

COMBATING FAKE NEWS ONLINE:

Exploring Australian Co-regulatory Framework For Malaysia

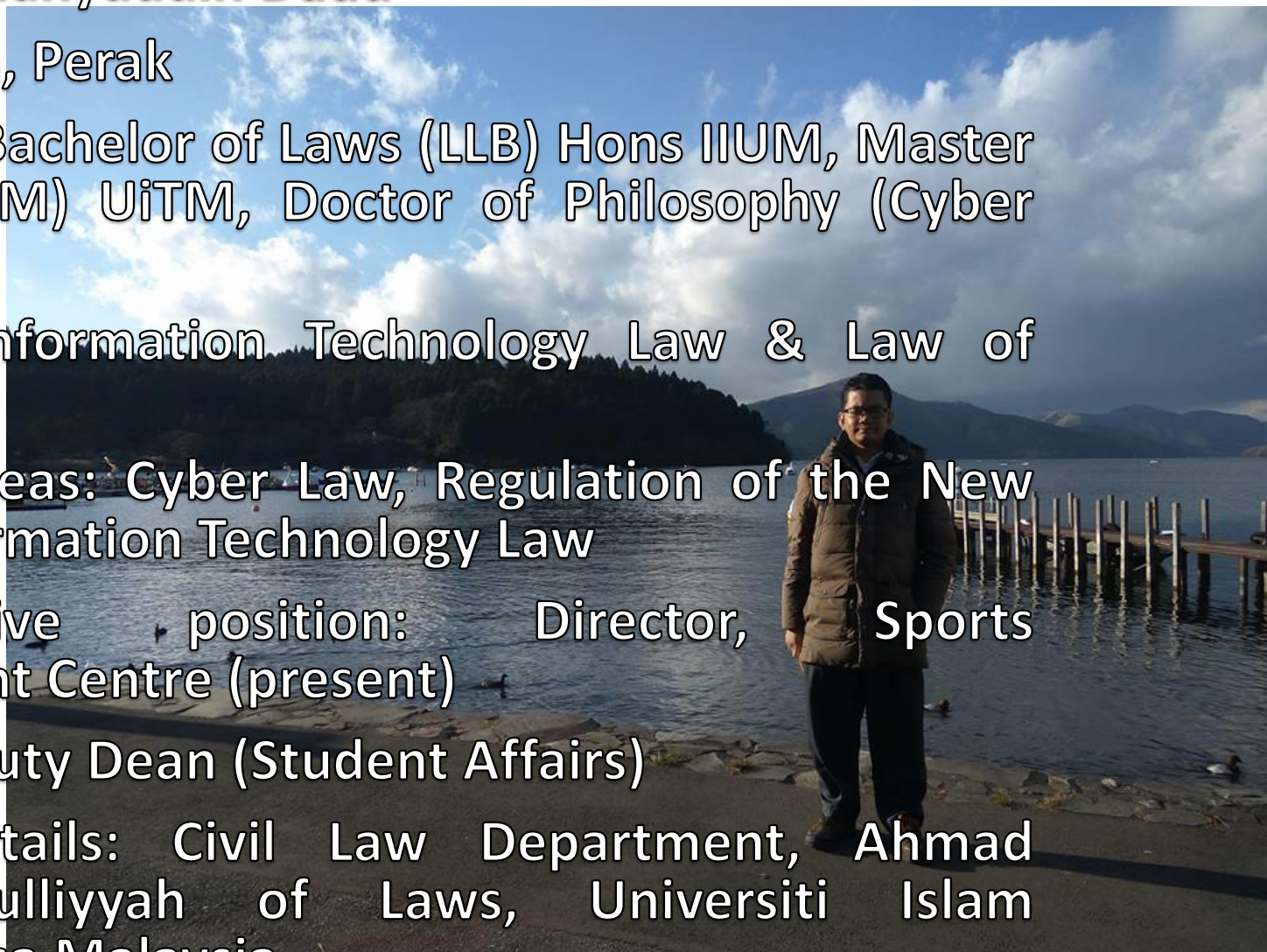
Asst. Prof. Dr. Mahyuddin Daud

Department of Civil Laws

Ahmad Ibrahim Kulliyyah of Laws

International Islamic University Malaysia

- Name: Dr. Mahyuddin Daud
- Born in Ipoh, Perak
- Education: Bachelor of Laws (LLB) Hons IIUM, Master of Laws (LLM) UiTM, Doctor of Philosophy (Cyber Law) IIUM
- Teaching: Information Technology Law & Law of Contracts
- Research areas: Cyber Law, Regulation of the New Media, Information Technology Law
- Administrative position: Director, Sports Development Centre (present)
- Former Deputy Dean (Student Affairs)
- Contact details: Civil Law Department, Ahmad Ibrahim Kulliyyah of Laws, Universiti Islam Antarabangsa Malaysia
- Email: mahyuddin@iium.edu.my



To analyse development on regulation of false content online

Malaysia

S211 &
S233 CMA
1998

Introduc
tion of a
website –
'Sebenarn
ya.my'

WhatsApp
group
admin to
monitor
false
content

Anti-Fake
News Act
2018

International

Joint Declaration on Freedom of
Expression and 'Fake News',
Disinformation and Propaganda

Countermeasures
by social media
platforms

Facebook

YouTube

Australian
co-regulation

Legal Framework for Internet Content Regulation in Malaysia

- Section 124 (e) of CMA 1998 entrusted the MCMC to monitor development and practice of IT industry self-regulation in Malaysia.
- What is self-regulation?
- No definition for Self-regulation in CMA1998 and Content Code

Issues



Fake news has been on the rise

Previous government used law & technology to restrict access to 'illegal' content

Before & after GE14, more fake news, ministers were busy correcting misstatements

Legal frameworks

International

United Nations Joint Declaration on
Freedom of Expression and 'Fake News',
Disinformation and Propaganda

National

Communications & Multimedia Act 1998

Content Code

Anti-Fake News Act 2018

Original: ENGLISH



JOINT DECLARATION ON FREEDOM OF EXPRESSION AND “FAKE NEWS”, DISINFORMATION AND PROPAGANDA

The United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information,

Having discussed these issues together with the assistance of ARTICLE 19 and the Centre for Law and Democracy (CLD);

Recalling and reaffirming our Joint Declarations of 26 November 1999, 30 November 2000, 20 November 2001, 10 December 2002, 18 December 2003, 6 December 2004, 21 December 2005, 19 December 2006, 12 December 2007, 10 December 2008, 15 May 2009, 3 February 2010, 1 June 2011, 25 June 2012, 4 May 2013, 6 May 2014, 4 May 2015 and 4 May 2016;

Taking note of the growing prevalence of disinformation (sometimes referred to as “false” or “fake news”) and propaganda in legacy and social media, fuelled by both States and non-State actors, and the various harms to which they may be a contributing factor or primary cause;

Expressing concern that disinformation and propaganda are often designed and implemented so as to mislead a population, as well as to interfere with the public's right to know and the right of individuals to seek and receive, as well as to impart, information and ideas of all kinds, regardless of frontiers, protected under international legal guarantees of the rights to freedom of expression and to hold opinions;

- **1. General Principles:**

- a. *States may only impose restrictions on the right to freedom of expression in accordance with the test for such restrictions under international law*, namely that they be provided for by law, serve one of the legitimate interests recognised under international law, and be necessary and proportionate to protect that interest.
- b. *Restrictions on freedom of expression may also be imposed, as long as they are consistent with the requirements* noted in paragraph 1(a), to prohibit advocacy of hatred on protected grounds that constitutes incitement to violence, discrimination or hostility (in accordance with Article 20(2) of the International Covenant on Civil and Political Rights).
- c. *The standards outlined in paragraphs 1(a) and (b) apply regardless of frontiers so as to limit restrictions not only within a jurisdiction* but also those which affect media outlets and other communications systems operating from outside of the jurisdiction of a State as well as those reaching populations in States other than the State of origin.
- d. *Intermediaries should never be liable for any third party content* relating to those services unless they specifically intervene in that content or refuse to obey an order adopted in accordance with due process guarantees by an independent, impartial, authoritative oversight body (such as a court) to remove it and they have the technical capacity to do that.

- E. **Consideration should be given to protecting individuals against liability for merely redistributing or promoting, through intermediaries,** content of which they are not the author and which they have not modified.
- f. **State mandated blocking of entire websites, IP addresses, ports or network protocols is an extreme measure which can only be justified where it is provided by law and is necessary** to protect a human right or other legitimate public interest, including in the sense of that it is proportionate, there are no less intrusive alternative measures which would protect the interest and it respects minimum due process guarantees.
- g. **Content filtering systems which are imposed by a government and which are not end-user controlled are not justifiable** as a restriction on freedom of expression.

CMA 1998

Content Requirements

Prohibition on provision of offensive content

211. (1) No content applications service provider, or other person using a content applications service, shall provide content which is indecent, obscene, false, menacing, or offensive in character with intent to annoy, abuse, threaten or harass any person.

(2) A person who contravenes subsection (1) commits an offence and shall, on conviction, be liable to a fine not exceeding fifty thousand ringgit or to imprisonment for a term not exceeding one year or to both and shall also be liable to a further fine of one thousand ringgit for every day or part of a day during which the offence is continued after conviction.

Section 233 CMA 1998

Improper use of network facilities or network service, *etc.*

233. (1) A person who—

(a) by means of any network facilities or network service or applications service knowingly—

(i) makes, creates or solicits; and

(ii) initiates the transmission of,

any comment, request, suggestion or other communication which is obscene, indecent, false, menacing or offensive in character with intent to annoy, abuse, threaten or harass another person; or

Malaysia's proposed Anti-Fake News Act 2018

What you really should know

Fake news — wholly or partly false news, information, data, reports

(Note: Not limited to social media, can be any written publication)

Law applies even if offence committed outside Malaysia by Malaysian/ foreigner

– **IF** fake news is about Malaysia/ affects a Malaysian.

Offences

Knowingly create, publish, distribute fake news

- Max RM500k fine / max 10-year jail / both.
- Max daily RM3k fine if offence continues after conviction.
- Punishable contempt of court if no apology as ordered by court.

Financial aid, intending/ knowing/ reasonable grounds to believe funds will be used for fake news

- Max RM500k fine / max 10-year jail / both



Failure to remove after knowing/ reasonable grounds to believe it is fake news

- Max RM100k fine
- Max daily RM3k fine if offence continues after conviction



If court orders removal of fake news, failure to comply:

- Max RM100k fine
- (Note: Court order can set time limit, can be served personally/ via post/ email/ social media account)
- Can apply to set aside court order, but it's not a stay of removal order.
- (Note: Can't apply to set aside if govt gets court order to remove fake news prejudicial to public order/ national security)
- Court can order police / authorised officer to remove publication if person fails to comply

Abetment (assisting) in any of the above offences

- Same punishment for actual offence

FAKE NEWS



Note:

- Police can arrest without warrant for all offences
- If offender is a body corporate (e.g. company), the director/ CEO/ manager/ secretary
 - can be charged severally or jointly in same trial
 - deemed guilty if body corporate is convicted, **UNLESS** he proves offence committed without his knowledge/ consent + had taken all reasonable precautions

Content Code

7.0 False Content

7.1 Content, which contains false material and is likely to mislead, due amongst others to incomplete information is to be avoided. Content providers must observe measures outlined in specific parts of this Code to limit the likelihood of perpetuating untruths via the communication of false content.

7.2 Content is false where prior to communications reasonable measures to verify its truth have not been adopted or taken.

7.3 Content which is false, is expressly prohibited except in any of the following circumstances:

- (a) Satire and parody;
- (b) Where it is clear to an ordinary user that the content is fiction.

7.4 Code Subjects must take all necessary steps outlined in the specific parts of this Code to limit the likelihood of provision of false Content.

Sebenarnya.my

SALUR KEPADA KAMI TENTANG KAMI

SEBENARNYA.MY
Tidak Pasti Jangan Kongsi

BUKAN SEMUANYA BENAR DI INTERNET!
TIDAK PASTI, JANGAN KONGSI

- ✓ Sentiasa periksa maklumat sebelum dikongsi.
- ✓ Saputlah maklumat dari sumber yang tepat, sahih.
- ✓ Legali sebenarnya.my untuk mengesahkan kesahihan maklumat.



UTAMA

NASIONAL ▾

SOSIAL ▾

INFO ▾

AKTIVITI



BERITA PILIHAN

300,000 MASYARAKAT INDIA DI
MALAYSIA TIDAK MEMPUYAI
KAD PENGENALAN?

TULIS: Februari, 2018

UNDANG-UNDANG BERITA
PALSU BAKAL DIGUBAL BAGI
TUJUAN PILIHAN RAYA UMUM
KE-14?

TULIS: Februari, 2018

KWSP TERLIBAT PERJANJIAN
MERAGUKAN DENGAN
BEBERAPA SYARIKAT KEWANGAN
ANTARABANGSA?

TULIS: Februari, 2018

LTAT HADAPI MASALAH
KEWANGAN UNTUK BAYAR
GRATUITI KEPADA PESARA
TENTERA?

TULIS: Februari, 2018

WASPADA PENYEBARAN
TARIKH PALSU
BERKENAAN PRU 14

TULIS: Januari, 2018

Impose liability on WhatsApp Group Administrator to Monitor False Content



- The government has also taken a step to curb dissemination of false content online through **online advisory warnings**, especially towards group administrators of mobile apps such as WhatsApp.
- The MCMC has issued an advisory for group administrators reminding them the Do's and Don'ts – which could also be understood as imposing certain responsibilities to monitor digital interactions happening within WhatsApp group.

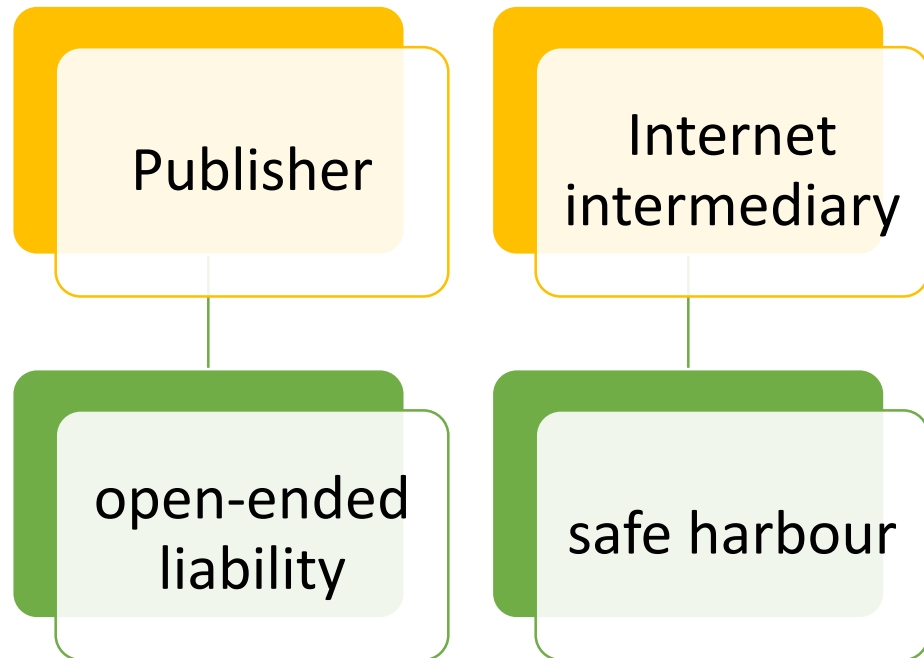


PERINGATAN UNTUK PENTADBIR KUMPULAN ADVISORY FOR GROUP ADMINS

PATUT DILAKUKAN/ DO	TIDAK PATUT DILAKUKAN/ DON'T
<p>Maklumkan kepada semua ahli kumpulan, sebab utama kumpulan aplikasi pesanan diwujudkan dan memikirkan samada peraturan asas diperlukan</p> <p><i>Inform members the reason for setting up the group and consider setting basic ground rules/</i></p>	<p>Menjadi pentadbir yang tidak aktif/senyap</p> <p><i>Be an absent administrator</i></p>
<p>Prihatin terhadap peraturan/ syarat-syarat yang digariskan oleh penyedia platform</p> <p><i>Be mindful of rules/conditions of use by platform provider</i></p>	<p>Melanggar standard komuniti/ garis panduan yang telah ditetapkan oleh penyedia platform (contohnya Facebook, Whatsapp, Telegram)</p> <p><i>Breach community standards/guidelines set by the platform provider (eg Facebook, Whatsapp, Telegram)</i></p>
<p>Menegur posting/komen untuk memastikan perbincangan berada di landasan yang betul</p> <p><i>Comment on posts to ensure discussions stay on track</i></p>	<p>Menggalak, menghasut atau bersubahat terhadap posting/komen yang jelik</p> <p><i>Encourage, incite or abet inappropriate posts</i></p>
<p>Sentiasa semak posting/ komen (dengan seorang moderator, jika boleh)</p> <p><i>Check posts regularly (possibly with a moderator)</i></p>	<p>Membenarkan perbincangan keluar daripada topik</p> <p><i>Allow discussions to wander off topic</i></p>
<p>Mempertimbangkan untuk mengeluarkan atau menyekat mereka yang terus membuat posting/ komen jelik</p> <p><i>Consider removing or blocking those who persist in making inappropriate posts</i></p>	<p>Takut untuk bertegas dengan troll atau dengan individu yang secara berterusan melanggar peraturan asas</p> <p><i>Be afraid to deal firmly with trolls or individuals who continuously breach the ground rules</i></p>
<p>Bekerjasama dengan agensi penguatkuasaan undang-undang yang menjalankan siasatan</p> <p><i>Cooperate in investigations by law enforcement agencies</i></p>	<p>Mengganggu/ menghalang siasatan yang sedang dijalankan oleh agensi penguatkuasaan undang-undang</p> <p><i>Hinder/obstruct investigations by law enforcement agencies</i></p>

How can liability be imposed to WhatsApp group admin under the self-regulation framework?

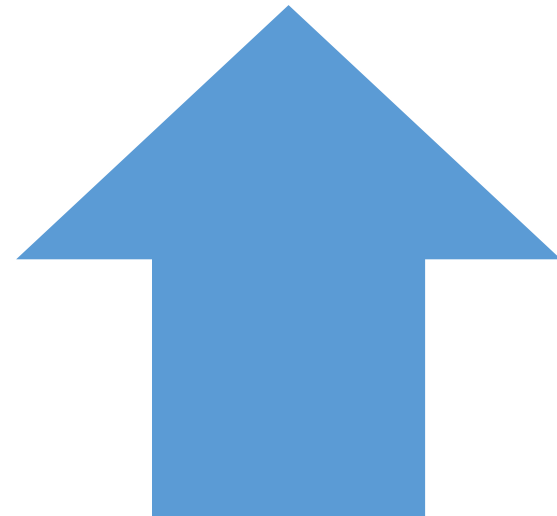
- Should WhatsApp group admin be considered a **publisher OR an Internet intermediary?**
- **See Section 114A of Evidence Act 1950**
- Part 5 Content Code





What is the position of a WhatsApp group administrator – who is an Internet user that contributes online content and at the same time, administers the online chat group

Would it be fair and reasonable to equate him similar to a publisher and impose liability on WhatsApp group administrator to monitor content in his group – failing which, may lead to legal consequences?



- If we were to consider the meaning of ‘publisher’ from Section 114A, a WhatsApp group administrator is arguably in the position to host, or becomes an administrator, or may indirectly facilitates to publish or re-publish content.
- However, he is not in the position to edit nor sub-edit the content that is contributed in the WhatsApp group that he administers.
- The best technical measure to prevent the spread of false content which a WhatsApp group admin may do is to delete any false or illegal content communicated by his group members. This may also be a self-regulatory instrument designed by WhatsApp developer to empower users to prevent false content from circulating further, which is a timely effort.
- Nevertheless, this does not absolve the admin from liability under Section 114A, to which he must then prove to court that he has taken initiative to delete or takedown the false content – to avoid from being presumed as a publisher under this pretext.


You can delete messages for everyone or just for yourself.

To delete messages for everyone

Deleting messages for everyone allows you to delete specific messages you have sent to either a group or an individual chat. This is particularly useful if you sent a message to the wrong chat or if the message you sent contains a mistake.

Messages you successfully delete for everyone will be replaced with *"This message was deleted"* in your recipients' chats (*). Similarly, if you see *"This message was deleted"* in a chat, it means that the sender deleted their message for everyone.

You can only delete messages for everyone for up to seven minutes after sending. Once seven minutes have passed, there is no way to delete messages for everyone. To delete messages for everyone:

1. Open WhatsApp and go to the chat containing the message you wish to delete.
2. Tap and hold the message. Optionally, tap more messages to delete multiple messages at once.
3. Tap **Delete**  at the top of the screen > **Delete for everyone**.

Other countries

Initiatives

Enactment of
fake news
legislations

Imposing
liability on
intermediaries

Social
awareness &
engaging 3rd
party fact-
checkers

Comparative perspectives...

- On the other part of the world, the United States have had their Presidential Election in 2016 which favours Donald Trump (The Organisation for Economic Co-operation and Development, 2010). Behind this election, there were many fake news published online and offline – and most of them sided on Trump.
- United States government's attitude in dealing with dissemination of false news have been a passive one – perhaps in honour of the right to free speech under the First Amendment of the US Constitution.
- As a result, studies confirmed that one of the factors leading to Trump's winning the election was because of influence from fake news (ARTICLE19, 2013).
- Facebook and Google have made efforts to remove false content on their platforms for violating its policies on objectionable and illegal content. Further, Facebook users may **flag** any false content as 'disputed by 3rd party fact-checker'.

- Other governments such as the United Kingdom and Russia have set-up website to list and verify any false content about the nation.
- Germany is preparing a bill to fight against the dissemination of fake news – including to fine social media sites for failure to promptly remove false content(Gu, Kropotov, Yarochnik, et al., 2017).
- The European Council established the European Union's External Action Service (EEAS) to review 'disinformation' content on weekly basis(Daud, 2016).

Countermeasures by social media platforms



excludes itself from any liability over 3rd party content
third-party fact checkers

doesn't remove false news, but significantly reduces its distribution by showing it lower in News Feed through machine learning

the company believes that it is important to empower netizens to decide "what to read, trust and share by informing them with more context and promoting news literacy".



Imposes the 'Community Guidelines Strikes'

YouTube Community Guideline does not ban 'fake news' alone, but is committed to ensure that the platform is free from spam, scams, and other deceptive practices .

Any users applied misleading metadata – such as misleading tags, titles or thumbnails that intend to boost the number of viewers, may cause content removal.

Australian responses on fake news

- Australia is in the early stages of responding to fake news and disinformation.
- The main Australian Government response so far has been the **creation of a taskforce to address threats to electoral integrity**, though the foreign interference laws, which passed the Parliament in June 2018, also have some relevance to the issue.
- In addition, the Australian Electoral Commission (AEC) commenced a social media literacy campaign and other activities to coincide with the 2019 federal election.
- There have also been several recent parliamentary inquiries and an inquiry by the Australian Competition and Consumer Commission (ACCC) examining issues related to fake news.

CO-REGULATION IN AUSTRALIA

Strong partnership between government, industry actors, and Internet users.

Internet industry develops its own code of practice, accreditation, or content rating schemes

Supported by government enforcement and statutes.

Regulatory measures

National Classification Scheme

Classifications
(Publications,
Films and
Computer
Games) Act 1995

National
Classification
Code

Guidelines for the
Classification of
Publication and
Guidelines for the
Classification of
Films and
Computer Games

Non-regulatory measures

Education
&
awareness

Hotlines

ISP
filtering



Australian Government
Attorney-General's Department

Use Classifications to Choose Computer Games and Films

Very mild impact



General

Mild impact



Parental
guidance
recommended

Moderate impact

Not recommended for people
under 15



Recommended
for mature
audiences



Not suitable for people
under 15. Under 15s must
be accompanied by a
parent or adult guardian

AGE
RESTRICTIONS
APPLY



Restricted
to 18 and over

Check the Classification

Strong impact

People under 15 must be accompanied
by a parent or adult guardian to hire or
buy these films or games or to see these
films in a cinema. These games cannot be
demonstrated in a public place

High impact

People under 18 are not permitted to
buy or hire these films or games or to
see these films in a cinema. These games
cannot be demonstrated in a public place

Films or computer games can be
advertised prior to being classified.
Check the classification at the cinema
or before hiring or buying films or
computer games

www.classification.gov.au or (02) 9289 7100

Recommendations

Co-regulation in Australia is streamlined towards regulation of content risks through classification

Australian co-regulation should be studied in detail as promising legal framework regulating the Internet in Malaysia.

Malaysia should to mandate content and service providers to classify and filter online content.

Future works should also involve development of a national classification scheme.

Recommendations



Internet industry should look into possibility to design a certified national filter.

If all parties concerned play more proactive roles, regulatory burden on the MCMC could be reduced.

In the end, we can develop more responsible netizens that cares for children online safety.

Findings on the issue of fake news

fake news
remains to be
visible and
readable
online

editorship
mechanism
was not
properly
developed
and
transparent
leading to
criticisms on
restriction of
freedom of
expression
and
information.

However, with
fake news
being a global
threat, the
status quo
cannot
withstand for
long.

Affirmative
action needs
to be taken in
order to face
the issues and
challenges
posed by fake
news – and
that could
start with
Internet
editorship or
classification

Conclusion and Recommendation

Self-regulation as practised in Malaysia needs to be **improved** to curb the spread of fake news online

In view of the initiatives to regulate false content, we must ensure that **constitutional right to freedom of speech & expression under Article 10 FC is safeguarded.**

This should not involve protecting those who spread false information – false content is not a valid form of protected expression

A balance must be struck between regulation and freedom in this sense, so that we will not be accused of committing censorship of information in the digital age.

Thank you for your attention

Kindly forward any queries to
mahyuddin@iium.edu.my