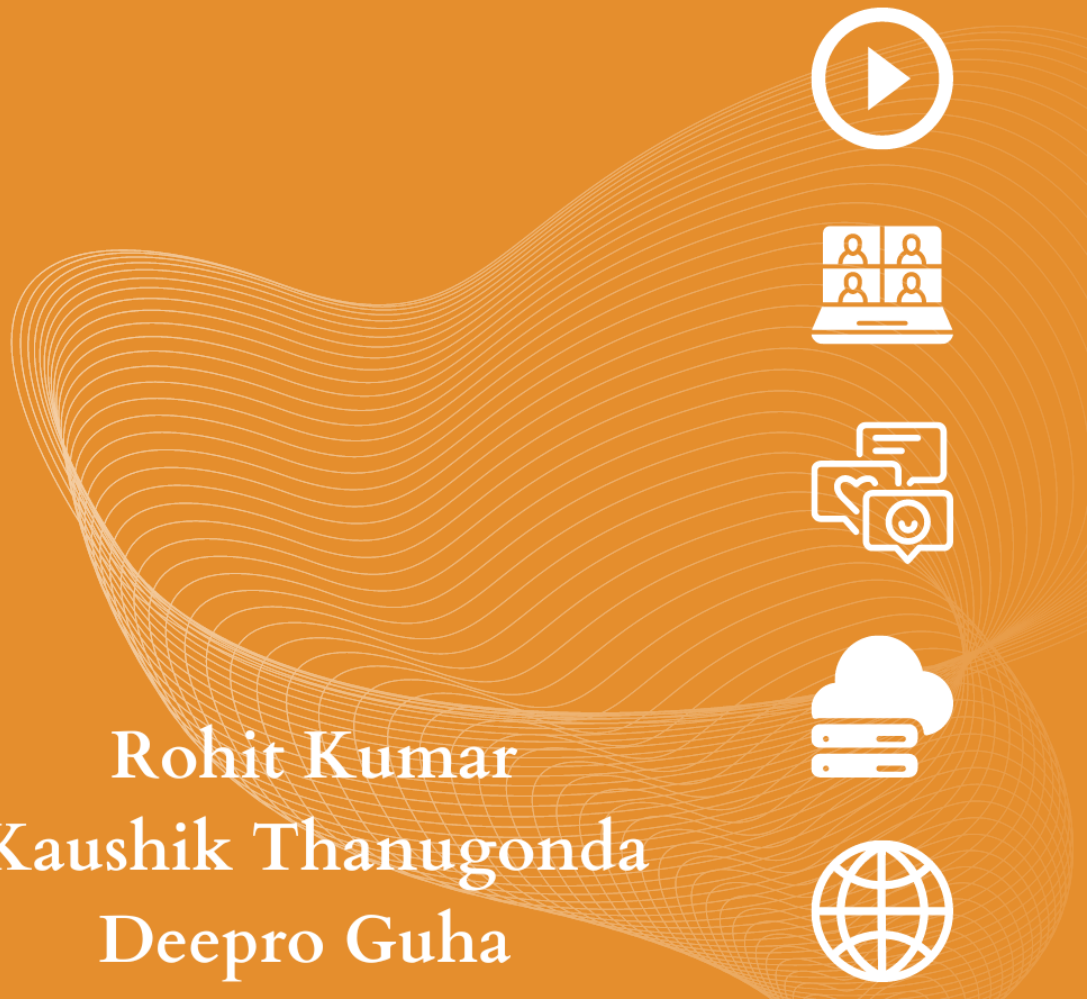
TQH

A Framework for Intermediary Classification in India

December 2022



Rohit Kumar
Kaushik Thanugonda
Deepro Guha

Note

This paper is an independent, non-commissioned piece of work by *The Quantum Hub (TQH)*, a public policy firm based in Delhi.

Authors

Rohit Kumar - Founding Partner at TQH
Kaushik Thanugonda - Analyst, Public Policy at TQH
Deepro Guha - Manager, Public Policy at TQH

I. Setting the context

As India prepares to replace its two-decades-old Information Technology Act with a new Digital India Act, it may be worth taking a fresh look at intermediary classification from a proportionate and risk-based approach to regulation

Intermediaries in India are defined under section 2(w) of the Information Technology Act 2000 ("IT Act").[1] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 ("IT Rules, 2021"), notified in February 2021, lay down the framework for their regulation. These rules introduce a new classification that categorizes intermediaries into different kinds and prescribes obligations for each category.

Although the classification is useful in providing differentiated obligations, the category definitions are still quite broad and not nuanced enough. In the digital world, there are various kinds of intermediaries, providing different types of services, and not all intermediaries inflict public harm or impact public discourse. Intermediaries providing enterprise solutions, for instance, are integral to the functioning of most organizations, but don't pose the same risks as a platform that allows for wider dissemination of information. Therefore, bucketing all such intermediaries into one unified category and imposing similar legal obligations on them may not be appropriate.

As the Government of India prepares to replace the IT Act 2000 with a new Digital India Act, it may be worth taking a fresh look at intermediary classification to recognise the complexity of the online space today. The new approach can take a proportionate and risk-based lens to regulation by considering a range of factors such as platform features, number and types of users, as well as the nature of risks involved to propose an alternative classification framework. If such a framework were to create space for participation by the industry, it may also allow service providers to come up with solutions that work to address platform-specific dynamics, without running the risk of overregulation.

Given the above context, this paper attempts to propose a new way of classifying intermediaries to help improve accountability and online safety, while also reducing legal obligations for intermediaries. It is hoped that the proposed framework can help achieve the government's policy goal of creating a safer internet ecosystem while also allowing businesses to thrive.

[1] 'Intermediary', with respect to any particular electronic record, means any person who on behalf of another person receives, stores or transmits that record or provides any service with respect to that record and includes telecom service providers, network service providers, internet service providers, web-hosting service providers, search engines, online payment sites, online-auction sites, online-marketplaces and cyber cafes

II. Current obligations on intermediaries

In India, the term intermediary is broad enough to cover a wide range of service providers, including internet service providers, cloud service providers, consumer-facing social media platforms, video sharing sites, etc.

Rule 3 of the IT Rules 2021 lays down new obligations for all intermediaries under the IT Act. These include periodic reminders to users about terms of service and privacy policies, 180-day retention of information related to users registering on a platform, 72-hour timelines for responding to law enforcement requests etc. Non-observance of these obligations can lead to loss of safe harbor protection guaranteed under Section 79 of the IT Act 2000.

Rule 2(w) of the IT Rules 2021 creates a new category of intermediaries called "social media intermediaries". These are defined as intermediaries that primarily or solely enable online interaction between two or more users and allow them to create, upload, share, disseminate, modify or access information using their service. The Rules also subclassify '*social media intermediaries*' ("SMIs") with over 50 lakh (5 million) registered users as '*significant social media intermediaries*' ('SSMIs').

SMIs do not have any special obligations under the IT Rules 2021, and are expected to adhere to the same norms as other intermediaries. However, SSMIs are subject to higher regulatory requirements under Rule 4. These include appointing India-based officers for ensuring compliance, coordination with law enforcement, and grievance redressal. SSMIs are also expected to have a physical contact address in India, publish monthly compliance reports, use automated tools to identify and remove illegal content while also constituting robust grievance redressal mechanisms and clearly identifying copyrighted/ sponsored content.

On 28th October 2022, the government notified the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2022 ("2022 Amendments"), which created new obligations for intermediaries. Among other things, the 2022 Amendments impose a legal obligation on intermediaries to make reasonable efforts to prevent users from uploading certain forms of harmful/ illegal content. It also establishes Grievance Appellate Committees (GACs) to which users can appeal decisions taken by grievance officers.

The amendments do not alter the classification of intermediaries prescribed under the 2021 Rules. Therefore, the current classification regime continues to disregard the diversity of forms among intermediaries, which can differ widely on the basis of the services provided, the level of access to sensitive information, functionalities etc. As it stands currently, similar obligations are laid down for all kinds of intermediaries despite fundamental differences in the way intermediaries transmit information or interact with users and their content.

III. Diversity of forms among intermediaries

1. Primary purpose of the service

The current framework for classification of intermediaries in India does not meaningfully differentiate between service providers on the basis of the services provided by them. Barring a few clauses on taking down content, network infrastructure players such as internet service providers, caching services, domain name registrars and other players such as cloud service providers are all required to comply with similar obligations as social media platforms or e-commerce websites, without regard to the intermediary's level of access to the information being transmitted.[2]

Even the 2022 Amendments continue this line of regulation. Although the press release accompanying a draft copy of the amendments had mentioned that they seek to specifically target social media platforms, the final amendments are applicable to all intermediaries. As a result, every intermediary under the IT Act is subject to a Grievance Appellate Committee's (GAC) jurisdiction and required to adhere to a timeline of 72 hours for removal of content in most cases.

This assumes that all intermediaries can always identify and remove unlawful information, which is not true. Many enterprise service providers such as cloud service providers have no visibility or access to the content they host; they often lease infrastructure to customers under contractual obligations that prevent them from accessing client data. Similarly, infrastructure providers such as content delivery networks and domain registrars only facilitate smooth functioning of internet infrastructure, without exercising control over the information passing through them. Therefore, any move to introduce a content moderation provision for such platforms may amount to violation of privacy and confidentiality under existing contracts.

The EU's Digital Services Act recognizes these distinctions, and divides intermediaries into three broad categories - conduits, caching and hosting services. While conduits merely transfer information, caching services involve automatic, intermediate and temporary storage of information for the sole purpose of efficient transmission. Hosting services, in comparison, are involved in storage of information provided by, and at the request of, recipients of the service.[3]

Conduits and caching services are exempted from being held liable for the information they transmit or temporarily store, as long as they do not modify the information. However, hosting services can be held liable if they have knowledge of hosted content, or do not take expeditious measures to remove it when notified of it.[4] Thus, we see that intermediaries acting

[2]Under 3(1)(d) of the IT Rules 2021, law enforcement and other designated agencies can request the intermediary to take down content. However, an exception is made under 3(1)(e) for intermediaries that only engage in "temporary or transient or intermediate storage of information".
[3]Article 2, Digital Services Act, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en
[4]Articles 3, 4, 5, Digital Services Act

as mere infrastructure providers, like conduit and caching services, have reduced liability for content, when compared to other intermediaries, like hosting services, which host information.

2. Type of user base

The type of user base is also a relevant consideration for classifying intermediaries. **Services primarily developed for internal use by organizations, educational institutions, businesses etc. are different from social media intermediaries and other public-facing services.** These services have different clients and different core functions.

Let's take the example of a business or enterprise software. Such software is designed specifically for use by businesses and companies. These can include unified communication tools like Microsoft Teams, Zoom etc. and SaaS services that provide HR management, collaboration, file sharing and customer management solutions (such as Zoho or Oracle).

Such tools are usually licensed to organizations or businesses, and are optimized for their use. These tools are also likely to have a closed user base, thus reducing the risk of harm through spam or misinformation going viral. Regulating such services in the same manner as public facing platforms is only likely to increase compliance burden for service providers, without meaningfully reducing risks presented by the internet.

3. Functionalities of the platform

Another important metric to distinguish between intermediaries is the functionality provided by their platforms. The current definition of intermediaries under section 2(w) of the IT Act 2000 provides a list of covered entities such as web hosting services and internet service providers, but does not distinguish or sub-classify them on the basis of their functions.

While social media intermediaries are now defined separately under IT Rules 2021, this definition is so widely worded as to include all sorts of communication services under its umbrella. The definition - "*primarily or solely enables online interaction between two or more users*" - not only includes major social media platforms such as Twitter, Facebook, Whatsapp etc. but also services like matrimonial apps and enterprise communication services such as Zoho or Webex. While it is true that the above mentioned services enable users to interact, the core purpose of these platforms is not social media intermediation and they don't pose the same risks as social media platforms.

For example, let us consider the viral spread of misinformation. This is usually possible on platforms that allow users to view or upload content such as videos, blogs, posts etc. and enable for such content to be discovered and shared with a wider audience. Sharing is usually facilitated through a content feed where users can see content from those they follow, or

through messaging services where large communities of users can access and disseminate information to both known and unknown users.[5]

Content moderation on such platforms involves both human content moderators who individually review content, as well as automated filters which detect and flag spam, copyright infringements and other illegal content. However, not all services that are classified as social media intermediaries in India facilitate viral spread of information. For instance, dating apps lack a "content feed" where user generated content can be seen by the public. Similarly, some online games only allow users to communicate for the purposes of the game, but do not provide access to records or allow sharing once the user exits the game. Business communication services and other unified communication platforms similarly lack public user-generated content, universal search, pages or public directories of any sort.

Services like these should therefore be regulated differently from social media platforms that provide "social" features for broader dissemination.

4. Access to information

While considering a framework for classification, we should also think about the kind of information that is held by an intermediary, and the manner in which it is held. Not all intermediaries retain or collect data from their users. Some intermediaries only have ephemeral access to information (they do not store information or allow users access to it in perpetuity), while others host content which is ephemeral in itself. For instance, some platforms like Snapchat and Instagram allow users to post "stories" which are pieces of content that disappear after 24 hours. Similarly, video conferencing platforms usually do not store all recordings unless a user actively chooses to record a meeting. Even if files, links etc. are shared between users during a meeting, they disappear afterwards unless the user specifically chooses to save them.

And while an enterprise software or a cloud service provider may allow users to store information on the cloud, it may only enable content discovery when specifically authorized by the file owner. A document made or hosted on Google Docs, for instance, cannot be accessed by other users unless the document is made publicly available or its unique link is disseminated through other channels. The platform itself does not allow for public access to information.

Differentiating between platforms by the level of access they give to content generated by other users, and the amount of data they retain is important. While there is potential for public harm on large social media platforms through circulation of problematic content (like CSAM[6] or misinformation), intermediaries that do not store information or only facilitate access to it by specific users/ within closed networks are less likely to lead to such harm.

---

[5]Filippo Menczer, Thomas Hills (2020, December 1). Information Overload Helps Fake News Spread, and Social Media Knows It. Scientific American. Retrieved December 16, 2022, from https://www.scientificamerican.com/article/information-overload-helps-fake-news-spread-and-social-media-knows-it/

[6]Child Sexual Abuse Material

Table I: How are other countries thinking about intermediary classification?

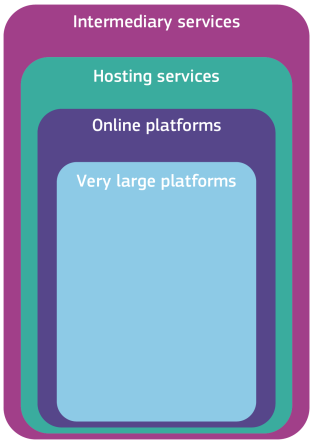
Country	Legal framework	Classification and corresponding obligations
EU	Digital Services Act	Intermediaries are first divided into conduit, caching and hosting services. Exemptions are created for conduit and caching services. Three subsequent tiers are created i.e. hosting services, online platforms and very large online platforms (online platforms reaching 10% or more of EU's population). Obligations are cumulative, proportionate to ability and size while ensuring intermediaries remain accountable. The first tier has the lowest obligations, increasing every subsequent tier.
Australia	Online Safety Act 2021 Consolidated Industry Codes of Practice for the Online Industry	eSafety, Australia's independent online safety regulator, took a co-regulatory approach. Draft codes were developed in association with the industry for eight key sections of the online industry that provide a wide range of services. Separate codes cover internet carriage services, hosting service providers, email, gaming, websites, search engines, messaging apps, social media platforms etc. Intermediaries have to **conduct mandatory risk assessments** to check for the risk of exposing users to Class 1 content (CSAM, terrorism etc.) to determine which code to adhere to. Violation of the code's provisions can lead to civil penalties or investigations by the online safety commissioner.
UK	Online Safety Bill (yet to be passed)	This Bill designates certain online services as "user-to-user services" and "search services". All such services, except those exempted by the Bill, are termed "regulated services". Certain services like email, one-to-one communication, internal business services etc. are exempted. All regulated services are required to follow certain compliances to ensure safety of users including removal of illegal content, introducing user empowerment tools, preventing fraudulent adverts etc. Regulated services are mandated to carry out **risk assessments** to assess the risk of harm related to illegal content and take proportionate steps to mitigate those risks. Obligations are higher for services accessible to minors.

IV. Potential models of regulation

Given the wide differences in product functionality, size, user base etc., there is an urgent need for more specificity in the regulation of intermediaries. **Specificity can help target policy concerns presented by different intermediaries, while also establishing a regulatory regime that is proportionate to the risks involved.**

In the text that follows, we explore the ways in which other jurisdictions classify intermediaries as well as the differences in legal obligations imposed on them.

European Union (EU)



The EU's Digital Services Act creates four layers of classification for intermediaries, with obligations cumulatively increasing for each layer. **It does not require intermediaries to monitor the information they transmit or store, or actively investigate illegal activity, but imposes obligations the moment an intermediary has any knowledge or notice of such activity**.[7] Some exemptions are given to micro and small enterprises.

All intermediary services are subject to a base level of obligations. The obligations increase cumulatively for each subsequent tier, with the highest level of obligations only imposed on very large online platforms (VLOPs). The tiers, ordered by increasing levels of compliance, are[8]:

- Online intermediary services offering network infrastructure: Internet access providers, domain name registrars, including also:
    - Hosting services such as cloud and web hosting services, including also:
        - Online platforms bringing together sellers and consumers such as online marketplaces, app stores, collaborative economy and social media platforms.
            - Very large online platforms (VLOPs), online platforms reaching more than 10% of the 450 million consumers in Europe.

Different due diligence obligations are laid out for each of these tiers. All intermediaries, irrespective of their size or function are mandated to establish a single point of contact for communication with competent authorities, to include in their terms and conditions any

[7]Article 7, Digital Markets Act, European Union.
[8]The Digital Services Act: ensuring a safe and accountable online environment. (n.d.). European Commission. Retrieved December 16, 2022, from https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en

restrictions that they may impose on users, and to comply with transparency reporting obligations (except micro and small enterprises).[9]

Hosting service providers i.e. intermediaries that store information at the request of a recipient of the service, have higher obligations than mere conduit or caching services; they have to provide mechanisms by which third parties can report illegal content, notify users about content take down, and report offenses as soon as these come to their knowledge.

Online platforms - except micro and small enterprises - in addition to the above, need to have a complaint and redress mechanism, out of court dispute settlements, ban targeted advertisements to children, and maintain transparency in recommender systems and online advertising.

VLOPs are classified in the highest category as they are seen as posing particular risks with respect to the dissemination of illegal content and societal harms. Online platforms qualifying for this threshold have to conduct risk assessments on the systemic risks regarding the use of their services, conduct mandatory external audits, and even provide data to legal authorities in specific cases.

Australia

Australia is taking an industry-led approach to improving online safety. It passed the *Online Safety Act 2021* to keep pace with advances in technology and the threats faced online from harmful behavior and toxic content. The Act requires the industry to develop new codes to regulate illegal and restricted content, such as those portraying acts of terrorism, and CSAM.

The Act allows for the establishment of codes or standards, which are to be developed by industry bodies or associations. eSafety, Australia's independent online safety regulator, is responsible for drafting and registering these industry standards. Illegal/ harmful content is classified into Class 1 and Class 2. Class 1 content (further divided into Class 1A and 1B) is considered most harmful and consists of child sex exploitation material, pro-terror content and content promoting, depicting or inciting extreme violence, drug misuse and violent crime.

eSafety has now come up with *The Consolidated Industry Codes of Practice for the Online Industry (Class 1A and Class 1B Material)*[10], in collaboration with industry associations. Breaches of the code can be reported to eSafety, which can investigate and enforce civil penalties, injunctions etc. In total, there are 8 codes which apply to different sections of the industry. Obligations are differentiated according to the types of services provided by intermediaries.

The eight types of services identified are:

---

[9]Tim Van Canneyt. The proposed DSA - part 1 - Transforming the delivery of online services through EU regulation. 12/01/2021. https://www.fieldfisher.com/en/services/technology-outsourcing-and-privacy/technology-and-outsourcing-blog/the-digital-services-act

[10]eSafety Commissioner, Australia. Industry Codes. https://onlinesafety.org.au/codes/

1. **Social media services** - including social networks, public media sharing networks, discussion forums and consumer review networks. The definition excludes software used for online business interaction/ purposes, and only includes services used for "social purposes".[11]

2. **Relevant electronic services** - this category includes email, instant messaging, SMS/ MMS, chat, online multiplayer games and online dating services. Because relevant electronic services are seen as facilitating private communication, measures for this category are "designed to be respectful of Australian end-users' legitimate expectations around the privacy and security of those communications".[12] Within this category, other subcategories are also defined. One such subcategory is 'closed communication services'; these are services which do not allow users to search/ recommend/ add other users without their communication details. As a result, these services have lower obligations. Similar exemptions apply to enterprise-relevant electronic services (like business software) and gaming services with limited communication functionality.

3. **Designated internet services** - this generally covers websites accessible in Australia.

4. **Search engine services** - all search engines such as Google, Bing, Yahoo etc.

5. **App distribution services** - app stores such as App Store and Google Play.

6. **Hosting services** - providers of hosting services, including data centers located in Australia.

7. **Internet carriage services** - all internet service providers who provide internet access to customers in Australia

8. **Manufacturing, supplying, maintaining or installing equipment** - manufacturers of devices used to browse, and connect to the internet - laptops, smartphones, TVs etc.

Each of these services has a separate code, with obligations clearly differentiated for each service. Regulation takes a risk-based approach, with due consideration given to the risk of services being able to store class 1 material, risk of exposure of such content to children, and possibility of access and distribution of such content.

Intermediaries can do their own risk assessment based on a prescribed criteria, to follow one of the 8 codes.[13]

[11]eSafety Commissioner, Australia. Development of industry codes: position paper. September 2021. https://www.esafety.gov.au/sites/default/files/2021-09/eSafety%20Industry%20Codes%20Position%20Paper.pdf, Pg. 80
[12]eSafety Commissioner, Australia. Industry Codes. Schedule 2 – Relevant Electronic Services Online Safety Code (Class 1A and Class 1B Material). August 2022. https://onlinesafety.org.au/wp-content/uploads/2022/08/2_RES-for-PC_Final.pdf, Pg. 3
[13]eSafety Commissioner, Australia. Development of industry codes: position paper. September 2021. https://www.esafety.gov.au/sites/default/files/2021-09/eSafety%20Industry%20Codes%20Position%20Paper.pdf, Pg. 50-52

Factors to be considered while undertaking a risk assessment:

- Functionality of a service - The role the service plays in facilitating access or exposure to, and/ or distribution of class 1 and class 2 material. Risk increases for services that enable communication, or allow users to generate, store, post or share material.
- Purpose of the service - Services developed for government use, for businesses, schools, universities etc. have a lower risk profile than social networks.
- Nature of the user base - A service targeted at children, for instance, is likely to have a higher risk profile than services targeted at adults.
- Number of end-active users - A user is more likely to be exposed to harmful content on a service with a large number of active end-users.
- Potential for virality - The ability of the service to enable rapid and widespread sharing or amplification of material can increase its risk profile.

Note: This list is not an exhaustive list of relevant factors but is intended to provide the industry in Australia with a guide to develop their risk assessment frameworks.

United Kingdom

The UK government published the "*Online Safety Bill*" last year, which suggests a proportionate and risk-based approach to regulating the online space. While the Bill is currently pending in the UK Parliament, it includes user to user services and search services within its ambit. Under the Bill, all regulated services are required to follow certain compliances including removal of illegal content, introducing user empowerment tools, preventing fraudulent adverts, carrying out risk assessments etc.

However, the Bill explicitly exempts the following types of services, provided certain conditions are met[14]:

- Email, SMS, and MMS services
- Services offering only one-to-one live aural communications
- Internal business services
- Limited functionality services
- Services provided by public bodies

Providers are also expected to conduct risk assessments and take steps to address the risks they have identified, as part of their duty of care towards users, especially children. The risk assessments are meant to ascertain potential for harm associated with

[14]UK Parliament. Draft Online Safety Bill. Retrieved 18 December 2022. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf. Schedule 1. Pg. 120-121

specific functionalities, including “algorithms used by the service, and how easily, quickly and widely content may be disseminated by means of the service”. Providers are also expected to re-assess risks if they are planning significant changes to their services.

V. Way Forward - New classification framework for India

As India looks to overhaul its information technology laws, creating a graded classification framework for intermediaries that imposes proportionate obligations based on functionalities, reach and potential harm would be important. This could help establish better accountability and online safety while minimizing compliance costs for intermediaries.

While looking at reach, the focus should also shift from total registered users to average monthly active users[15] (or another such metric that may be suitable). Many platforms such as matrimonial websites, used car sales, property rental/ sale websites etc. often have a large number of registered users but only a small fraction of these users are active users. Estimates suggest that in many such cases, the monthly active user base is only 10-20% of the total registered user base. Therefore, considering actual reach would be critical for establishing proportionate legal obligations on intermediaries.

It is also important to establish a framework of regulation that is not very rigid or one that forcefits intermediaries into a particular bucket based on a narrow interpretation of some criteria. With innovation and changes in technology, new products may blend features from across different types of intermediaries; this means that any category contours we define today may not work in the future. **Therefore, what we ideally need is a framework that can easily adapt to the changing nature of products. One approach to building such a framework will be adopting a co-regulatory model that provides for service specific risk-assessment before imposing specific obligations.**

Based on the learnings from EU[16], Australia and UK and factoring in lessons from the Indian context, we propose one such tiered classification framework below. The key identifiers for each category as well as the nature of obligations that should apply to the category are outlined in the text that follows the table.

[15]Monthly active users (MAUs) are usually calculated based on the number of unique users engaging with an online service at least once a month, usually through logging in, visiting, making queries etc. For the purposes of classification, an intermediary's average MAU over the preceding 12 months could be considered.

[16]The EU's Digital Services Act is broader in scope than the IT Rules 2021, and includes a wider range of digital platforms, including app stores and e-commerce sites. It lays down obligations such as barring targeted advertisements to children, and algorithmic transparency. In India, other legislations such as the Digital Personal Data Protection Bill are looking at data-related and other obligations. Hence, we have left those out of consideration for the current classification.

Table II: Proposed classification framework

Category	Examples of Services	Level of Obligations	Nature of Obligations
A. Exempted Intermediaries: • Micro and small enterprises • Caching and conduit services	Micro and small enterprises include platforms that are relatively new, and wherein user base and/ or revenue are under specified thresholds. Caching and conduit services include CDN, Internet Exchanges, DNS service providers etc. that only involve transfer or transient storage of information, with the aim of enabling or improving the functioning of other intermediaries.	Level 0	Basic obligations for accountability, including a single point of contact for communication with law enforcement, clearly stated terms and conditions of use, and a reasonably responsive grievance redressal system tailored to the nature of service being offered by the service provider.
B. Intermediaries other than Communication Services	All intermediaries that are involved in hosting, transmission or curation of user-generated information, except those that are primarily designed as communication services. Examples: data centers, cloud services, web-hosting, search engines, online-marketplaces etc.	Level 1	Broad obligations, over and above *Level 0*, including appointment of a grievance officer[17], cooperation with law enforcement, prescribed timelines for grievance redressal, identification of advertised/ copyrighted content etc. Obligations can also include creating a reporting mechanism through which third parties can flag illegal content, with reasonable timelines for content takedown. Shorter timelines can be prescribed but must be restricted to very sensitive cases such as child sexual abuse, morphed nude images, etc.
C. Communication Services	Intermediaries that primarily enable communication/ online interaction between end-users		Risk assessments to be undertaken by all communication services to assess if they should be classified as “contained”, “open” or “designated open” depending on the risk of virality linked with their platforms.

[17]Rule 3(2)(a) of the IT Rules 2021 already mandates all intermediaries to have a grievance officer to respond to user complaints etc.

C1. Contained Communication Services	Communication services that do not lend themselves to widespread sharing or amplification of information. Examples: enterprise suites (such as email, video communications), matrimonial, property renting websites, etc.	Level 1	Similar obligations as *Intermediaries other than Communication Services*
C2. Open Communication Services	Intermediaries that enable communication/ online interaction between two or more end-users but also pose a risk for viral spread of information. Examples: social media platforms	Level 2	Obligations, over and above *Level 1*, can include automated tools to identify and mitigate sharing of illegal content, shorter timelines to address complaints regarding problematic content that may go viral, based on certain virality thresholds (explained below).
C3. Designated Open Communication Services	*Open Communication Services* with active users above a certain threshold	Level 3	Obligations over and above *Level 2*, including India based grievance and nodal officers, physical address in India, periodic compliance reporting, and an in-house grievance appellate mechanism with independent external stakeholders to increase confidence in the grievance process.

A. Exempted Intermediaries (Level 0)

Taking a proportionate approach to regulation, some exceptions should be created for intermediaries that pose minimal risk. Two such categories are the following:

- Micro and small enterprises: The government can create suitable exemptions for micro and small enterprises to make sure that obligations are proportionate to an enterprise's ability and size. A threshold can be created in the context of the IT Rules by considering a combination of criteria such as annual turnover and/ or active user base. This could help safeguard users while also reducing obligations on new startups and other small firms.

- Conduit and caching services: Intermediaries providing conduit and caching services can be exempted from liability, as they only involve transfer or transient storage of information, with the aim of enabling or improving the functioning of other intermediaries without modifying the information in any way. Rule 3(e) of the IT Rules 2021 already creates a similar exemption; it must be retained and expanded.

Obligations on this category of intermediaries can be limited to the bare minimum necessary to ensure accountability. This could involve identifying a single point of contact for communication with law enforcement, clearly stating terms and conditions of use, and setting up a reasonably responsive grievance redressal system tailored to the nature of service being offered by the enterprise (meaning that the specifics of the grievance process can be left to the intermediary).

B. Intermediaries other than Communication Services (Level 1): All intermediaries excluding communication services and exempted categories

This can be the broadest tier that can include all intermediaries that are involved in hosting, transmission or curation of user-generated information, except those that are primarily designed as communication services. Communication services pose specific risks with regard to spread of information and are, therefore, best categorized separately. This category would then include all services including data centers, cloud services, web-hosting service providers, search engines, online-marketplaces etc.

Broad obligations can be imposed on this tier, including appointment of a grievance officer, reasonable timelines for grievance redressal, cooperation with authorities in case of law enforcement requests, identifying advertised/ copyrighted content, and clearly providing terms of use/ service, privacy policy etc., as the case may be. **However, Level 1 intermediaries must not be mandated to prospectively monitor content, but remove illegal content as and when brought to their notice.**[18] The timeline for content takedown should be reasonable in

[18]The 2022 Amendments to the IT Rules asks intermediaries to make *'reasonable efforts'* to cause a user not to upload certain categories of harmful content. However, enforcing such a provision may prove difficult, as intermediaries may not always be aware of content uploaded to their platforms. It would be better instead to set timelines for intermediaries to take down harmful content as soon as they are made aware of it.

all cases, with a shorter timeline only for very sensitive circumstances such as with respect to child sexual abuse, morphed nude images, etc.

Obligations can also include creating a reporting mechanism through which third parties can flag illegal content, and reporting such content to law enforcement if required. Some of these requirements are already mandated under the IT Rules 2021, and the CERT-In Directions u/s 70B (6) of the IT Act 2000 (dated 28.04.2022).

C. Communication Services

Services that enable communication/ online interaction between users can include a wide range of platforms and products, including e-mail, instant messaging, business/ enterprise communication, gaming platforms etc. But all of these services do not present the same risk of harm. Therefore, we recommend sub-classifying communication services into three categories.

C1. Contained Communication Services (Level 1)

The first sub-category would include all intermediaries that allow online interaction between two or more end-users but do not lend themselves to widespread sharing or amplification of information i.e. they present a low risk of virality.[19]

By adding the qualifier on virality, such a definition will cover all services which: (1) are not designed to facilitate or promote large scale sharing of information (examples: matrimony services, email); (2) are used in closed networks for internal communication within organizations including universities and businesses (examples: Zoho, internal messaging boards in universities etc.); or (3) are smaller in size (such as new services with a relatively contained/ small active user base).[20]

This definition builds on the classification presented in the Australian context, and adds learnings from India's experience with social media regulation. As viral spread of misinformation and illegal content is a key concern, contained communication services can be defined in a way to include only those platforms that present a very low risk of such activity.

India can require service providers to undertake periodic risk assessments like those prescribed under the UK Online Safety Bill to determine if the functionalities offered by their platforms present a risk of harm through viral spread of information. The government can also prescribe some indicative criteria for platforms to undertake such risk assessments while also requiring such assessments to be submitted to it/ made public. The criteria for assessing risk could include size of user base, features offered by the platform, patterns of usage, history of previous incidents, if any etc.

Once a risk assessment has been completed and if the risk of harm is low, a service provider can be categorized as a "contained communication service". Such contained communication

[19]The FAQs released with the IT Rules 2021, defined virality as the tendency of any content to be circulated rapidly and widely from one internet user to another. https://www.meity.gov.in/writereaddata/files/FAQ_Intermediary_Rules_2021.pdf
[20]This will cover intermediaries that are bigger than those excluded by the micro and small enterprises category.

services can then be regulated as general intermediaries, without any additional obligations applying to them.

C2. Open Communication Services (Level 2)

Open Communication Services would include all intermediaries that enable online social interaction between two or more end-users but also pose a risk for viral spread of information.

Given the risks accompanying such services, additional obligations may be imposed on them. These can include obligations to create and deploy automated tools to identify and mitigate sharing of illegal content[21], as well as shorter timelines to address complaints regarding problematic content that may go viral.

Under the 2022 Amendments to the IT Rules, the government has prescribed a requirement to remove content within 72 hours of complaint if it is unlawful or harmful under six prescribed (but, broadly defined) categories. This change is specifically meant to address the issue of virality to ensure that the spread of unlawful or harmful content, including misinformation, is curtailed before it causes significant damage.

But, it is important to note that even on large social media platforms, all complaints may not require expeditious redressal if there is limited risk of virality. Therefore, even while prescribing obligations for quicker redressal of complaints, an attempt should be made to undertake a graded approach to minimize compliance burden on service providers. **Given this, India should consider explicitly defining virality in terms of the width and pace of spread. In the event of a complaint, content that crosses prescribed thresholds could be reviewed on priority.** Such a provision is likely to help reduce compliance costs and also increase the efficacy of the grievance process.

Open Communication Services can also be nudged to employ best practices such as **'virality circuit breakers' to nip the spread of information without actual content removal**.[22] This approach can involve temporary suspension of algorithmic amplification for trending topics, and other such measures that may help improve overall outcomes without actual content removal.

C3. Designated Open Communication Services (Level 3)

Open Communication Services over a certain threshold of active users can be categorized as **Designated Open Communication Services (DOCS),** on the same lines as the significant social media intermediary classification under the current IT Rules framework.

---

[21]Rule 4(4) of the IT Rules already lays down this obligation and asks SSMIs to endeavour to deploy automated tools or other mechanisms to proactively identify content depicting rape, CSAM etc.

[22]Bak-Coleman, J.B., Kennedy, I., Wack, M. *et al.* Combining interventions to reduce the spread of viral misinformation. *Nat Hum Behav* 6, 1372–1380 (2022). https://doi.org/10.1038/s41562-022-01388-6

Such designated services can be subjected to the highest level of obligations, somewhat similar to what is currently being done, but perhaps with a few important changes.

Like under the current IT Rules regime, obligations can include appointment of India based grievance and nodal officers, identification of a physical address in India and periodic compliance reporting. However, to improve trust in the neutrality of the grievance redressal process and to enhance its efficacy, DOCS can also be required to set up a tiered grievance redressal system with appeals to decisions taken by grievance officers lying with **in-house appellate committees that include independent stakeholders from outside** (such as in the case of the Sexual Harassment Act). Such a tiered grievance redressal system with voices from outside the intermediary can help bring in more uniformity in the application of terms of service. **This is also likely to help build more confidence in the grievance process, while reducing the need for a government appointed Grievance Appellate Committee (GAC) to appeal platform decisions**.

Finally, as in the case of *Open Communication Services*, DOCS should also be required to review only those complaints on priority which cross certain **prescribed virality thresholds** (instead of all complaints under the 6 listed categories under the 2022 Amendments to the IT Rules). They should also be encouraged to deploy **'virality circuit breakers'** to nip the spread of illegal/ misinformation without actually having to remove it.

VI. Concluding Remarks

Through this paper, our goal has been to propose a risk-based framework which classifies intermediaries based on functionalities, reach and potential harm. Such a framework could help establish accountability and online safety, while reducing legal obligations for a large number of intermediaries. Additional obligations - while useful - can impose significant economic costs on businesses, and dent their growth potential. Any regulation should, therefore, be a balancing act between costs and benefits. In prescribing the alternative framework, we have tried to propose a classification that will, hopefully, achieve the government's policy goal of creating a safer internet ecosystem while also allowing businesses to do what they do best in terms of service provision. This classification is also likely to be better suited to the context of the thriving Indian startup ecosystem.

For the proposed approach to be effective, metrics for risk assessment and appropriate thresholds would have to be defined and reviewed on a periodic basis in consultation with the industry. Eventually, a risk accreditation market may emerge wherein third parties are engaged to undertake assessments and help classify intermediaries.

The government can also work with the industry to explore alternative tools such as circuit breakers and other algorithmic methods to stem the damage from harmful content. This can help develop new measures that do not involve content takedown, but can nevertheless contain its spread.

In conclusion, we would like to re-emphasise that the suggested classification is a potential alternative route to regulation and we hope this leads to more discussion and deliberation.

www.thequantumhub.com

The Quantum Hub (TQH)
No. 29-30, Lala Lajpat Rai Marg, Lajpat Nagar III
Delhi – 110024

+91-11-4084-5940

office@thequantumhub.com