a preservation notice or a judicial warrant, and not under a verbal or written authorisation issued under the *TIA Act 1979*, or a request under section 280 of the *Telecommunications Act 1997*, or an assistance request and notice under the *Access and Assistance Act 2018*.

Other non-law enforcement bodies, such as ASIC, that have the power under sections 280 and 313 (3) of the *Telecommunications Act 1997*, to block malicious URLs and domains that aim to defraud individuals with fake websites, given the privacy implications, must also be subject to a governance framework that is equal to the domestic preservation notices and warrants, issued to

[155] However, under section 317B of the Assistance and Access Act 2018 (Cth), this law is restricted to serious offences.

the Agencies under the *TIA Act 1979*. In this manner the governance and oversight is harmonised for all.

XIII. Conclusion

Malicious URL-based attacks are on the increase. URLs are used as vectors and must be collected and analysed to investigate cyber risks. As such new and dynamic detection methods are required, which may include pro-actively scanning and collecting URLs to scan and detect attack types. This obviously has legal implications on collecting URLs, based on whether URLs are regarded legally as content or as non-content. Given the dynamic nature of malicious URLs, the effective measure may be to allow for its collection as telecommunications data, and to state as such in no uncertain terms. However, this has major privacy implications. This route may not be advisable given the analysis of this submission that URLs are essentially the contents or substance of a communication. Collecting URLs without a judicial warrant, to preserve the privacy of the ordinary Australian citizen and resident may therefore result in public objections. It was given this public objection that URLs were specifically excluded from the mandatory data retention regime, but without it being clarified that it was because web browsing history was content or non-content. This issue must now be resolved given the cybersecurity implications.

Under the telecommunications data retention regime, Telcos and ISPs are not required to retain web browsing history. Telco's may retain, for whatever period, any URLs in order to investigate any cyber security risks, and thus will have it in their possession. However, any information the Telco or ISP holds, that is regarded by the Agencies to be telecommunications data and not considered to be the contents of a communication may be authorised by the Agencies to be collected, if the Telco and ISP has the telecommunications data in its possession. In order to hand over suspected malicious URLs sent in an email, clicked on by an individual employee, or a Telco or ISP customer, or a private individual, to the Agencies, it is not legally clear whether URLs are to be disclosed as classified as telecommunications data or as the contents of a communication. The classification will determine how the privacy of the individual Australian citizen or resident will be protected.

For the envisaged cyber security strategy and the resulting legal framework to work effectively, this primary issue must be settled: the policy and legal decision must be made whether URLs are the contents of a communication or are regarded as telecommunications data. If URLs are content, then only via judicial warrants can URLs be accessed and used to investigate cyber threats and incidents. URLs contain the query portion that is typed in by the individual and therefore is the contents of a communication, for which a judicial warrant must first be obtained, for: i.) real-time (intercepted access while in transit); ii.) prospective; and iii) historic access to URLs.

This submission has argued that URLs can be both content and non-content, but that the content nature of URLs must be preferred in the interest of public trust. Any envisaged cybersecurity regulatory framework should be mapped to the *TIA Act 1979* to establish this in principle. The preservation notice process and the judicial warrant process to access content should be revised and applied to cybersecurity risks and the relevant and already existing governance measures be further strengthened and strictly followed to collect, access and use the content of communications.

Granting individuals access to their web browsing history may be a worthwhile exercise as it would allow individuals to donate their records for cybersecurity research purposes as part of a strategy to effectively combat cybercrime.

A. Questions, answers and key recommendations

Following the discussions above, and recommendations already outlined therein, the questions raised as part of the strategy development process are briefly answered below:

Table 3 Questions, Answers and Recommendations

Questions	Brief Answers	Recommendations as possible strategy options
1 What is your view of the cyber threat environment? What threats should Government be focusing on?	URL-based threats have increased globally at the rate of 26% since the last quarter of 2018.	**Prioritising URL-based HTTPS phishing attacks:** The Government should focus on detecting and preventing URL-based HTTPS phishing attacks.
2 Do you agree with our understanding of who is responsible for managing cyber risks in the economy?	The government is responsible for showing leadership, setting the tone, encouraging and incentivising public and private sector to better manage the areas of the cyberspace they have control over.	**Leadership:** This consultative process was a good demonstration of taking charge and of leadership. The government now has the role to call each of the stakeholders to task in this regard, and lead by example. The government can develop a responsibility matrix in the 2020 strategy that maps out the role every major stakeholder must play and the measurable metric to ensure compliance.
3 Do you think the way these responsibilities are currently allocated is right? What changes should we consider?	ASIO and the AFP are the investigative arms and they interface with relevant stakeholders to collect and act on cyber threats. This must be maintained. The ASD is part of the military and ACSC is the national cybersecurity hub. The	**Central cyber situational awareness:** The ACSC should be further strengthened as the hub that will coordinate cyber situational awareness at the national level, with

	relationship between the ACSC, ASIO and the AFP can be clearly outlined so that the data collection, sharing and processing arrangements are clearly understood. The review roles any oversight bodies have in this regard, including the independent judiciary should be clearly spelt out.	threat intelligence collected from private and the public sector.
5 How can Government maintain trust from the Australian community when using its cyber security capabilities?	The collection and sue of URLs should be subject to preservation notices and judicial warrant warrants.	**Legal certainty and predictability:** URLs should be legally classified as the contents of a communication, in a clear and unambiguous manner.
10 Is the regulatory environment for cyber security appropriate? Why or why not?	The *TIA Act 1979* contains the primary building blocks for the regulatory environment. These elements can be applied beyond the telecommunications sector to private and public bodies, to assist the Agencies in the same fashion.	**Harmonisation of existing laws:** The right of review and right to object under the *Assistance and Access Act 2018* should be harmonised with the amended *TIA Act 1979*, for when its jurisdiction is extended to apply to private and public business, as regards cybersecurity.
11 What specific market incentives or regulatory changes should Government consider?	Strengthening the judicial oversight powers over the exercise of the cybersecurity capabilities. Private and public enterprises can be incentivised to share and trade threat intelligence in return for cybersecurity assistance and advice form the government.	**Independent Judiciary:** Given that private devices are targeted and used as vectors in unprecedented fashion, and government requires access to all types of data to effectively counter threats, this access must be negotiated through a robust judiciary, to ensure public trust and transparency.
14 How can Australian governments and private	By real world data philanthropy with universities for masters and	**R&D:**

entities build a market of high quality cyber security professionals in Australia?	PhD studies to analyse URLs and their patterns and trends, to be one step ahead and to develop appropriate preventative tools, such as the similarity-mapping tool, because blacklisting is not going to work.	The government can sponsor PhDs and master's level research for vulnerable entities that are managing critical infrastructure, as well as strategic small and medium enterprises, to work on solutions that can potentially be applied universally and be commercialised. Similarity matching tools for URL-based threats have great potential for commercialisation.
18 How can governments and private entities better proactively identify and remediate cyber risks on essential private networks?	Boards of directors can be tasked to prepare annual threat reports, tabled to shareholders. Any request for threat intelligence by the government must be mediated via the independent judiciary.	**Protect I&CT:** Under section 313 of the *Telecommunications Act 1997*, the Telco is required to detect and prevent cyber-attacks. This obligation can be imposed on essential private networks as well.
21 What are the constraints to information sharing between Government and industry on cyber threats and vulnerabilities?	The uncertainty regarding whether URLs are content or telecommunications data, or anther category of alien data. Given this uncertainty, the Telco's may not be sure how to comply with requests for URLs, and oversight bodies like the Commonwealth Ombudsman may not know how properly supervise such information disclosures under a future cybersecurity strategy that would rely on the proactive collection and analysis of URLs as part of the bigger program to detect and prevent cyber threats.	**URLs legally classified as content:** URLs should be classified legally and under government policy as the contents of a communication. Category A and category B Entities should be granted the express legal right, if they choose to, retain and analyse URLs for malicious activity. Category A and category B Entities should be allowed to voluntarily share suspected malicious URLs, that has been de-anonymised, with the Agencies.

		Content sharing: The sharing of this content should be made subject to an application, lodged with an independent court, staffed with suitably qualified judges and personnel, notifying known individuals of the application, and informing them of their right to object to the application, and allowing for a review of any decision made.
		Create a legal and policy environment that is conducive to sharing information between stakeholders: Individuals should be able to request access to their URLs, unlike the 2015 AATA decision in the case of Mr. Grubb, and be allowed to share it with the Agencies, in the performance of their functions, and with researchers in academia and industry. This will be helpful in cases where individuals suspect malicious activity.