netsafe

Netsafe submission on Strengthening the Cyber Security of Critical Infrastructure in New Zealand

About Netsafe

1. Netsafe is New Zealand's independent, non-profit online safety charity. Taking a technology-positive approach to the challenges digital technology presents, we work to help people in New Zealand take advantage of the opportunities available through technology by providing practical tools, support and advice for managing online challenges.

2. We are an independent non-profit organisation adjacent to Government and law enforcement, supported by the public and private sector and with a focus on online safety. Netsafe provides free support, advice and education seven days a week through a helpline, our website and face to face service delivery across New Zealand

3. Netsafe is also the Approved Agency under the Harmful Digital Communications Act 2015 (HDCA). One of the purposes of the HDCA is to deter, prevent, and mitigate harm caused to individuals by digital communications. Netsafe's functions as the Approved Agency are set out in section 8 of the HDCA. Those functions include:

    a. to receive and assess complaints about harm caused to individuals by digital communications

    b. to investigate complaints

    c. to use advice, negotiation, mediation, and persuasion (as appropriate) to resolve complaints

    d. to establish and maintain relationships with domestic and foreign service providers, online content hosts, and agencies (as appropriate) to achieve the purpose of the Act

    e. to provide education and advice on policies for online safety and conduct on the Internet.

4. Netsafe welcomes the opportunity to provide feedback on the proposed framework to strengthen the cyber security of New Zealand's critical infrastructure.

5. Netsafe's frontline support to individuals experiencing online harm, scams, and digital exploitation provides a unique, system-wide view of how cyber threats manifest in

practice for people who report to Netsafe. Each year we handle more than 25,000 reports from New Zealanders and are uniquely placed to comment on the experiences of individuals who report online harm.

6. We strongly support the intent of the proposed regime, including the establishment of minimum standards, improved information sharing, and alignment with international frameworks. However, we consider that the current approach remains primarily technically framed, while the threat landscape has evolved to become increasingly human-mediated and harm-driven.

7. Cyber resilience is no longer solely a technical challenge. It is a human and societal one. We recommend that the framework explicitly incorporates this dimension to ensure it is fit for purpose.

International Context and Opportunity for Leadership

8. The proposed regime appropriately aligns with international approaches, including:

    - Australia's Security of Critical Infrastructure Act 2018 (SOCI)
    - The European Union's NIS2 Directive
    - Singapore's Cybersecurity Act 2018
    - Canada's proposed Bill C-8

9. These frameworks share common features:

    - asset identification and registration
    - minimum cyber risk management requirements
    - incident reporting obligations
    - regulatory oversight and enforcement

10. They are underpinned by a recognition that market incentives alone have not delivered sufficient cyber resilience, necessitating government intervention.

11. However, across jurisdictions, there remains a consistent gap and that is limited explicit integration of human vulnerability, user harm, and public trust as core components of cyber security.

12. At the same time, global threat patterns show that:

    - social engineering is one of the primary initial attack vectors
    - fraud and scams and impersonation attacks generate large-scale economic and social harm
    - incidents and breaches increasingly propagate through users rather than systems alone

13. New Zealand has the opportunity to lead internationally by embedding a human-centred cyber resilience model alongside technical controls.

The Evolving Nature of Risk to Critical Infrastructure - from system compromise to human exploitation

14. The consultation emphasises increasing activity from persistent threat actors, hacktivists, and organised crime. Netsafe agrees with this assessment and notes an additional shift. Cyber threats are increasingly characterised by high-volume, low-cost exploitation of people at scale. These include:

    - phishing and impersonation of people and organisations
    - bad actors leveraging trusted infrastructure platforms
    - manipulation of users to bypass security controls

15. In many cases, critical infrastructure is not directly compromised but is instead used as a channel for harm.

Critical infrastructure as both target and vector

16. Critical infrastructure should be understood as:

    - a target of disruption
    - a dependency for essential services
    - a vector through which harm can be delivered to users

17. Examples include:

    - banking systems facilitating fraudulent and scam transactions
    - telecommunications networks enabling impersonation
    - digital services exploited to target customers

18. These harms:

    - scale rapidly
    - bypass technical controls
    - erode trust in essential services

19. Recommendation: explicitly recognise social engineering, scams, and misinformation as threats to critical infrastructure resilience.

Trust as a critical dependency

20. Cyber incidents increasingly undermine public trust, which is itself a foundational component of infrastructure resilience.

21. Loss of trust can lead to:

    - reduced uptake of essential services
    - reluctance to engage with digital systems
    - amplification of harm through misinformation

22. A system that is technically secure but not trusted is not resilient.

23. Recommendation: recognise public trust and user harm as core dimensions of critical infrastructure risk.

Establishing a Minimum Baseline — Including the Human Layer

24. The consultation identifies that there is currently no common baseline for cyber security across the system. Netsafe strongly supports the establishment of minimum standards. However, we recommend that this includes both:

    - a technical baseline (existing focus), for example:
        - system controls
        - detection and response capability
        - governance and assurance

    - and a human-layer baseline (currently missing), for example:
        - protection against scams and social engineering
        - safer by design of user-facing systems
        - behavioural risk mitigation

25. Recommendation: define a minimum human-layer baseline, requiring entities to identify and mitigate vulnerabilities arising from user interaction with systems.

Risk Management — Integrating Behavioural and Harm Dimensions

26. The proposed framework outlines a risk management approach (scope, understand, evaluate, treat, monitor). We support this structure and recommend expanding its application.

Human-layer risk assessment

27. Entities should assess:

- how users may be manipulated
- how systems can be misused for harm
- user behavioural pathways that bypass controls

Safety-by-design

28. Critical infrastructure providers should:

- reduce frictionless scam pathways
- implement protective prompts and warnings
- design services anticipating misuse

Addressing market failure

29. We agree that current market behaviour does not match the level of cyber risk. In particular, the costs of scams, estimated to be more than $3billion per year to the New Zealand economy (GASA and Netsafe report 2025) and social engineering are often externalised to individuals and communities. Incorporating user protection into minimum standards helps correct this imbalance.

Incident Definition, Reporting, and Response

30. The consultation raises important questions regarding incident definitions and reporting thresholds.

Expanding incident thresholds

31. We recommend that reporting thresholds consider not only system impact, but also:

- scale of user harm
- financial loss to customers
- widespread targeting or exposure

Strengthening response expectations

32. Current proposals focus on technical containment. We recommend including:

- timely and clear user notification consistent with requirements under the Privacy Act 2020 for data breach notifications
- coordinated public communication
- access to post incident support pathways for victims of cyber incidents

33. From Netsafe's experience, harm escalates rapidly where these elements are absent.

Learning from real-world incidents

34. Events such as the CrowdStrike outage demonstrate that:

- disruption may occur without malicious intent
- impacts are shaped by timing and user dependency

35. This reinforces the need to incorporate human impact scenarios into resilience planning.

Information Sharing and System Visibility

36. We strongly support the development of improved information-sharing mechanisms, including a trusted information network.

37. However, current reliance on voluntary reporting is constrained by:

- reputational risk
- commercial sensitivity
- legal uncertainty

Integrating public-facing harm intelligence

38. We recommend greater awareness is needed of:

- consumer reports of scams and harm
- frontline case data
- behavioural threat indicators

39. The public often acts as an early warning system for emerging threats.

Reducing reporting barriers

40. We recommend consideration of mechanisms for:

- anonymised data sharing in a privacy enhancing way
- safe disclosure pathways

41. This will improve participation and system visibility.

Governance, Accountability, and System Roles

Director accountability

42. We support strengthened governance expectations and recommend that these explicitly include:

    - customer harm
    - fraud and scam exposure
    - user safety outcomes

43. This ensures cyber security is treated as a trust and service issue, not solely an IT function.

Role of trusted intermediaries

44. There is currently a gap between:

    - technical agencies
    - infrastructure operators
    - the public

45. More should be considered to close the gap between who knows what about an incident and the role each sector needs to play to create safer critical infrastructure environments.

46. Recommendation: formally recognise the role of public-facing organisations in:

    - threat detection
    - incident response
    - public communication

Exercising, Scenario Planning, and System Preparedness

47. We support the proposal to deepen exercises and scenario testing.

48. Recommendation: include human-layer scenarios, such as:

    - large-scale scam campaigns targeting infrastructure users

- misinformation during service outages
- coordinated social engineering attacks

Proportionality, Criticality, and Equity

49. The consultation considers how to define critical infrastructure and nationally significant entities.

50. We recommend incorporating into criticality assessments:

    - potential for widespread consumer harm
    - impact on public trust

    <u>Equity considerations</u>

51. Cyber incidents disproportionately affect:

    - older adults
    - digitally excluded populations
    - Māori and Pacific communities

52. Recommendation: require inclusive approaches to resilience, including:

    - accessible communication
    - targeted harm prevention

Emergency Powers and Public Trust

53. We acknowledge the rationale for emergency intervention powers in significant cyber events. To maintain public confidence, we recommend:

    - a test of necessity and proportionality
    - clear safeguards
    - transparency of use
    - careful handling of sensitive personal data
    - the more intrusive the power, the further from the agency itself the approval should sit
    - strict time limits
    - what alternatives were considered and why they haven't worked
    - if a system is taken over there needs to be a full forensics trail
    - post incident reporting to the public or Parliament

54. If these powers are going to exist, they need to be credible, trusted, and built to withstand misuse- otherwise they will fail in practice and lose public legitimacy.

Conclusion

55. Netsafe strongly supports the development of a strengthened cyber security framework for critical infrastructure in New Zealand.

56. To ensure the framework reflects the realities of the modern threat landscape, we recommend that it:

- explicitly recognise human-mediated cyber threats
- incorporate user harm and public trust into risk definitions
- establish a human-layer baseline alongside technical standards
- strengthen incident response to include harm mitigation
- better join up public-facing threat intelligence
- formalise the role of trusted intermediaries

57. Critical infrastructure resilience depends not only on secure systems, but on protected, informed, and confident users. New Zealand has a unique opportunity to lead internationally by embedding this perspective from the outset. Netsafe stands ready to support this work as part of a coordinated, system-wide approach.

Netsafe

19 April 2026