

A South African Legal Framework to Govern Online Harm

What Might We Learn from the EU's Digital Services Act?

Nicole Fritz and Justine Limpitlaw



A South African Legal Framework to Govern Online Harm

What Might We Learn from the EU's Digital Services Act?

By Nicole Fritz and Justine Limpitlaw



Published by Konrad-Adenauer-Stiftung
Media Programme Sub-Saharan Africa

Telephone: + 27 (0)11 214-2900

www.kas.de/mediaafrica

Facebook: @KASMediaAfrica

Instagram: @KASMediaAfrica

LinkedIn: @KASMediaAfrica

ISBN: 978-1-0483-7323-3 (print)

ISBN: 978-1-0483-7324-0 (e-book)

© Konrad-Adenauer-Stiftung, 2026

This publication is an open source publication. Parts thereof may be reproduced or quoted provided the publication is fully acknowledged as the source thereof.

Download an electronic copy of *A South African Legal Framework to Govern Online Harm: What Might We Learn from the EU's Digital Services Act?* from www.kas.de/en/web/medien-afrika/einzeltitel/detail/-/content/new-study-making-digital-platforms-accountable-in-south-africa

Cover photo: AI image

Proofreading: Bruce Conradi
bruce.conradie@theresearchfaculty.com

Design, layout
and production: Heath White, [ihwhiteDesign](http://ihwhiteDesign.com)
heath@ihwhitedesign.com

Printing: Typo Printing Investments,
Johannesburg, South Africa

Foreword

Digital platforms have become one of the most important arenas of democratic life. Elections, political debate, journalism and civic engagement increasingly take place on platforms operated by global technology companies. These platforms connect people, enable access to information and create new opportunities for participation. At the same time, they exert significant influence over what information we see, what gains visibility and which voices are amplified.

Also in South Africa, we have seen clearly that the impact of digital communication is no longer confined to the online world. Hate speech, disinformation, intimidation, incitement to violence and the manipulation of public debate can have serious consequences for social cohesion and democratic institutions. Yet any discussion about regulation must be guided by an equally important principle: the protection of freedom of expression and open democratic discourse.

It is against this backdrop that we asked Nicole Fritz and Justine Limpitlaw to explore what South Africa might learn from the European Union's Digital Services Act. The strength of this study lies in its balanced and practical approach. It neither advocates importing the European model wholesale nor calls for state control of online content. Instead, it offers thoughtful proposals for strengthening transparency, accountability and oversight, while remaining firmly grounded in constitutional principles and democratic freedoms.

What resonates most with me is the report's central argument: that the largest platforms should be accountable for the impact of their systems. They should be accessible to users, transparent in their decisions, open to independent scrutiny and capable of assessing and addressing the risks their services create. At the same time, regulation must remain proportionate, focusing on those platforms whose scale and influence give them the greatest power to shape public discourse.

I hope this publication contributes to a constructive conversation among policymakers, regulators, civil society, researchers, journalists and technology companies. South Africa does not need a copy of Europe's Digital Services Act. But

we can learn from international experience and adapt those lessons to our own constitutional values, institutional realities and democratic aspirations. This report is an important contribution to that effort.

Hendrik Sittig
Media Programme Sub-Saharan Africa
Konrad-Adenauer-Stiftung
Johannesburg, South Africa
October 2026

Table of contents

Foreword	v
Introduction	9
SECTION 1	16
Scope and differentiated application: Which digital services should be regulated?	
SECTION 2	21
Intermediary liability and safe harbours: Modernising South Africa's existing framework	
SECTION 3	29
Establishing jurisdiction and effective local accountability	
SECTION 4	34
Accessible complaint, notice and redress mechanisms	
SECTION 5	45
Transparency and reporting	
SECTION 6	52
Access to platform data and independent research	

SECTION 7	58
Systemic risk assessment adapted to South African circumstances	
SECTION 8	65
Protection of minors online	
SECTION 9	71
Local linguistic, cultural and contextual capacity	
SECTION 10	74
Funding effective oversight: A proportionate supervisory fee	
SECTION 11	76
Institutional architecture and co-regulation	
SECTION 12	83
Effective enforcement	
Conclusion	89
Acknowledgements	91



Introduction

The European Union's (EU) Digital Services Act (DSA)¹ arguably represents the most ambitious attempt globally to ensure public accountability from the private companies—social media giants like Meta, Alphabet, X, and ByteDance—that provide the channels and platforms through which much of our information is communicated, encountered, amplified and contested.

The DSA's reach is extensive: from intermediary liability and mechanisms for reporting illegal content, through transparency obligations and access to platform data, to requirements that the largest platforms assess and mitigate the systemic risks generated by their services.

South Africa plainly cannot replicate that framework exactly. The European Union represents a market of some 450 million people, backed by the regulatory institutions and collective economic leverage of 27 Member States. South Africa is a single state with significantly more limited regulatory, technical and financial capacity. The question propelling our research into the applicability of the DSA to the South African context, and this resulting report, was accordingly never whether South Africa should enact its own version of the DSA. It was more practical: what has Europe learnt about regulating digital intermediaries that might be useful to South Africa, and which elements of that experiment could sensibly be adapted to our very different circumstances?

To explore that question, we studied the architecture and implementation of the DSA and undertook a research visit to Berlin, Brussels and Dublin,² engaging with regulators, policymakers, researchers, civil-society actors, platform representatives and others directly involved in, or affected by, Europe's emerging digital regulatory framework. We were interested not simply in what the DSA says on paper, but in how its provisions are beginning to operate: which mechanisms appear valuable,

1 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277, 27 October 2022, 1–102. The DSA entered into force on 16 November 2022 and became generally applicable on 17 February 2024. Certain provisions applied earlier, and designated VLOPs and VLOSEs became subject to their additional obligations four months after designation: arts 92–93.

2 We travelled to Europe from 5–13 July 2026: Berlin (5–8 July), Brussels (8–10 July), and Dublin (12–13 July).

where difficulties have emerged, how the regulatory burden is distributed, and what institutional machinery is required to make apparently attractive legal obligations meaningful in practice.

The research took place against two developments that gave those questions a particular urgency.

A South African warning

As this report was being prepared, South Africa was experiencing a marked upsurge in xenophobia and hatred directed at foreign nationals, of almost exclusively African origin.

In the months preceding the so-called 30 June deadline announced by vigilante organisation March and March and its associated backers, anti-migrant incitement and mobilisation became especially prominent. Much of it unfolded online, with key influencers promoting this content.³ Of course, xenophobia was not created in South Africa by digital platforms, nor can online activity alone explain the social, economic and political forces that sustain it. But social media provided an extraordinarily effective infrastructure through which hate could be organised, repeated and amplified and through which what might otherwise have appeared to be the views of particular groups or individuals could acquire the appearance of widespread public consensus.

No one can seriously maintain any longer that digital platforms simply provide places in which people speak. Their systems determine what is recommended, repeated, ranked and made visible; what reaches audiences far beyond the speaker's immediate network; and what forms of engagement are rewarded with still greater distribution. The resulting impression of ubiquity has political consequences. Views that appear marginal in one context can, through repetition and amplification, acquire the appearance of an irresistible popular tide.

The events surrounding 30 June followed us, unexpectedly, on our research trip to Europe. In Berlin, Brussels and Dublin, people we encountered repeatedly wanted to know what had happened—the impending deadline had been widely covered in the international press. Our interlocutors were not only those professionally engaged with digital regulation and South African affairs, but people we encountered along our journey. Those with whom we spoke, from people transporting us between airports and hotels to senior European political analysts, expressed alarm and revulsion at scenes emerging from South Africa.

3 See Institute for Security Studies, 'Lucrative online influence detected during South Africa's 30 June protests' (2 July 2026), reporting analysis by Murmur Intelligence of coordinated online amplification around the protests; Yossabel Chetty, 'How the anti-migrant 30 June campaign was disrupted by civil society groups', *Daily Maverick* (12 July 2026); Gudani Tshikota and William Bird, 'When misinformation becomes a weapon — Afrophobia, social media and 30 June', *Daily Maverick* (16 June 2026).

Those concerns have since received a far more formal expression. On 27 August 2026, the United Nations Committee on the Elimination of Racial Discrimination publicised a decision adopted under its early-warning and urgent-action procedure specifically addressing xenophobia, racial discrimination and hate crimes against migrants, asylum seekers, refugees and persons perceived as such in South Africa.⁴ The Committee expressed grave concern about developments in South Africa, including killings and other forms of violence directed at migrants, and drew particular attention to racist hate speech and harmful stereotypes disseminated through the internet and social media. Among the measures it called for was engagement with media, internet-service providers and social-media platforms to combat online hate speech and harmful stereotypes.

These developments must compel the gravest concern and abhorrence at the xenophobia and raw hatred unleashed. They represent a rejection of the fundamental values at the heart of South Africa's Constitution of equality, dignity and freedom. Equally, South Africa's constitutional democracy depends upon institutions and norms that are not only vigilantism's opposite; they are destroyed by vigilantism: law rather than mob power determines right and entitlement; only constitutionally authorised institutions exercise public power; coercive force may not be privately appropriated by groups claiming to speak for "the people"; and unpopular and vulnerable minorities remain entitled to the protection of the law.

When online mobilisation legitimates or facilitates the exclusion of people from communities, workplaces, schools or healthcare; when it contributes to an environment in which private groups purport to decide who may remain in particular spaces; or when relentless digital vilification helps make physical violence against a class of people appear admissible, the problem is not confined to the internet. The integrity of the democratic order itself is implicated.

But if the current xenophobic mobilisation lays out the stakes at their most extreme, it is only one manifestation of the challenge. South Africa has experienced online harassment and threats against journalists and women in public life; coordinated information manipulation; threats to children; electoral disinformation; an explosion of scams and deception; and other forms of digitally enabled mobilisation capable of producing serious offline consequences. But different harms require different responses, and not everything harmful should be made unlawful. Indeed, an overbroad regulatory response may itself endanger the very freedom of expression needed to call out these harms and that any legitimate framework must protect.

It is the need to strike this balance that makes the regulatory question such a fraught one. The choice cannot simply be between an unregulated internet and a state empowered to decide what South Africans may say. The difficult and delicate task

4 UN Committee on the Elimination of Racial Discrimination, 'Xenophobia, racial discrimination and hate crimes against migrants, asylum seekers, refugees and those perceived as such in South Africa', Decision 1 (2026), CERD/C/ZAF/DEC/1, adopted under the Committee's Early Warning and Urgent Action Procedure.

is whether rules can be devised that make the companies operating the principal infrastructures of digital public life accessible, transparent and accountable; that provide an effective means of redress when unlawful conduct occurs; that permit independent scrutiny of how their systems operate; and that require the largest platforms to understand and respond to the serious risks their own systems may help create or intensify.

At present, South Africa possesses pieces of such a framework, scattered across different statutes, regulators and areas of law.⁵ What it lacks is a coherent architecture for platform accountability.

A global contest over who gets to regulate the digital public sphere

The second context in which this report has been prepared is geopolitical. Because the DSA represents such an ambitious European regulatory experiment, and because it seeks to cover such a vast, prominent market, it has also become an object of unusually intense transatlantic political conflict.

The administration of United States President Donald Trump has characterised aspects of European digital regulation as restrictions upon free expression and discriminatory burdens upon American technology companies. In February 2025, President Trump directed US officials to investigate foreign policies or practices that require or incentivise American technology companies to moderate content in ways said to undermine freedom of speech and political engagement.⁶ By August 2025, the administration had reportedly instructed American diplomats in Europe to campaign against the DSA and was considering visa restrictions against European officials involved in its implementation.⁷

US pressure subsequently became more intense. Following DSA enforcement action against X, senior US officials denounced the European regime as constituting censorship and discriminatory treatment of American companies. In December 2025, the Trump administration imposed visa restrictions upon five Europeans associated with digital regulation or organisations working on online hate and disinformation, including former European Commissioner Thierry Breton, one of the architects of the DSA.⁸ European institutions and governments characterised

5 See Annexure A, 'A Detailed Consideration of the Existing Legislative Environment Regarding Content Shared Online', available alongside the electronic version of this report.

6 Donald J Trump, Presidential Memorandum, '[Defending American Companies and Innovators from Overseas Extortion and Unfair Fines and Penalties](#)' (21 February 2025) s 3(d).

7 Humeyra Pamuk, '[Rubio orders US diplomats to launch lobbying blitz against Europe's tech law](#)', Reuters (7 August 2025); Humeyra Pamuk, '[Trump administration weighs sanctions on officials implementing EU tech law, sources say](#)', Reuters (25 August 2025).

8 Humeyra Pamuk and Kanishka Singh, '[US targets former EU commissioner, activists with visa bans over alleged censorship](#)', Reuters (23 December 2025).

the measures as an attempt to intimidate Europe into abandoning its regulatory autonomy.⁹

The resistance is not confined to the US government. Some of the largest American technology companies and their owners have increasingly aligned themselves with the argument that European regulation constitutes censorship. Meta chief executive Mark Zuckerberg, for example, pledged in January 2025 to work with the Trump administration to push back against governments he said were pursuing American companies and institutionalising censorship, while specifically criticising Europe's regulatory approach.¹⁰

Of course, none of the fierce opposition from the US itself signals that the DSA is good regulation. A law is not wise or even sound because powerful people oppose it. And there are legitimate questions about the DSA's complexity, the administrative burdens it establishes, its effects upon freedom of expression, the enormous discretion involved in concepts such as systemic risk, and whether a regulatory model built for the European Union can or should influence jurisdictions with very different constitutional traditions and institutional capacities.

But the intensity of the resistance is nevertheless revealing. It demonstrates that the DSA is not simply another layer of technical compliance. Its requirements reach matters of considerable importance to the business models and governing autonomy of the world's largest platforms: how they moderate content; what they disclose about advertising and recommender systems; what data they must make available to regulators and independent researchers; how they assess risks generated by their services; and whether regulators can compel compliance and impose penalties meaningful even to corporations of extraordinary global scale.

The conflict concerns something larger than the interpretation of particular provisions of European law. It concerns who is entitled to set the terms upon which the digital infrastructure of contemporary public life operates.

For much of the history of the commercial internet, those terms have been determined overwhelmingly by the companies themselves. Platforms wrote the rules, designed the systems through which information was distributed, decided what outsiders could know about those systems, determined how complaints would be handled and, to a remarkable extent, chose the jurisdictions in which they were prepared to regard themselves as meaningfully accountable.

The DSA represents a direct challenge to that status quo. Its foundational proposition is that companies may build and operate digital services, and for the purposes of innovation should indeed do so, but when those services become

9 European Commission, 'Statement by the European Commission on the US decision to impose travel restrictions on certain EU individuals' (24 December 2025); French Ministry for Europe and Foreign Affairs, 'Sanctions américaines à l'encontre de cinq personnalités européennes' (24 December 2025).

10 Meta, 'More Speech and Fewer Mistakes' (7 January 2025); see also Reuters, 'Meta shelves fact-checking in policy reversal ahead of Trump administration' (7 January 2025); N Fritz, 'Decimation of Meta's Moderation is Zuckerberg's Craven Sucking Up to Trump', *Daily Maverick* (13 January 2025).

sufficiently consequential to the lives, rights and democratic environments of millions of people, their governing power cannot remain entirely self-defined and self-supervised.

Whether Europe has struck the correct balance is properly open to debate. But the principle that democratic societies are entitled to establish that balance for themselves should not be controversial.

What South Africa can learn

South Africa enters this debate very differently situated from either the European Union or the United States. We in no way possess Europe's market power. Nor are the world's dominant platforms South African companies whose commercial interests our government might seek to defend abroad. We are overwhelmingly rule-takers in a digital environment designed elsewhere.

Yet South Africans live with the consequences of those design choices. Our elections and public debates occur on these services. Our journalists, activists and political representatives depend upon them. South African children use them. Businesses are built on their services and advertise through them. Communities organise through them. And harm occurring through them can spill rapidly into physical life.

This creates a fundamental imbalance. The platforms are global, but the consequences are often local. That imbalance provides the starting point for this report. As already indicated, the object is not to recommend wholesale transplantation of the DSA into South African law. Some elements of the European framework depend upon institutional resources South Africa simply does not possess. Some respond to the distinctive architecture of the European Union. Others may be unnecessarily elaborate for South African purposes. And every proposed intervention must be tested against South Africa's own Constitution, legal system, institutional capacity and particular commitment to freedom of expression.

Instead, our report asks what might usefully be borrowed, adapted or learnt from. Some of the DSA's most valuable lessons are surprisingly elementary. Regulation cannot work if global platforms serving millions of South Africans can deny that they are meaningfully subject to South Africa's jurisdiction and laws and require litigants to pursue foreign corporate entities abroad. Users require accessible mechanisms for reporting allegedly unlawful conduct and challenging consequential moderation decisions. Platforms should be required to disclose enough about their South African operations for their conduct to be scrutinised. Regulators need access to the information necessary to perform their functions. Independent researchers can substantially enlarge the capacity of a resource-constrained state to understand complex digital systems. The largest platforms can reasonably be expected to assess the systemic risks their services generate for South Africans and to demonstrate adequate linguistic and contextual competence. And none of these obligations will mean very much unless the institutions responsible for enforcing them possess adequate powers and resources.

While the DSA provides no ready-made answer to South Africa's digital challenges, what it does provide is a large-scale, evolving experiment in moving beyond voluntary platform governance towards public accountability.

At a moment when South Africa is witnessing, in immediate and alarming ways, the capacity of online mobilisation to intensify social division and facilitate real-world harm — and when the world's most ambitious regulatory attempt to constrain platform power has itself become the subject of a major geopolitical struggle — the question of what we might learn from that experiment is not merely theoretical. It is a question about the kind of digital public sphere that allows South Africa's democracy to survive.

The introduction has sought to explain why the question of digital regulation has acquired such urgency in South Africa and what is at stake in determining how the power exercised by global digital platforms should be governed. The sections that follow are necessarily more technical in focus. They examine particular features of the DSA, how they operate, and whether the principles underlying them might usefully be adapted to South African circumstances. The report does not follow the DSA's provisions sequentially. Instead, it is organised around a series of connected regulatory questions that build upon one another. It begins by considering which digital services should fall within a South African regulatory framework and how obligations might be differentiated according to their function and systemic significance. It then considers the existing framework for intermediary liability and safe harbours; the jurisdictional and local-accountability requirements necessary to make regulation effective; and the complaint, notice and redress mechanisms through which users can engage with platforms and challenge their decisions.

From there, the report turns to the transparency and data-access obligations necessary for meaningful external scrutiny, before considering the particular responsibilities that may be warranted by different forms of risk and different categories of service. These include the enhanced obligations that might appropriately attach to the largest and most systemically consequential services; measures directed specifically at protecting children online; and the need for platforms operating in South Africa to possess adequate linguistic and contextual capacity. The final sections address the institutional foundations necessary to make such a framework work in practice: how effective supervision might be funded, where regulatory responsibility might be located, and what powers would be required to secure compliance. Taken together, these sections ask not how South Africa might reproduce the DSA, but what elements of its emerging regulatory architecture might contribute to a coherent and constitutionally appropriate system of platform accountability in South Africa.



Section 1

Scope and differentiated application: Which digital services should be regulated?

Any South African framework modelled in part on the DSA must begin by identifying the services to which it applies. This is not simply a matter of statutory definition. The nature of the service determines both the risks it is capable of generating and the obligations that can reasonably be imposed upon it.

The DSA does not treat every entity participating in the digital ecosystem as though it performs the same function or exercises equivalent power. Instead, it establishes a graduated structure in which obligations become progressively more demanding according to the nature and, ultimately, the scale and systemic significance of the service concerned.

1.1 The DSA's basic taxonomy

At the broadest level, the DSA applies to providers of “intermediary services”. Article 3 distinguishes three familiar categories. A “mere conduit” service transmits information provided by a recipient of the service, or provides access to a communications network. A “caching” service also transmits information but involves its automatic, intermediate and temporary storage for the purpose of making onward transmission more efficient. A “hosting” service stores information provided by, and at the request of, a recipient of the service.

These categories describe functions rather than particular types of companies. This is important in a rapidly changing technological environment. The DSA's recitals make clear that, depending upon the function actually performed, intermediary services may encompass internet-access providers, virtual private networks, domain-name services, content-delivery networks, cloud infrastructure and web-hosting services, messaging and webmail services and other components of the internet's technical infrastructure.

Within this broad category, however, the DSA identifies particular services that warrant additional regulation. An online platform is a hosting service which, at the request of a user, stores and disseminates information to the public, unless dissemination is merely a minor and ancillary feature of another service. Social-media services, online marketplaces and certain other services through which user-provided information is made publicly available consequently occupy a different regulatory position from a company that simply transmits communications or provides cloud storage.

The DSA then distinguishes Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs) according to their reach within the European Union.¹¹ These services bear the most extensive obligations because their scale means that risks associated with their operation may themselves become systemic.

The result is not one regulatory category but a series of nested ones. All online platforms are intermediary services and are generally hosting services, but not all hosting services are online platforms. And only a relatively small number of platforms and search engines are subject to the DSA's most demanding obligations.

1.2 South African law already contains part of this architecture

This basic taxonomy is not foreign to South African law. Chapter XI of the Electronic Communications and Transactions Act 25 of 2002 (ECTA) establishes limitations upon the liability of providers performing functions substantially corresponding to mere conduit, caching and hosting. Sections 73, 74 and 75 respectively address these functions, while section 76 separately addresses "information location tools".

South African legislation therefore already recognises an important underlying proposition: different kinds of intermediaries perform different functions and should not necessarily attract identical legal consequences.¹²

What ECTA does not provide is the more developed regulatory architecture required for the contemporary platform environment. Enacted in 2002, it necessarily predates the emergence of social-media platforms and other digital services in their present form and scale. Its taxonomy is employed principally for determining limitations upon intermediary liability. It does not build upon those categories to establish graduated positive obligations concerning matters such as content-moderation procedures, advertising transparency, recommender systems, systemic risk or researcher access.

The DSA offers a useful development of this earlier model. It retains the basic technical categories, while recognising that some hosting services have evolved

¹¹ DSA arts 3(g), 3(i), 3(j) and 33; see also recitals 5–7 and 13–14.

¹² Electronic Communications and Transactions Act 25 of 2002 (ECTA), ss 73–76.

into platforms exercising qualitatively different forms of power over information, markets and public discourse.

1.3 Regulation should follow function rather than corporate identity

A South African framework should retain this functionally based and technologically neutral approach. Whether a provider falls within a category should depend principally upon what the relevant service does, rather than upon the label adopted by the company providing it. A large technology company may simultaneously provide communications, cloud-storage, hosting, search and social-media services. Different functions may appropriately attract different legal obligations.

This approach also provides greater resilience against technological change. Legislation framed around particular companies or existing business models risks becoming obsolete as technologies develop. Functional categories are more capable of accommodating new services that perform substantially equivalent roles.

At the same time, South Africa should review whether ECTA's existing definitions remain adequate for the contemporary digital ecosystem. Any reform should make clear that intermediary services may encompass modern infrastructure services where they perform the relevant functions, while avoiding the assumption that every participant in the technical infrastructure of the internet should be regulated as though it were a social-media platform.

1.4 Differentiated obligations are essential

The most important principle to borrow from the DSA is therefore not necessarily its precise definitions, but its graduated allocation of responsibility. An internet-access provider that merely transmits information should not be required to undertake a systemic-risk assessment concerning public discourse. A cloud-storage provider that stores material privately for a customer should not necessarily be subject to rules designed for platforms that algorithmically recommend content to millions of members of the public. A small South African discussion forum should not bear the same compliance burden as a global social-media service used by millions of South Africans.

The obligations considered in this report should therefore attach selectively. Basic obligations, such as maintaining an appropriate point of contact and complying with lawful orders, may reasonably apply across a broad category of intermediary services. Notice-and-action requirements are particularly relevant to hosting services because those providers have the technical capacity to remove or disable access to stored information. Requirements concerning the explanation of content-moderation decisions, advertising transparency and recommender systems are most relevant to online platforms performing those functions.

The most demanding requirements considered in this report — including systemic-risk assessment and mitigation, enhanced researcher access, independent audit and potentially a supervisory levy — should principally be reserved for platforms and search engines whose reach or significance within South Africa makes genuinely systemic consequences possible.

1.5 Determining systemic significance in South Africa

South Africa should not mechanically reproduce the DSA's threshold for designation as a VLOP or VLOSE. The European threshold of 45 million average monthly active recipients¹³ reflects the scale of the European Union and corresponds broadly to ten percent of its population. A South African threshold would need to be independently determined.

User numbers are an obvious starting point, but should not necessarily be the sole criterion. A service with fewer users may nevertheless occupy an unusually important position in a particular part of the information ecosystem. Relevant considerations could therefore include South African user numbers and reach; the nature of the service; its role in the dissemination or discovery of public information; its market position; the extent to which information is algorithmically ranked, recommended or amplified; and its capacity materially to affect South African public discourse, markets or the exercise of constitutional rights.

Any designation process should nevertheless be based upon sufficiently objective statutory criteria. A discretionary power to designate particular services as “systemically significant” without clear criteria could create uncertainty and risks of selective or politically motivated regulation.

1.6 A South African regulatory pyramid

The resulting framework would therefore resemble the DSA in structure without necessarily reproducing its categories or thresholds exactly. At its base would be a broad and technologically neutral category of intermediary services, incorporating the familiar functions of transmission, caching and hosting already recognised by ECTA. Within this would sit hosting services and, more particularly, services that disseminate user-provided information to the public. Additional obligations would attach where the provider operates an online platform or search engine, and the strongest obligations would apply only where the service reaches a defined level of scale or systemic significance in South Africa.

This differentiated approach is especially important in a jurisdiction with limited regulatory resources. It directs the greatest regulatory attention towards the services capable of producing the greatest societal consequences, while avoiding unnecessary burdens upon smaller providers and the technical infrastructure on which the functioning of the internet itself depends.

¹³ DSA art 33(1) and recital 76.

The central principle should therefore be proportionality of regulatory responsibility to function, capacity and systemic significance. South Africa need not regulate every digital intermediary in the same way in order to achieve effective regulation of the digital sphere.



Section 2

Intermediary liability and safe harbours: Modernising South Africa's existing framework

A framework looking to regulate digital platforms and reduce the risk they pose of societal harm should not proceed from the assumption that intermediaries should become generally liable for everything their users say or do.

On the contrary, limitations upon intermediary liability — or “safe harbours” — were not only important for the internet's early development but remain a fundamental component of an open internet and of the protection of freedom of expression.¹⁴ The challenge is to preserve appropriate protection for intermediaries in relation to third-party content, while ensuring that such protection does not relieve platforms of responsibility for their own conduct, systems and regulatory obligations.

In this respect, South Africa already possesses an established statutory framework. Chapter XI of ECTA creates limitations upon the liability of service providers performing specified intermediary functions. The question is therefore not whether South Africa should introduce safe-harbour provisions — it already has them — but whether its existing regime remains appropriate more than two decades after its enactment and what might usefully be learnt from the DSA's modernisation of the same basic regulatory tradition.

2.1 The existing ECTA framework

Chapter XI of ECTA distinguishes several circumstances in which qualifying service providers may obtain protection from liability. Section 73 protects providers performing a transmission or access function, broadly corresponding to “mere conduit”, provided that specified conditions are satisfied, including that the

¹⁴ See, in the EU context, Directive 2000/31/EC (E-Commerce Directive) arts 12–15; DSA arts 4–8. South Africa adopted a broadly comparable conditional safe-harbour structure in Chapter XI of ECTA.

provider does not initiate the transmission, select the addressee or modify the data contained in the transmission.

Section 74 establishes a limitation upon liability for automatic, intermediate and temporary storage undertaken for the purpose of making onward transmission more efficient: the familiar caching function.

Section 75 addresses hosting. Subject to its requirements, a service provider is not liable for damages arising from information stored at the request of a recipient where the provider does not have actual knowledge that the information or associated activity infringes the rights of a third party, and is not aware of facts or circumstances from which such infringement is apparent; or where, upon receiving a qualifying notification, it acts expeditiously to remove or disable access to the information.

Section 76 provides further protection in relation to information-location tools, including referring or linking users to locations containing infringing information, subject again to conditions concerning knowledge and expeditious action.

Section 78 provides that, when supplying the relevant services, a provider is under no general obligation to monitor the data it transmits or stores or actively to seek facts or circumstances indicating unlawful activity — for which the DSA provides a clear analogue in article 8. The architecture is therefore already recognisably one of conditional intermediary immunity rather than blanket immunity.¹⁵

As discussed below, however, Chapter XI does not exhaust the obligations imposed upon intermediaries under South African law, and ECTA itself preserves duties arising under other statutory regimes.

2.2 An unusual threshold requirement: membership of a recognised representative body

There is, however, a distinctive feature of the South African regime that warrants reflection. Under section 72 of ECTA, the Chapter XI limitations upon liability are available only where the service provider is a member of an industry representative body recognised by the Minister under section 71 and has adopted and implemented that body's official code of conduct.¹⁶

The High Court has accordingly described Chapter XI as creating “safe harbour defences” for qualifying service providers.¹⁷ But access to those defences is conditioned not simply upon the provider's conduct in relation to the relevant information, but upon participation in a particular co-regulatory structure.

¹⁵ ECTA ss 73–78.

¹⁶ The Internet Service Providers' Association (ISPA) was recognised as an industry representative body for purposes of s 71 of ECTA by GN 588, Government Gazette 32252 (22 May 2009).

¹⁷ *Ketler Investments CC t/a Ketler Presentations v Internet Service Providers' Association* 2014 (2) SA 569 (GSJ) para 74.

There may have been good reasons for this model when ECTA was enacted. Membership of a recognised representative body could encourage adherence to common standards and provide an institutional mechanism for ensuring responsible conduct. But it is unclear whether membership of such a body should still determine the availability of a fundamental rule governing intermediary liability.

The DSA's approach is different. Its liability exemptions arise directly from the statutory conditions applicable to the relevant intermediary function. A provider performing a mere-conduit, caching or hosting function does not have to join an industry association before it can rely upon the applicable limitation of liability.

There is a strong argument for South Africa to consider separating these two questions. As we propose in section 11 of the report, the creation of a co-regulatory body has much to recommend it, especially considering South Africa's institutional and resource constraints. But the basic question of whether an intermediary should be held liable for information created by somebody else should arguably turn upon the nature of the service and the intermediary's own knowledge and conduct, rather than upon voluntary membership of an approved association.

2.3 Safe harbours protect more than intermediaries

The case for retaining intermediary safe harbours is not principally about protecting large technology companies from legal liability. It is also about protecting the communications environment in which users exercise freedom of expression. If a provider hosting millions of communications faces liability whenever any of those communications subsequently proves unlawful, the rational response is likely to be precautionary removal. Faced with uncertainty, an intermediary has little commercial incentive to defend controversial or borderline expression on behalf of a user if leaving the material available exposes it to substantial damages.

Overbroad intermediary liability can therefore produce private censorship by risk aversion. Conditional safe harbours respond to this problem by recognising that an intermediary should not automatically be treated as the publisher or author of information merely because its technical service transmitted or stored that information.

This principle is especially important for services performing genuinely neutral technical functions. An internet-access provider should not ordinarily have to determine the legality of every communication passing through its network. Nor should providers be compelled to monitor all user activity in an attempt to discover possible unlawful conduct.

Both ECTA and the DSA consequently reject a general monitoring obligation. The DSA similarly preserves liability exemptions for mere conduit, caching and hosting, while allowing judicial and administrative authorities to require action in particular cases. Any South African reform should preserve these principles.

2.4 Hosting requires a different balance

Hosting services raise more difficult questions because they store information and are capable, once appropriately notified, of disabling access to particular material. Both ECTA and the DSA consequently make hosting immunity conditional upon the provider's knowledge and response.

There are, however, differences worth examining. ECTA section 75 is framed principally in terms of knowledge that material or activity is infringing the rights of a third party.¹⁸ Article 6 of the DSA uses the broader concepts of "illegal activity" and "illegal content".¹⁹ The latter is defined by reference to information or activity that does not comply with applicable EU or Member State law, irrespective of the subject matter concerned.

South Africa should not simply replace the ECTA formulation with the DSA wording without examining the consequences. The appropriate scope of the hosting safe harbour requires consideration of the different forms of civil and criminal unlawfulness potentially arising online and of the relationship between the safe harbour, court orders, statutory reporting obligations and the notice-and-action procedure discussed in section 4 of this report.

The underlying principle should nevertheless remain that mere receipt of an allegation should not automatically establish illegality or liability. A notice alleging that content is defamatory, for example, does not determine that it is defamatory. Questions of truth, public interest, privilege and other defences may be highly contestable. A system that effectively requires immediate removal whenever a complainant alleges unlawfulness would create powerful incentives for abuse and would be especially dangerous for journalism and public-interest expression.

18 Section 75(1) of ECTA provides: "A service provider that provides a service that consists of the storage of data provided by a recipient of the service, is not liable for damages arising from data stored at the request of the recipient of the service, as long as the service provider:

- (a) does not have actual knowledge that the data message or an activity relating to the data message is infringing the rights of a third party; or
- (b) is not aware of facts or circumstances from which the infringing activity or the infringing nature of the data message is apparent; and
- (c) upon receipt of a take-down notification referred to in section 77, acts expeditiously to remove or to disable access to the data."

19 Article 6(1) of the DSA provides: "Where an information society service is provided that consists of the storage of information provided by a recipient of the service, the service provider shall not be liable for the information stored at the request of a recipient of the service, on condition that the provider:

- (a) does not have actual knowledge of illegal activity or illegal content and, as regards claims for damages, is not aware of facts or circumstances from which the illegal activity or illegal content is apparent; or
- (b) upon obtaining such knowledge or awareness, acts expeditiously to remove or to disable access to the illegal content."

2.5 Notice, knowledge and removal

The relationship between notice and knowledge is consequently critical. Article 16 of the DSA creates a structured notice-and-action mechanism for hosting providers. Importantly, the DSA provides that notices meeting specified requirements may give rise to actual knowledge or awareness for purposes of the Article 6 hosting exemption where they enable a diligent provider to identify the illegality of the relevant activity or information without a detailed legal examination.

This is substantially more nuanced than a rule under which every complaint automatically creates knowledge and triggers a legal obligation to remove. A revised South African framework should similarly distinguish between notification of alleged illegality and determination of illegality. A sufficiently precise and substantiated notice should require responsible consideration and may, depending upon the clarity of the circumstances, affect the provider's entitlement to rely upon a safe harbour. But the framework should not encourage automatic takedown merely because a notice has been submitted.

The procedural requirements governing notices, reasons, appeals and safeguards against abusive complaints are considered in section 4 of this report. For purposes of intermediary liability, the important point is that notice-and-action and safe harbour are connected but distinct: the former provides a process through which alleged illegality is brought to the provider's attention; the latter determines the circumstances in which the provider may itself become legally exposed in relation to third-party information.

2.6 Voluntary moderation should not destroy the safe harbour

A further DSA provision may be particularly valuable in encouraging responsible platform behaviour. Article 7 provides that an intermediary should not lose the benefit of the liability exemptions solely because it voluntarily undertakes good-faith and diligent investigations or other measures aimed at detecting, identifying, removing or disabling access to illegal content, or takes necessary measures to comply with legal requirements.

This addresses a potential regulatory perverse incentive. If voluntarily searching for unlawful content automatically increased a provider's exposure to liability, providers might rationally choose to know as little as possible about activity occurring through their services. South African reform should consider adopting an equivalent principle. Responsible voluntary measures undertaken in good faith should not, merely by virtue of having been undertaken, deprive an intermediary of statutory protection.

Such a provision would require appropriate safeguards. It should not immunise a provider from liability for its own unlawful conduct or permit claims of voluntary moderation to defeat obligations otherwise imposed by law. Its purpose would

simply be to ensure that the safe-harbour regime does not reward deliberate ignorance.

2.7 No general monitoring obligation

The prohibition upon general monitoring deserves explicit preservation. The internet could not function on the basis that every intermediary must inspect every communication passing through or stored by its systems in order to determine whether it might be unlawful. Such a requirement would be technically burdensome, intrusive of privacy and potentially devastating for freedom of expression.

A prohibition upon general monitoring does not mean that platforms can ignore specific information brought to their attention. Nor does it preclude targeted monitoring required by a lawful order in appropriate circumstances, or the imposition of specific statutory duties concerning identified forms of serious unlawful conduct. South African law already contains examples of the latter, particularly in relation to child sexual abuse material, where the Films and Publications Act and the Sexual Offences Act impose obligations that extend beyond ECTA's general safe-harbour conditions.

The distinction is between such targeted or subject-specific duties and a general obligation to inspect all communications for possible illegality. Section 78 establishes the latter proposition: the state should not convert intermediaries into general surveillance mechanisms charged with continuously searching all user communications for possible unlawful activity. That principle should remain a central safeguard in any reformed South African framework.

2.8 Safe harbour does not mean regulatory immunity

Perhaps the most important conceptual lesson from the DSA is that protection from liability for third-party content does not entail immunity from regulation. These are different questions. The first asks when an intermediary should itself be legally liable for information or conduct originating with a user. Safe-harbour provisions principally answer that question.

Importantly, ECTA's safe-harbour regime is not exhaustive of the obligations that South African law may impose upon intermediaries. Section 79 expressly preserves obligations imposed under other regulatory regimes and obligations imposed by law or court order to remove, block or deny access to data. In particular contexts, therefore, sector-specific legislation may impose duties that go beyond the conditions attaching to ECTA's limitation of liability.

This is particularly apparent in relation to child sexual abuse material. The Films and Publications Act²⁰ imposes obligations upon internet service providers to take reasonable steps to prevent their services from being used for the hosting

20 65 of 1996.

or distribution of child pornography and, where the requisite knowledge exists, to prevent access, report the material to SAPS and preserve evidence. It imposes still more proactive obligations upon providers of child-oriented services, including obligations to moderate those services and take reasonable steps to prevent their use for offences against children.

The Criminal Law (Sexual Offences and Related Matters) Amendment Act²¹ similarly requires an electronic communications service provider that becomes aware that its service or network is being used in the commission of specified child-pornography offences to report the offence, preserve relevant information and take reasonable steps to prevent access to the material.

South African law therefore already recognises that particular categories of especially serious unlawful content may justify more onerous intermediary obligations than those arising under the general ECTA safe-harbour regime. The difficulty is that these obligations have developed piecemeal across different statutes, employ different definitions of the intermediaries to which they apply, and do not necessarily extend readily to the contemporary online platforms through which the relevant content may be hosted and disseminated.

The second question, distinct from when an intermediary should be liable for content originating from a user, asks what obligations attach to the intermediary because of its own operation of the service. A platform may enjoy a safe harbour in relation to a user's unlawful post, while simultaneously breaching statutory duties because it failed to operate a proper notice mechanism, provide reasons for a moderation decision, comply with transparency requirements, permit regulatory access to information or undertake a required systemic-risk assessment.

Attempts to respond to platform power by simply expanding secondary liability for user speech carry substantial risks for freedom of expression. They incentivise platforms to remove material whenever legal uncertainty arises and place private companies in the position of making defensive judgments about the legality of speech. A better, constitutionally compliant framework would retain robust protection against automatic liability for third-party expression, while imposing much clearer primary responsibilities upon platforms for their own systems, processes and conduct.

2.9 Modernisation rather than abandonment

South Africa therefore does not need to abandon the basic approach adopted in Chapter XI of ECTA. In important respects, that legislation anticipated principles that remain fundamental to contemporary intermediary regulation. What requires reconsideration is the framework surrounding those principles.

A modernised regime should consider whether safe-harbour eligibility should continue to depend upon membership of a recognised industry representative

21 32 of 2007.

body; whether the existing provisions adequately encompass the range of contemporary intermediary services; whether the formulation of hosting liability remains appropriate; how knowledge should arise through a modern notice-and-action procedure; and how statutory protection can be preserved for responsible voluntary moderation.

At the same time, the foundational safeguards should remain: liability should follow the intermediary's actual function and conduct; neutral intermediaries should not automatically be liable for information originating with users; and providers should not be subject to a general obligation to monitor everything transmitted or stored through their services. The result should be neither blanket immunity nor general platform liability. It should be a conditional safe harbour combined with enforceable duties directed at the intermediary's own conduct.

That combination provides an important foundation for the sections that follow. Having examined which services might fall within a South African regulatory framework, and the circumstances in which they may be protected from liability for third-party content, the report now proceeds to examine the positive obligations that should accompany the increasingly consequential role of digital platforms in South African life.



Section 3

Establishing jurisdiction and effective local accountability

The most fundamental requirement for any meaningful South African framework for the regulation of online intermediaries is logically prior to the substantive obligations considered later in the report: South Africa must be able to assert regulatory and legal authority effectively over significant digital services intentionally offered within the country.

This presents a particular difficulty in relation to global platforms. A provider may have millions of South African users, direct services and advertising towards the South African market, derive substantial commercial benefit from activity in South Africa and exercise considerable influence over the country's information environment, while the corporate entity providing the service remains established offshore. Where such an entity relies upon its status as a *foreign peregrinus*²² to resist the jurisdiction of South African courts or regulators, substantive regulation — no matter how considered — risks being of limited practical value.

South African law already provides a conceptual foundation for addressing this problem. In *Bid Industrial Holdings (Pty) Ltd v Strang and Another*,²³ the Supreme Court of Appeal reconsidered the traditional common-law rules governing jurisdiction over *foreign peregrini*. It held that, in the circumstances considered there, jurisdiction could be exercised where the foreign defendant was served while in South Africa and there was an “adequate connection” between the suit and the court's area of jurisdiction from the perspective of the appropriateness and convenience of the matter being determined there. The Court ultimately formulated the requirement as a “sufficient connection” between the suit and the relevant jurisdiction.

22 A person not domiciled or resident within the area of jurisdiction.

23 2008 (3) SA 355 (SCA).

More recently, the Constitutional Court's decision in *BNP Paribas v Competition Commission of South Africa*²⁴ illustrates the continuing importance of sufficient connection in determining the reach of South Africa's legal jurisdiction over foreign actors operating in an increasingly borderless economy. The case concerned allegations of manipulation of the USD/ZAR exchange rate involving foreign banks, including "pure peregrini" with no physical presence in South Africa. In the earlier proceedings, the Competition Appeal Court had developed an "adequate connecting factors" test for establishing personal jurisdiction over such entities, alongside a separate "qualified effects" test for subject-matter jurisdiction. Although the Constitutional Court was not called upon to reconsider that jurisdictional framework (the relevant earlier order had not been appealed), it proceeded on the basis that adequate connecting factors between the foreign respondent, the dispute and the South African forum were required. In doing so, referencing *Bid Industrial Holdings*, it emphasised that the inquiry concerns real connections between the suit and the local forum.

The analogy with digital intermediaries should not be pressed too far: *BNP Paribas* concerned the particular statutory framework of competition law and does not itself establish jurisdiction over foreign online platforms. It is nevertheless instructive. The litigation reflects the increasing difficulty of applying jurisdictional rules developed around territorial presence to economic actors whose activities transcend national borders. Indeed, the Competition Appeal Court had expressly observed in the litigation that modern technology has produced a global economy in which business is no longer necessarily based on "bricks and mortar" and national borders may readily be transcended.²⁵ For present purposes, the significance of the case lies less in supplying a ready-made jurisdictional rule than in reinforcing the legitimacy of asking whether there are real and adequate connections between a foreign actor's activities and South Africa.

A South African digital-services regulatory framework could give considerably greater certainty to that inquiry. Rather than requiring courts and regulators repeatedly to determine, through litigation, whether adequate connecting factors exist, legislation could specify the circumstances in which a foreign intermediary will be regarded as sufficiently connected to South Africa to attract defined regulatory obligations.

The DSA provides a useful comparative model for doing so: its application does not depend upon where an intermediary provider is incorporated or headquartered. It applies to intermediary services offered to recipients located within the European Union, irrespective of the provider's place of establishment. The DSA uses the concept of a "substantial connection" with the EU to distinguish genuine participation in the European market from the mere technical accessibility of a

24 *BNP Paribas v Competition Commission of South Africa; Credit Suisse Securities (USA) LLC v Competition Commission of South Africa; Competition Commission of South Africa v Bank of America Europe DAC* [2026] ZACC 28.

25 *Competition Commission of South Africa v Bank of America Merrill Lynch International and Others* [2024] 1 CPLR 1 (CAC) (8 January 2024) para 85.

service there. Relevant factual criteria include a significant number of users within the Union and the targeting of activities towards one or more Member States.²⁶

A South African framework could similarly provide expressly that specified obligations apply to an intermediary provider established outside South Africa where there are adequate connecting factors between the provider or the provision of its services and South Africa. Rather than leaving those connections wholly to case-by-case determination, legislation could identify relevant indicia. These might include:

- ▶ a significant number of recipients of the service in South Africa;
- ▶ deliberate targeting or adaptation of the service towards South African users;
- ▶ providing the service or interface in South African languages, particularly languages whose use demonstrates deliberate localisation for this market, such as isiZulu, isiXhosa or Afrikaans;
- ▶ the sale of advertising directed towards South African audiences;
- ▶ commercial activity or revenue derived from the South African market;
- ▶ the use of South African domains, currencies, payment systems, marketing or other locally adapted functionality; and
- ▶ other conduct demonstrating sustained and purposeful participation in the South African digital environment.

Language is potentially an especially useful indicator in the digital context. The fact that a globally available English-language service can be accessed in South Africa establishes relatively little about whether South Africa has specifically been targeted. A provider that takes deliberate technical and commercial steps to make its service available in isiZulu or isiXhosa, by contrast, is much more obviously designing and offering that service for South African users. That should at least constitute strong evidence of purposeful participation in the South African market.

The precise statutory threshold would require careful formulation. Mere accessibility of a foreign website from South Africa should plainly not suffice. The objective would instead be to identify a meaningful degree of purposeful engagement with the country proportionate to the regulatory obligations that follow.

3.1 The importance of a local legal representative

For significant intermediary providers satisfying that threshold but lacking an establishment in South Africa, legislation should require the appointment of a South African legal representative. The representative should be expressly authorised to receive legal process and regulatory notices, respond to information

²⁶ DSA art 2(1) and recital 8. See also art 3(e).

demands and engage with relevant authorities concerning compliance. The DSA's separate requirement for accessible points of contact could similarly be adapted.

Recent South African experience demonstrates why such a requirement is vital. In 2025, proceedings brought against Meta sought urgently to stop a network of anonymised WhatsApp channels and Instagram accounts from soliciting and disseminating intimate and pornographic images of South African schoolchildren. A High Court order was successfully obtained, requiring Meta, among other things, to remove identified accounts, prevent further accounts from being established and disclose identifying information concerning those responsible.²⁷

Yet the litigation exposed a fundamental difficulty in obtaining effective accountability from a multinational platform operating extensively within South Africa. The order was met with only partial compliance, at best. This was despite Meta having personnel in South Africa and being represented by South African lawyers. In the exchange of correspondence prior to the hearing of the urgent application, Meta's South African attorneys replied that they had received no instructions, that they were not authorised to accept service on behalf of any Meta entity and that correspondence should instead be directed to Meta's legal department in California. This occasioned delay at a time when the gravest possible harm was being inflicted.

Only after the institution of contempt proceedings, and the possible arrest of its local personnel, did Meta finally engage and agree to a consent order. This required that, among other measures, it remove reported accounts and channels as far as is technically feasible, provide identifying information concerning their creators and establish a direct reporting channel with The Digital Law Company for urgent child-protection matters.

The case illustrates the problem starkly. Where the most serious harm is occurring through a service extensively used in South Africa, the effectiveness of South African law should not depend upon litigants first navigating the faraway corporate architecture of a multinational technology company to identify an offshore entity and establish how valid legal process can reach it.

Nor is the appropriate response merely to require platforms to nominate a postal address. A local representative must possess a legally defined mandate sufficient to make the arrangement consequential. Legislation should therefore specify the representative's authority to receive service and regulatory demands and should determine what consequences follow when the provider fails to comply. Otherwise, the requirement risks producing a nominal local presence while leaving substantive responsibility offshore.

27 For an account of this case, see The Digital Law Company's own report available on their home page <https://www.thedigitalawco.com/> accessed 24 August 2026. See also Nicole Fritz, 'Big Tech in SA Exposed in Chilling WhatsApp, Instagram Child Sex Accounts Case', *News24* (21 July 2025) <https://www.news24.com/opinions/columnists/nicole-fritz-big-tech-in-sa-exposed-in-chilling-whatsapp-instagram-child-sex-accounts-case-20250720-0886> accessed 24 August 2026.

The Meta proceedings also illustrate a broader point. It is not that South Africa should insist that every global digital company incorporate a South African subsidiary or maintain extensive operations here. Rather, it is this: a company that intentionally and substantially participates in the South African digital market should not be permitted to structure that participation so that the legal entity controlling the service remains practically inaccessible to South African courts and regulators.

This is the jurisdictional infrastructure upon which almost every other proposed adaptation of the DSA depends. Obligations for accessible complaint, notice and redress procedures; to publish South Africa-specific transparency reports; disclose advertising information; provide access to data; undertake systemic-risk assessments or comply with regulatory directions have limited value unless there is an identifiable entity upon whom those obligations can effectively be imposed and against whom they can ultimately be enforced.



Section 4

Accessible complaint, notice and redress mechanisms

Once jurisdiction and local accountability have been established, a second foundational requirement is to ensure that persons affected by conduct on online services have accessible and effective means of engaging the intermediary concerned. A regulatory framework is of limited practical value if a person confronted with unlawful content cannot readily report it, cannot determine whether the complaint has been considered, and cannot obtain reasons for the platform's response. Equally, users whose content is removed, or who have their accounts suspended or the visibility of their content restricted, should have some means of understanding and challenging those decisions.

This is an area in which the DSA offers a particularly compelling model for South Africa. Rather than attempting to prescribe in advance how platforms must decide every content dispute, it establishes procedural rules governing how those disputes must be received, considered, decided and, where appropriate, reviewed. Articles 16, 17 and 20 of the DSA operate together to create a chain of procedural accountability: a mechanism for notifying hosting services of allegedly illegal content; an obligation to explain significant moderation decisions; and, for online platforms, the requirement of an internal mechanism through which specified decisions can be challenged.

It is necessary, however, to distinguish two related processes within this framework. The first concerns notices submitted to a hosting provider alleging that particular content is illegal. Article 16 requires hosting providers to establish a statutory notice-and-action mechanism for this purpose. The second concerns the platform's own exercise of content-moderation powers. Platforms may restrict content not only because it is unlawful but because it contravenes the platform's own terms and conditions, which may prohibit categories of content that are perfectly lawful. Where platforms exercise this broader contractual power, the DSA does not prescribe the substantive standards they must adopt, but it does impose procedural requirements concerning transparency, reasons and opportunities to challenge specified decisions.

The DSA regulates illegal content and platform-defined impermissible content differently. In the former case, the law determines the substantive standard (by reference to existing law) and requires a mechanism through which alleged illegality can be brought to the provider's attention. In the latter, the platform principally determines what additional categories of content it will not permit on its service, while the DSA regulates the fairness and transparency with which those rules are applied.

4.1 A standardised mechanism for notifying platforms of illegal content

Article 16 of the DSA requires hosting services to establish mechanisms through which any individual or entity can notify them of specific information considered to constitute illegal content. The mechanisms must be electronic, easy to access and user-friendly. They must enable sufficiently precise and adequately substantiated notices, including an explanation of why the information is alleged to be illegal and information enabling the provider to locate it.

A platform should obviously not be entitled to make itself practically inaccessible to people attempting to report unlawful conduct occurring through its service. But at the same time, the mere assertion that content is unlawful should not automatically prompt its removal. The notice must contain sufficient information to permit a proper assessment, and the provider must make a decision about the particular content concerned.

This distinction is especially important in South Africa, given the constitutional protection afforded to freedom of expression. A South African requirement of a notice-and-action system should look to facilitate effective action against content alleged to be unlawful under South African law, rather than creating a general mechanism for demanding removal of content because it is considered offensive, harmful, false or objectionable.

The distinction also helps preserve the proper allocation of responsibility to Parliament, courts and platforms. The underlying law determines what is unlawful; the intermediary provides a mechanism through which allegedly unlawful material can be brought to its attention and addressed. A particularly salient feature of the DSA is that it contemplates complaint notices that have procedural and potentially liability-related consequences without turning the complainant into the adjudicator of illegality.

Under Article 16, sufficiently precise and adequately substantiated notices may give a hosting provider actual knowledge or awareness relevant to its conditional exemption from liability under Article 6. The provider is therefore placed in a position where it must properly consider the notification and decide whether action is warranted. It must process notices in a timely, diligent, non-arbitrary and objective manner.

For South Africa, a comparable framework could bring considerably greater certainty to the existing intermediary-liability regime. Instead of users navigating different and sometimes opaque reporting systems maintained at platforms' discretion, legislation could prescribe minimum characteristics of a valid notification process and the consequences that follow once an intermediary has received a sufficiently substantiated notice of allegedly unlawful content.

There should, however, be no assumption that receipt of a notice invariably requires removal. Such a rule would create powerful incentives towards over-removal: faced with legal uncertainty, the rational response of a platform would often be to remove lawful but contested speech rather than risk liability for leaving it online. The objective should instead be responsible decision-making following notice.

A South African adaptation should pay particular attention to what constitutes an accessible complaint mechanism. It would not be sufficient for a global intermediary technically to provide a reporting form buried within an international help centre, particularly if the process is difficult to navigate, available only in English or disconnected from any meaningful human or locally informed review.

The DSA requires notice mechanisms to be easy to access and user-friendly. Its recitals go further, envisaging mechanisms that are clearly identifiable and at least as easy to find and use as mechanisms for reporting breaches of a platform's own terms and conditions. In the South African context, accessibility might require attention to:

- ▶ availability through the platform or service itself;
- ▶ clear and intelligible instructions;
- ▶ mobile accessibility, given patterns of internet access in South Africa;
- ▶ reasonable accommodation of South African languages;
- ▶ accessibility for persons with disabilities; and
- ▶ appropriately expedited channels where the alleged illegality involves imminent or particularly serious harm.

The final consideration is particularly important. Not every complaint requires the same response time. A disputed defamatory statement versus, for example, the active dissemination of child sexual abuse material, or credible threats of violence or non-consensual intimate images plainly presents different degrees of urgency. A South African framework could require providers to maintain mechanisms capable

of triaging complaints according to the nature and immediacy of the alleged harm, while preserving appropriate safeguards against abuse of urgent processes.²⁸

4.2 Complainants should know what happened to their complaints

A complaint mechanism is of little value if a notice disappears into a platform's systems without acknowledgement or explanation. The DSA therefore requires hosting providers, where contact information has been supplied, to confirm receipt without undue delay and subsequently notify the complainant of the decision taken and the possibilities for redress.²⁹

This apparently modest requirement has considerable practical significance. It converts reporting from one-way communication into an accountable process. A South African framework could similarly require platforms to acknowledge qualifying notices, provide some indication of their status and communicate the eventual outcome. Where no action is taken, the complainant should be able to understand, at least in broad terms, why: for example, because the information could not be located; because insufficient information was provided; because the platform concluded that the material was not unlawful; or because some alternative measure was considered appropriate.

In instances where a complaint is acted upon, platforms must be accountable to users whose content they restrict. Article 17 of the DSA requires hosting providers to give affected users a clear and specific statement of reasons for specified restrictions imposed because content is considered illegal or incompatible with the provider's terms and conditions. Among other things, the user must be told what measure has been imposed, the facts and circumstances underlying the decision, whether automated means were used, the legal or contractual grounds relied upon, and what mechanisms exist for challenging the decision.

4.3 Trusted flaggers and prioritisation of notices

The DSA supplements its general notice-and-action mechanism with a system of "trusted flaggers". Article 22 requires online platforms to give priority to notices submitted by designated trusted flaggers and to process and decide upon them without undue delay. Trusted-flagger status is conferred by the relevant national digital services coordinator upon entities that have demonstrated particular expertise and competence in detecting, identifying and notifying illegal content;

28 Such a system might, for example, require complaints involving an apparent imminent threat to life, child sexual abuse material or the ongoing dissemination of non-consensual intimate imagery to be immediately escalated for specialist review, while an allegation of defamation ordinarily proceeds through the standard notice-and-action process. Triage should determine the urgency and level of scrutiny afforded to a complaint, rather than predetermine whether the complained-of content is unlawful or must be removed.

29 Article 16(4) & (5).

are independent from online platforms; and carry out their activities diligently, accurately and objectively.

Importantly, trusted-flagger status does not confer authority to determine that particular content is unlawful, nor does a notice from a trusted flagger automatically require its removal. The platform remains responsible for deciding how to respond to the notice. The effect of the designation is principally to ensure that notices from demonstrably reliable and expert sources receive priority and expedited consideration.

The concept of trusted flaggers is not novel to South Africa. Certain platforms already maintain voluntary arrangements under which selected South African organisations receive privileged or expedited reporting channels. Moxii Africa (formerly Media Monitoring Africa) and the Electoral Commission have been accredited as trusted flaggers by Meta to facilitate the escalation of election-related complaints through the Real411 process.³⁰ Childline South Africa has similarly been identified as a participant in YouTube's Trusted Flagger programme in the child-safety context.³¹

These arrangements demonstrate the potential utility of differentiated treatment for notices submitted by organisations possessing particular expertise. They also, however, highlight an important difference between voluntary platform practice and the DSA model. In South Africa, trusted status is presently conferred by platforms themselves, apparently according to their own criteria and on terms determined by them. Article 22 instead places accreditation within a statutory framework: status is conferred by an independent digital services coordinator according to prescribed criteria of expertise, independence, diligence, accuracy and objectivity, and carries corresponding transparency and accountability obligations.

A South African adaptation might therefore build upon, rather than replace, existing practices. The case for regulation would lie not simply in creating expedited reporting channels—some already exist—but in making the criteria for trusted status more transparent, consistent and independent of unilateral platform discretion, while establishing safeguards against abuse. Consideration would also need to be given to whether accreditation should be undertaken by a public regulator or through some other suitably independent mechanism, particularly given South Africa's institutional-capacity constraints.

Importantly, neither existing platform practice nor any future statutory trusted-flagger regime should confer upon designated organisations authority to determine

30 Electoral Commission of South Africa, 'Electoral Commission partners with social media giants to combat disinformation in 2024 National and Provincial Elections' <https://www.elections.org.za/content/About-Us/News/Electoral-Commission-partners-with-social-media-giants-to-combat-disinformation-in-2024-National-and-Provincial-Elections/> accessed 15 September 2026.

31 Google and YouTube, 'Aotearoa New Zealand Code of Practice for Online Safety and Harms' page 5, https://technewzealand.org.nz/wp-content/uploads/sites/8/2022/12/Google-Australia-and-New-Zealand_Baseline-Report_New-Zealand-Code-of-Practice-for-Online-Safety-and-Harms_2022.pdf accessed 15 September 2026.

what content is unlawful. The value of trusted status lies in reliable identification and expedited consideration, not delegated censorship. A trusted flagger's notice should receive priority because of the organisation's demonstrated expertise and record of responsible reporting; the ultimate determination of how to respond remains subject to applicable law and the appropriate decision-making process.

As with the general notice-and-action mechanism, the distinction between priority of consideration and determination of illegality is crucial. A trusted flagger should be able to ensure that a properly substantiated notice reaches the front of the queue; it should not be able, merely by virtue of its designation, to determine what South Africans may lawfully say or access online.

4.4 A second track: accountability for platforms' own moderation decisions

The notice-and-action process described above concerns content alleged to be illegal. It should be distinguished from a second form of complaint and redress addressed by the DSA: accountability for decisions taken by platforms in the exercise of their own content-moderation powers.

Platforms are not limited to restricting content that is unlawful. They ordinarily establish their own terms and conditions or community standards and may decide that certain categories of content are not permitted on their services, even though the content is lawful. The DSA largely leaves the substantive content of those additional rules to the platforms themselves. It does, however, regulate how those rules are communicated and enforced.

This distinction is reflected in Article 17. Where a hosting provider imposes specified restrictions because it considers information either illegal or incompatible with its terms and conditions, it must provide the affected user with a clear and specific statement of reasons. The user must be told what measure has been imposed, the facts and circumstances underlying the decision, whether automated means were used, whether the decision rests upon illegality or the platform's contractual rules and what mechanisms exist for challenging it.

This is an important distinction for possible South African adaptation. A South African regulatory framework need not require the state to determine all categories of content that a private platform may or may not carry. Platforms could remain free, within applicable law, to establish additional community standards governing their services. But where they exercise that private rule-making and enforcement power in ways that significantly affect users' expression, regulation could require that power to be exercised transparently and accountably.

4.5 A meaningful right to challenge moderation decisions

The obligation to provide reasons is complemented, for online platforms, by a right to seek reconsideration. Article 20 of the DSA requires an internal

complaint-handling system through which affected users can challenge specified moderation decisions, including decisions based either upon alleged illegality or upon incompatibility with the platform's own terms and conditions.

The system must be electronic, free of charge, easy to access and user-friendly. Complaints must be handled in a timely, diligent, non-discriminatory and non-arbitrary manner, and the resulting decision may not be taken solely through automated means. Where the complaint demonstrates that the original decision was unfounded, the platform must reverse it without undue delay.

The significance of this requirement is not that every moderation decision becomes the equivalent of an administrative decision susceptible to elaborate legal proceedings. Rather, it establishes a minimum principle of contestability. Platforms will inevitably make mistakes, as will automated systems. That is perhaps especially true for moderators working at enormous scale and often without sufficient local context. A regulatory framework must therefore assume the possibility of error and require a practical mechanism for correcting it.

For South Africa, this could be especially important for journalists, media organisations, political actors, civil-society organisations and others whose ability to communicate with large audiences may depend heavily upon a small number of platforms. Erroneous account suspension or content removal can have substantial consequences, even where it is subsequently reversed, particularly during elections, breaking-news events or periods of political mobilisation.

The DSA's insistence that decisions on internal complaints are not to be taken solely by automated means is also relevant to South Africa. Automated moderation systems may be particularly susceptible to error where they encounter languages, dialects, idioms and political or cultural references poorly represented in their training or design. The requirement for meaningful reconsideration therefore intersects with the separate need, discussed below, for platforms operating at scale in South Africa to possess adequate linguistic and contextual competence.

Significantly, the requirement is not that every initial moderation decision be made by a human being. Given the volume of content handled by major platforms, that would be unrealistic. But where a user meaningfully challenges a consequential decision, there is a strong argument that the platform should provide access to genuine reconsideration rather than merely submitting the same content to the same automated process for a second time.

4.6 Complaint systems should not become mechanisms for private censorship

Notice mechanisms can themselves be weaponised. Political actors, corporations and organised groups may submit large numbers of complaints in an attempt to suppress criticism, investigative journalism or unpopular expression. The DSA recognises this problem and permits action against persons who frequently

submit notices or complaints that are manifestly unfounded, subject to procedural safeguards.³²

This concern is particularly acute if the submission of a notice has consequences for intermediary liability. A poorly designed system could create an inexpensive mechanism through which powerful actors exert pressure for removal simply by alleging unlawfulness.

South African regulation should therefore protect both sides of the process: it should make it substantially easier for victims of genuinely unlawful online conduct to obtain a response, while ensuring that the mechanism cannot readily be converted into a tool for suppressing lawful public-interest expression.

4.7 Escalation of suspected serious criminal conduct

The DSA mandates the design of an effective notice-and-action system that also recognises that some information coming to the attention of a hosting provider raises concerns extending beyond whether particular content should remain available on the service. Article 18 requires hosting providers that become aware of information giving rise to a suspicion of criminal activity involving a threat to the life or safety of persons to inform the relevant law-enforcement or judicial authorities promptly and provide the relevant information available to them.

A comparable principle is readily applicable in South Africa and is already reflected, in more limited forms, in existing legislation. Section 54 of the Cybercrimes Act³³ creates reporting and information-preservation obligations for electronic communications service providers that become aware of specified cybercrimes involving their services or networks, while expressly providing that the obligation does not require providers generally to monitor communications or actively seek evidence of unlawful activity.³⁴

There may therefore be value in considering whether South African law should provide a more coherent reporting obligation applicable to hosting providers that become aware, in the ordinary operation of their services or complaint mechanisms, of information providing reasonable grounds to suspect specified serious offences or imminent threats to life or physical safety. Such an obligation should be carefully delimited. It should arise from information of which the provider has actually become aware rather than creating a general monitoring obligation, and legislation

32 Article 23(2) of the DSA provides: "Providers of online platforms shall suspend, for a reasonable period of time and after having issued a prior warning, the processing of notices and complaints submitted through the notice and action mechanisms and internal complaints-handling systems referred to in Articles 16 and 20, respectively, by individuals or entities or by complainants that frequently submit notices or complaints that are manifestly unfounded."

33 Cybercrimes Act 19 of 2020, s 54.

34 See also the discussion at section 2.8 of this report of the obligations imposed by the Films and Publications Act and the Criminal Law (Sexual Offences and Related Matters) Amendment Act.

should specify the relevant threshold, categories of conduct, competent authority, reporting period and information-preservation requirements.

This would create an important escalation pathway within the broader complaints architecture: some matters require a platform moderation decision; sufficiently serious matters may additionally require referral to the institutions empowered to investigate crime and protect those at risk.

4.8 Independent review of unresolved platform disputes

Article 21 of the DSA addresses an important weakness that remains even where platforms are required to provide reasons for moderation decisions and maintain effective internal complaint procedures. A user who unsuccessfully challenges a decision internally may otherwise be left with an unattractive choice: accept the platform's decision or pursue potentially costly and time-consuming litigation.

The DSA therefore entitles affected users, including persons whose notices concerning allegedly illegal content have been rejected, to refer disputes concerning specified platform decisions to a certified out-of-court dispute-settlement body. Such bodies must satisfy requirements concerning independence, expertise, accessibility, procedural fairness and efficiency. The parties are required to engage with the process in good faith, although the resulting decision is not binding and the right to approach a court remains unaffected.

The problem Article 21 seeks to address is highly relevant to South Africa. An internal appeal determined ultimately by the same company responsible for the original decision does not provide genuinely independent redress. Yet requiring an ordinary user to approach a South African court to challenge, for example, the erroneous removal of content or suspension of an account will in most instances be neither practical nor proportionate. There is consequently a significant redress gap between internal platform processes and judicial proceedings.

The DSA's particular institutional solution is, however, less readily transferable. Establishing and supervising a system of certified dispute-resolution bodies would itself require regulatory resources and sufficient independent expertise in platform governance, freedom of expression, South African law and digital technologies. Significantly, even the DSA does not require EU Member States themselves to establish such bodies; rather, it creates a framework within which qualifying independent bodies can be certified.

For South Africa, the more immediately applicable lesson may therefore be the principle rather than the precise institutional model. Consideration should be given to providing an accessible and inexpensive means of independent review between the exhaustion of a platform's internal complaint process and resort to the courts. This might ultimately involve the accreditation of suitable existing alternative-dispute-resolution bodies, an ombud-type mechanism, or another comparatively lightweight arrangement rather than the creation of an entirely new institution.

Whatever model is adopted, it should be independent of the platforms concerned, sufficiently expert to adjudicate disputes involving both South African law and platforms' own terms and conditions, capable of acting expeditiously, and readily accessible to ordinary users. Particular care would also be required to ensure that independent dispute resolution does not displace the courts or confer upon a private body an inappropriate power to determine the boundaries of lawful expression finally.

The central insight of Article 21 is nevertheless valuable for South Africa: meaningful procedural fairness requires more than allowing a platform to review its own decisions. If platform power is to be genuinely contestable, there should ultimately be some realistic avenue of independent redress short of High Court litigation.

4.9 The relationship with courts and regulators

Finally, platform complaint procedures should supplement rather than displace South African legal institutions. Platforms should not become the final arbiters of whether expression is lawful. A person dissatisfied with a platform's response must remain able, where the law permits, to approach a court or relevant regulatory body. Conversely, where a competent South African court or regulator has lawfully determined that particular action must be taken, a platform's internal complaint procedure cannot substitute for compliance with that order.

The objective is therefore to create an effective first line of recourse, capable of resolving large numbers of matters rapidly and inexpensively without requiring every dispute to reach a court. This is another reason why the preceding requirement of local legal accountability matters. Particularly serious or unresolved complaints must ultimately be capable of escalation to an identifiable entity subject to South African law.

4.10 A foundational element of effective platform regulation

Taken together, these mechanisms establish a framework of procedural accountability operating in several directions. Those seeking action against allegedly illegal content must have an accessible means of bringing that content to a platform's attention and receiving a reasoned response. Those subjected to the platform's exercise of its own moderation powers must be able to understand significant decisions affecting their content or accounts and have a meaningful opportunity to challenge them. Where information coming to a hosting provider's attention indicates sufficiently serious criminal conduct or threats to life or safety, there should be an appropriate pathway for escalation to law-enforcement authorities. And where internal platform processes fail to resolve a dispute, there is value in providing some realistic avenue of independent redress short of costly court proceedings.

For South Africa, the distinction between illegal content and content that platforms choose to prohibit under their own terms and conditions is to be carefully and conscientiously observed. A statutory notice-and-action mechanism should

principally concern illegality as determined by South African law. It should not become a general mechanism through which complainants can invoke the coercive force of law to secure action against material merely because they regard it as harmful, false, offensive or otherwise objectionable. Platforms may choose to maintain broader content standards, but the state's principal concern in relation to those standards should be that they are transparent and that their enforcement is procedurally fair and meaningfully contestable.

This approach also provides an important means of protecting freedom of expression while improving accountability for online harms. It does not require the state to adjudicate every content dispute or prescribe the entirety of platforms' content standards. Nor does it leave platforms wholly free to determine how complaints and moderation disputes are handled. Instead, legislation can establish minimum procedural requirements governing accessibility, acknowledgement, timely consideration, reason-giving, opportunities for challenge and, where appropriate, escalation or independent review.

The attraction of this model for South Africa is also institutional. The primary responsibility for establishing and operating these systems rests with the intermediary itself. A South African regulator need not become a first-instance tribunal for the enormous volume of individual content disputes arising in relation to digital services. Its role can instead be to establish minimum standards, supervise whether providers' systems comply with them, identify persistent or systemic failures and intervene where necessary.

There is an important connection here with the jurisdictional requirements discussed in the preceding section. Complaint and redress mechanisms are meaningful only if there is ultimately an identifiable provider that is answerable to South African law. Particularly serious matters, repeated failures of platform processes and lawful directions issued by courts or regulators must be capable of escalation to an entity over which South African authorities can exercise effective jurisdiction.

Accessible notice, complaint and redress mechanisms should therefore not be regarded as peripheral consumer-protection measures. They constitute part of the basic infrastructure of effective platform regulation. Before requiring sophisticated algorithmic audits or systemic-risk assessments, South Africa should be able to insist upon something more elementary: when people in South Africa encounter allegedly unlawful conduct through a digital service, or when a service exercises significant power over their expression, there must be an accessible process through which the matter can be raised, someone required to respond, reasons capable of being understood, and a realistic means of challenging or escalating the resulting decision.



Section 5

Transparency and reporting

Once online intermediaries are effectively subject to South African jurisdiction and accessible complaint and redress mechanisms have been established, a further foundational requirement is transparency concerning how platforms exercise their powers and operate their services in South Africa.

This is one of the central themes of the DSA. Its transparency requirements do not serve a single purpose. Some enable individual users to understand decisions affecting them; others enable regulators, researchers and the public to assess how platforms operate at scale. Collectively, they seek to reduce a fundamental information asymmetry: platforms possess extensive information about how their systems operate and affect users, while those subject to, studying or regulating those systems ordinarily have very little.

This problem is especially acute in South Africa. Major platforms may publish substantial global transparency reports, but these frequently provide limited insight into their operations specifically within South Africa. It may therefore be difficult to determine how much content is removed or restricted in South Africa; on what grounds; how frequently automated systems are responsible; whether complaints and appeals are successful; what resources platforms devote to South African languages; or how advertising and recommender systems operate in the South African information environment.

A South African adaptation of the DSA should therefore focus not simply on requiring more information, but on requiring information that is locally relevant, sufficiently standardised, accessible and capable of meaningful comparison.

5.1 South Africa-specific transparency reporting

Article 15 of the DSA requires providers of intermediary services to publish periodic reports concerning content moderation in which they engage. The information required varies according to the type of service but includes matters such as orders received from public authorities, notices concerning allegedly illegal content,

content moderation undertaken on the provider's own initiative, use of automated systems, complaints received and the outcomes of those complaints.

Additional reporting requirements apply to online platforms under Article 24 and to Very Large Online Platforms and Very Large Online Search Engines under Article 42.

The underlying principle is readily adaptable to South Africa. Significant intermediary providers operating in South Africa could be required periodically to publish prescribed information concerning their South African operations specifically, rather than relying upon global or regional reports from which South African experience cannot readily be isolated.

Depending upon the nature of the service, relevant information could include:

- ▶ the number and categories of notices received concerning content alleged to be unlawful under South African law;
- ▶ the number and nature of requests or orders received from South African courts, regulators and law-enforcement authorities;
- ▶ the action taken in response to those notices and orders;
- ▶ content-moderation measures taken on the platform's own initiative, including the categories of content concerned and whether action was taken because of illegality or violation of the platform's terms and conditions;
- ▶ the number and categories of account suspensions, content removals, visibility restrictions and demonetisation decisions;
- ▶ the extent to which automated systems were used to detect content or make moderation decisions;
- ▶ the number of internal complaints or appeals lodged by South African users and their outcomes;
- ▶ information concerning the time taken to process notices and complaints, including urgent complaints;
- ▶ where trusted-flagger arrangements exist, the number of notices received through those channels and the action taken upon them; and
- ▶ information capable of demonstrating the adequacy of the platform's linguistic and contextual capacity in South Africa.

The objective should not be the indiscriminate production of data. Reporting requirements should be designed around information necessary to answer meaningful regulatory questions: Are platforms responding to complaints? How quickly? What are they removing and why? How much moderation is automated? How often are decisions successfully challenged? Are South African users and languages receiving adequate attention?

5.2 Standardisation and comparability

A significant lesson from the implementation of the DSA is that transparency is of limited value if every provider decides for itself what to report and how to categorise the information.

The EU has subsequently adopted harmonised templates for DSA transparency reports, standardising their format, content and reporting periods.³⁵ This followed experience demonstrating that variations in platforms' reporting practices and interpretations of data categories made their moderation practices difficult to compare. This lesson is particularly relevant to South Africa. A requirement that platforms simply "publish a transparency report" could produce documents containing extensive information, while permitting each provider to select categories and methodologies that make meaningful comparison impossible.

Any South African framework should therefore prescribe common definitions, reporting categories, periods and formats, appropriately differentiated according to the nature and size of the service. Where practicable, data should be published in machine-readable form so that regulators, researchers and civil-society organisations can analyse it without having first to extract information manually from lengthy corporate reports.

This need not require South Africa to reproduce the detail of the European reporting regime. A smaller number of carefully selected and consistently defined indicators may be more useful than imposing elaborate reporting obligations that South African regulators lack the capacity to analyse.

5.3 Transparency about individual moderation decisions and their aggregation

The requirement discussed in the preceding section that users receive reasons for significant moderation decisions has a wider transparency function. Under the DSA, statements of reasons supplied to users under Article 17 are, in the case of online platforms, also submitted in anonymised form to the European Commission's DSA Transparency Database. This creates a publicly accessible and machine-readable repository of platform moderation decisions. The database allows moderation activity to be examined not merely through periodic aggregate statistics but at a far more granular level.

This is an important regulatory innovation created by the DSA. A reason supplied only to the affected user provides individual procedural fairness. The aggregation of these reasons creates an extensive repository of evidence concerning how platforms actually exercise moderation power.

³⁵ Commission Implementing Regulation (EU) 2024/2835 of 4 November 2024 laying down templates concerning the transparency reporting obligations of providers of intermediary services and providers of online platforms under Regulation (EU) 2022/2065, especially recitals 1–6 and arts 1–2.

The scale of the resulting information has been substantial. The Commission reports that more than nine billion moderation decisions were reported to the database during the first half of 2025 alone, with approximately 99% taken proactively on the basis of platforms' own terms and conditions rather than following findings of illegality.³⁶ South Africa need not necessarily reproduce a database built on identical requirements. But there is considerable merit in the underlying principle. Significant online platforms could be required to submit standardised, anonymised information concerning consequential moderation decisions to a central repository or other publicly accessible system.

That information could permit researchers, journalists, regulators and civil society to identify patterns that individual users cannot see: whether particular forms of expression are disproportionately restricted; whether one platform relies unusually heavily upon automated moderation; whether certain categories of decisions are frequently reversed on appeal; or whether significant changes in moderation behaviour coincide with elections or other major public events.

Appropriate privacy safeguards would plainly be necessary. The objective would be transparency concerning platform decisions, not the creation of a government database of users' speech.

5.4 Transparency concerning advertising

The DSA also recognises that transparency is required not only where platforms restrict information but where they are paid to amplify it. Article 26 requires online platforms displaying advertisements to ensure that users can identify, clearly and in real time, that particular content is, in fact, advertising; the person on whose behalf the advertisement is presented; who paid for it where that is a different person; and meaningful information concerning the principal parameters used to determine why the advertisement was shown to that recipient.

These requirements are highly susceptible to South African adaptation. Users should be able to distinguish paid content from organic expression and know who is seeking to influence them and who is financing that attempt.

The principle has particular importance in the political and electoral context. Political communication may increasingly be channelled through paid influencers, sponsored posts and other forms of compensated content that do not resemble conventional advertisements. Transparency requirements could therefore extend beyond formally purchased platform advertisements to ensure appropriate

³⁶ European Commission, 'The impact of the Digital Services Act on digital platforms' <https://digital-strategy.ec.europa.eu/en/policies/dsa-impact-platforms> accessed 15 September 2026.

disclosure where political or electoral communication is disseminated for remuneration.³⁷

For the largest platforms, consideration could also be given to requiring publicly searchable advertising repositories containing prescribed information about advertisements displayed to South African users, their sponsors, periods of dissemination, broad targeting criteria and reach. This would provide considerable public benefit without requiring the state to determine the truth or acceptability of the political message itself. It regulates the transparency of influence rather than the content of political opinion.

5.5 Transparency concerning recommender systems

A further important DSA innovation concerns the systems through which platforms determine what information users actually see. Article 27 requires online platforms using recommender systems to explain in plain and intelligible language the main parameters used by those systems and the options available to users to modify or influence them. The explanation must identify the most significant criteria determining what information is recommended and their relative importance. This requirement reflects the recognition that platforms do not merely host information but that recommender systems determine the relative visibility, prominence and reach of information and therefore exercise considerable influence over the information environment.

A South African framework could similarly require significant platforms to provide intelligible explanations of the principal factors determining recommendations and ranking. The objective would not be disclosure of source code or proprietary algorithms but to provide users and regulators with meaningful information about the logic governing what is amplified or prioritised: for example, the role of previous engagement, popularity, recency, location, paid promotion or other significant signals.

37 South Africa's Electoral Commission has already recognised this concern in its Draft Electoral Code of Conduct on Countering Disinformation, which seeks, among other things, to require greater transparency concerning paid influencers and other compensated political communication. The limitation is institutional: the Electoral Commission's regulatory authority under the electoral framework is directed principally at political parties and candidates, rather than at the digital platforms through which such communications are disseminated. The Draft Code therefore seeks to achieve transparency from the political-actor side, including by requiring parties and candidates to disclose relevant paid relationships. A DSA-style approach would provide an important complement by placing transparency obligations directly upon platforms, which are better positioned to identify paid promotion occurring through their services and to provide systematic information concerning its sponsor, payer, targeting and dissemination. The distinction illustrates the additional regulatory leverage that a platform-focused framework could provide. Electoral Commission of South Africa, Draft Electoral Code of Conduct on Countering Disinformation, GN 3995, Government Gazette 54925 (30 June 2026).

This is especially relevant where public debate focuses upon whether platforms amplify inflammatory, deceptive or otherwise harmful material. Without some transparency concerning recommendation systems, it is difficult to distinguish claims about platform amplification from evidence of how those systems are actually designed to operate.

5.6 Transparency should reveal South African capacity as well as South African outcomes

For South Africa, transparency should extend beyond counting moderation decisions. There is a legitimate regulatory interest in knowing whether platforms have sufficient capacity to operate responsibly in the South African environment. This might include appropriate reporting concerning the languages in which complaint and moderation systems operate; the extent to which automated moderation tools have been tested for relevant South African languages; the availability of personnel possessing appropriate linguistic and contextual expertise; and the arrangements in place for responding to significant local events.

Care would be required not to compel disclosure of personal information about individual employees or commercially sensitive operational information unnecessarily. The objective is instead to enable platforms to demonstrate that their South African operations are not simply an afterthought within global trust-and-safety structures. A platform serving millions of South Africans should be capable of demonstrating, rather than merely asserting, that its systems can understand and respond appropriately to the environment in which those users participate.

5.7 Proportionality and differentiated obligations

As elsewhere in the DSA, transparency obligations should be proportionate to the nature, scale and impact of the service. It would make little sense to impose the same reporting burden upon a small South African hosting provider as upon a global social-media platform serving millions of South Africans. Requirements should therefore be graduated, with relatively basic obligations applicable more generally and more extensive reporting requirements reserved for services exceeding appropriate thresholds of South African users, reach or systemic significance.

This is also important to avoid creating regulatory barriers that entrench the very largest platforms by imposing compliance costs that smaller competitors cannot bear. South Africa need not adopt the DSA's numerical thresholds or categories mechanically. The relevant thresholds should reflect the size of the South African market and the particular regulatory objectives being pursued.

5.8 Transparency as regulatory infrastructure

Transparency requirements should ultimately be understood as more than an obligation upon platforms to publish corporate reports. Properly designed, they create the information infrastructure upon which the remainder of platform regulation depends.

Complaint statistics reveal whether notice-and-action mechanisms actually work. Moderation data reveal how platforms exercise their private rule-making powers. Advertising disclosures reveal who pays to influence users. Recommender-system transparency begins to expose the mechanisms through which information acquires visibility and reach. Information about local resources helps determine whether platforms are capable of responding adequately to South African circumstances.

For a country with substantially more limited regulatory capacity than the European Union, this function is especially important. South African authorities cannot realistically observe or independently reconstruct everything occurring within the systems of multinational platforms. Regulation should therefore place the initial burden upon platforms to generate, preserve and disclose standardised information about their own conduct.

At the same time, transparency should not be confused with accountability itself. Vast quantities of published data are of little consequence if they cannot be understood, compared or acted upon. South African requirements should therefore prioritise information that is locally disaggregated, standardised, machine-readable where appropriate and responsive to clearly identified regulatory questions.

One of the most compelling lessons of the DSA is consequently not simply that platforms should disclose more. It is that entities exercising substantial power over the digital information environment should be required to make the exercise of that power sufficiently visible to permit meaningful public scrutiny and effective regulatory oversight.



Section 6

Access to platform data and independent research

Transparency and reporting requirements can significantly improve public understanding of how online platforms operate. They nevertheless have an inherent limitation: the platform largely remains the source of the information disclosed. Even where legislation prescribes what must be reported and how, public transparency cannot provide access to all the data necessary to understand how complex platform systems operate, what risks they generate, or whether measures adopted to address those risks are effective.

The DSA responds to this information asymmetry through a more ambitious mechanism: requiring the largest platforms and search engines to provide access to certain data for regulatory oversight and independent research. Article 40 therefore represents an important complement to the transparency requirements discussed above. It recognises that meaningful scrutiny of platform power may require appropriately qualified outsiders to examine information that platforms would not otherwise make publicly available.

6.1 Regulatory access to platform data

Article 40 first establishes the powers of regulatory access. Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs) must provide the digital services coordinator of the establishment or the European Commission, at their reasoned request and within a reasonable period, with access to data necessary to monitor and assess compliance with the DSA.

The significance of this requirement is straightforward. A regulator cannot effectively supervise obligations concerning matters such as systemic risk, recommender systems, advertising or content moderation if the regulated entity possesses virtually all the relevant information and can determine what information the regulator is permitted to see.

A similar principle would be important in South Africa. Any authority tasked with supervising significant online intermediaries would require statutory powers to obtain information reasonably necessary to determine compliance. This should extend beyond information that platforms are already required to publish. Subject to appropriate safeguards, a regulator should be capable of requiring access to relevant internal data where this is necessary to perform its statutory functions.

This would also reinforce the importance of the jurisdictional and local-representation requirements discussed earlier. A statutory power to demand information has limited value unless it can be effectively exercised against the foreign provider controlling that information.

6.2 Independent researcher access

More innovative is the DSA's recognition of a role for independent researchers in scrutinising the operation and societal effects of the largest platforms. Article 40 creates two distinct routes of researcher access. First, Article 40(12) requires VLOPs and VLOSEs to provide qualifying researchers with access, without undue delay, to data that are publicly accessible through their online interfaces, including real-time data where technically possible. This access must be for research contributing to the detection, identification and understanding of the systemic risks addressed by the DSA.

Second, Article 40(4) creates a mechanism through which vetted researchers can obtain access to non-public platform data. A Digital Services Coordinator may make a reasoned request requiring a VLOP or VLOSE to provide such data where access is necessary for research contributing to the detection, identification and understanding of systemic risks or to the assessment of the adequacy, efficiency and impact of risk-mitigation measures.

This latter mechanism is potentially transformative. Independent research into platforms has historically depended heavily upon information platforms choose to make publicly available, access to public interfaces, or voluntary cooperation from the companies themselves. Article 40 converts access to relevant platform data, in defined circumstances and subject to safeguards, from a matter substantially within corporate discretion into a regulatory obligation.

6.3 Access is qualified, not unrestricted

The DSA does not create a general entitlement for academics or civil-society organisations to demand access to platform data. The safeguards around researcher access are substantial. To qualify as a vetted researcher for access to non-public data, researchers must be affiliated with a recognised research organisation; be independent from commercial interests; disclose the funding of their research; demonstrate that they can comply with data-security and confidentiality requirements and protect personal data; justify that the requested data are necessary and proportionate for the proposed research; and commit

themselves to making their research results publicly available free of charge within a reasonable period.

The research itself must also serve specified purposes connected to the DSA's systemic-risk framework. Article 40 is therefore not a mechanism for obtaining platform data for any research project regarded as interesting or socially useful. The access must contribute to understanding defined systemic risks or assessing measures intended to mitigate them.

Providers may seek amendment of a data-access request where they do not possess the data or where providing access would create significant vulnerabilities for the security of their service or the protection of confidential information, particularly trade secrets. The regulatory authority must then determine whether alternative means of access can satisfy the research purpose. The framework accordingly attempts to reconcile two legitimate concerns: platforms should not be able to invoke confidentiality or commercial sensitivity as a blanket shield against independent scrutiny, but neither should researcher access compromise users' privacy, data security, trade secrets or the security of the service itself.

6.4 The emerging European implementation model

The practical significance of Article 40 has been strengthened by the adoption of Commission Delegated Regulation (EU) 2025/2050, which entered into force on 29 October 2025.³⁸ It establishes detailed procedures and technical conditions governing access to non-public data by vetted researchers, including application procedures, data-access modalities and requirements concerning security and confidentiality.

The European Commission has also established a central DSA data access portal through which researchers can submit applications and communicate with digital services coordinators and data providers. Platforms must make available data catalogues describing the data assets that may potentially be requested. The purpose is not simply to create a legal entitlement in the abstract but to provide practical infrastructure through which qualifying researchers can identify potentially relevant datasets and seek access to them.

These arrangements are still relatively new, and their effectiveness will depend significantly upon implementation. They nevertheless demonstrate the considerable institutional machinery required to translate a statutory right of researcher access into a functioning system.

38 Commission Delegated Regulation (EU) 2025/2050 of 1 July 2025 supplementing Regulation (EU) 2022/2065 by laying down the technical conditions and procedures under which VLOPs and VLOSEs are to share data with vetted researchers.

6.5 Why researcher access may be particularly valuable in South Africa

The principle underlying Article 40 has considerable potential for application in South Africa. South Africa faces an especially pronounced information and capacity asymmetry in its relationship with global technology companies. Platforms possess vast quantities of data concerning how information is produced, distributed, recommended and consumed by South African users. By contrast, South African regulators, universities, journalists and civil-society organisations have only limited visibility into those systems.

At the same time, it would be unrealistic to expect a South African regulator to develop internally all the expertise required to investigate the operation and effects of complex digital platforms. Relevant expertise is distributed across universities, research institutes, civil-society organisations and independent technical researchers. A carefully constructed researcher-access regime could therefore enable South Africa to mobilise expertise that exists outside the state. Qualifying researchers could contribute to understanding matters such as coordinated information manipulation; electoral disinformation; recommender-system amplification; online gender-based violence and harassment; risks to children; the treatment of South African languages by automated systems; the prevalence and distribution of unlawful content; and the effectiveness of platform measures intended to address these risks.

Research of this kind could in turn strengthen regulatory oversight, public policy and democratic debate without requiring the state itself to conduct every investigation.

6.6 A proportionate South African model

South Africa should nevertheless be cautious about attempting to reproduce the institutional architecture set out in Article 40 wholesale. The EU system depends upon digital services coordinators capable of vetting researchers, assessing technically complex data requests, determining appropriate access modalities and resolving disputes concerning confidentiality, security and proportionality. South Africa would need to consider whether an existing regulator could realistically perform these functions and whether the volume of potential applications justified comparable machinery.

A more proportionate model might begin with two principles. First, the competent South African regulator should possess a clear statutory power to obtain data necessary for its own supervisory and enforcement functions. Second, legislation could create a controlled mechanism through which appropriately qualified independent researchers may obtain access to specified platform data for defined public-interest purposes. Such access could be mediated by the regulator rather than creating an unrestricted right enforceable directly against platforms.

The criteria for researcher qualification could draw upon Article 40: demonstrated expertise; institutional affiliation where appropriate; independence from relevant

commercial interests; transparency concerning research funding; adequate data-security and privacy arrangements; a clearly defined research question; proportionality between the data requested and that question; and an undertaking to publish resulting research in the public interest. South Africa might also initially limit enhanced data-access obligations to the largest and most systemically significant platforms, rather than imposing them across the intermediary sector.

6.7 Who should qualify as a researcher?

The South African context may warrant particular attention to the range of institutions eligible to participate. The DSA's provisions are not confined exclusively to conventional university academics. Article 40(12), for example, expressly contemplates qualifying researchers affiliated with not-for-profit bodies, organisations and associations. Any South African adaptation should similarly avoid defining eligible researchers so narrowly that significant expertise within civil society is excluded.

South Africa has a substantial community of organisations engaged in monitoring elections, digital rights, media freedom, misinformation, hate speech, online harms and other aspects of the information environment. Where such organisations possess genuine research expertise, methodological competence and adequate safeguards for handling data, there is no obvious reason in principle why their work should be excluded merely because they are not universities.

The appropriate emphasis should be upon competence, independence, methodological integrity and capacity to protect the data concerned, rather than institutional label alone.

At the same time, access to non-public platform data can itself confer considerable power. A vetting system would therefore need to guard against data access becoming a means of advancing commercial, partisan or other private interests under the guise of public-interest research.

6.8 Data protection, privacy and security

Any South African researcher-access regime would have to be constructed consistent with the Protection of Personal Information Act³⁹ and other applicable laws. Researcher access should not be understood as creating an entitlement to unrestricted personal information about platform users. Data minimisation, anonymisation or pseudonymisation, secure processing environments and restrictions upon onward disclosure may all be necessary depending upon the nature of the data and research.

39 Protection of Personal Information Act 4 of 2013 (POPIA), particularly ss 8–19 and, where special personal information is implicated, ss 26–33.

The DSA model is useful precisely because it treats the mode of access as capable of adaptation to the sensitivity of the information concerned. In some instances, data might be supplied directly; in others, access might occur through a secure processing environment or other controlled mechanism. This suggests that the regulatory question should not always be framed simply as whether data must be disclosed. It may instead be: what form of access permits legitimate independent scrutiny while adequately protecting users and genuinely confidential information?

6.9 From transparency to independent scrutiny

Researcher access represents an important progression from the transparency obligations discussed in the preceding section. Transparency requires platforms to explain aspects of their conduct and publish prescribed information about their operations. Researcher access addresses the limits of that model by recognising that meaningful accountability cannot depend entirely upon platforms deciding what outsiders are permitted to know about them.

This distinction is particularly important where the regulatory concern is systemic rather than individual. Aggregate transparency statistics may reveal how many items of content a platform removed, but they may not reveal how recommender systems contributed to the spread of particular material, whether coordinated networks manipulated visibility, whether moderation systems perform differently across South African languages, or whether measures introduced by a platform actually reduced an identified risk.

For South Africa, Article 40 therefore offers something potentially more valuable than another reporting requirement. It offers a model through which the profound informational advantage enjoyed by global platforms can be partially redistributed to regulators and independent researchers capable of subjecting their systems to evidence-based scrutiny.

Given South Africa's limited regulatory resources, that possibility warrants serious consideration. The objective should not be to reproduce the EU's elaborate data-access machinery, but to establish the principle that platforms operating at sufficient scale in South Africa cannot retain exclusive control over information necessary to understand the significant public risks generated by their services. Properly safeguarded independent access to such information could substantially enlarge South Africa's capacity for platform oversight without requiring that all the necessary expertise and analytical capacity reside within the state itself.



Section 7

Systemic risk assessment adapted to South African circumstances

The DSA's systemic-risk regime represents a significant departure from regulatory approaches focused primarily upon individual items of unlawful content. Articles 34 and 35 require Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs) periodically to identify, analyse and assess systemic risks arising from the design, functioning and use of their services and related systems, and to adopt reasonable, proportionate and effective measures to mitigate the risks identified.

The importance of this approach lies partly in where it locates regulatory attention. Many of the most consequential effects of digital platforms cannot adequately be understood by examining individual posts, accounts or moderation decisions in isolation. They may arise instead from the interaction between enormous numbers of users and the systems through which platforms rank, recommend, amplify, monetise and moderate information. The DSA therefore asks not simply whether particular content is unlawful, but whether the design and operation of a service creates or intensifies risks at a societal scale.

7.1 The DSA's categories of systemic risk

Article 34 identifies four broad categories of systemic risk that the largest platforms must assess. The first concerns the dissemination of illegal content through their services. The second concerns actual or foreseeable negative effects upon the exercise of fundamental rights, including freedom of expression and information, privacy, non-discrimination, children's rights and consumer protection. The third category concerns actual or foreseeable negative effects upon civic discourse and electoral processes, and public security. The fourth encompasses serious negative consequences for physical and mental well-being, including gender-based violence and the protection of public health and minors.

The assessment is expressly directed not merely at the presence of problematic content but at the contribution made by the platform's own systems. Providers must

consider, among other things, the design of recommender and advertising systems, content-moderation systems, terms and conditions and their enforcement, data-related practices and possibilities for intentional manipulation of the service.

This is an important conceptual shift. It recognises that a platform may contribute to harm without itself creating the underlying content. Its architecture may determine which material receives prominence, how rapidly it spreads, which audiences encounter it, whether those responsible can monetise it and whether coordinated actors can manipulate the system to manufacture apparent popularity or consensus.

7.2 Why a systemic-risk approach has particular relevance to South Africa

There is considerable potential value in adapting this approach to South Africa.

Existing legal frameworks are generally better equipped to respond to identifiable unlawful acts than to examine how the architecture of a digital service may contribute to broader societal harms. Yet many contemporary concerns about platforms arise precisely at this systemic level: coordinated information manipulation; amplification of inflammatory or deceptive material; online mobilisation directed at vulnerable communities; gendered harassment; risks to children; interference with electoral discourse; and the possibility that engagement-driven systems disproportionately reward divisive or sensational content.

A systemic-risk framework provides a way of asking different questions. Rather than requiring the state to determine the truth of millions of individual statements, it can ask whether a large platform has adequately investigated how its systems contribute to the production, amplification or acceleration of defined risks, and what reasonable measures it has adopted in response.

This distinction is especially important in relation to disinformation. Broad legal prohibitions on false information present obvious difficulties for freedom of expression, including questions about who determines truth and how rapidly contested claims can legitimately be adjudicated. A systemic-risk approach does not eliminate those difficulties, but it provides an alternative regulatory focus. The question becomes whether platforms understand and appropriately respond to the manipulation of their services and the role their own design, recommendation and monetisation systems play in amplifying demonstrably harmful patterns of conduct.

7.3 South African risks should be defined in South African terms

South Africa should not simply reproduce the DSA's systemic risk categories verbatim. Any statutory risk framework should identify a relatively narrow set

of concerns grounded in South African constitutional values and circumstances. These could include significant risks to:

- ▶ the dissemination of content unlawful under South African law;
- ▶ constitutional rights, including freedom of expression, equality, dignity and privacy;
- ▶ electoral integrity and democratic participation;
- ▶ media freedom and pluralism;
- ▶ public safety and the risk of violence;
- ▶ children and other particularly vulnerable users;
- ▶ coordinated manipulation of the information environment; and
- ▶ serious forms of online gender-based violence and harassment.

Careful statutory definitions would be essential. “Systemic risk” should not become an open-ended justification for governmental intervention in lawful expression merely because that expression is unpopular, disturbing or politically inconvenient. The South African Constitution should operate as a central constraint upon both the identification of risks and the measures adopted to mitigate them. A platform should not be encouraged to suppress lawful political speech simply because doing so represents the easiest means of reducing a broadly formulated societal risk.

7.4 Assessment should include the platform’s own contribution to risk

A South African framework should preserve one of the DSA’s most valuable features: the requirement that providers examine how their own systems and design choices contribute to identified risks. Relevant factors could include recommender systems, ranking and amplification, advertising systems, monetisation incentives, virality features, account-creation processes, moderation systems, automated tools and the effectiveness with which terms and conditions are enforced.

Platforms should also assess whether their services are susceptible to coordinated or inauthentic manipulation, including the use of networks of accounts to manufacture visibility or apparent public support. This shifts some of the inquiry away from the familiar question, “What objectionable content did users post?” towards the equally important question, “What characteristics of the service affected the reach, impact and incentives associated with that content?”

7.5 Risk mitigation

Risk assessment without corresponding obligations would have limited value. Article 35 of the DSA therefore requires VLOPs and VLOSEs to adopt reasonable,

proportionate and effective mitigation measures tailored to the risks they have identified. Potential measures under the DSA include adapting the design or functioning of the service; changing recommender systems; modifying terms and conditions or their enforcement; adjusting advertising systems; strengthening content-moderation processes; cooperating with trusted flaggers; and improving internal resources and testing.

A South African analogue should direct a similarly risk-specific response. What action is appropriate to an identified risk will depend upon its nature and the service concerned. The legal obligation should principally be one of reasonable and proportionate risk mitigation, accompanied by an obligation to demonstrate why the measures selected are appropriate and how their effectiveness is assessed.

The focus on the effect upon fundamental rights must remain central. Measures designed to mitigate one risk may create another. Aggressive automated moderation, for example, may reduce the circulation of some harmful material, while disproportionately suppressing lawful expression in languages for which automated systems perform poorly. Effective risk regulation therefore requires consideration not merely of whether intervention reduces a specified harm, but of its consequences for the wider rights environment.

7.6 Proportionality is especially important for South Africa

The DSA reserves its most demanding systemic-risk obligations for VLOPs and VLOSEs. South Africa should follow the principle of differentiated obligations without necessarily adopting the EU's numerical thresholds. Systemic risk assessment is resource-intensive. Requiring every hosting service or small online platform to undertake elaborate assessments would be disproportionate and could create barriers to entry that favour incumbent global platforms.

The enhanced obligations should therefore apply only to services of sufficient scale or significance within South Africa. Relevant criteria might include the number of South African users, reach, market position, the nature of the service and its capacity to affect the South African information environment significantly. The objective should be to concentrate the strongest obligations upon those providers whose scale and architecture make genuinely systemic consequences possible.

7.7 Independent scrutiny

Platforms should not be the sole judges of whether their risk assessments are adequate. This is where the systemic-risk regime connects directly with the transparency and researcher-access provisions discussed above. Risk assessments should be available to the competent regulator, with appropriately redacted versions or summaries made public. Independent researchers should, subject to appropriate safeguards, be able to examine relevant platform data and test claims about the existence of risks and the effectiveness of mitigation measures.

A further feature of the DSA's systemic accountability framework is the requirement that Very Large Online Platforms and Very Large Online Search Engines undergo independent audits at least annually. Under Article 37, those audits assess compliance with specified obligations under the DSA, as well as relevant commitments undertaken through codes of conduct and crisis protocols. The auditing organisation must satisfy requirements concerning independence, absence of conflicts of interest, expertise and professional ethics, and its report must provide an audit opinion and, where appropriate, operational recommendations.⁴⁰

The significance of an independent audit extends beyond the particular European procedures adopted to implement it. Much of platform regulation necessarily depends upon information and assessments generated by the regulated provider itself. Platforms undertake their own systemic-risk assessments, select mitigation measures and report upon their operation. If external scrutiny ends there, a regulatory framework risks allowing platforms effectively to mark their own homework.

Independent audits can provide an additional layer of verification. It can test whether prescribed processes have been followed, whether the evidence relied upon supports the provider's conclusions and, in appropriate cases, whether measures said to mitigate identified risks appear reasonably capable of doing so. Audits should not replace regulatory oversight or independent research. Rather, it can provide the supervising authority with an additional source of technically informed evidence with which to assess compliance.

South Africa should nevertheless approach the transplantation of the European audit model cautiously. The DSA's audit regime is technically elaborate and potentially costly, and effective auditing of platform systems may require expertise spanning law, data science, algorithmic systems, human rights and content moderation. A requirement of this kind should therefore be confined to the largest or most systemically significant services and should not necessarily reproduce the European model in full.

The more important principle is one of external verification. Where a platform is required to assess systemic risks and adopt measures to mitigate them, legislation should allow for appropriately qualified and genuinely independent persons or organisations to test compliance with those obligations against standards prescribed by the supervising authority. Platforms should not be permitted to identify the risks, determine the adequacy of their response and then remain the sole judges of whether that response has worked.

7.8 Crisis preparedness and response

Systemic risks may change dramatically during periods of crisis. A platform whose ordinary operation presents manageable risks may assume a very different

40 DSA art 37; Commission Delegated Regulation (EU) 2024/436 of 20 October 2023 supplementing Regulation (EU) 2022/2065 by laying down rules on the performance of audits for VLOPs and VLOSEs.

significance during armed conflict, widespread civil unrest, a public-health emergency or another event in which rapidly circulating information, coordinated manipulation or online incitement can contribute to grave offline consequences. The DSA accordingly supplements its ordinary systemic-risk framework with a specific crisis-response mechanism.

Article 36 provides that, where extraordinary circumstances give rise to a serious threat to public security or public health in the European Union or significant parts of it, the European Commission, acting upon a recommendation of the European Board for Digital Services, may require one or more Very Large Online Platforms or Very Large Online Search Engines to assess whether and how the functioning or use of their services significantly contributes, or is likely to contribute, to that threat. Where such a contribution is identified, providers may be required to adopt specific, effective and proportionate measures to prevent, eliminate or limit it and to report on the measures taken and their impact.

Importantly, the mechanism does not give the Commission unconstrained power to direct platforms during periods labelled as crises. The DSA requires the measures imposed to be strictly necessary, justified and proportionate, taking account of the gravity and urgency of the threat and of their implications for the rights and legitimate interests of those affected. The potential impact of measures upon fundamental rights must expressly be considered. The mechanism is therefore exceptional not only in the circumstances that trigger it, but in the safeguards that accompany its use.

There is a strong case for South Africa to incorporate at least the principle of heightened platform preparedness and accountability during periods of genuine crisis. Recent experience demonstrates how rapidly online communications can assume heightened importance during periods of civil unrest, xenophobic mobilisation, threats to public safety or other fast-moving events. In such circumstances, the ordinary tempo of platform risk assessment and regulatory engagement may be inadequate. The largest and most systemically significant platforms could reasonably be required to maintain crisis protocols, effective escalation channels and sufficient surge capacity to reassess emerging risks and respond rapidly where the operation of their services is contributing to serious harm.

South Africa's constitutional framework also recognises that exceptional circumstances may justify exceptional legal responses. Section 37 of the Constitution permits a state of emergency to be declared where the life of the nation is threatened by war, invasion, general insurrection, disorder, natural disaster or another public emergency and the declaration is necessary to restore peace and order. It even permits legislation enacted in consequence of such a declaration to derogate from certain rights in the Bill of Rights. But precisely because exceptional powers carry exceptional risks, the Constitution subjects them to stringent conditions: derogation must be strictly required by the emergency; certain rights remain non-derogable; declarations are time-limited and their extension subject

to parliamentary control; and the declaration, its extension and measures taken pursuant to it remain subject to judicial scrutiny.

That constitutional structure offers a useful principle for thinking about platform regulation in times of crisis. The existence of extraordinary circumstances should not be treated as a reason for abandoning ordinary constitutional commitments, but nor does the Constitution require government to regulate as though extraordinary circumstances make no difference. Rather, exceptional powers must correspond to an exceptional and objectively identifiable threat, be demonstrably necessary to address it, remain proportionate to that threat, operate for no longer than required and be subject to effective oversight.

This does not necessarily mean that a South African equivalent of the DSA's article 36 should be triggered only by the formal declaration of a state of emergency under the Constitution's section 37. That threshold is exceptionally high, and circumstances capable of producing acute systemic risks through digital services may arise well before the life of the nation is threatened. There is an important distinction between a constitutional state of emergency, permitting derogation from rights, and a statutory mechanism requiring systemically significant platforms temporarily to intensify risk assessment, mitigation and reporting. The latter could operate within the ordinary constitutional order and would remain subject to the limitation analysis required by section 36 of the Constitution.

Any South African crisis-response mechanism should nevertheless be carefully circumscribed. Legislation should define the circumstances capable of triggering enhanced obligations; distinguish genuine threats to public safety or security from ordinary political controversy or dissent; limit enhanced requirements to what is necessary and proportionate; provide for their termination once the triggering circumstances have passed; and ensure appropriate transparency and review.

Particular caution would be required before empowering an administrative authority to direct platforms to remove or suppress particular categories of otherwise lawful expression. A crisis mechanism should principally be concerned with requiring platforms to assess and mitigate exceptional systemic risks, rather than providing the government with an emergency power to control online discourse.

The broader lesson from Article 36 is therefore valuable even if its institutional machinery is not reproduced. Platform regulation should be capable of responding to the fact that systemic risks are not static. A framework designed only for ordinary circumstances may prove least effective at precisely the moment when digital platforms have their greatest capacity to magnify serious offline harm. South Africa should consider requiring its largest digital services both to prepare in advance for foreseeable crises and, subject to carefully defined safeguards, to assume temporarily enhanced risk-assessment, mitigation and reporting obligations when exceptional circumstances arise.



Section 8

Protection of minors online

One of the most important innovations of the DSA is that it treats children not merely as a vulnerable class requiring protection from unlawful content, but as rights-bearing participants in the digital environment whose privacy, safety and security must actively be protected through the design and operation of online platforms. Consistent with the DSA's broader approach, the emphasis is not principally upon creating new categories of unlawful content, but upon imposing obligations concerning the way in which digital services are designed and operated. Article 28 applies that approach specifically to minors, requiring online platforms accessible to them to put in place appropriate and proportionate measures to ensure a high level of privacy, safety and security. The objective is not simply to protect children from particular forms of harmful content, but to ensure that they can participate in the digital environment while continuing to access the educational, social, creative and communicative opportunities it provides.

8.1 The DSA's approach: privacy, safety and security by design

Article 28 requires providers of online platforms accessible to minors to put in place appropriate and proportionate measures to ensure a high level of privacy, safety and security for minors using their services. The obligation is deliberately framed as a general duty rather than as an exhaustive list of prescribed interventions. Different platforms present different risks: a social networking service, a video-sharing platform, an online gaming environment and a messaging service expose children to materially different forms of harm. The DSA therefore adopts a risk-based approach in which the measures expected of a provider depend upon the nature, scale and characteristics of the service and the reasonably foreseeable risks to children arising from its operation.

This is significant because it rejects a one-size-fits-all conception of child protection. The legal obligation is not simply to remove prohibited material after it has been reported. Rather, providers must consider whether the design and functioning of their services expose children to avoidable risks and must adopt measures proportionate to those risks. In 2025, the European Commission issued detailed

guidelines under Article 28 explaining how this obligation may be fulfilled. Although those guidelines are non-binding, they provide an authoritative indication of the kinds of measures the Commission considers capable of ensuring a high level of protection for minors and of how regulators may assess compliance with Article 28.⁴¹

The resulting approach is best understood as one of ‘safety by design’. Child protection is not treated as an afterthought addressed only through complaints and takedowns, but as a consideration capable of influencing privacy settings, recommender systems, account design, reporting tools, commercial practices and the governance arrangements through which platforms identify and mitigate risks to younger users.

8.2 Children’s rights and the positive obligations of platforms

An important feature of the European approach is that the protection of minors is framed positively rather than paternalistically. The Commission’s guidelines begin from the recognition that children derive substantial benefits from participation in online life. Digital platforms provide access to educational resources, opportunities for creativity, communication with peers, participation in public discussion and the development of skills and confidence. Measures intended to protect children must therefore avoid unnecessarily restricting their access to information, participation and freedom of expression.

This balance resonates strongly with South Africa’s constitutional framework. Section 28 affords every child special protection, including the right to be protected from maltreatment, neglect, abuse and degradation, while providing that a child’s best interests are of paramount importance in every matter concerning that child. At the same time, children are independently entitled to the rights contained in the Bill of Rights, including dignity, privacy, equality, freedom of expression and access to information, subject to the ordinary limitations clause. The constitutional objective is therefore not to exclude children from digital spaces, but to ensure that those spaces are structured in a manner consistent with their rights and developmental interests.

A South African framework inspired by Article 28 should accordingly resist the temptation to treat child protection as synonymous with censorship. The more valuable lesson from the DSA is that platforms themselves bear the responsibility to anticipate and reduce foreseeable risks through the way in which their services are designed and operated.

41 European Commission, ‘Guidelines on measures to ensure a high level of privacy, safety and security for minors online’, pursuant to Article 28(4) of Regulation (EU) 2022/2065, C/2025/5519, OJ C, 10 October 2025.

8.3 Privacy, account design and unwanted contact

One of the clearest themes emerging from the Commission's guidelines concerns the importance of privacy-enhancing default settings for minors. Children should not be required to navigate complex privacy menus in order to avoid unwanted exposure. Rather, platforms should consider making minors' accounts private by default, limiting the discoverability of younger users and reducing opportunities for unsolicited contact from strangers. The underlying principle is that platform architecture should reduce foreseeable opportunities for grooming, harassment and exploitation before harm occurs rather than relying exclusively upon children to report abuse afterwards.

The guidelines similarly recommend that platforms accessible to minors provide effective reporting and blocking tools that are understandable and readily available to children, together with appropriate support mechanisms for users who encounter harmful behaviour. This emphasis upon usability is important. A reporting mechanism that technically exists but is incomprehensible to younger users provides considerably less protection than one designed around the actual capacities and experiences of children and adolescents.

For South Africa, these principles appear readily adaptable. They do not require the creation of entirely new substantive offences or categories of unlawful content. Instead, legislation could impose differentiated design obligations upon platforms likely to be accessed by minors, requiring them to demonstrate that account settings, privacy controls, reporting mechanisms and escalation processes provide an appropriately high level of protection for younger users. Such obligations would complement, rather than replace, the existing protections contained in South African child-protection, criminal and privacy legislation.

8.4 Recommender systems, addictive design and harmful exposure

Perhaps the most innovative aspect of the Commission's 2025 guidelines is its attention to the ways in which platform design itself may intensify risks to children. Rather than focusing solely upon harmful content, the guidelines address features capable of encouraging compulsive engagement or repeatedly directing minors towards increasingly harmful material. These include recommender systems, engagement-driven design and other interface features that may stimulate problematic or addictive patterns of use.

The recommendations do not prescribe a single technological solution. Instead, they encourage providers to assess how their recommender systems affect younger users and to provide greater user control. Among the measures contemplated are allowing minors to reset recommended feeds, reducing reliance upon profiling when appropriate, and encouraging exploration of content beyond highly repetitive recommendation loops. The broader principle is that systems optimised exclusively for engagement should also be evaluated according to their effect upon children's privacy, safety and healthy participation online.

This aspect of the DSA deserves particular attention in South Africa because it illustrates the difference between regulating speech and regulating incentives. A recommendation algorithm may amplify perfectly lawful material in ways that nevertheless produce harmful developmental or psychological consequences for younger users. Addressing that problem does not necessarily require determining that the underlying expression is unlawful. It may instead require greater transparency, meaningful user choice and design measures intended to reduce foreseeable risks created by the platform's own systems.

8.5 Age assurance without unnecessary surveillance

Age assurance presents one of the most difficult policy questions in contemporary digital regulation. There is an obvious public interest in preventing children from accessing pornography, gambling and other services intended exclusively for adults. Equally, indiscriminate age-verification systems can create substantial privacy risks by requiring users to disclose identity documents or other highly sensitive personal information merely to participate online.

The DSA itself adopts an important balancing principle. Article 28 expressly provides that compliance with the obligation to protect minors does not require platforms to process additional personal data merely in order to determine whether a recipient is a minor. This reflects the recognition that child protection should not become a justification for unnecessary surveillance or indiscriminate collection of personal information.

The Commission's subsequent guidelines distinguish between different forms of age assurance and recommend that measures should be accurate, reliable, robust, proportionate and privacy-preserving. In higher-risk contexts, such as pornography, gambling or services legally restricted to adults, the guidelines contemplate stronger age-verification measures where less intrusive alternatives would not adequately protect minors. At the same time, the Commission has developed a privacy-preserving age-assurance prototype intended to demonstrate that proving eligibility need not require revealing a person's complete identity or unnecessary personal information.⁴²

For South Africa, the transferable principle is not that a particular European age-verification technology should be adopted. Rather, any future legislative framework should recognise both sides of the equation: children require effective protection from genuinely age-inappropriate services and content, while age-assurance mechanisms should themselves comply with constitutional and privacy principles. The objective should be to establish minimum standards for reliability, proportionality and data minimisation rather than mandate unnecessarily intrusive identity verification across the internet.

42 European Commission, 'The EU approach to age verification' <https://digital-strategy.ec.europa.eu/en/policies/eu-age-verification> accessed 15 September 2026.

8.6 Commercial practices and profiling-based advertising

Article 28 also addresses the commercial exploitation of children's personal data. Online platforms may not present advertisements based upon profiling where they are aware with reasonable certainty that the recipient is a minor. This prohibition recognises that children occupy a particularly vulnerable position within behavioural advertising ecosystems and that personalised commercial persuasion based upon extensive profiling raises concerns extending beyond ordinary consumer protection.

This provision is noteworthy because it regulates the platform's business model rather than the content of speech. The concern is not that advertising is inherently unlawful, but that behavioural advertising directed at children may exploit information asymmetries and developmental vulnerability in ways that justify heightened protection. It therefore complements the DSA's wider emphasis upon transparency and accountability in digital advertising discussed earlier in this report.

South Africa should consider whether existing privacy and consumer-protection laws adequately address profiling-based advertising directed at children or whether a future digital-services framework should introduce additional protections for minors. Any such intervention would require careful alignment with POPIA and the constitutional protection of privacy, but Article 28 offers a valuable example of a targeted obligation directed specifically at platform practices rather than broad restrictions upon commercial communication.

8.7 Governance, reporting and accountability for child safety

The Commission's guidelines make clear that protection of minors is not achieved through isolated product features alone. Effective child safety requires governance structures within platforms themselves. Providers are encouraged to undertake periodic risk reviews, involve appropriate expertise, evaluate the effectiveness of protective measures, provide accessible support and reporting mechanisms, and integrate consideration of children's rights into relevant organisational decision-making.

This governance dimension aligns closely with the broader architecture proposed elsewhere in this report. Requirements concerning transparency, researcher access, systemic-risk assessment and independent audit all become particularly valuable where they enable meaningful scrutiny of whether platforms are genuinely protecting children rather than merely asserting that they do so. For the largest services, child safety should therefore form an identifiable component of systemic-risk assessment and of the information made available to the supervising authority and, where appropriate, independent auditors or researchers.

There is also considerable value in requiring platforms to demonstrate that they possess adequate local knowledge and capacity when addressing risks to children. The experience of a South African child encountering grooming, cyberbullying or harmful content may differ materially according to language, socio-economic

circumstances and patterns of platform use. Protective systems developed exclusively around European or North American users cannot simply be assumed to function effectively in South African conditions. This reinforces the importance of the linguistic and contextual obligations proposed in the following section.

8.8 Applicability to South Africa

South Africa already possesses an extensive body of law directed towards the protection of children. The Constitution affords children special protection and makes their best interests paramount; the Children's Act provides the principal legislative framework for the care and protection of children; privacy legislation regulates the processing of personal information;⁴³ and criminal and other legislation addresses exploitation and child sexual abuse material. The difficulty is not the absence of substantive child-protection norms, but the fact that many contemporary online harms arise through the design and operation of global digital platforms whose systems are regulated only indirectly by existing South African law.

The DSA suggests several principles that South Africa could realistically adapt. First, platforms accessible to minors could be placed under a general statutory duty to adopt appropriate and proportionate measures ensuring a high level of privacy, safety and security for children. Secondly, that duty could be elaborated through regulatory standards dealing with privacy-by-default settings, effective reporting and blocking tools, protection against unwanted adult contact, age-appropriate recommender-system design and other foreseeable risks associated with particular categories of services. Thirdly, profiling-based advertising directed at minors warrants serious consideration as an area in which enhanced statutory protection may be justified. Finally, the largest and most systemically significant platforms should be required to demonstrate through risk assessments, transparency reporting and external scrutiny that the measures adopted to protect children are effective in practice rather than merely aspirational.

Careful adaptation is nevertheless essential. South Africa should avoid importing detailed European implementation rules without regard to domestic constitutional requirements, institutional capacity or technological realities. In particular, age assurance must be approached with sensitivity to privacy, equality and accessibility; obligations imposed upon platforms should be proportionate to the risks presented by particular services; and measures intended to protect children should preserve, rather than unnecessarily curtail, children's own rights to expression, participation and access to information. The objective is not to create a more restrictive internet for young people, but a digital environment in which children can participate with substantially greater safety and dignity.

43 Children's Act 38 of 2005; Protection of Personal Information Act 4 of 2013.



Section 9

Local linguistic, cultural and contextual capacity

The requirement that systemic risks be assessed in their actual social context has particular significance in South Africa. Article 34 expressly requires VLOPs and VLOSEs, when undertaking systemic-risk assessments, to take account of specific regional or linguistic aspects, including where these are specific to an individual Member State. This reflects an important reality: systems developed and operated globally may perform very differently across countries, languages and social contexts.

For South Africa, this should be elevated from a subsidiary consideration within systemic-risk assessment into a significant regulatory concern in its own right.

9.1 A multilingual information environment

South Africa has twelve official languages and an exceptionally diverse linguistic and cultural environment. Platforms operating at scale cannot credibly claim to moderate content, identify risks or provide effective complaint mechanisms if their systems lack adequate capacity to understand the languages in which South Africans actually communicate.

The issue extends beyond literal translation. Meaning frequently depends upon idiom, slang, code-switching, historical references, political context, humour and culturally specific forms of insult or threat. The same words may carry markedly different significance depending upon who uses them, towards whom, and in what circumstances. Automated systems trained predominantly upon English-language or Global North datasets may therefore perform unevenly when applied to South African languages and contexts.

9.2 Capacity should be demonstrable

South African regulation should not attempt to dictate precisely how many moderators a platform must employ or where they must be located. It could instead require the largest providers to demonstrate that their linguistic and contextual capacity is proportionate to the nature and scale of their South African operations and the risks associated with their services.

Relevant information might include the languages supported by complaint and appeal mechanisms; languages in which human moderation is available; testing of automated systems across South African languages; arrangements for escalation where local context is required; and the availability of personnel or external expertise capable of understanding significant South African events and social dynamics. This connects directly with the transparency obligations proposed in section 5.6 of this report. Assertions of adequate local capacity should, where appropriate, be supported by reportable information rather than simply accepted at face value.

9.3 Local capacity as a freedom-of-expression issue

Linguistic competence should not be understood only as a mechanism for identifying more harmful content. It is equally important to protect lawful expression from erroneous restrictions. A moderation system that cannot understand isiZulu, isiXhosa or another South African language may fail to identify genuine threats or unlawful content. But it may also misunderstand satire, political rhetoric, reclaimed language, journalism or ordinary conversation and remove lawful material incorrectly.

The requirement for local competence therefore serves both safety and freedom of expression. It improves the ability of platforms to identify genuinely serious risks while reducing the likelihood that crude or context-insensitive moderation will suppress legitimate speech.

9.4 Capacity during elections, crises and rapidly developing events

Platforms should also be capable of increasing local capacity during periods in which risks predictably intensify. Elections provide an obvious example. So too may outbreaks of xenophobic mobilisation, civil unrest or other events producing rapidly changing online activity with potentially grave offline consequences.

Large platforms could be expected to demonstrate appropriate preparedness for foreseeable high-risk periods, including effective escalation mechanisms and access to relevant local expertise. This should not mean that the government directs platforms' treatment of political speech. Rather, the platform should be required to demonstrate that its systems are capable of understanding and responding responsibly to the environment in which they operate.

For South Africa, the broader principle is straightforward: a platform should not be able to claim responsible stewardship of a South African information environment that it lacks the linguistic and contextual capacity to understand adequately.



Section 10

Funding effective oversight: A proportionate supervisory fee

A regulatory framework is only as effective as the institutions capable of implementing it. This presents a particularly acute challenge for South Africa. The obligations considered above would require expertise in law, technology, data science, platform governance, content moderation, advertising systems and algorithmic systems. Regulators would need to analyse transparency reports, assess data requests, examine risk assessments, investigate complaints and, where necessary, enforce compliance against some of the world's largest and best-resourced companies.

It would be unrealistic to assume that these responsibilities can simply be added to the mandates of existing South African institutions without corresponding resources.

10.1 The DSA model of funding

Article 43 of the DSA requires the European Commission to charge VLOPs and VLOSEs an annual supervisory fee designed to cover the estimated costs incurred in exercising its supervisory functions under the Regulation. The Commission's implementation of the system demonstrates that the fee operates as assigned revenue contributing directly to the costs of DSA supervision. The principle is significant: the costs associated with specialised supervision of the largest digital services need not be borne entirely by the general taxpayer.

10.2 The case for a South African supervisory levy

There is a strong argument for adopting some version of this principle in South Africa. Global platforms operating at enormous scale possess resources vastly exceeding those available to many South African regulatory institutions. A regulatory framework that imposes complex obligations without equipping the supervising

authority with the technical and professional capacity to test compliance risks becoming largely performative.

A proportionate levy upon the largest regulated services could contribute to funding the specialised capacity required for effective oversight. The purpose should not be punitive. The fee should be understood as a means of financing the additional regulatory costs generated by the supervision of complex and systemically significant services.

10.3 Design safeguards

Any South African levy would require careful design. It should apply only to providers of sufficient scale or significance and should be calculated according to transparent criteria. Consideration could be given to relevant South African turnover, user numbers, regulatory costs or some combination of factors.

There should also be a clear relationship between the revenue raised and the costs of digital-platform supervision. Ideally, funds should be ring-fenced or otherwise reliably appropriated for the relevant regulatory functions, rather than disappearing into general revenue while the supervising institution remains under-resourced.

Transparency should operate in both directions. Just as platforms should account for their conduct, the regulator should publish information concerning the fees collected, the regulatory functions funded and the expenditure incurred. Care would also be required to avoid a fee structure that imposes disproportionate costs upon smaller providers or creates barriers to entry. Enhanced supervisory fees should principally attach to those providers whose scale and complexity generate the greatest supervisory demands.

10.4 Regulatory independence

Funding design must also protect institutional independence. A regulator should not become financially dependent upon discretionary payments from particular platforms or develop incentives to maximise fee revenue. The levy should therefore be determined according to law and objective criteria, rather than negotiated between regulator and provider. Platforms should have appropriate mechanisms to challenge incorrect assessments, but should not be capable of withholding funding because they object to the regulator's conduct.

The broader lesson from the DSA is compelling for South Africa: serious regulation requires serious regulatory capacity, and it is legitimate to require the largest regulated entities to contribute proportionately to the exceptional costs of supervising them.



Section 11

Institutional architecture and co-regulation

The preceding sections have identified a range of mechanisms drawn from the DSA that could usefully inform the development of a South African framework for the regulation of online intermediaries. Those recommendations do not depend upon responsibility for their implementation being assigned to any particular institution. Questions concerning the substance of the regulatory framework and its institutional location are distinct, and the value of mechanisms such as transparency obligations, accessible complaint procedures, regulatory access to data and systemic-risk assessment does not turn upon which South African authority is ultimately entrusted with their supervision.

Any new framework will nevertheless require an institutional architecture capable of giving those obligations practical effect. South Africa must therefore consider whether responsibility should be entrusted to an existing regulator, distributed among several institutions, or assigned to a newly established body. In our view, there are strong reasons to build upon existing institutions rather than create an entirely new regulator and, within such a model, for the Independent Communications Authority of South Africa (ICASA) to assume principal responsibility for supervising compliance with the obligations imposed upon digital intermediaries.

The nature of this proposed role requires emphasis at the outset. ICASA would regulate platform compliance with the regulatory framework. It would not ordinarily determine whether individual pieces of content are unlawful, nor should it become a general appellate tribunal for individual content-moderation decisions. The mechanisms discussed in section 4 would continue to provide the principal pathway for such disputes: initial determination by the platform; internal platform review; access to an independent out-of-court dispute-resolution mechanism; and, where appropriate, recourse to the courts or another institution possessing jurisdiction under the substantive law concerned.

ICASA's function would be to ensure that platforms maintain and properly operate the systems required by law, comply with transparency, data-access,

risk-assessment and other statutory obligations, and remedy systemic or repeated regulatory failures. This distinction between content adjudication and platform supervision is central to the institutional model proposed here.

11.1 Building upon existing South African institutions

Digital-platform regulation intersects with numerous existing fields of South African law. Depending upon the issue concerned, questions may arise involving electronic communications, data protection, equality and discrimination, consumer protection, competition, elections, criminal law, children's rights and freedom of expression. Attempting to consolidate substantive responsibility for all of these areas within a new digital regulator would create considerable duplication. It would also require South Africa to recreate expertise already present within established public institutions.

A more realistic approach is therefore a networked regulatory model. Existing institutions should retain responsibilities falling within their statutory mandates, while a principal authority assumes responsibility for supervising the cross-cutting obligations imposed upon digital platforms themselves. This requires an important distinction between the function of regulating platforms and regulating every field of law implicated by content carried on those platforms.

Whether a platform has established an accessible notice-and-action mechanism, operated an effective internal complaint system, supplied required reasons, complied with transparency requirements, provided prescribed access to data, undertaken an adequate systemic-risk assessment or demonstrated appropriate South African linguistic and contextual capacity is principally a question of platform regulatory compliance.

Whether a particular statement constitutes prohibited hate speech, processing violates POPIA, conduct contravenes electoral law, material amounts to prohibited content under another statute, or behaviour constitutes a criminal offence is a different question. Those issues arise under substantive South African law and should ordinarily continue to be determined through the institutions and processes to which that law assigns responsibility. The regulatory framework should connect these functions where necessary, but it should not collapse them into one institution.

11.2 ICASA as the principal supervisory authority

We propose that ICASA serve as the principal supervisory and coordinating authority under a South African digital-platform framework. ICASA has several institutional advantages. It is an established independent communications regulator with substantial experience regulating communications services and broadcasting, developing and enforcing regulatory codes, monitoring compliance and supervising entities whose services constitute important communications infrastructure.

Its independence is particularly significant. Section 192 of the Constitution requires national legislation to establish an independent authority to regulate broadcasting in the public interest and to ensure fairness and a diversity of views broadly representing South African society. The ICASA Act in turn establishes the Authority as independent, subject only to the Constitution and the law, and requires it to act impartially and without fear or favour.⁴⁴

It must be appreciated that section 192 relates specifically to broadcasting and does not itself give ICASA constitutional authority to regulate online platforms. A digital-platform mandate would have to be conferred by Parliament. But ICASA's existing statutory independence, communications-sector expertise and institutional experience provide a substantial foundation upon which such a function could be built.

There is a further practical advantage. ICASA already operates in regulatory environments in which mandates intersect and has experience of formal cooperation with other public and private regulatory bodies.⁴⁵ This provides a useful institutional basis for the networked model proposed here. ICASA would nevertheless require substantially expanded specialist capacity. Platform regulation involves expertise extending beyond conventional electronic communications regulation into fields including data science, algorithmic systems, content moderation, online advertising, child safety, information integrity, freedom of expression and South Africa's particular linguistic and social environment. As discussed in section 9, an expanded mandate would therefore have to be accompanied by corresponding resources.

11.3 What ICASA would regulate

ICASA's central responsibility should be to supervise compliance with the direct obligations imposed upon regulated intermediaries and platforms by the new framework. Depending upon the final statutory design, these would include obligations concerning:

- ▶ accessible notice-and-action mechanisms for allegedly illegal content;
- ▶ internal complaint and appeal procedures;
- ▶ reason-giving for significant moderation decisions;
- ▶ transparency reporting;
- ▶ advertising and recommender-system transparency;
- ▶ regulatory and researcher access to platform data;

⁴⁴ Constitution s 192 and ICASA Act 13 of 2000.

⁴⁵ See, for example, ICASA Act s 4(3A)(b) and s 4B(8)(b), and Chapter 10 of the Electronic Communications Act 36 of 2005.

- ▶ systemic-risk assessment and mitigation by systemically significant services;
- ▶ appropriate linguistic and contextual capacity;
- ▶ cooperation with trusted flaggers and other recognised mechanisms;
- ▶ compliance with requirements concerning local representation and regulatory contact; and
- ▶ other procedural and systemic obligations imposed directly upon providers.

This focus is important. ICASA's concern would not ordinarily be whether the platform reached the legally correct conclusion about every individual notice or moderation decision. Its concern would be whether the provider has established and operates systems that comply with the law.

Individual complaints may nevertheless provide important evidence of regulatory failure. A single erroneous decision may require correction through the dispute-resolution processes considered in section 4. A pattern of similar complaints may reveal something different: an inadequate notice mechanism, persistent failure to provide reasons, deficient human review, systematic discrimination against particular languages, failure to implement successful internal appeals, or some other structural weakness.

ICASA should be able to investigate and address the latter. In this way, individual experience can inform regulatory supervision without converting the regulator into the ordinary adjudicator of individual content disputes.

11.4 A co-regulatory model

Assigning principal responsibility to ICASA should not mean that the authority itself performs every regulatory function. South Africa's institutional and resource constraints make co-regulation particularly attractive. Properly designed, co-regulation occupies a space between two unsatisfactory extremes. It neither leaves platforms to govern themselves according to purely voluntary standards nor requires the state to assume direct responsibility for every moderation decision, complaint, audit or technical assessment.

South African communications regulation already provides useful precedents. Broadcasting regulation permits qualifying broadcasters, subject to an approved code and disciplinary mechanism, to operate within a co-regulatory system, reflected most prominently in the role played by the Broadcasting Complaints Commission of South Africa (BCCSA).⁴⁶ ICASA has also developed formal cooperative

46 Electronic Communications Act 36 of 2005, s 54.

relationships with other bodies where regulatory responsibilities intersect.⁴⁷ A future platform framework could develop this model further. Parliament would establish statutory obligations, minimum standards and safeguards. ICASA would supervise compliance, prescribe or approve appropriate regulatory codes and retain the ability to intervene where co-regulatory arrangements prove inadequate. Platforms would bear primary responsibility for operating the complaint, moderation, transparency, risk-assessment and other systems required by law.

Independent actors could perform specified functions within that framework. Trusted flaggers can contribute specialised expertise concerning particular forms of allegedly unlawful content. Independent researchers can help identify and understand systemic risks. Independent auditors can test aspects of platform compliance. And, as discussed in section 4, appropriately qualified independent dispute-resolution bodies can provide a means of resolving individual disputes without either leaving the platform as the final decision-maker or requiring ICASA to adjudicate their merits.

The resulting model might best be understood as regulated co-governance. Responsibility is distributed where doing so improves expertise, accessibility and efficiency, but within a statutory framework establishing binding obligations and ultimate public accountability. Co-regulation must not, however, become another name for industry self-regulation. ICASA would require sufficient authority to determine whether relevant codes and mechanisms satisfy statutory requirements, supervise their operation and take enforcement action where they fail to do so.

11.5 Individual disputes and independent redress

The distinction between platform supervision and individual content adjudication has particular implications for the complaint architecture considered in section 4. The first line of recourse should remain the platform itself. A person alleging that content is unlawful should have access to an effective notice-and-action mechanism and receive a reasoned response. A user affected by a significant moderation decision should similarly be entitled to reasons and meaningful internal reconsideration.

Where that process fails to resolve the dispute, an accessible independent mechanism should be available. As discussed in section 4, the DSA's system of certified out-of-court dispute settlement offers a useful model, although its precise institutional machinery need not be reproduced. South Africa might accredit suitable existing bodies, establish an ombud-type mechanism or develop another comparatively lightweight system capable of providing independent and expeditious review. The important point for present purposes is that ICASA need not itself perform this adjudicative function.

47 ICASA, 'Memoranda of Understanding' <https://www.icasa.org.za/pages/mous> accessed 15 September 2026, confirming existing arrangements with the Advertising Regulatory Body, the Film and Publications Board, the Competition Commission and the National Consumer Commission, among others.

Its role should instead be to establish or supervise the regulatory conditions within which independent dispute resolution operates: prescribe minimum requirements where appropriate, recognise or accredit qualifying mechanisms if legislation so provides, monitor whether platforms participate as required, and intervene where repeated disputes reveal non-compliance with the platform's statutory obligations.

The decisions of an independent dispute-resolution mechanism should not displace access to the courts. Nor should such a body be given inappropriate final authority to determine the lawful boundaries of expression. As under the DSA model, independent dispute settlement can provide an accessible means of resolving many disputes without transforming it into the final source of binding law.

11.6 Relationship with specialist institutions and courts

Existing South African institutions would continue to exercise the jurisdiction conferred upon them by substantive law. Thus, privacy and data-protection matters falling within POPIA remain within the statutory framework administered by the Information Regulator. Questions concerning prohibited discrimination and hate speech remain subject to PEPUDA and the relevant roles of the Equality Courts and SAHRC. Electoral disputes remain subject to the institutions and procedures created by electoral law. Matters falling within the statutory mandate of the Films and Publication Board remain subject to that framework. Apparent criminal conduct remains a matter for law-enforcement authorities and the criminal courts. Competition questions remain within the competition-law regime.

The precise interaction between these institutions and ICASA would require detailed statutory design. Formal cooperation arrangements should facilitate referrals, information sharing and coordination where the same factual circumstances raise both platform-compliance questions and issues falling within another body's mandate.

But these institutions should not be conceptualised as appellate bodies reviewing ICASA's decisions on the legality of individual content, because ICASA should ordinarily not make such determinations in the first place.

Nor should every dispute be required to pass through ICASA before reaching the competent institution. Where a person has a statutory right to approach the Information Regulator, Equality Court, Electoral Court or ordinary courts, the new platform framework should not unnecessarily obstruct that right. Similarly, where existing law requires a platform to report specified conduct directly to a competent authority, that obligation should remain. The objective is coordination, not the construction of an additional jurisdictional obstacle.

11.7 Institutional independence and safeguards for freedom of expression

Institutional independence is especially important because platform regulation inevitably intersects with public discourse. An authority supervising systems capable of affecting the availability and distribution of expression must be insulated as far as possible from partisan political direction. ICASA's existing statutory independence provides an important institutional foundation, but any expansion of its mandate should reinforce rather than dilute that protection.

The proposed allocation of functions itself provides an additional safeguard. ICASA would principally supervise platform systems and compliance with statutory processes rather than become a general arbiter of the truth, desirability or acceptability of lawful online speech. Authoritative determinations of illegality would continue to be made through the legal institutions to which South African law assigns that function.

Regulatory decisions should be reasoned and, subject to legitimate privacy and confidentiality interests, decisions of broader significance should be publicly accessible. ICASA's exercise of its new functions should remain subject to ordinary requirements of legality, reasonableness and procedural fairness and to judicial review. The institutional proposition advanced here is consequently narrower than the creation of a general online-content regulator. ICASA should serve as the principal supervisor of obligations imposed upon digital platforms, operating within a co-regulatory and networked system in which independent dispute-resolution mechanisms address appropriate individual disputes, existing specialist institutions retain their substantive mandates, and courts remain the ultimate guardians of legality and constitutional rights.



Section 12

Effective enforcement

The regulatory framework proposed in this report will have little value unless the authority responsible for supervising it has the powers and resources necessary to determine whether regulated providers are complying with their obligations and to respond effectively where they are not.

It is hard to imagine a context in which this is more imperative than one involving global digital platforms. The disparity in information, technical expertise and resources between major technology companies and South African regulatory institutions is considerable. Platforms possess detailed knowledge of their own systems, while regulators are necessarily dependent upon access to information and data held by those they regulate. Moreover, penalties that may be significant for domestic firms can be commercially immaterial to multinational companies operating at an enormous scale.

Effective regulation therefore requires more than the creation of substantive obligations. ICASA would need sufficient investigatory powers to assess compliance; remedial powers to require deficiencies to be corrected; and meaningful sanctions where providers repeatedly or deliberately fail to comply. These powers should be graduated and proportionate, with the principal objective of securing compliance rather than punishment for its own sake.

12.1 Investigatory and supervisory powers

Effective enforcement begins with information. As discussed elsewhere in this report, one of the fundamental difficulties of platform regulation is the informational asymmetry between providers and those attempting to assess their conduct. Platforms possess extensive information concerning their systems, moderation practices, advertising, recommender systems, risk assessments and internal operations that regulators ordinarily cannot obtain independently. ICASA should therefore possess statutory powers sufficient to determine whether regulated providers are complying with the obligations imposed upon them.

Subject to appropriate limitations and differentiated application according to the nature and scale of the service, these should include powers to:

- ▶ require information and documents reasonably necessary to assess compliance;
- ▶ obtain access to relevant platform data where statutorily authorised;
- ▶ require explanations concerning the operation of relevant systems and processes;
- ▶ receive and scrutinise transparency reports, systemic-risk assessments and other prescribed regulatory information;
- ▶ investigate complaints or other information indicating possible regulatory non-compliance;
- ▶ conduct or commission appropriately defined audits and technical assessments;
- ▶ require information concerning steps taken to remedy identified regulatory deficiencies; and
- ▶ cooperate and exchange relevant information with other competent authorities, subject to appropriate privacy, confidentiality and data-protection safeguards.

These powers should be calibrated to the nature and significance of the regulated service. Particularly demanding requirements concerning access to data, independent audits or technical systems should principally attach to providers whose scale, function or systemic significance justifies the regulatory burden. They should also be tied to ICASA's statutory functions. Regulatory oversight should not provide an unlimited entitlement to commercially sensitive information, personal data or private communications unrelated to legitimate supervision. Requirements to produce information or provide access to data should be necessary and proportionate to the regulatory purpose for which they are exercised and subject to appropriate safeguards concerning confidentiality, privacy and security.

12.2 Compliance and remedial powers

The principal objective of enforcement should ordinarily be to secure compliance. Many regulatory failures are likely to concern deficient systems or processes rather than deliberate defiance of the law. A platform may operate an inadequate complaints mechanism, fail to provide prescribed reasons, publish incomplete transparency information, produce an insufficient systemic-risk assessment, lack appropriate processes for South African languages or fail to provide access to data in the manner required by law.

In such circumstances, the appropriate regulatory response will ordinarily be to require the deficiency to be corrected. ICASA should therefore possess a

graduated range of remedial powers. Depending upon the nature and seriousness of the breach, these could include recommendations, compliance plans, formal compliance directions and enforceable undertakings. The Authority might, for example, require a provider to correct deficiencies in its notice-and-action system; improve the operation of its internal complaint procedures; comply with prescribed reason-giving requirements; publish missing transparency information; undertake or reconsider an inadequate systemic-risk assessment; provide regulatory or researcher access to data required by law; improve deficient linguistic or contextual capacity; or comply with requirements concerning its South African legal representative.

This compliance-oriented approach is particularly appropriate to a co-regulatory framework. Not every deficiency needs to attract a financial penalty immediately. Engagement between the regulator and provider may often provide the most efficient means of correcting a problem. But such engagement can operate effectively only against the background of credible enforcement. Persistent failure to remedy identified deficiencies, repeated non-compliance or deliberate obstruction should attract progressively stronger measures.

12.3 Responding to urgent regulatory failures

Some regulatory failures may require intervention on a substantially shorter timescale. This may arise, for example, where a provider fails to operate a prescribed urgent-escalation mechanism for reports involving particularly serious unlawful activity; where a required safety process breaks down during a rapidly developing crisis; where information that a provider is legally required to preserve is at risk of loss; or where another serious failure of the provider's systems creates an immediate risk of substantial harm.

ICASA should therefore possess appropriately circumscribed interim powers to require urgent compliance with obligations imposed by the regulatory framework. Depending upon the circumstances, these might include requiring a platform immediately to activate a prescribed escalation procedure, deploy appropriate human review, preserve relevant information where legally authorised, remedy a failed complaint mechanism, or comply with an urgent lawful direction issued by a competent authority.

The threshold for such intervention should be clearly defined in legislation. The nature of the measure should be proportionate to the urgency and seriousness of the regulatory failure, and particularly intrusive measures should attract correspondingly stronger procedural and judicial safeguards. Urgent powers are important precisely because digital systems operate at extraordinary speed and scale. A regulatory direction delivered weeks after a serious systems failure may provide little meaningful remedy. Effective oversight therefore requires an ability to act rapidly in exceptional circumstances without allowing urgency to become a justification for unconfined regulatory discretion.

12.4 Compliance with lawful orders

A coherent digital regulatory framework must also address the relationship between platforms and lawful orders issued by competent South African authorities. Where a court or other competent authority has lawfully directed a provider to take specified action, the platform should be subject to clear obligations concerning receipt, processing and compliance with that order. The framework should prescribe appropriate points of contact and procedures so that lawful orders are not frustrated by uncertainty about corporate structure, jurisdiction or the identity of the entity responsible for compliance.

This is one reason why the local representation requirement considered in section 3 is so fundamental. A regulatory framework cannot operate effectively if South African courts and public authorities struggle to identify an authorised person upon whom process or regulatory directions can be served.

The framework should similarly require providers to communicate appropriately with the issuing authority concerning the action taken in response to an order and, where compliance is not possible, or the order is challenged, the basis for that position. Failure by a regulated provider to comply with these procedural and regulatory requirements would itself constitute a matter capable of enforcement under the framework, independently of the substantive legal proceedings giving rise to the order.

12.5 Meaningful and proportionate penalties

Where remedial measures do not secure compliance, meaningful sanctions must be available. As discussed in section 2, safe-harbour protection in respect of third-party content should not shield a provider from regulatory consequences arising from a breach of obligations imposed directly upon it under a new regulatory framework. Failure to comply with requirements concerning transparency, complaint procedures, risk assessment, data access or other statutory duties would constitute a regulatory breach in its own right and should be enforceable as such.

Fixed monetary penalties that appear substantial in South African terms may be commercially insignificant to the world's largest technology companies. There is therefore merit in considering administrative penalties linked to turnover, subject to an appropriate statutory ceiling, rather than relying exclusively upon fixed Rand amounts.

The DSA provides a useful model of the underlying principles. It requires penalties to be effective, proportionate and dissuasive and permits fines for non-compliance of up to six percent of the provider's worldwide annual turnover. South Africa need not reproduce either that percentage or the relevant turnover base. Those matters should be determined according to domestic circumstances and constitutional requirements. But the principle of economic proportionality is important where regulated entities may range from relatively small domestic providers to some of the world's largest corporations.

The seriousness of the breach should also determine the appropriate sanction. Relevant considerations might include its nature, gravity and duration; whether it was repeated; whether the provider acted deliberately or negligently; the number of users affected; the consequences of the breach; the provider's cooperation with the regulator; steps taken to remedy the problem; and its history of previous non-compliance.

Penalties should principally attach to failures of regulatory compliance: persistent deficiencies in required systems, failure to provide prescribed information, obstruction of lawful investigations, failure to undertake required assessments, refusal to remedy identified breaches or non-compliance with lawful regulatory directions.

There is less obvious justification for creating broad criminal liability for ordinary regulatory non-compliance. Administrative penalties and judicial enforcement should generally provide adequate mechanisms. Parliament might nevertheless consider narrowly defined offences for deliberate obstruction, knowingly supplying materially false information or similarly serious conduct where administrative sanctions would be inadequate.

12.6 Judicial enforcement and procedural safeguards

A regulator must ultimately be able to secure compliance where a provider refuses to obey a lawful direction. ICASA should therefore be empowered, where necessary, to approach a court for appropriate orders compelling compliance with obligations or regulatory directions issued under the framework. Judicial enforcement would be particularly important where coercive measures are required or where enforcement substantially affects protected rights.

The availability of effective enforcement should be matched by appropriate procedural safeguards. A provider facing regulatory action should ordinarily receive adequate notice of the alleged breach, an opportunity to respond, a reasoned decision and sufficient clarity concerning the steps required for compliance. Regulatory action should be proportionate to the seriousness of the breach and the consequences of continued non-compliance.

Particularly intrusive investigatory or remedial measures require correspondingly strong safeguards. Where a measure would substantially interfere with expression, privacy, confidential information or other protected rights, legislation should specify the conditions for its exercise with sufficient precision. Judicial authorisation should be required where the nature of the intervention makes this constitutionally appropriate. ICASA's own decisions should remain subject to judicial review.

Regulatory transparency should also form part of this framework. Significant enforcement decisions should ordinarily be accompanied by publicly available reasons, subject to legitimate confidentiality, security and privacy considerations. Over time, a body of published decisions would provide platforms, users and the

regulator itself with greater certainty about the application of the framework and permit public scrutiny of the regulator's exercise of its powers.

12.7 An enforceable rather than aspirational framework

The question of enforcement ultimately returns the analysis to a foundational requirement of this report. A South African digital regulatory framework would fail if it contained sophisticated obligations but no practical means of enforcing them against providers operating on a global scale. Jurisdiction and local representation must therefore be matched by a supervising authority possessing the information, expertise, resources and powers necessary to determine whether statutory obligations are being met and to respond effectively where they are not.

The objective should not be enforcement for its own sake, or the construction of an unnecessarily punitive regulatory regime. It should be to ensure that obligations imposed by Parliament are capable of producing compliance in practice. A graduated system of investigation, remedial directions, administrative penalties and, where necessary, judicial enforcement can provide that capacity while remaining subject to requirements of legality, proportionality, procedural fairness and judicial review. Effective enforcement is, in this sense, the final condition necessary to ensure that the regulatory framework considered in this report amounts to more than a set of aspirations.



Conclusion

South Africa does not need its own Digital Services Act. Attempting to reproduce the European framework provision by provision would be neither realistic nor desirable. The DSA reflects the particular legal and institutional architecture of the European Union and is supported by regulatory resources and economic leverage that South Africa does not possess. Its value for South Africa lies instead in the regulatory principles and mechanisms it offers for consideration and adaptation.

Perhaps the most important lesson is that meaningful regulation of digital platforms need not require intermediaries to become generally liable for what their users say, nor the state to assume responsibility for determining what information may appear online. Appropriate protection from liability for third-party content can coexist with direct obligations upon the companies that design and operate digital services. Platforms may remain protected from automatic liability for users' speech, while being required to maintain accessible complaint procedures, explain consequential decisions, disclose meaningful information about their operations, permit appropriate external scrutiny, assess and mitigate serious risks generated by their systems, protect children using their services and comply with lawful regulatory requirements.

That distinction is particularly important in South Africa. Our constitutional commitment to freedom of expression counsels strongly against regulatory models that incentivise indiscriminate removal of contested speech or confer broad powers upon public authorities to determine what may be said online. But freedom of expression does not require that the private infrastructures through which an ever-greater share of public communication takes place remain opaque, inaccessible or effectively beyond the reach of public law.

The analysis in this report therefore points towards a framework in which responsibility is differentiated according to what a service does, the risks it creates and the power it exercises. Regulation must begin by ensuring that providers substantially serving South Africans are within the effective reach of South African law. It should make consequential exercises of platform power more transparent and contestable, provide regulators and independent researchers with the information necessary for meaningful scrutiny, and impose greater responsibilities where the scale or nature of a service creates greater risks. Those responsibilities must also be meaningful in South African conditions, including for children and across the country's diverse linguistic and social contexts.

These measures depend upon one another. A right to complain means little if the provider is legally unreachable. Transparency is of limited value if there is no capacity to interrogate what is disclosed. Risk assessments accomplish little if platforms alone determine whether their responses are adequate. And statutory obligations will remain aspirational if the authority charged with supervising them lacks the expertise, resources and powers necessary to secure compliance. South Africa's limited regulatory capacity therefore argues for a selective and proportionate framework, not for continued reliance upon voluntary platform governance: one that concentrates the greatest obligations upon the most consequential services, requires providers themselves to undertake much of the primary compliance work, draws upon independent expertise and builds upon existing institutions wherever possible.

Acknowledgements

This report benefited enormously from the generosity of those who gave us their time, shared their expertise and helped us understand how the Digital Services Act is working in practice. During our research visit to Berlin, Brussels and Dublin in July 2026, we had the privilege of meeting people working on the DSA from very different perspectives — as regulators, policymakers, researchers, civil society organisations and platforms.

Our particular thanks go to Dr Julian Jaursch; Jakob Ohme of the Weizenbaum Institute for the Networked Society; Johannes Heidelberger of the Bundesnetzagentur; Clara Helming of AlgorithmWatch; Sara Tengo, Alejandro Menéndez Fernández and their colleagues at the European Commission's Directorate-General for Communications Networks, Content and Technology (DG CNECT); Paolo Cesarini of the European Digital Media Observatory; Manon Letouche and Giulia Torchio of the 5Rights Foundation; Eva Morris, Susan Moss, Brie Pegum and Patrick Callaghan at TikTok; and Jessica Gaynor, Nadia Mahfoud, Claudia Gichuhi, Rachel Cunningham and Kevin Whooley of Coimisiún na Meán. We are grateful for the candour with which they shared their experiences and perspectives. The conclusions and recommendations in this report are, of course, our own. While we did not get to meet Giacomo Mazzone of the European Digital Media Observatory in person, he was extraordinarily generous in facilitating contacts for our visit, and we are indebted.

This research would not have been possible without the generous support of the Konrad-Adenauer-Stiftung (KAS), and we are especially grateful to Hendrik Sittig and Cailan Ferreira-Mokoka. Dr Olaf Wientzek, Sven Nicolay and Meike Lenzner of the KAS Brussels office generously made time for us during our visit there.

We have also benefitted greatly from conversations closer to home. Our thanks go to our colleagues in the Free Expression Legal Network (FELN), whose collective experience in defending freedom of expression in South Africa continually informs CFE's thinking about how regulation can respond to genuine online harms without unnecessarily restricting lawful expression. We are similarly grateful to our friends and colleagues at the Campaign on Digital Ethics (CODE) and Moxii Africa, whose work and insights into the South African digital environment have enriched our understanding of these issues.

Finally, our thanks go to the entire team at the Campaign for Free Expression for their support throughout this project, and especially to Deborah Petersen, whose formidable logistical support made the research trip—and much else besides—possible.

And to those who held the forts while we were away: you know who you are, and we are very grateful.

September 2026



Download an electronic copy of
*A South African Legal Framework
to Govern Online Harm: What
Might We Learn from the EU's
Digital Services Act?* from [www.kas.de/en/web/medien-afrika/
einzeltitel/detail/-/content/new-
study-making-digital-platforms-
accountable-in-south-africa](http://www.kas.de/en/web/medien-afrika/einzeltitel/detail/-/content/new-study-making-digital-platforms-accountable-in-south-africa)
e-book ISBN: 978-1-0483-7324-0

ISBN: 978-1-0483-7323-3