

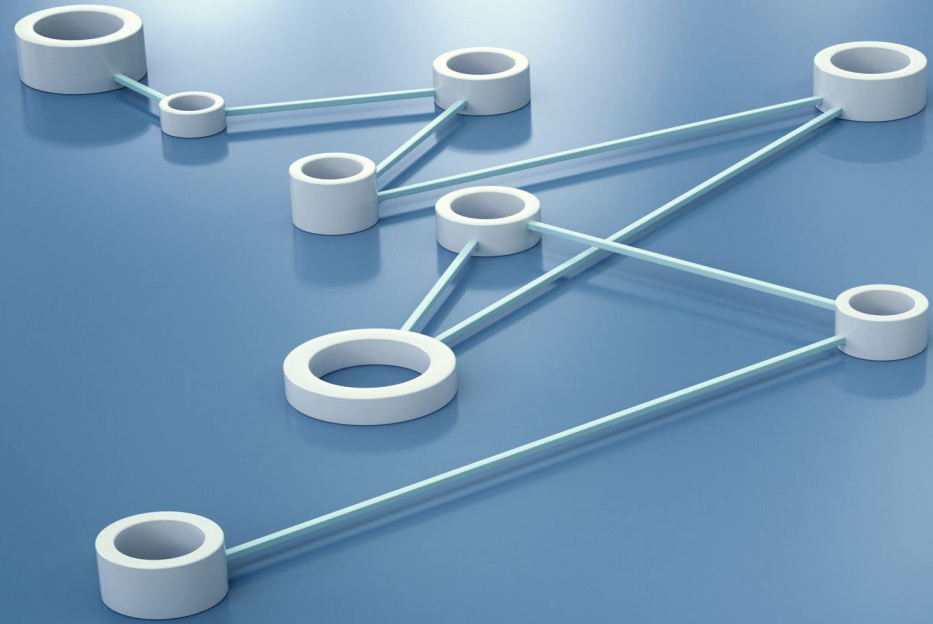
---

# Legal and Policy Choices in Making Platforms Responsible for Online Mis-Information

iGYRO Symposium - Innovation 4.0

Professor Jeannie Marie Paterson

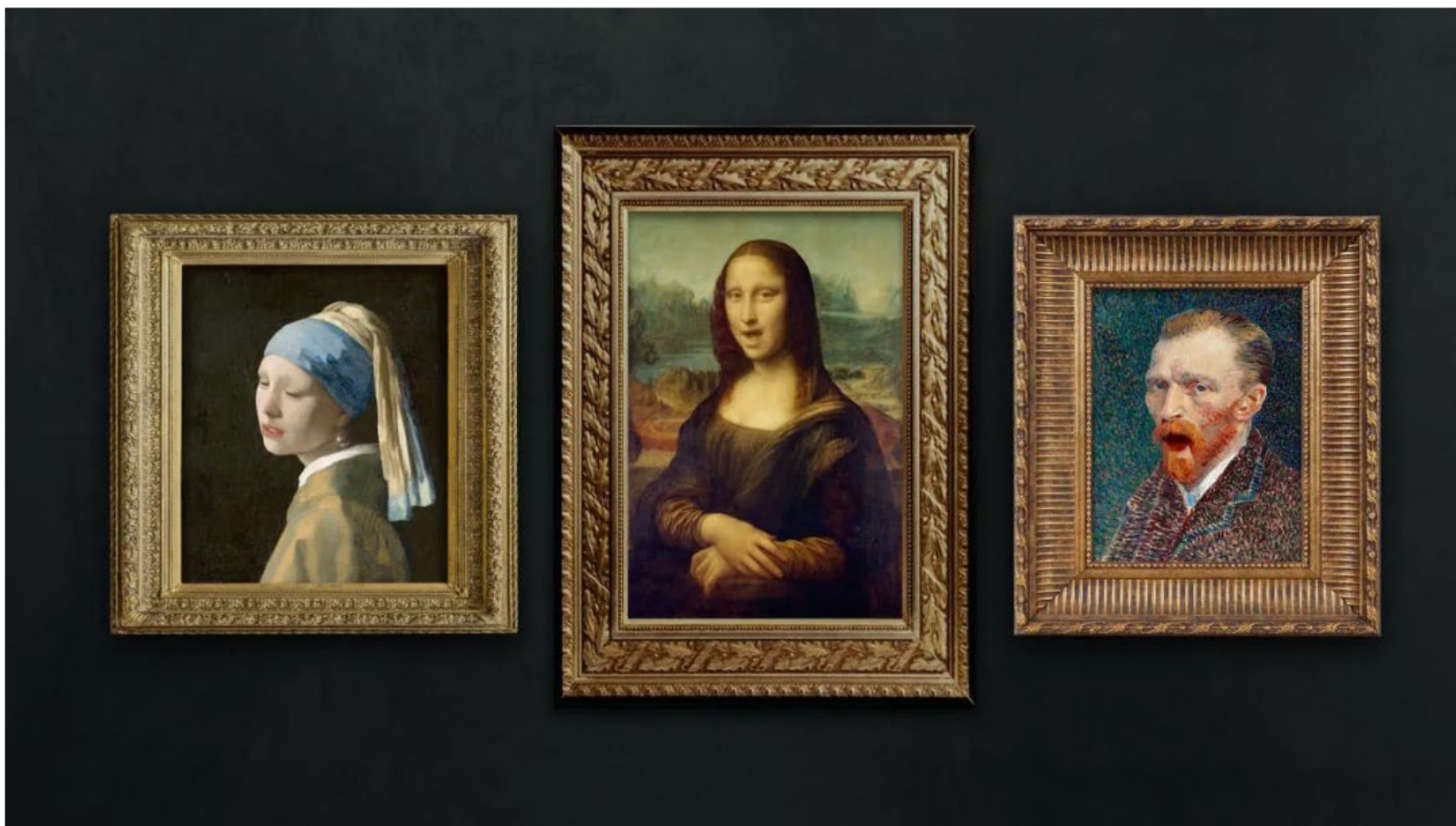
University of Melbourne



THE UNIVERSITY OF  
**MELBOURNE**

---

**Centre for  
Artificial Intelligence  
and Digital Ethics**



This is amusing – but not all AI images are...

Concerns about  
misinformation  
they have been  
amplified by AI

# Imagen

Our highest quality text-to-image model

Try in Gemini >

Try in ImageFX >



Create



Edit



Personalize **New!**

Release

## DALL·E 2

DALL·E 2 is an AI system that can create realistic images and art from a description in natural language.

Try DALL·E 2 Follow on Instagram



DALL·E 2

the Firefly

Now in Adobe

s, audio and now video all in  
commercially safe Firefly generative  
production with seamless integr  
ing you complete creative c

tions ▾

Learn ▾

For Teams

tivity, Unleash

generative AI with a unique su  
convey your ideas to the wor



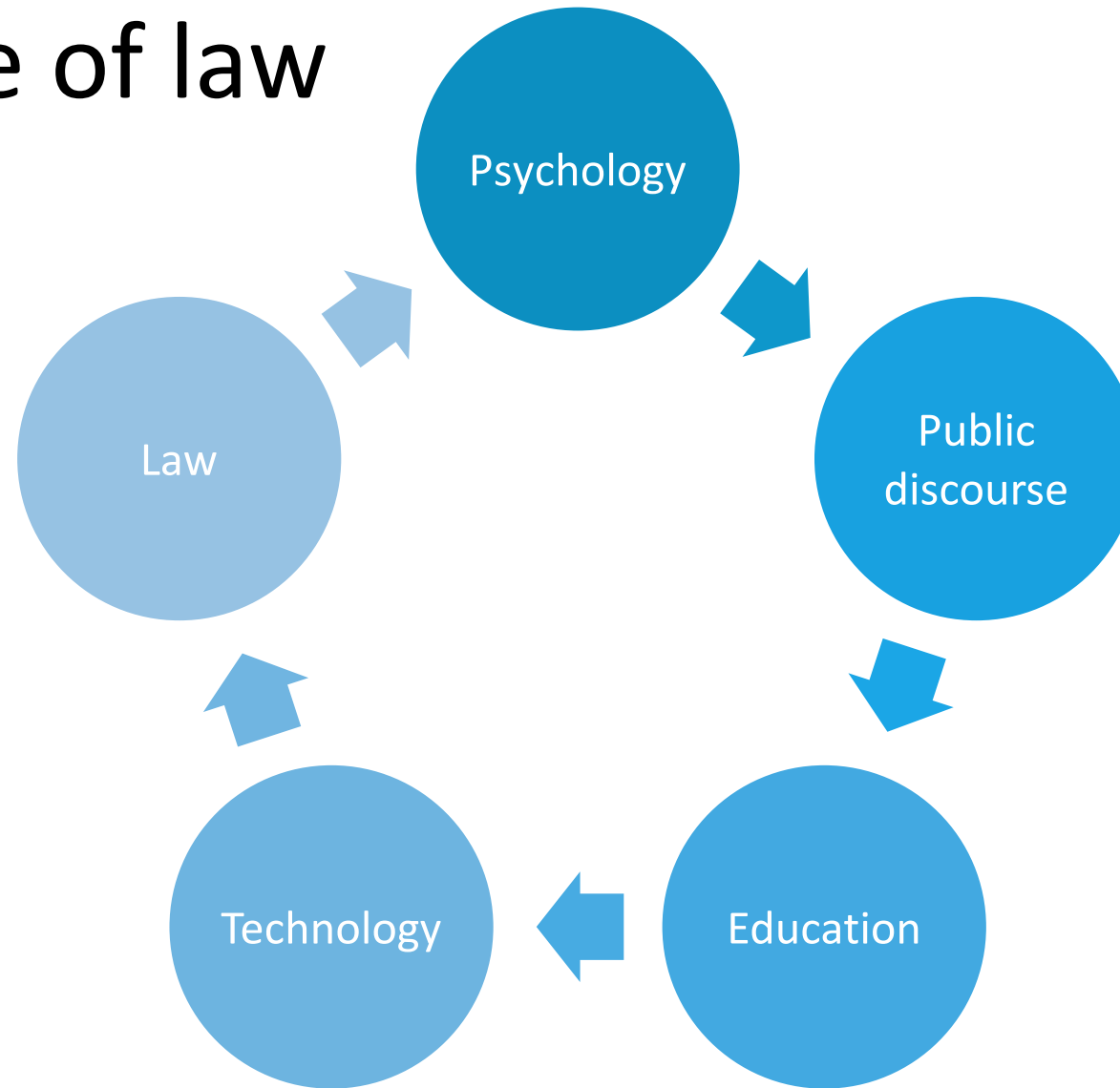
And spread via social media

# Outline

1. Role of 'the law'
2. Regulatory design
3. Regulatory networks



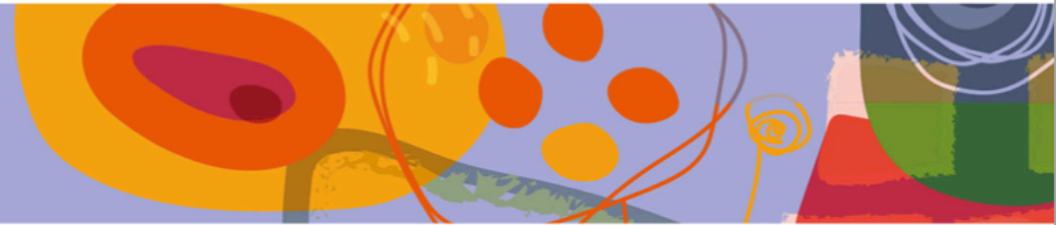
# 1. The role of law





# Psychology

Humanities & Social Sciences  
Communications




REVIEW ARTICLE



<https://doi.org/10.1057/s41599-023-01813-9>

OPEN

## Distractions, analytical thinking and falling for fake news: A survey of psychological factors

Adrian Kwek <sup>1✉</sup>, Luke Peh<sup>2</sup>, Josef Tan<sup>3</sup> & Jin Xing Lee<sup>4</sup>

# Education

## Stop and Consider

During a federal election a lot of things are said and written about the voting process. **Some of it may be false or misleading.**

Stop and consider the content and source of what you see, hear or read about the process.

AI-generated text, images, video and audio could be used to try and deceive you about where to vote, how to complete ballot papers or why the process may not be trustworthy. **Anyone can be affected.**

Apply healthy scepticism. Check or ignore it, but don't share or repost. You can also report it to the place where you saw it. Support friends and family to do the same.

Australia has one of the most trusted electoral systems in the world. It belongs to voters and we all have a responsibility to protect it.

**Don't risk what we have.**

### Check the content

#### Look for:

- words that cause strong feelings, like shock, anger or excitement
- a small number of choices being presented, instead of all of them
- blaming of a particular group or person who is not the only person/group responsible
- attacking a person rather than focusing on an argument.

### Check the source

#### Ask yourself:

- Where did the information come from? Is it from a trustworthy place?
- Is it from Australia or overseas?
- Could it be AI? AI technology makes it easy to create false text, images, video and audio.
- Why was it published? Is there a reason that someone could be trying to convince you of something that isn't true?

### Check our website

#### Get the true info:

- when to vote
- how votes are counted
- how a result is decided
- how the AEC keeps votes safe.

**GET  
TIPS**

**aec.gov.au** 13 23 26




Follow us to hear from us directly


Authorised by the Electoral Commissioner, Canberra



# Technology














Coalition for  
Content Provenance  
and Authenticity


[About](#) [FAQ](#) [Guiding Principles](#) [Specification](#) [Membership](#) [News](#) [Contact](#) 

An open technical standard providing publishers, creators, and consumers the ability to trace the origin of different types of media.

[Learn more >](#)





# Law

## Singapore

- Online Safety (Miscellaneous Amendments) Act 2022
- Protection from Online Falsehoods and Manipulation Act 2019

## Australia

*Online Safety Act 2021* (Cth) (inquiry and amended)

*Criminal Code Amendment (Deepfake Sexual Material) Act 2024* (Cth)

*Online Safety Amendment (Social Media Minimum Age) Act 2024* (Cth)

Digital Platform Levy / news bargaining incentive (ongoing)

*Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2024* (Cth) (abandoned)



## 2 Regulatory design

With a goal of complexity but not incoherence

Good regulatory  
design includes  
clarity about:

Harms addressed

The regulatory target

Obligations imposed

Enforcement

# What kinds of harms should be addressed ?

- Illegal content
- Online bullying
- Intimate image abuse
- Scams
- Fake news
- Political manipulation
- Political satire
- Conspiracy theories

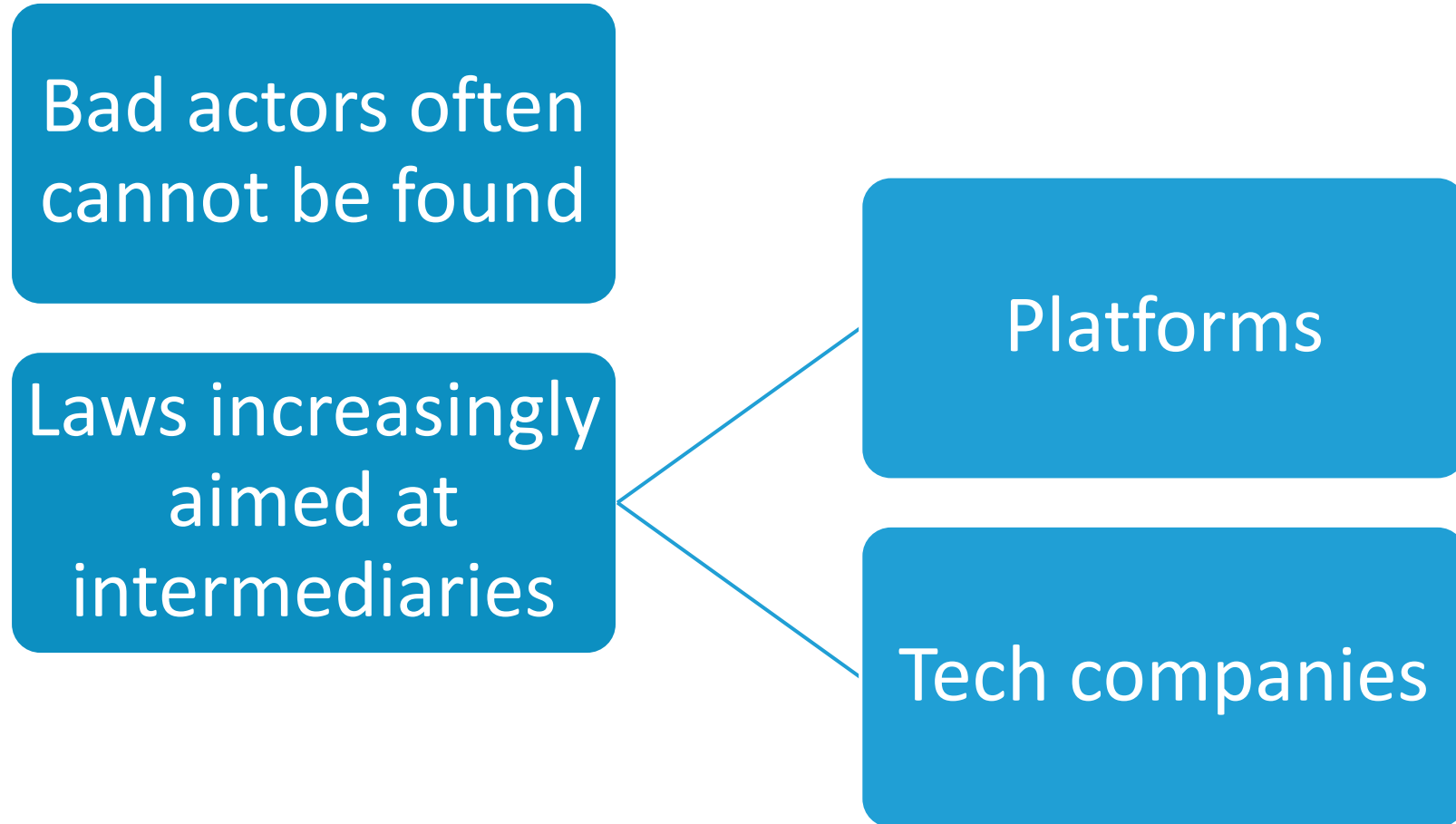


# Policy choices about the response to different kinds of harms ...





# Who is the regulatory target?

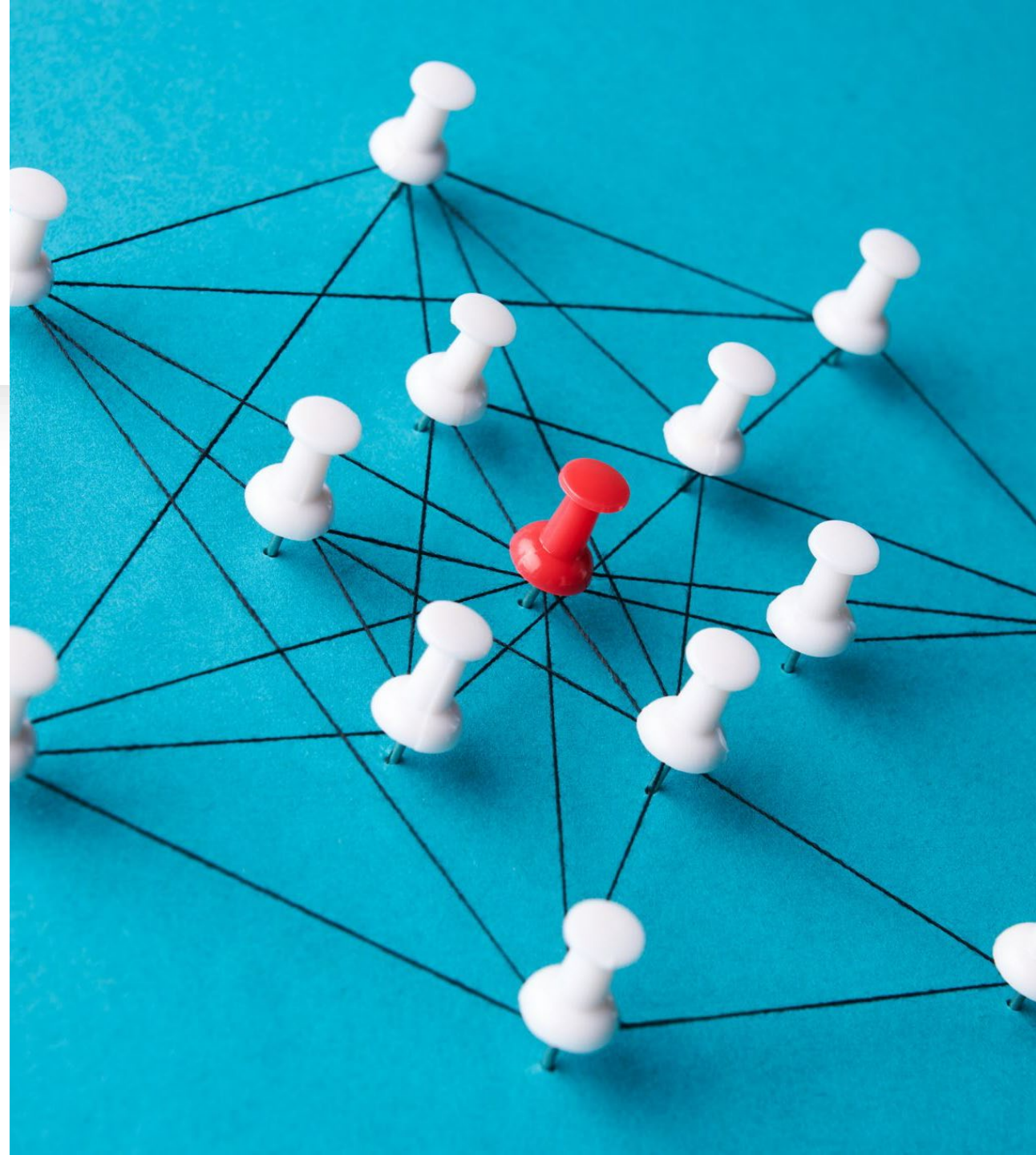


# Platforms

Protection from Online Falsehoods and Manipulation Act Part 4 Directions:

For internet intermediaries:

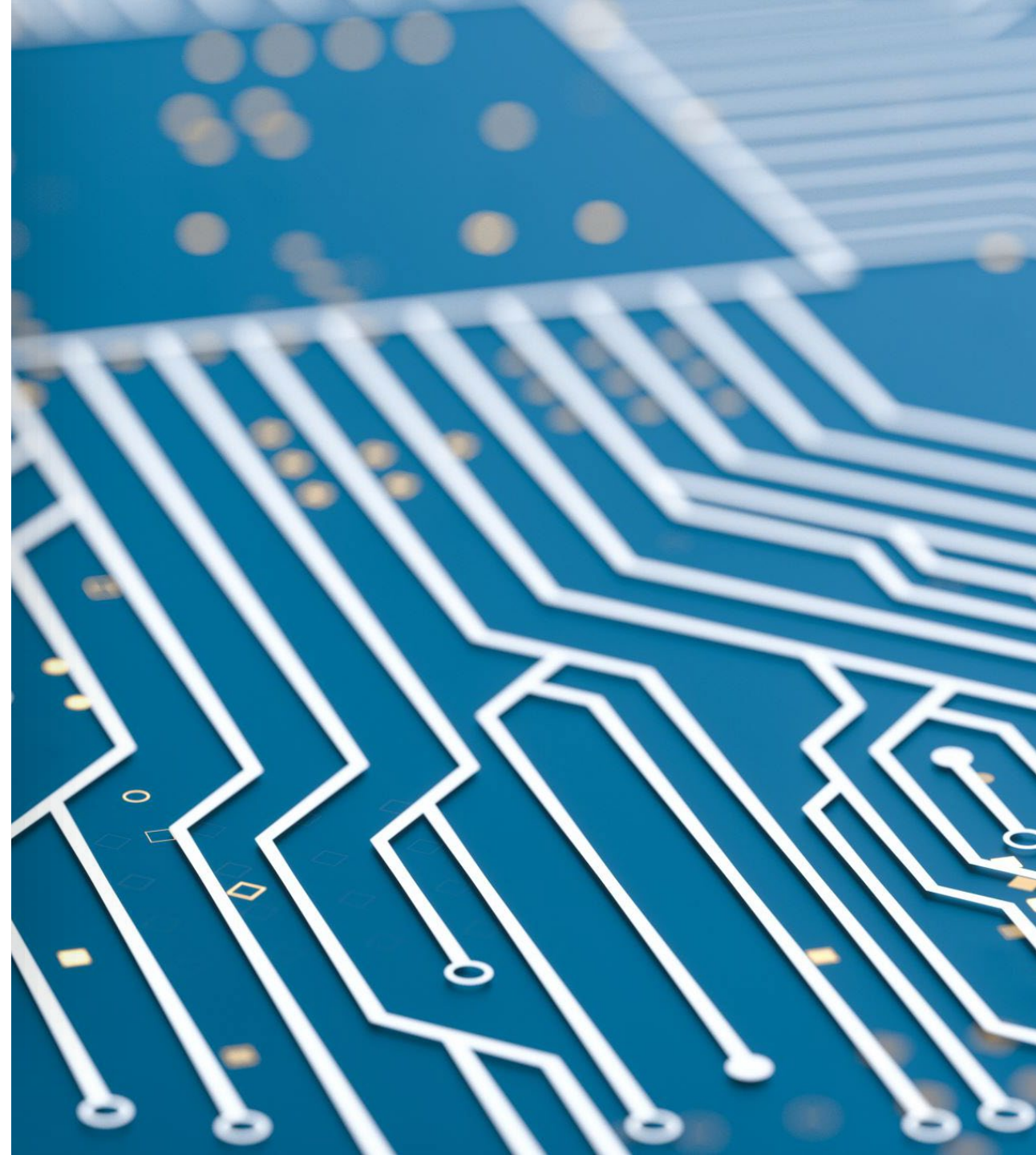
- Targeted Correction Direction
- Disabling Direction
- General Correction Direction



# Tech companies and transparency

EUAI Act article 55

Providers AI systems that create synthetic content (like deepfakes) must mark their outputs as artificially generated.



# What obligations should be imposed?

We may want to tell a regulated entity precisely what they must do

But that goes quickly out of date

Outcome based standards may be more flexible in face of changing technical capacity

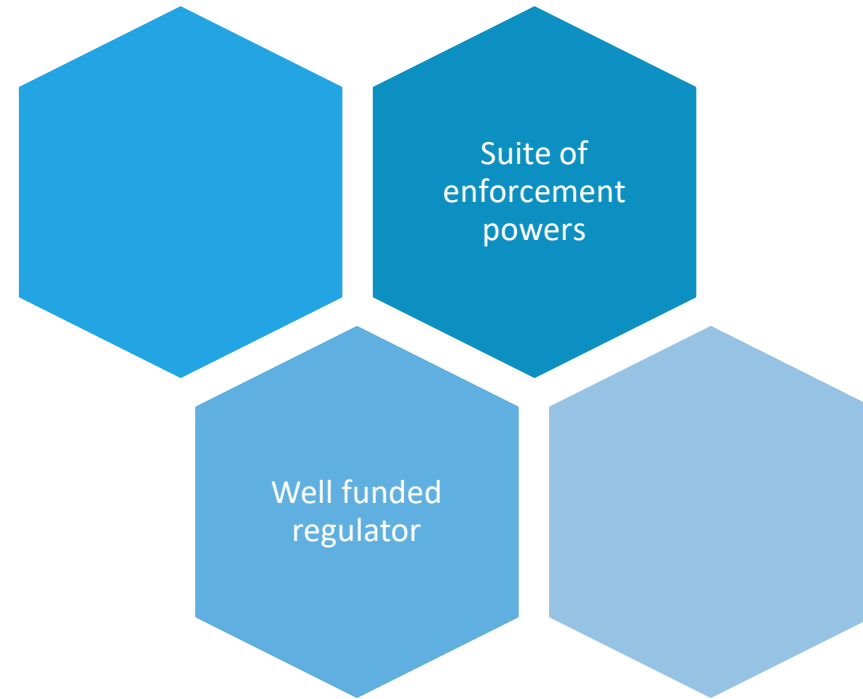


# Eg Online Safety Act – Duty to Comply

**45J.—**(1) Every provider of an online communication service or an internet access service to whom is given a section 45H direction or section 45I blocking direction, has the duty to take all reasonably practicable steps to comply with the direction given to the provider.

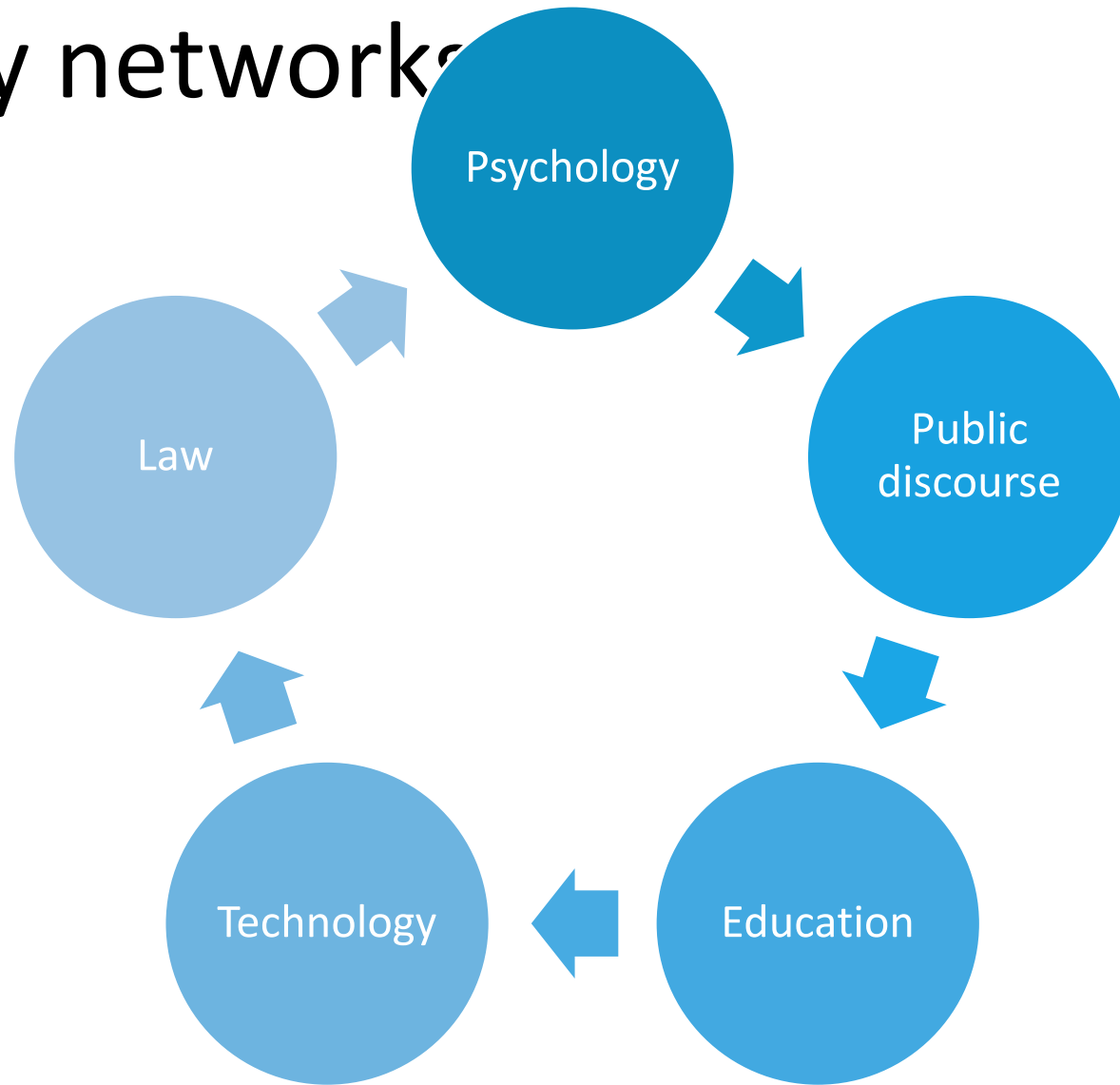
(2) No civil or criminal liability is incurred by a provider of an online communication service or an internet access service provider, or an officer, employee or agent of such a provider, for anything done or omitted to be done with reasonable care and in good faith in complying with a section 45H direction or section 45I blocking direction given to the provider.

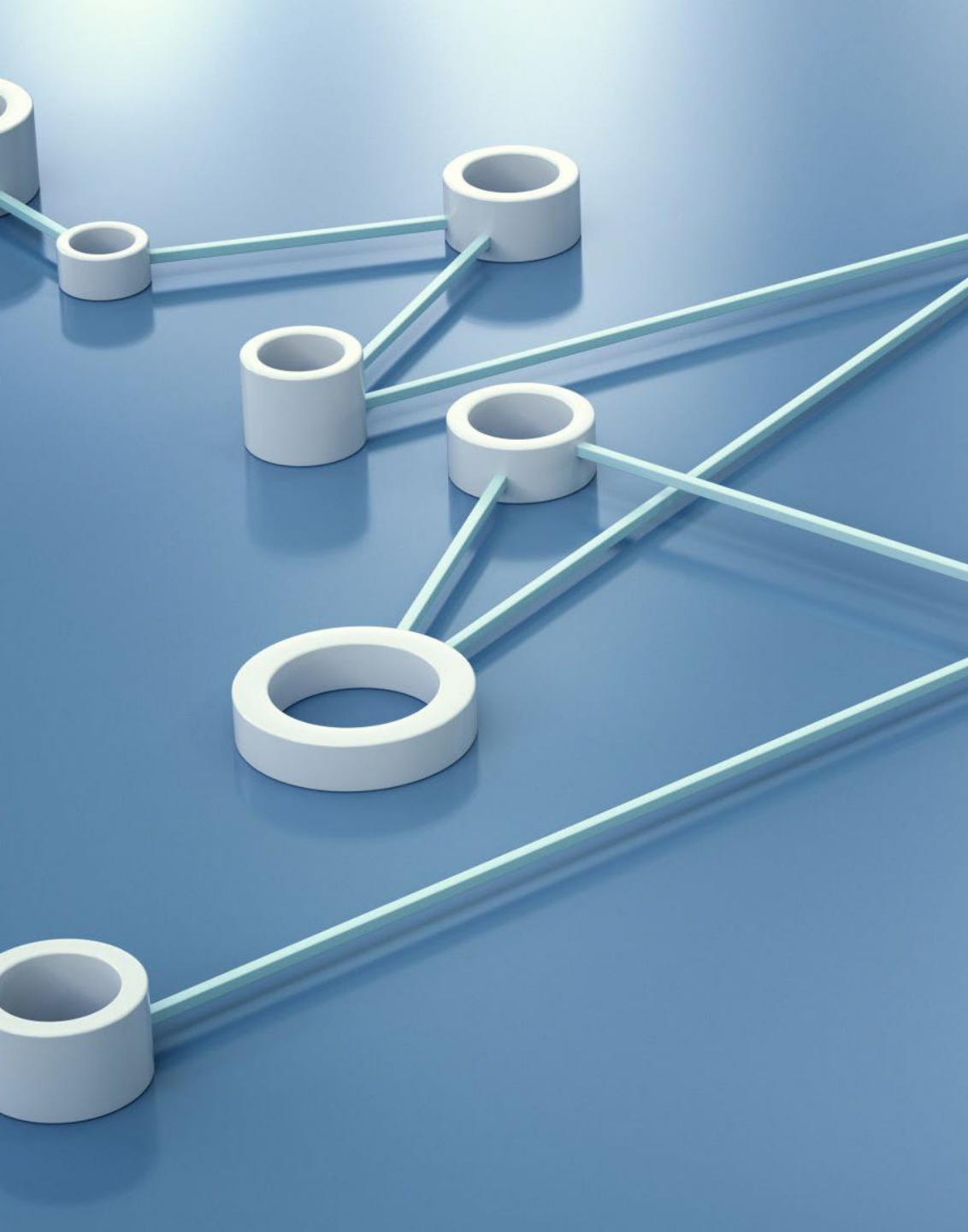
What  
enforcement  
powers are  
available





# Regulatory networks





# Concluding thoughts

Responding to mis, mal and dis information:

Law is only part of the response

In the legal regulatory response :

- complexity but not incoherence

This means:

- coordinated regulatory networks,
- proportionate regulatory targets and
- good regulatory design overall.