Table of Contents

Draft bill views

I'll opine on some of the bullet points on page 7 of the guidance note. If there seem to be things missing it's because they aren't worth answering as expounding on my own solution to this problem will render them useless anyway.

the definitions of misinformation and disinformation

I note that on page 7 of the guidance note it is explicitly stated that: *The Bill does not seek to curtail freedom of speech.*

In that context, how can the definition of misinformation also include the phrase: *reasonably likely to cause or contribute to serious harm?*

Reasonably likely, suggests there is an existing motive for prosecuting, one that can be applied even before. A prosecution on the basis that something is "reasonably likely" is not sound... There should be actual evidence, not correlations.

If freedom of speech is to be protected then in an active case, distinguishing between misinformation by analysis of "intent" and "serious harm" should *only* be done retroactively after events constituting serious harm have occurred.

While the guidance note states it doesn't seek to curtail freedom of speech [through any explicit means], if turned into legislation this will end up happening as digital services themselves take preventative measures to preempt any and all information with *any possibility* of causing serious harm. Why? To avoid bad public relations or potential liabilities that comes with inadvertently disseminating misinformation.

At best extreme content moderation, at worst, digital services would just refuse to operate in Australia.

Furthermore the definition of disinformation is redundant:

"Disinformation is intended to capture 'misinformation that has been disseminated with the intention of deceiving another person'."

The definition of misinformation already includes "*deception*". If there is no cause or contribution to "serious harm" what possible pretext could one have for prosecuting a case? I note we already have reforms to defamation law in relation.

This also calls into question the broader purpose of this bill, as if serious harm was in fact caused or contributed to, there would be other charges to levy against the defense (defamation, etc).

Also who gets to decide what constitutes serious harm?

I note "hatred" is an example given (table). What about hatred against a group that are pro-eugenics via genocide? Is expressing that hatred legitimate? Or perhaps hatred of a political party that condemns refugees seeking asylum to being imprisoned unjustly for dozens of years?

When is it justified to express extreme discontent *without* being punished for doing so? I wonder if Robodebt victims had a way to do so, would some still be alive?

What if 2 kinds of "harm" are opposing each other, yet a choice must be made in the name of pragmatism? Does that constitute a charge of disinformation even though there was no other choice but to tell a "white lie"? If so how can a punitive action be ethically justified?

the definition of digital platform services and the types of services we propose be subject to the new framework

I think this is also functionally useless.

For example, in all 3 definitions of content aggregation services, connective media services, and media sharing services, all of them use the words "*primary function*".

In that case what category does video game chat fall into? Communication between players in many games with a multiplayer online component is a *secondary function*. Yet it can still be used to communicate.

Another example:

"The new powers will cover all digital platform services highlighted above with some exclusions."

"The services carved out from the new laws are email services and media sharing services that do not have an interactive feature, such as Broadcast Video On Demand Services (BVODS) and Subscription Video On Demand Services (SVODS)."

This makes no sense, suggesting these services also cannot spread misinformation? So a fox news based BVODS couldn't possibly disseminate misinformation? [redacted]

"The Minister may also by legislative instrument specify a digital platform service to be exempt from the ACMA powers. These services are excluded services for misinformation purposes"

So anyone who pays enough lobbying money gets exempt? And this isn't anti-capitalist or anti-free speech at all? What an absolute farce.

How should instant messaging services be brought within the scope of the framework while safeguarding privacy?

It should *never* be applicable to private messaging services for *any* reason whatsoever because there is no way to have something be public (for review) and yet also maintain privacy, they are mutually exclusive.

I note this conclusion is already supported in other areas, for example how legal proceedings are conducted (attorney/client privilege, no media in order to not prejudice jurors), and anti-corruption bodies (ICAC, NACC).

Either something is digitally secure, or it is not. It is binary in nature much like the medium it uses, and there is no middle ground. Australia isn't Russia i.e. arbitrarily looking at everyone's private Telegram (app) messages, and I hope to god we never *ever* become like that.

Under the Telecommunications and Postal Services Act 1989, it is an offense to open mail (view private content) that you are not authorized to open and/or if it is not addressed to you

Technology provides certain conveniences over postal mail: speed (real-time), bi-directional use, reliability (guarantees on who it gets sent to), security (encryption), even multi-party communication between more than 2 correspondents.

But none of that should undermine the principle that without the consent of at least 1 correspondent (preferably all involved), private messaging (instant or otherwise), stays private. End of story.

Ignoring this would have far reaching consequences for cyber security, well beyond what the scope of this proposed legislation should, or could entail.

How can the digital platforms industry operationalise the Bill and various content exemptions (e.g. professional news, satire, authorized electoral content)?

They cannot. Not only because it seems to be self defeating in language and structure, but it also fails to recognize some practical elements of what the technological implementation would require.

Even with advancements in AI, the outcome of this legislation is creating a never ending game of whack-a-mole. It's not very efficient nor sustainable, and could be viewed as anti-competitive / monopolistic if we consider the compute power that would be required to launch a new service as compared to what vested interests already possess.

Overall my opinion is this bill is fundamentally the wrong approach to be tackling the problem at hand (misinformation / disinformation) as it neglects the factors contributing to its existence in the first place.

Historical Overview & Comparison

To understand what factors contribute to misinformation in the virtual environments we interact with, it's worth looking at the history that led us to this point.

Tim Berners-Lee in 1989 created the protocols that formed "the web" for the purpose of communications between universities and research entities. Examined in context of that time period, and contrast with the present day, reveals the key areas that need to be addressed.

Academia

1. Identity

Content published at that time (early 90's), were likely to be accompanied by a persons *real* identity. If a scientific discovery was made, verified, or disparaged, the author would want to be credited.

Even in the case it wasn't being used for scientific pursuits (e.g. programmers using it in the development of Linux). Due to the fact it was limited access initially (to those institutions) individuals using it had a high enough pedigree of academic excellence to understand from other aspects of their life, authorship was still important. Thus were self conscious to the point of being self motivated and consistent about it.

As such it can be summarized that in most cases: the impetus to be *forthcoming* about ones identity was present, if not using a real identity, at least being consistent.

It's true in the present some people have enough self integrity to be consistent about representing themselves accurately online, as well as financial motivations i.e. creating a consistent user name "handle" (pseudonym) on all digital services.

But it's too easy to make accounts for online services. "Bots" (robots / automated programs) can do so at the rate of hundreds per minute, and all of them aren't tied to anything other than a phone number or email (insecure systems in themselves) and can be obtained anonymously.

2. Integrity

The institutions involved (universities, etc.) had other mechanisms of ensuring integrity. For example the way science is conducted, and the way open source programming is done with commit reviews. Both have high levels of verification as a feature. If caught lying / publishing fraudulent results, authors would face quite heavy consequences to their reputation, and negative impacts to their career relative to the severity of the offense.

As such there was additional motivation to be *truthful* and correct in ones publications, but to reiterate: *This was still dependent on the fact authors were being forthcoming with their identity.* In some respects it pays homage to the philosophy of Kant: "I think, therefore, I am".

By contrast to the present day, it is simple and even profitable to make anonymous accounts to spread content that is popular with no substance, including that which constitutes misinformation.

Commercial Interests

Originally the internet began as a productive tool enabling communication of scientific / research data, academic and educational in nature, subsequently it became public domain in 1995.

In the present, the internet is profits and influence driven via money from the **advertising industry**.

The lax standards within this industry emphasizes quantity [of web traffic] at all cost over quality. Which translates as online services being unlikely to care about the actual nature of accounts (are they even real? Bots?), and may even lie or permit lying in some cases for profit:

- https://adalytics.io/blog/invalid-google-video-partner-trueview-ads
- https://www.cnbc.com/2023/04/25/amazon-reviews-are-being-written-by-ai-chatbots.html

Individual end users of said platforms ("influencers", media company accounts, etc) will be just as apathetic as some of the profits are handed off to them. That is, someone collecting revenue from an advertiser or sponsor can artificially boost the numbers for an increase in profit, for example:

- using a large number of bot accounts to "pad" the numbers.
- saying something offensive or contrarian ("click bait").

Both digital platforms and users profiting off ads are part of this self-reinforcing model driven by funds from advertising industry, and it's largely due to the fact advertising techniques, associated legalities, and policy surrounding advertisers in business has failed.

Provided financial incentive which depends on overall numbers in web traffic exists, so does the motive for misinformation.

But I also reiterate, this environment only exists because identity is fluid online.

If there were more reliable means of tying an end user account to an actual identity, would digital services and some of their users still engage trying to use "click bait" and/or bots?

Some may, but the majority I'm willing to bet would not. Because it would have a negative impact on their professional identity and social repute in reality. Not to mention there would be lasting consequences for repeat offenses for the user account in question.

Generally speaking, users do *not* want "sponsored content" shoved in their face. Ironically the reason the internet gained popularity in the first place over broadcast mediums (TV, cable, movies) was the ability to avoid ads, finding exactly the content you need.

Problem

Out of these 3 things, identity, integrity, and commercial interests, this proposed legislation to give ACMA more powers seems to only target the "integrity" area.

That is, ACMA would fill the traditional role of "ministry of education". An authoritative body that creates standards, expects all institutions under its governance follow/enforce rules, and penalize those who break them...

But as was the case before the internet became public domain, this is still dependent being able to identify users in the first place. In the state of the current internet there is no motivation or practical mechanism to identify users reliably, thus there can be no meaningful ways of enforcing policy, codes of conduct, etc.

Even the guidance note acknowledges this (guidance note *page 11/12*):

"*It can be difficult to attribute an anonymous piece of content to a person or a bot, and the powers will treat misinformation the same way no matter how it was created...*"

Yet despite this acknowledgment *there is no practical method* offered in solving it? So who is being held accountable? The digital service in question? That seems ethically wrong, since while they do enable users with this capability, they are not responsible for how users themselves behave. Or are we going to start applying this principle universally and penalizing car manufacturers every time there's an incident on the road?

Even in the case a code of conduct is developed by ACMA and/or digital platforms, in the case of a violation and a user or group of users accounts are penalized with a ban. They can simply make a different account? Which means all of this amounts to nothing?

All 3 areas need to be addressed if misinformation for digital services is going to be dealt with and have a net positive outcome. The contents of this bill while perhaps should be assessed in future, should *not* be the focus in the present. We should be addressing the other 2 areas first (identity, advertising).

Why? Because currently we can't even clearly define the shape of the actual problem, that is, there's no way to observe the true nature and scope of misinformation / disinformation.

This is mostly due to bots. How are we to know if it's 100,000 real users propagating misinformation, or 1 person with 99,999 bot accounts?

Eliminating any and all potential bot accounts *first* for those digital services with "high discoverability" operating within Australia, and gaining an actual appreciation for the shape and scope of the issue, will affect how it is addressed.

It's true that perhaps ACMA may need new powers, but this is putting the cart before the horse. It could also be that sigint just needs one or two new directives to operate under.

Solution Part 1: Australian Identity System (AIS)

Digital services need to have some authoritative means of identifying users, without relying on a mechanism as weak as a phone number or email.

How do we know it'll work? Even without the deconstructions and scientific tests of Orwell's 1984, we know that all thinking agents (even pets) behave differently not just if they're being observed, but if they *think* they're being observed.

If AIS is set up correctly (see subsequent details), I believe large improvements in digital hygiene, etiquette and civility in general can be achieved, *without* government having to legislate explicitly regarding content.

It is an indirect but much more sophisticated approach.

AIS Policy Direction

It is absolutely critical that AIS is *not imposed* on all online service providers by default.

First because it'd be impossible anyway, No country has the influence to coerce the rest of the internet to follow them, and even in the case they could, new protocols would be developed in response (example veilid.org).

I recommend this be a domain based approach. Make the root level domain (.au) available, then mandate that any service using this domain name (including 2nd and 3rd level domains .com.au .edu.au .gov.au etc) *must* implement AIS for user identification. This could be verified as part of the SSL certificate issuing process.

In this way people can know simply be observing the domain name if the digital service has AIS implemented. Native apps (phone or otherwise) would also need to mention this domain and/or have some marker for identification.

But second and more important, freedom of speech is dependent on the existence of privacy. That is, if one wants to express themselves publicly that also comes with the price of being identified. If one wants to maintain their privacy (and possibly safety) whilst *still* being able to express themselves freely, that necessarily requires the capability to maintain anonymity.

If a digital platform wishes to permit accounts representing non-human entities (e.g. companies), or even anonymous accounts, this will be down to company policy in how it's handled. But at the very least as a condition for being able to use AIS, a significant distinction between an AIS account and non-AIS account (for the purposes of people) should be mandated.

All that said, even though AIS would be voluntary, as a digital service grows (in the number of end users / traffic) it is highly likely service providers themselves would *want* to implement AIS once certain thresholds are reached.

There are many impetus for doing so, including more accurate internal analytics for themselves, but the biggest one is simple to understand. Operating costs.

The ability to check a key, guaranteeing a person can only have 1 user account, and subsequently enforcing codes of conduct that would have permanent consequences in their violation. As compared to;

Moderating all content all the time, that could come from repeat offenders who have multiple anonymous accounts or even bot networks at their disposal, that can (currently) be created without limit.

The latter is *extremely* difficult and expensive($) from development, to implementation and operation, and even with the capabilities of Ai will not be 100% effective.

While some would object, the vast majority of users I suspect would appreciate this guarantee.

People want to be apprised of *people*. The reason why digital platforms such as twitter and facebook gained such popularity is this mechanism to connect with other people in a more direct manner, *not* via media outlets or PR departments.

Being able to provide guarantees to Australians that an account on a digital platform with an Australian domain name (ending in .au) does in fact belong to a person that has been verified to a high level of confidence is an extremely *valuable* proposition. Particularly true for digital services that facilitate events in reality (e.g. dates, hobby groups, etc).

I believe the benefits of AIS outweigh the risks, and it would set Australians and the Australian digital economy up for decades to come with relatively minimal expense from government.

User accounts will be valued more, with an unwillingness to lose or misuse them (including spreading misinformation) as users could only have 1 account with an AIS supported service. That is, losing access for whatever reason (violation of platforms code of conduct), there would be no possibility of re-registering under a different account. A user would have to “appeal” the ban and justify why their account should be reinstated.

Bots for Australian digital services can be eliminated virtually overnight. The net impact of this on misinformation would be significant. Often times bots are employed to artificially boost such content into the foreground of results, searches, algorithms for home feeds, etc. The fact a piece of content has a lot of “likes” or is “re-shared” tends to give it credibility when it should not. Yet without bot accounts, it’s likely much of this misinformation would never gain the initial traction required for it to “go viral” and bring it into the public consciousness at large.

AIS Operational Overview

In my professional opinion (programmer + network engineer + tech aficionado) AIS should operate based on an asymmetric protocol such as Gnu Privacy Guard (GPG) if not employ it directly.

I'll be grossly oversimplifying for brevity in this description, but any sufficiently senior level programmer / system administrator / security expert should be able to expound further, as GPG has been a reliable mechanism in use for at least 3 decades.

An Australian citizen can generate an asymmetric primary key pair (public / secret). The public portion is registered with a government run service, the secret key must never be disclosed to anyone. Note the scheme will need updating as "noiseless" quantum computers with large numbers of qubits become available in future.

Registration and/or fixing issues with a registered primary public key should use the already established and mature "point system" to prove one's identity (e.g. birth certificate, drivers license, bank statements, passport, etc).

This would require going into an authorized venue (post office, RTA, centerlink) and only these venues should have any access to perform (create / update / revoke) operations on the public key data bank. As such they should be a fortress in terms of security, and undergo regular audits.

The primary key pair can also be used to generate sub key pairs (also public / secret format). A digital service may request a public sub key be provided when a user registers, which can be verified against the public key listed on the government service, thereby guaranteeing a users identity, and most importantly can ensure each person only has 1 account for a participating service.

It also means citizens will be responsible in part for their own digital security (i.e. keeping the secret keys, a secret). It's for this reason I suggest using the point system and venues for management, as there needs to be an authoritative short-circuit to "reset an identity" in the unlikely event the secret part of a primary key is compromised.

Software can automate much of all this, in fact asymmetric mechanisms are used by some of the recent standards of security protocols (FIDO):

https://fidoalliance.org/

AIS Threats to Freedom of Speech / Privacy

There is some danger of AIS being a threat to freedom of speech.

For example revoking the public primary key or even simply denying access to it (depending on individual company policy), would be catastrophic for an end users freedom of speech as they wouldn't be able to register for anything new, or could also be prohibited to from posting to digital services they had already registered to.

There is also some danger of AIS being a threat to privacy. Since there's a relationship between primary and sub keys, logging details on which services validate sub keys, as well as gaining additional information about users from them could allow someone to map an individuals internet footprint with accuracy that could be qualified as invasive.

This is why we'd also need extremely well defined and strong protections in the legislation defining AIS and preventing its misuse. For example:

- AIS must *not* be mandated for use on all digital services operating within Australia.

- AIS must not store sub key backlinks with primary keys, instead using a multi-challenge based model.

- AIS should only store the minimal amount of detail necessary in a primary key e.g. keystring, name, contact details, social status (warrants, criminal charges), etc. and certainly no details that would be adversely affect users and/or *cannot* be changed easily (e.g. subkey backlinks, age, photo, blood type, etc).

- The government API facing the public internet should only respond to validation requests. There should be no form of request logging.

- AIS should be federated in that, while government may run the primary key servers, other entities may mirror them for redundancy, performance, and versioning purposes.

- No digital service (including government run ones) should be able to use the primary public key itself. They *must* all use a sub key generated by the users.

Etc

Generally speaking AIS would be the manifestation of a trilateral bond of trust between government, commercial digital services, and the people. As such I believe this bond should be well defined in whatever legislation is conceived.

I'm sure privacy advocates and security experts would also like to weigh in and give their professional views on details of the architecture, threat modeling, etc.

To this end an independent process should be initiated, with submissions to be evaluated by experts in the field as well as those designing the system, and it should *not* be anyone from a consulting firm (PwC, Deloitte, Ernst & Young, KPMG, etc). This *must* be a government initiative the stakes are too high for it to be done any other way.

For those ornery, anti-authoritative naysayers arguing that this would be equivalent to government surveillance.

If implemented poorly I agree it could be catastrophic. But it should be mentioned Government *already* does this kind of thing: passports, car registration, proof of identity cards, centerlink, opal cards, firearm registration, medicare numbers, tax file numbers / bank accounts, etc, etc.

If implemented properly as detailed, this is actually *an improvement* to security for Australians, while simultaneously addressing the problem of misinformation.

Solution Part 2: Advertising Reforms

Why they're needed

Advertisers are still using the tactics of broadcast mediums. "Force feeding" people ads, when in reality people become disenfranchised if not annoyed. Ad-blocking software (Pi-Hole, browser extensions, etc) is not without popularity *for a reason*.

These marketing tactics can be viewed as an exploitation of the "rule of reciprocity" a psychological trait in all people that comes from us being a social species, and guarantees all of us implicitly understand the concept of "debt".

Primitive humans 20,000 years ago
Person A: You can sit by my fire to keep warm.
Person B: Thanks, I owe you... Here's some food I hunted.

Digital services
Service: We're providing this service Z, this is what it does.
User: Wow neat! Z's great. Oh ads?... you can have my time.

The problem is *time* is a true non-fungible asset, it can't be equated to anything, and once it's gone you can't get it back.

In the context of both advertising and misinformation, this equates to the rule of reciprocity being abused. Because it's facilitating what should be considered an "illegal trade" i.e. people are giving up something they can *never* get back.

In the case of advertising, people don't want to view advertisements (often video), yet much of the time they are shanghaied into it, as the content / information they actually want to access is obscured by said advertisement until it has fully been presented.

In the case of misinformation, the offense is even more egregious as people must waste *even more* time having to verify and/or disprove the content in question, and that's assuming they have any doubt of it in the first place.

It may be argued that this is voluntary, after all if you don't want exposure to ads, or be mislead by misinformation, just don't use the digital service, right?

Perhaps this was the case initially, but given how much digital platforms permeate society now, and how centralized they become in order to facilitate high discoverability, I don't think participation on such platforms is voluntary any longer.

This is also compounded through other mechanisms such as "biased algorithms" i.e. ways digital platforms can inject sponsored content *without* informing users, resulting in such content (including sponsored misinformation) "bubbling to the top" of search results and home feeds, while other content is pushed down as a result, despite the fact the latter may have better (more truthful / accurate) substance.

Precedent

In the United States there is precedent for dealing with such predatory marketing tactics. Hare Krishna's would (almost aggressively) get you to take a flower, then ask for a donation.

You would have to sneak around through airports as nobody actually wanted the flower. Yet once it's in your hands it becomes more difficult to refuse the donation request because reciprocity is so "hard-wired" into the human psyche.

Consequently solicitations of this kind were banned by multiple US port authorities beginning in the 1980's, and appeals by ISKCON have (mostly) failed.

Society does not like being taken advantage of in this way, because it undermines the fundamental confidence we have in our fellow humans.

Reforms

I think limiting advertising techniques that may be used on platforms with high discoverability in combination with AIS, should have a net positive outcome, both for the general public in terms of UX, and for stemming the tide if not eliminating misinformation altogether.

For example, consider if there was a ban on the ability of social platforms to use biased algorithms i.e. no manipulation of users home feeds or content being displayed at all. The only forms of advertising would be either product placement and/or endorsement *by users* (content creators).

Even in the case an entity wanted to promote misinformation, they'd have to find a content creator willing to put their professional and social reputation at stake to do it (since AIS ties user accounts to real identities).

In the case it still happens, the only way that content creator would be able to escape the consequences, would be to become an ex-pat, leave the country and assume a new identity in reality. Which is not something easy to do, nor would it completely erase their connection to the promoted misinformation if it was done in a half-hearted manner (e.g. they kept the same appearance).

Digital platforms would be losing out in terms of revenue, so things would have to be re-balanced perhaps via contractual obligations. For example if a product or service wants to advertise through a user on a digital platform, an obligatory fee + commission (%) goes to that platform as part of the transaction. Which is yet another reason advertisers would have to be more selective about who they choose to represent their products and services (someone truthful) and not simply grant money to platforms hand over fist.

I think that this should be part of a broader reform for a legal framework around advertising conditions in general (including political ads).

Secondary Outcomes

Eliminating some cyber attacks

A message encrypted with a sub key (provided on user registration) can only be decrypted with the corresponding secret key. Therefore provided the initial key exchange has integrity (no man in the middle), any message claiming to be from a digital service which has *not* been encrypted, can immediately be discarded.

This would eliminate a significant number of phishing type attacks (i.e. sender pretending to be a trusted entity).

Minors vs Adult Content

It would be possible to address the issue of underage persons being exposed to adult content. Consider if instead of one government key service it was split virtually in two. Same software architecture and hardware but with a virtually "sharded" key data bank. One part for underage, one for overage (of consent).

Any adult digital services that wish to operate in Australia would have to implement AIS verifying a sub key against the *overage* key shard. Underage sub keys would fail to work thus denying access, and without having to store or specify an actual age number with keys. This would require a person having to refresh their public primary key once the age of consent is reached.

It wouldn't completely eliminate the problem. But combined with other measures, some implemented by government and ISP's, some implemented by parents. It should be possible to significantly reduce the footprint of situations where this is ever an issue.