TAKSHASHILA
INSTITUTION

Governance of DCNs I: Categorisation of Harms

Prateek Waghre, Vishal Ramprasad

TAKSHASHILA WORKING PAPER 2021 - 01

V 1.0, 26 July 2021

Abstract

This paper undertakes a review and categorisation of the harms attributed to Digital Communication Networks (DCNs). In lieu of descriptors such as 'social media', 'Big Tech', etc, the paper defines DCNs as a combination of services, products and companies that enable instant communication at scale via the internet, as well as the societies that adopt them. The harms are classified based on competitive, data, and narrative effects to identify common themes within and across these spheres.

The paper then categorises these harms as potential market failures, social problems, and cognitive biases, with the aim of contextualising them. It also identifies whether the harms are emergent or are a function of the scaling effects of DCNs. Based on this categorisation, the paper concludes that further study of the psychosocial effects of the harms attributed to DCNs are required in the Indian context. It contends that competition policy alone will not address the wide set of social problems and biases identified, especially since it will be utilised in a limited set of jurisdictions, and the current effects of DCNs are seen globally.

I. Introduction

Since 2010, the estimated percentage of the global population with access to the internet has risen from ~30% to ~66%[1]. This period has also coincided with the growth of user-generated content (UGC), as well as the usage of various social media platforms and messaging services. In 2021, the 4.2 billion users of social media worldwide are projected to spend a collective 3.7 trillion hours on such services[2]. India has amongst the largest userbases for many of these social media platforms and messaging services[3].

At current levels of usage and adoption, concerns have been raised about impacts on global collective behaviour due to these rapid changes in the scale and structure of human networks, as well as their impact on the quality of information in circulation, and the role played by algorithms in directing the flow of information[4]. The initial optimism surrounding these platforms and services has since given way to pushback commonly referred to as 'the techlash'[5]. In a narrow framing of the term, it is directed at the "growing power and influence that large technology companies hold"[6]. Various harms have been attributed to social media platforms, the communication networks they enable, and the companies that run them. These include the facilitation of interference in democratic processes, spreading information disorder, increasing polarisation, eroding faith in media and government institutions, abusing market dominance, the exploitation of user data, and so on.

Despite their rapid proliferation over the last decade, there is still significant headroom for these platforms to grow their respective userbases. Approximately 47% of the global population have yet to adopt social media platforms[7]. In India, even with an estimated 795 million internet connections, the average connection density stands at 59 connections per 100 people[8]. This implies that further amplification of the harms attributed to social media platforms remains a strong possibility. Therefore, it is essential to articulate and categorise these perceived harms to inform appropriate policy interventions in the future.

This paper undertakes a review of the harms that have been attributed to Digital Communication Networks, defined and described in Section II. The authors of this paper hold that these networks also hold significant potential for good, which should be harnessed while simultaneously seeking to minimise their harms. We aim to explore these benefits in a subsequent working paper.

In Section II, we define Digital Communication Networks (DCNs) as composite entities consisting of services, products that enable instant communication at scale via the internet, the firms/groups that design/create them, and the networks that adopt them. We also outline the scope of the types of harms considered and criteria for their selection. Section III includes commentary on themes observed in the spheres of competition, data, and narratives, based on our review. In Section IV, we tabulate the harms from our review as market failures or social problems that need to be addressed collectively, i.e. through a combination of policy interventions and behavioural changes. We also identify cognitive biases contributing to these harms, and whether they can be attributed to the emergence of DCNs or their scaling effects.

Section V concludes that competition policy is likely to offer remedies for only a subset of the identified harms, and even that would only be effective in limited geographies. The breadth of social problems and biases identified highlight the need for further studies to understand the psychosocial effects on DCNs in the Indian context.

II. Digital Communication Networks

Descriptors such as 'information and communication technologies' (ICTs)[9], 'social media platforms', 'social media networks', 'social messaging applications', 'Big Tech', 'digital platforms', and so on have all been used to address some or all parts of the technologies that enable communication over the internet. However, these terms do not sufficiently delineate the scope of the networks they aim to describe.

Furthermore, these terms are also used to refer to different applications or corporations, depending on the context and source. 'ICTs' include live/recorded broadcasting technologies, telephony, and networking technologies. 'Digital platforms' encompass e-commerce, various types of aggregators (food-delivery, concierge services, cabs), government service delivery portals, and so on: the term is thus too broad for the purposes of this paper. 'Social media platforms', 'social media networks', and 'social media apps' can exclude closed-messaging services such as Whatsapp, Telegram, Signal, or audio-driven platforms like Clubhouse; 'social messaging apps', on the other hand, can exclude platforms such as TikTok, YouTube, Sharechat, or Twitch, making the term too narrow. The terms 'Big Tech' or 'tech giants' group together a small set of (generally American) multinationals that operate across different kinds of markets, making these labels unsuitable. Such distinctions are further complicated by the overlapping and fluid nature of feature sets on offer, essentially requiring that any collective descriptor constantly hits a moving target.

Therefore, we define Digital Communication Networks (DCN) as composite entities consisting of the following components:

- **Capability:** Internet-based products/services that enable instantaneous low-cost or free communication across geographic, social, and cultural boundaries. This communication may be private (1:1), limited (1:n e.g. messaging groups), or broad (Twitter feeds, Facebook pages, YouTube videos, live streaming), and so on.
- **Operator(s):** Firms/groups that design/operate these products and services.
- **Networks:** The entities/groups/individuals that adopt/use these products and services, and their interactions with each other.

Like social media platforms, two defining aspects of DCNs are low entry costs and reliance on user-generated content[10]. The network component of DCNs can exist across products and services offered by different operators, rather than being centred around the capabilities developed by any single operator. The purpose of this framing is to study them from the perspective of their *effects on societies as a whole*, rather than focussing on specific companies, technologies, sharing mechanisms, user dynamics, and so on— attributes which are constantly evolving. In this paper, we focus on the effects of DCNs in the spheres of competition in markets, data gathering, and narratives in public discourse. To investigate their effects on narratives, we have considered whether DCNs allow for user-generated content and whether the network component utilises them for instantaneous communication, as defined in the 'Capability' component above.

Settling on this definition of DCNs brings inevitable trade-offs. In prioritising combined effects across competition, data and narratives, policy issues such as competition in e-commerce marketplaces/service aggregators, and labour conditions for gig-economy workers are not considered in scope. This creates grey zones, such as in the case of employment policies and working conditions concerning content moderation staff, while content moderation practices, in general, remain within the scope of this paper. This should not be considered a comment on the relative importance of these policy issues. Instead, the existence of such grey zones implies that they have different effects, and may require a different set of responses than those required to address the harms being considered in the context of DCNs.

Table 1: Illustrative examples of DCNs

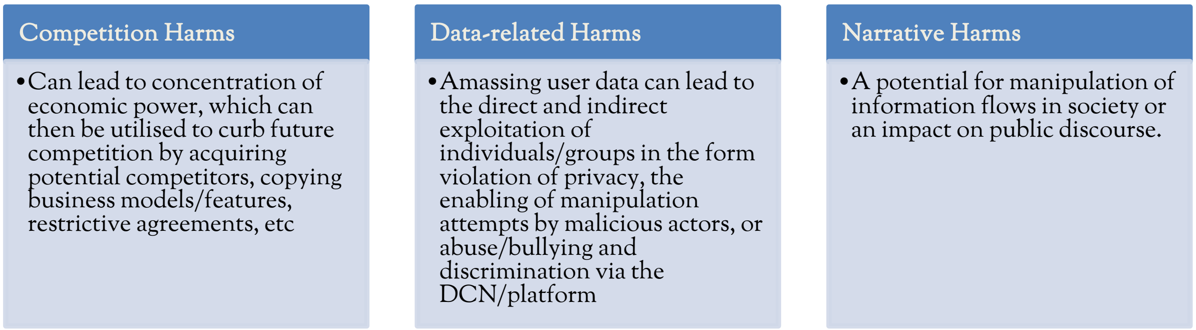
Description	Component of DCN	Reason
Newsfeed/Timeline-centric social media networks: Facebook, Twitter, Sharechat etc.	✅	Competition Effects: ✅ Data Effects: ✅ Narrative Effects: ✅ • User-generated content: ✅ • Used for/enables instantaneous communication at scale: ✅
Messaging services/applications: WhatsApp, Signal, Telegram, etc.	✅	Competition Effects: ✅ Data Effects: ✅ Narrative Effects: ✅ • User-generated content: ✅ • Used for/enables instantaneous communication at scale: ✅
Short-form video sharing: TikTok, Instagram Reels, Moj, Chingari, etc.	✅	Competition Effects: ✅ Data Effects: ✅ Narrative Effects: ✅ • User-generated content: ✅ • Used for/enables instantaneous communication at scale: ✅
Live video broadcast: Facebook Live, YouTube Live, Twitch, etc.	✅	Competition Effects: ✅ Data Effects: ✅ Narrative Effects: ✅ • User-generated content: ✅ • Used for/enables instantaneous communication at scale: ✅
Live audio rooms: Clubhouse, Twitter Spaces, Facebook Live Audio Rooms, etc.	✅	Competition Effects: ✅ Data Effects: ✅ Narrative Effects: ✅ • User-generated content: ✅ • Used for/enables instantaneous communication at scale: ✅
Single/multi brand e-commerce firms/marketplaces:	❌	Competition Effects: ✅ Data Effects: ✅ Narrative Effects: ❌

Amazon, Flipkart, Urban Ladder, etc.		- User-generated content: ✅; comments, reviews, etc - Used for/enables instantaneous communication at scale: ❌
Online curated content platforms: Netflix, Disney Hotstar, etc.	❌	Competition Effects: ✅ Data Effects: ✅ Narrative Effects: ❌ - User-generated content: ✅; comments, reviews, etc - Used for/enables instantaneous communication at scale: ❌
Taxi-aggregator services: Uber, Ola, etc.	❌	Competition Effects: ✅ Data Effects: ✅ Narrative Effects: ❌ - User-generated content: ❌ - Used for/enables instantaneous communication at scale: ❌
Food delivery, concierge services: Zomato, Swiggy, Dunzo, etc.	❌	Competition Effects: ✅ Data Effects: ✅ Narrative Effects: ❌ - User-generated content: ✅; comments, reviews, etc - Used for/enables instantaneous communication at scale: ❌

III. Review of Harms

To enumerate the harms being attributed to DCNs, we relied on a review of select national and multilateral reports and investigations, as well as outputs from academic or civil society initiatives and organisations. While there is an abundance of resources from journalistic and academic sources focusing on specific and individual harms, the sources reviewed aggregated several of these. In keeping with the long-term policy guidance goals of this paper, the harms listed in the sources reviewed are also taken as indicative of relative priorities in public policy–related discourse.

The individual harms were collated and categorised as having competitive, data or narrative effects. This categorisation was not mutually exclusive, since one harm can have effects in one or more of these spheres. The considerations for each of these effects are outlined below:



Competition Harms

- Can lead to concentration of economic power, which can then be utilised to curb future competition by acquiring potential competitors, copying business models/features, restrictive agreements, etc

Data-related Harms

- Amassing user data can lead to the direct and indirect exploitation of individuals/groups in the form violation of privacy, the enabling of manipulation attempts by malicious actors, or abuse/bullying and discrimination via the DCN/platform

Narrative Harms

- A potential for manipulation of information flows in society or an impact on public discourse.

Figure 1: Classification of Harms

1. Competition Harms

Governments across the world are paying close attention to the growing dominance of Digital Communication Networks, and are attempting to understand the implications for competition. In 2019, the European Commission published a report on adapting existing competition laws given the increasing importance and influence of digital platforms[11]. In the same year, a government–instituted Expert Panel on Digital Competition in the UK published its report, recommending policy measures to harness opportunities and address challenges posed by DCNs[12].

The reports argue that Digital Communication Networks are predisposed to become natural monopolies because of features unique to their business models, which centre on network effects, switching costs, and economies of scale, as well as the importance of aggregating user data for a platform to be successful.

The network effect is a phenomenon wherein a platform becomes more appealing to users as more people participate. Platforms like Facebook, Twitter and WhatsApp have a large userbase and enjoy strong network effects. This poses a challenge for new entrants providing similar services, as switching costs are high. For instance, a new platform competing with Facebook will find it difficult to persuade users to switch from Facebook. The platforms thus enjoy demand-side economies of scale, wherein the value of a platform is higher to new users if more people they want to connect with are already on the platform. This was observed during the attempted migration from WhatsApp in India to Telegram and Signal in early 2021 over privacy concerns. However, this did little to dent the 400 million–strong userbase of WhatsApp in the country.[13]

Further, DCNs benefit from extreme returns to scale. While large factories and retailers are also more efficient than their smaller counterparts because of supply-side economies of scale, the marginal cost of serving an additional user on digital networks is virtually nil.

Data collected from their large userbases benefit DCNs through user feedback loops. Intelligence generated from the data allows platforms to understand online behaviour and preferences, which is used to personalise and improve their services. Further, user data is also monetised by targeted advertising revenue streams.

These characteristics have meant that DCNs prioritise growth in the number of users in the short or medium term over profitability. In effect, firms compete *for* the market and not *in* the market. This leads to winner-take-all outcomes, influenced by factors such as first or early entry, access to large pools of capital that finance short-term and medium-term losses in the pursuit of user growth, and regulatory frameworks that lack the tools to deal with such behaviour. Existing competition laws protect consumers' interests from predatory pricing and address monopolistic practices resulting from the abuse of market power. Pricing, however, might not give accurate cues to determine platform dominance, as most DCNs offer their services at no monetary cost to consumers. Users, essentially, pay for these services with their data. Competition laws, therefore, need to adapt to tackle competition considerations posed by the business models of firms that operate DCNs.

The subsections that follow attempt to identify how dominant DCNs affect competition, and defines three ways in which the operation of DCNs can adversely impact other participants. The first is the efficiency with which dominant networks can identify threats and take preemptive measures, which might restrict potential competition and threaten innovation, leading to fewer options for consumers to choose from. The second addresses the implications of the power imbalance between dominant networks and other businesses that depend on them to access a large userbase. The third covers vertical integration, wherein DCNs enter new sectors in which they compete directly with firms that depend on them.

Stifling Innovation and Competition

It is common for DCNs to collect and analyse user data to assess internet behaviour. While these are done largely to personalise their services and to serve targeted advertising, DCNs also inevitably track and collect information about rival or complementary platforms that are fast gaining popularity among users[14]. User data inaccessible to others, is utilised to inform decision-making with respect to acquisitions of nascent competition, or to develop a clone service/product and scale its use across its network. Google is said to have acquired one company every ten days in 2010, while Facebook has acquired 92 startups since 2007, including Instagram and WhatsApp[15].

New entrants find it difficult to compete with large platforms because of investor apprehension. In what is referred to as "innovation kill-zone", the US Sub-Committee on Antitrust reported that investors refrain from funding companies that compete directly with dominant networks, as they are perceived as risky investments[16]. This creates an economic disincentive for new entrants that compete with large firms that operate DCNs.

Dictating Terms of Business

Many businesses depend on dominant networks to access their large userbase. For example, third-party app developers, ad-tech companies, and publishers rely on some of the largest social media platforms to reach a wide consumer base. This absence of competitive options creates a severe power imbalance between firms that operate DCNs and the businesses that rely on them. This might reflect in the form of forced compliance to changes in terms of engagement or a heavy dependence on changes to various algorithms. Another manifestation of this imbalance can be in the form of unequal access to user data, by giving preference to companies that align with the business interests of the firms that operates the DCN. For example, in a lawsuit filed in 2015, Facebook was accused of exclusionary conduct by providing preferential access to user data to Amazon while refusing the same level of access to other companies.

Rivalling Platform Participants

Collecting user data also helps DCNs make inroads into adjacent sectors. DCNs are known to develop services/products that directly rival businesses that operate on their platform. This can lead to self-preferencing practices, where dominant networks rank their rivalling products higher than similar competing products on their platform[17].

2. Data-related Harms

Digital Communication Networks rarely charge end users for their services. Users pay for the platform with their data, which is used to develop behavioural profiles, which then inform the curation of personalised content and monetisation of userbases through targeted advertising. Users usually agree to the privacy policies of platforms while signing up, but rarely understand what data is being collected and how it is being used. The concentration of data in the hands of a few dominant corporate entities, with little clarity of how it is being used, has triggered responses from different governments to protect user privacy. The most well-known is the General Data Protection Rules (GDPR) that came into effect in 2018 in the EU, which is binding on platforms that collect data in EU member countries. The GDPR attempts to put users in control of their own data. Similar policy responses are being contemplated in other countries as well, including India. The Personal Data Protection Bill in India emerged as a response to the Supreme Court judgement declaring privacy a Fundamental Right in K. S Puttaswamy (Retd.) & Anr. v. Union of India & Ors[18]. The Bill seeks to protect the personal data of individuals, and attempts to provide a framework for the processing of such data by corporate entities.

There has been little oversight on data collection practices and the opaque algorithms used to monetise data. This can lead to many undesirable effects on individuals, societies and polities. Since data is intrinsically linked to the business models of DCNs, there is an incentive for dominant platforms to employ different and sometimes unethical methods to gather individual information. It is therefore essential to understand the implications of users surrendering their information to DCNs wittingly or unwittingly. This section identifies issues pertaining to the collection and use of data that need to be considered while developing frameworks for DCNs in the information economy along the themes of privacy, individual agency, the attention economy, and the data value chain.

It is worth mentioning here that intelligent algorithms that analyse large volumes of user data can also be abused by malign actors to influence societies. While some of these categories cover such effects as well, they are addressed in more detail in other sections of this document. This section focuses only on the practices employed by dominant platforms for collecting and utilizing user data — practices which urgently require the attention of policymakers.

Privacy

As discussed in the previous section on competition, dominant firms operating DCNs stifle competition through early acquisitions, introducing clone products, and other tactics. The dependence of their business models on user data and the absence of competitive alternatives means that, in the absence of regulatory pressure, there is little incentive for dominant platforms to develop robust and user-friendly privacy policies. Given the increasing role DCNs play in enabling access to information and networks, consumers are likely to accept policies that erode their privacy rather than give up the platform entirely.

In some cases, data collection and processing have pervasive implications on individual privacy. For instance, Facebook has patents on highly intelligent algorithms that, from users' data, draw inferences about their romantic relationships, gender, economic status, sexuality, inclination towards extroversion, and emotional stability[19]. Users also have little clarity on how their data is being shared.

The concentration of large volumes of data in the hands of a few firms makes them attractive targets for cyber-attacks. Despite claimimg to deploy strict security protocols, DCNs are not immune to breaches and leaks[20]. Further, the availability of demographic or identity information of users of DCNs enables the targeting of certain groups for online harassment. An international survey of female journalists revealed that two-thirds of respondents experienced online abuse in the form of death and rape threats, sexist comments, cyberstalking, account impersonation, and obscene images[21]. Children, people with disabilities, religious minorities, and the LGBTQIA+ community are also more likely to encounter bullying[22]. This might lead to the withdrawal of, or reduced participation of such groups in, digital public spaces.

Individual Agency

The privacy policies of firms operating DCNs do a poor job of explanaining to users what data of theirs is being collected and how, how it is being used, and who it is being shared with. Firms are known to give users little choice to restrict or reject access to their personal data.

Smartphone apps are increasingly being used to access DCN services. In 2016, the standard terms of agreement allowed social media platforms unfettered access to user data, usually collected from smartphones, including private chats and emails, GPS location, IP address, and Wi-Fi points[23]. They also reserved the right to modify, delete and share user data with third parties without specific consent or intimating the user. Users had no option to opt out of these terms. Apps also invariably require permissions to access the data of users' friends (second-degree access) and friends of friends (third-degree access), therefore denying even prudent users the ability to control their own data[24]. However, regulation in different countries, coupled with a general increase in consumer awareness, has compelled corporations to tighten their privacy policies.

Another compromise of individual agency is the perpetration of behavioural nudges using dark patterns. Dark patterns are mechanisms that are employed by digital platforms to manipulate or trick users' decision making to serve their economic interests over consumer interests. Dark patterns capitalise on the System One thinking of the human brain, which triggers spontaneous or impulsive responses from people[25]. Some common dark patterns that compromise individual agency include user interface design (such as associations with colour, giving an exaggerated sense of threat, deceiving framing of sentences, and so on); the presentation of information to reflect a false sense of scarcity, making users opt-out rather than opt-in for services; and making it easy for users to opt-in and extremely difficult to opt-out of services[26].

DCNs employ opaque algorithms to create behavioural profiles and hyper-personalise a content feed that reflects the consumer's preferences. With traditional media, users had the option to choose and consume content. On social media platforms, algorithms display curated content (via feed algorithms) that users might not have sought out voluntarily in order to keep them engaged. Consumers, therefore, lose the choice to accept or reject content, or to control the sequence or organization of content that is provided to them.

The Attention Economy

In a worldwide experiment involving ten countries, students were asked to go an entire day without engaging with social media platforms. It was observed that most students participating in the experiment were unable to go 24 hours without these platforms and used expressions like anxiety, irritability, and insecurity to describe their emotions[27]. While this experiment was limited to social media platforms, most other DCNs also operate in the attention economy, where they provide services to consumers, often for free, in exchange for their data and attention. This data and attention are sold (directly, or via ad exchanges) to companies for targeted advertising.

To win users' scarce time and attention and keep them engaged, DCNs use different methods to make themselves addictive. Hyper-personalization of content curated to consumers' interests is a way to sustain engagement. Another example is compulsion loops, where platforms use a system of unpredictable rewards to trigger the dopamine circuit of the user's brain. Other common design methods such as red dots on app icons, banner notifications, bell sounds, and vibrations are also meant to trigger the release of dopamine.[28]

Data Value Chain

The first three issues have identified individual or societal implications of data–related considerations like collection, analysis and sharing. This issue takes a step back to understand the global implications of the concentration of user data in the hands of a few corporations.

Firms that operate Digital Communication Networks, as defined in this paper, are large corporations mainly from the United States and China that have successfully penetrated other parts of the world, including heavily populated countries of the Global South. The data value chain is therefore inevitably controlled by corporations from very few geographies, and user data is often routed back to these geographies and processed there. For instance, 16 out of the 17 data centres owned by Facebook are in developed economies, despite India being its largest user base[29]. The value generated by the digital economy is captured by a few companies in select countries, while developing economies, where a large part of the data is generated are unable to optimise these value chains in the pursuit of digital industrialisation.

Storing and processing domestic data in other countries raises questions about its ownership and about user privacy, and has implications on national security. During recent geopolitical tensions between India and China, national security implications were invoked to justify banning apps and services from China-based corporations.[30]

3. Narrative Harms

In this paper, the identification of narrative harms caused by DCNs was made on the basis of whether there was potential to impact public discourse or manipulate information flows. We identified five partially overlapping themes across these narrative harms. Examples provided under each theme are indicative, and not exhaustive. Within each theme in the narrative harms category, there is significant intersection with competitive and data harms.

Consequences of the Actions of Firms that Operate DCNs, and the Scaling Effects of DCNs.

It is possible to classify some narrative effects based on whether they are a result of direct actions of the companies/corporations that operate DCNs. For example, the chilling effect on freedom of expression through content takedowns or inaction, resulting in the silencing of vulnerable groups[31]; regulatory capture through political lobbying[32]; boosting or stifling particular viewpoints as a result of design choices, with the potential to exercise power over content dissemination[33]; acceleration of the attention economy[34]; and so on. Conversely, many harms are a result of the scaling effects of DCNs, for example, the virality of mis/disinformation[35]; coordinated attacks and inauthentic behaviour[36]; proliferation of CSAM and terrorist propaganda material[37], among others. This classification centres around the actions of the firms that operate DCNs.

Incentives

DCNs result in attention markets due to an economic imperative to maximise time-spent and to boost engagement metrics[38], touched upon in 'The Attention Economy' subsection in the Data-related Harms section above. These rely on creating a social-validation feedback loop (activating desires for connection, competition, FOMO-avoidance and habit creation), and facilitating digital behaviour contagion[39]. (The data-gathering required for these purposes also implies an intersection with data harms.)

The role of engagement as currency increases the tendency among participants to 'perform' – sensationalism and tribalism are rewarded. Conversely, unwillingness to indulge in the performative aspects of DCNs may also be penalised[40]. This perpetuates a system where the benefits accrue largely to participants near the top. In politics especially, this has manifested as the 'celebritisation' of politicians[41] and the growth of ideological populism[42]. More generally, the combination of such incentives and affordances for users of DCNs can result in the amplification of harmful messages and behaviours[43] (for example, hate speech, fear speech, or dangerous speech) and the unchecked virality of low-quality information[44].

Control and Manipulation of Information

Some of the most visible narrative harms pertain to the control and manipulation of information and information flows. The extensive use of automated, machine-learning-based methods to moderate hate and extremist speech can also lead to the suppression of legitimate voices[45]. Limited transparency can exacerbate such risks. Cooperation with, kowtowing to, or inadvertently arming authoritarian regimes with knowledge power (power gained by accumulating vast amounts of data about people and turning it into knowledge)[46] results in the suppression of dissent. Similarly, coordinated attacks in the form of harassment, targeting, trolling, swarming, or doxxing by financially, ideologically, or politically motivated actors (especially against women, sexual minorities, religious minorities and vulnerable groups) may stifle viewpoints to the point that they are no longer expressed publicly. Political weaponisation of mis/disinformation or other forms of information manipulation through distractions, diversions, or information overload have become easier[47]. Some of these harms rely on tracking and targeting or microtargeting individuals and groups, thereby intersecting with data-related harms[48]. The overlap of data and narrative harms are also a feature of DCNs with newsfeeds/timelines that generate personalised and curated information diets[49].

Trust and Integrity

An adverse impact on levels of trust in information itself is a logical consequence of the the manipulation of information flows through the mechanisms discussed in the previous paragraph. The intersection of these harms with the dominance of individual DCN firms have implications for the production of reliable information, since the news media ecosystem is significantly affected[50] by a reduction in trust. The shift of news consumption to a content aggregation model has also resulted in the atomisation of even high-quality journalistic content, which requires significant investment in order to compete in attention markets with lower quality information[51]. Simultaneously, the ability of the publishers of high-quality news to monetise content has also reduced, since the dynamics of digital advertising markets have eliminated the exclusivity of the relationship between publishers and their readers[52]. Due to these business model disruptions in the news media ecosystem, their incentives have shifted. This has resulted in the adoption of more partisan positions or following partisan coverage (thereby participating in performance politics, discussed in the 'Incentives' subsection above)[53], and an even greater emphasis on speed over accuracy[54]. As a result, overall trust in news media itself appears to have reduced[55]. Combined with an information environment with a low signal-to-noise ratio, incentives to engage in tribalism, and the availability of affordances for inauthentic activities, there is also an impact on institutional trust — especially democratic and scientific institutions. The latter of which is observable during COVID-19[56]. These effects can be both short-term and long-term, and they can be particularly abused by domestic and international adversaries to subvert the integrity of civic or democratic processes.

Social Cohesion

Similar people tend to connect (homophily). Design choices made by firms operating DCNs, which recommend connections based on fewer degrees of separation, can result in denser clusters of similar people (clustering). Clustering, homophily[57], hyper-personalised information diets driven by algorithmic feeds, human biases, and context collapse combine to foster lower levels of empathy[58], greater tribalism and polarisation, radicalisation of public discourse though mis/disinformation, incivility and hate speech[59]. Differences between various groups go beyond considerations of evidence (attitudinal polarisation), into disliking individuals or groups holding an opposing or different opinion/view (affective polarisation), and towards a disagreement on shared realities (knowledge polarisation)[60] — for example, disputing whether COVID-19 is real or a hoax. Performative incentives of digital nationalism[61] lead to escalatory and sometimes competitive cycles of hate speech, fear-mongering, the radicalisation of public discourse, the spread of conspiracy theories, incitement and even physical violence.

IV. Categorisation of Harms

In this section, we attempt the situate the harms identified in our review within broad policy areas. In order to do this, we categorise harms on the basis of potential market failures, social problems, and cognitive biases. Given the absence of regulation in response to many of the attributed harms, 'government inaction' is a potential government failure that we could have identified. However, since regulatory responses are still being framed, we chose not to categorise them on the basis of government failure at this stage. An assessment of the rollout of the various regulatory interventions is planned for a subsequent review.

Market Failures

In Table 2, we categorise harms into potential market failures. A market failure occurs when there is an inefficient distribution of goods and services in the market[62]. We used the following kinds of market failure to categorise harms with competitive effects (Table 2).

- Market power: Market power has been subdivided into gatekeeper power (platforms serving as infrastructure for digital markets), leveraging power (integration across markets enabling the creation of an advantageous position in another) and information exploitation power (gathering data from consumers and business users) [63].

- Information asymmetry: When some participants in the market have more or better information than others.
- Externalities: Situations when the production of goods or services imposes additional costs on others, that are not reflected in the prices of the goods or services produced.

In one instance, we have deviated from the market failure categorisation to identify a potential government failure.

Table 2: Categorisation of harms as potential market failures

Harm(s) / Effect (s) Attributed	Potential Market Failure
Acquiring Market Power	
Acquisition of nascent competition	Market Power - Gatekeeper
Leveraging existing market power to gain market share in other sectors	Market Power - Leveraging
Preserving Market Power	
Economic disincentive to fund products/services directly/indirectly competing with platforms	Negative Externality
Lock-in costs / High switching costs preventing users from migrating to potential competitors	Market Power - Gatekeeper
Political lobbying to influence favourable policies that keep competition in check	– *(Government Failure: Regulatory Capture)*
Persuasive product design (vanity metrics, dopamine rewards through notifications, etc.) to incentivise extended usage.	Information Asymmetry
Tracking/cloning potential competitors/threats	Market Power - Information Exploitation
Aggregator model enabling disintermediation of 'service provider' - 'consumer' relationship	Market Power - Gatekeeper
Leveraging user data in development of new technologies, and improvement of existing ones	Market Power - Information Exploitation
Abusing Market Power	
Exclusionary conduct by forcing businesses to enter restrictive agreements / contracts	Market Power - Gatekeeper
Forcing business users to comply with changes in business practices	Market Power - Gatekeeper
Opacity of transaction flows in digital advertising markets (particularly automated and programmatic)	Information Asymmetry
Rationing access to data for third party entities to favour own business interests	Market Power - Gatekeeper
Self-preferencing own verticals / products while competing with other businesses	Market Power - Leveraging
Uncertainty for businesses whose access to markets can be affected by algorithmic changes	Information Asymmetry

Information Consumption and User Behaviour	
Commodification of attention to support ad-driven business models	Negative Externality
Digitally contagious behaviours powered by automatic curation	Information Asymmetry
Disintermediation of public discourse leading to algorithmic gatekeeping, fragmentation	Negative Externality
Rising levels of mistrust in news media and loss of trustworthy credible news sources leaves a gap filled in with false / misleading and low-quality information	Negative Externality
Reduced monetisation of news diminishing ability to adequately support high-effort reporting	Negative Externality
Privacy-related	
Consumers/users unaware of data collection practices and extent, making it difficult to compare privacy costs	Information Asymmetry
Use of behavioural nudges, dark patterns, and manipulative design interfaces to increase likelihood of users consenting to being tracked	Information Asymmetry
User tracking and micro-targeting at scale	Market Power - Information Exploitation
Weaker privacy protections for consumers combined with the lack of alternatives and high switching costs	Market Power - Gatekeeper

Social Problems and Cognitive Biases

In Table 3, we categorise the harms as social problems and cognitive biases. We also include our assessment a distinction between whether the harm is a function of the scaling effect of DCNs, or an emergent one that can be attributed to actions of corporations/companies that constitute DCNs or the existence of DCNs themselves.

We use the term social problem in the context of “less than ideal situations” affecting many individuals/groups within a society, which need to be addressed through “collective effort or action”[64]. While specific problems such as crime, poverty, and sexual abuse are recognised as social problems, many of the harms we identified do not fall neatly into these categories. Therefore, we defined a set of problems pertaining to DCNs (Table 3): addiction, child abuse, civic integrity, crime, harassment, sexual abuse/harassment, inequity, information integrity, misaligned incentives, privacy impact, radicalisation, social division and suppression. Short descriptions are included in Appendix 1.

Cognitive biases are systematic, unconscious lapses of the human mind that occur while performing mental tasks related to memory (formation, retention, recall), attention (top-down, bottom-up), and other tasks (numerical, and temporal estimates). One of the leading hypotheses proposed to explain the existence of cognitive biases is that they may aid in faster decision-making when the human brain is overwhelmed by information[65]. For this paper, we have chosen a subset relevant to information processing and human social networks (Table 3): anchoring bias, availability cascade, bandwagon effect, blind-spot bias, confirmation bias, conservatism bias, information bias, recency, and stereotyping. Definitions are included in Appendix 1.

Table 3: Categorisation of harms based on potential social problems and cognitive biases

Harm(s) / Effect (s) Attributed	Emerging Harm	Scaled Harm	Social Problem(s)	Cognitive Biases
Actions of DCN firms				
Chilling effect on Freedom of Expression due to inconsistent / incorrect automated and manual content moderation and policy enforcement	✅		Inequity, Social Division, Suppression	-
User tracking and micro-targeting at scale	✅		Privacy Impact	-
Weaker privacy protections for consumers combined with lack of alternatives and high switching costs	✅		Misaligned Incentives, Privacy Impact	-
Impact on mental health, especially minors	✅		Addiction	-
Incentives				
Persuasive product design (vanity metrics, dopamine rewards through notifications, etc.) to incentivise extended usage	✅		Addiction, Misaligned Incentives	-
Consumers / users unaware of data collection practices and extent, making it difficult to compare privacy costs		✅	Misaligned Incentives, Privacy Impact	-
Commodification of attention to support ad-driven business models		✅	Misaligned Incentives	-
Digitally contagious behaviours powered by automatic curation	✅		Misaligned Incentives	-
Use of behavioural nudges, dark patterns, manipulative design interfaces to increase likelihood of users consenting to being tracked.	✅		Misaligned Incentives, Privacy Impact	-
Skewing incentives towards participation / performance		✅	Misaligned Incentives	Anchoring Bias, Conservatism Bias, Stereotyping
Information Control / Manipulation				
Boosting or stifling certain viewpoints based on platform affordances or product design	✅		Information Integrity	Bandwagon Effect,

				Availability Cascade
Virality of low-quality information / misinformation / disinformation		✅	Information Integrity	Anchoring Bias, Confirmation Bias, Recency Bias
Political weaponisation of misinformation / disinformation / malinformation		✅	Civic Integrity, Information Integrity, Misaligned Incentives, Social Division	Bandwagon Effect, Availability Cascade, Recency Bias
Amplifying messages/behaviours by manipulation, coordination, and inauthentic behaviour		✅	Information Integrity, Misaligned Incentives	Bandwagon Effect, Availability Cascade, Recency Bias
Information ecosystem with low Signal-to-Noise ratio		✅	Information Integrity	Information Bias
Intentionally flooding the information ecosystem with low quality information		✅	Information Integrity	Recency Bias
Knowledge power accrued utilised by authoritarian regimes to monitor, suppress dissent		✅	Inequity, Suppression	-
Trust and Integrity				
News consumption shifting to an aggregation model	✅		Information Integrity	Recency Bias
News content 'atomising', appearing alongside other content, removed from its source	✅		Information Integrity	-
Rising levels of mistrust in news media and loss of trustworthy credible news sources, leaving a gap filled in with false / misleading and low-quality information		✅	Information Integrity	Information Bias
Reduced monetisation of news, diminishing ability to adequately support high-effort reporting		✅	Information Integrity	Blind-spot Bias
Interference in elections and civic processes		✅	Civic Integrity, Information Integrity, Social Division	-
Manipulation of public discourse by foreign adversaries		✅	Civic Integrity, Information Integrity, Social Division	-

Social Cohesion				
Presence of terrorism-related propaganda material		✅	Radicalisation, Social Division	-
Accelerating clustering and homophily through design choices, leading to echo chambers		✅	Social Division	Bandwagon Effect, Anchoring Bias, Availability Cascade
Targeted abuse/trolling/swarming/hate speech to suppress viewpoints, harassment especially of minority/protected groups		✅	Crime, Harassment, Social Division, Suppression	Stereotyping, Blind-spot Bias
Targeted abuse/trolling/swarming of public figures		✅	Harassment, Social Division, Suppression	Stereotyping
Increased attitudinal, affective and knowledge polarisation		✅	Social Division	Anchoring Bias, Conservatism Bias, Stereotyping, Bandwagon Effect
Greater incentives for public figures to amplify polarisation / tribalism / sensationalism / incivility		✅	Misaligned Incentives, Radicalisation, Social Division	Anchoring Bias, Conservatism Bias, Stereotyping
Exposure to multiple types of propaganda and conspiracy theories		✅	Information Integrity, Radicalisation, Social Division	-
Algorithmic gatekeeping leading to changes in individual and public opinion formation processes resulting in filter bubbles and echo chambers, radicalisation of public discourse though mis/disinformation, incivility and hate speech	✅		Information Integrity, Social Division	Anchoring Bias, Conservatism Bias, Stereotyping, Bandwagon Effect, Recency Bias
Crime / Abuse / Harassment				
Conduits for sharing and exchange of child sexual abuse material (CSAM)		✅	Child Abuse, Privacy Impact, Sexual Abuse/Harassment	-
Conduits for sharing and exchange of non-consensual sexually explicit material or non-sexually explicit material in a sexualised context		✅	Privacy Impact, Sexual Abuse/Harassment	-

Prevalence of content encouraging self-harm and suicide		✅	Crime, Harassment	-
Doxxing, blackmail, encouraging harassment		✅	Crime, Harassment, Privacy Impact	-
More opportunities for scams and sale of illegal/dangerous goods		✅	Crime	-
Privacy risks due to data leaks, hacks, cyber attacks	✅		Privacy Impact	-

V. Conclusion and Discussion

Digital Communication Networks have had mixed effects on society. The combination of low entry barriers and large amounts of user-generated content have enabled a significant amount of surplus by creating new possibilities for interaction and economic activity. Simultaneously, they have also created new harms and worsened many existing problems. Current global adoption levels and internet connection density in India imply that there remains substantial scope for an even greater impact on the scale and structure of human networks and information integrity. In this paper, we undertook a review of the various harms attributed to DCNs, primarily based on national and multilateral reports and investigations, as well as outputs from academic and civil society initiatives or organisations.

After classifying and categorising the harms in Sections III and IV, specific themes emerged. On the competition front, the harms included adverse effects on innovation; appropriation of user data or cloning of feature-sets; enforcement of unfair contracts; and coercion of business partners. In terms of data-related harms, misaligned incentives between firms that operate DCNs and their users result in actions that are considered to be violative of individual and collective privacy. Many of these harms could be attributed to the existence of DCNs themselves. For narratives, the majority of the harms were amplified by the scaling effects of DCNs. The combination of incentives to 'perform and participate' in attention markets, affordances for control and manipulation of information, result in the lowering of trust in institutions and news media, leading to a negative impact on social cohesion. This resonates with the higher occurrences of cognitive biases associated with tribalism (anchoring, stereotyping, bandwagon), and the recency/availability of information based on our classification. The implications for public discourse and information ecosystems extend beyond the consideration of just the truth and falsity of a particular message. The long-term effects on individual mental health and reasoning ability due to DCN participation require further study if we are to govern DCNs to maximise the benefits while minimising the harms.

The categorisation of attributed harms based on market failures indicated that certain harms (mainly those with competitive effects, and some with data-related effects) can be defined using the vocabulary of market failures. However, many of the data and narrative harms represent a wide range of social problems ranging from crime, addiction, sexual harassment to misaligned incentives, inequity and information integrity, civic integrity and social division that will require a combination of policy interventions, the evolution of new social norms, and behavioural changes to address. Further study on the psychosocial effects of these harms and broader impact in the Indian context is required. Together, these considerations also raise the questions of how, and whether, the antitrust interventions currently being pursued in some of the developed economies will affect DCNs across competition, data and narrative spheres in the rest of the world.

In subsequent work, we plan to identify the benefits that DCNs enable, assess overlaps and contradictions between proposed/enacted DCN governance measures, and explore the role of global internet governance mechanisms with the aim to define appropriate frameworks to govern DCNs.

VI. Acknowledgements

The authors would like to thank Nitin Pai, Mihir Mahajan, Pranay Kotasthane and Anupam Manur for their guidance and feedback. We also want to acknowledge inputs and contributions from Ruturaj Gowaikar for identifying and defining the cognitive biases used, and Suchir Kalra during the research collation phase of this paper.

VII. References

[1] Internet World Stats, 'Internet Growth Statistics 1995 to 2021 - the Global Village Online', Internet World Stats, 2021, https://www.internetworldstats.com/emarketing.htm.

[2] Simon Kemp, 'Digital 2021: The Latest Insights into the "State of Digital"', We Are Social, 27 January 2021, https://wearesocial.com/blog/2021/01/digital-2021-the-latest-insights-into-the-state-of-digital.

[3] PIB Delhi, 'Government Notifies Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021', Press Information Bureau, 25 February 2021, https://pib.gov.in/pib.gov.in/Pressreleaseshare.aspx?PRID=1700749.

[4] Joseph B. Bak-Coleman et al., 'Stewardship of Global Collective Behavior', *Proceedings of the National Academy of Sciences* 118, no. 27 (6 July 2021), https://doi.org/10.1073/pnas.2025764118.

[5] Robert D. Atkinson et al., 'A Policymaker's Guide to the "Techlash"—What It Is and Why It's a Threat to Growth and Progress' (Information Technology and Innovation Foundation, 28 October 2019), https://itif.org/publications/2019/10/28/policymakers-guide-techlash.

[6] Robert D. Atkinson et al., 'A Policymaker's Guide to the "Techlash"—What It Is and Why It's a Threat to Growth and Progress' (Information Technology and Innovation Foundation, 28 October 2019), https://itif.org/publications/2019/10/28/policymakers-guide-techlash.

[7] Kemp, 'Digital 2021'.

[8] Telecom Regulatory Authority of India, 'The Indian Telecom Services Performance Indicators October – December, 2020', 27 April 2021, https://www.trai.gov.in/sites/default/files/QPIR_27042021_0.pdf.

[9] 'Information and Communication Technologies (ICT)', 22 June 2020, http://uis.unesco.org/en/glossary-term/information-and-communication-technologies-ict.

[10] Ekaterina Zhuravskaya, Maria Petrova, and Ruben Enikolopov, 'Political Effects of the Internet and Social Media', *Annual Review of Economics* 12, no. 1 (2 August 2020): 415–38, https://doi.org/10.1146/annurev-economics-081919-050239.

[11] European Commission. Directorate General for Competition., 'Competition Policy for the Digital Era.' (LU: Publications Office, 2019), https://data.europa.eu/doi/10.2763/407537.

[12] Digital Competition Expert Panel, 'Unlocking Digital Competition, Report of the Digital Competition Expert Panel', n.d., https://www.gov.uk/government/publications/unlocking-digital-competition-report-of-the-digital-competition-expert-panel.

[13] Ananya Bhattacharya, 'Will India's Signal Fever Make a Dent in WhatsApp's 400 Million User Base?', Quartz, accessed 16 July 2021, https://qz.com/india/1955262/will-indians-ditch-facebooks-whatsapp-for-signal-telegram/.

[14] Bennett Cyphers and Gennie Gebhart, 'Behind the One-Way Mirror: A Deep Dive Into the Technology of Corporate Surveillance', Electronic Frontier Foundation, 2 December 2019, https://www.eff.org/wp/behind-the-one-way-mirror.

[15] 'A Balancing Act : The Promise and Peril of Big Tech in India', Tandem Research, accessed 23 April 2021, http://tandemresearch.org/publications/a-balancing-act-the-promise-and-peril-of-big-tech-in-india.

[16] Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary, 'Investigation of Competition in Digital Markets', 2020, https://judiciary.house.gov/uploadedfiles/competition_in_digital_markets.pdf.

[17] Sam Shead, 'Google Agrees to Change Global Advertising Practices as France Imposes Unprecedented $268 Million Fine', CNBC, 7 June 2021, https://www.cnbc.com/2021/06/07/google-fined-by-france-for-abusing-online-advertising-position.html.

[18] Software Freedom Law Centre, India, 'Supreme Court Upholds Right to Privacy as a Fundamental Right', SFLC.in, 24 August 2017.

[19] Ronald Deibert, 'The Road to Digital Unfreedom: Three Painful Truths About Social Media', Journal of Democracy, n.d., https://journalofdemocracy.com/articles/the-road-to-digital-unfreedom-three-painful-truths-about-social-media/.

[20] Zack Whittaker, 'A Huge Database of Facebook Users' Phone Numbers Found Online', *TechCrunch* (blog), n.d., https://social.techcrunch.com/2019/09/04/facebook-phone-numbers-exposed/.

[21] Department for Digital, Culture, Media & Sport and Home Office, 'Online Harms White Paper', 15 December 2020, https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper.

[22] Department for Digital, Culture, Media & Sport and Home Office, 'Online Harms White Paper', 15 December 2020, https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper.

[23] Internet Society, 'Issue Paper, Asia-Pacific Bureau: Social Media', November 2017, https://www.internetsociety.org/resources/doc/2017/issue-paper-asia-pacific-bureau-social-media/.

[24] Gabriel J. X. Dance, Nicholas Confessore, and Michael LaForgia, 'Facebook Gave Device Makers Deep Access to Data on Users and Friends', *The New York Times*, 3 June 2018, sec. Technology, https://www.nytimes.com/interactive/2018/06/03/technology/facebook-device-partners-users-friends-data.html, https://www.nytimes.com/interactive/2018/06/03/technology/facebook-device-partners-users-friends-data.html.

[25] Daniel Kahneman, *Thinking, Fast and Slow*, 1st pbk. ed (New York: Farrar, Straus and Giroux, 2013).

[26] Beni Chugh and Pranjal Jain, 'Unpacking Dark Patterns: Understanding Dark Patterns and Their Implications for Consumer Protection in the Digital Economy', *RGNUL Student Research Review Journal* 7 (2021): 23.

[27] Deibert, 'The Road to Digital Unfreedom'.

[28] John Herrman, 'How Tiny Red Dots Took Over Your Life', *The New York Times*, 27 February 2018, sec. Magazine, https://www.nytimes.com/2018/02/27/magazine/red-dots-badge-phones-notification.html.

[29] 'Data Centers', *Facebook Sustainability* (blog), n.d., https://sustainability.fb.com/data-centers/.

[30] Press Information Bureau, 'Government Bans 59 Mobile Apps Which Are Prejudicial to Sovereignty and Integrity of India, Defence of India, Security of State and Public Order', Press Information Bureau, n.d., https://pib.gov.in/pib.gov.in/Pressreleaseshare.aspx?PRID=1635206.

[31] European Commission, 'Impact Assessment of the Digital Services Act', 15 December 2020, https://digital-strategy.ec.europa.eu/en/library/impact-assessment-digital-services-act.

[32] Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary, 'Investigation of Competition in Digital Markets'.

[33] Francis Fukuyama et al., 'Report of the Working Group on Platform Scale', 17 November 2020, https://cyber.fsi.stanford.edu/publication/report-working-group-platform-scale.

[34] Sinan Aral, *The Hype Machine: How Social Media Disrupts Our Elections, Our Economy, and Our Health--and How We Must Adapt*, First edition (New York: Currency, 2020).

[35] Aral.

[36] European Commission, 'Impact Assessment of the Digital Services Act'.

[37] Department for Digital, Culture, Media & Sport and Home Office, 'Online Harms White Paper'.

[38] Michael H. Goldhaber, 'View of The Attention Economy and the Net | First Monday', 7 April 1997, https://firstmonday.org/article/view/519/440.

[39] Aral, *The Hype Machine*.

[40] 'Undress or Fail: Instagram's Algorithm Strong-Arms Users into Showing Skin', *AlgorithmWatch* (blog), n.d., https://algorithmwatch.org/en/instagram-algorithm-nudity.

[41] Mattias Ekman and Andreas Widholm, 'Twitter and the Celebritisation of Politics', *Celebrity Studies* 5, no. 4 (2 October 2014): 518–20, https://doi.org/10.1080/19392397.2014.981038.

[42] Angelos Kissas, 'Performative and Ideological Populism: The Case of Charismatic Leaders on Twitter', *Discourse & Society* 31, no. 3 (1 May 2020): 268–84, https://doi.org/10.1177/0957926519889127.

[43] Joan Donovan and danah boyd, 'Stop the Presses? Moving From Strategic Silence to Strategic Amplification in a Networked Media Ecosystem', *American Behavioral Scientist* 65, no. 2 (1 February 2021): 333–50, https://doi.org/10.1177/0002764219878229.

[44] Soroush Vosoughi, Deb Roy, and Sinan Aral, 'The Spread of True and False News Online', *Science* 359, no. 6380 (9 March 2018): 1146–51, https://doi.org/10.1126/science.aap9559.

[45] Jillian York, *Silicon Values* (Brooklyn: Verso Books, 2021).

[46] Guy Schleffer and Benjamin Miller, 'The Political Effects of Social Media Platforms on Different Regime Types', *Texas National Security Review*, 1 July 2021, https://tnsr.org/2021/07/the-political-effects-of-social-media-platforms-on-different-regime-types/.

[47] Samuel Woolley and Philip N. Howard, eds., *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*, Oxford Studies in Digital Politics (New York, NY, United States of America: Oxford University Press, 2019).

[48] Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Paperback edition (London: Profile Books, 2019).

[49] Amber Sinha, *The Networked Public: How Social Media Is Changing Democracy* (New Delhi: Rupa Publications India Pvt. Ltd, 2019).

[50] Australian Competition and Consumer Commission, 'Digital Platforms Inquiry - Final Report', Text, 25 July 2019, https://www.accc.gov.au/publications/digital-platforms-inquiry-final-report.

[51] Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary, 'Investigation of Competition in Digital Markets'.

[52] Dina Srinivasan, 'Why Google Dominates Advertising Markets: Competition Policy Should Lean on the Principles of Financial Market Regulation', *Stanford Technology Law Review* 24, no. 1 (7 December 2020): 55–175, https://law.stanford.edu/publications/why-google-dominates-advertising-markets/.

[53] Chris J. Vargo and Lei Guo, 'Networks, Big Data, and Intermedia Agenda Setting: An Analysis of Traditional, Partisan, and Emerging Online U.S. News', *Journalism & Mass Communication Quarterly* 94, no. 4 (1 December 2017): 1031–55, https://doi.org/10.1177/1077699016679976.

[54] Sinha, *The Networked Public*.

[55] Reuters Institute for the Study of Journalism, 'Reuters Institute Digital Report 2020', n.d., https://www.digitalnewsreport.org/survey/2020/.

[56] Danielle M. McLaughlin, Jack Mewhirter, and Rebecca Sanders, 'The Belief That Politics Drive Scientific Research & Its Impact on COVID-19 Risk Assessment', ed. Anat Gesser-Edelsburg, *PLOS ONE* 16, no. 4 (21 April 2021): e0249937, https://doi.org/10.1371/journal.pone.0249937.
[57] Eun-Mee Kim and Jennifer Ihm, 'Online News Sharing in the Face of Mixed Audiences: Context Collapse, Homophily, and Types of Social Media', *Journal of Broadcasting & Electronic Media* 64, no. 5 (1 December 2020): 756–76, https://doi.org/10.1080/08838151.2020.1835429.
[58] Lewen Wei and Bingjie Liu, 'Reactions to Others' Misfortune on Social Media: Effects of Homophily and Publicness on Schadenfreude, Empathy, and Perceived Deservingness', *Computers in Human Behavior* 102 (1 January 2020): 1–13, https://doi.org/10.1016/j.chb.2019.08.002.
[59] Dr. Birgit Stark et al., 'Are Algorithms a Threat to Democracy? The Rise of Intermediaries: A Challenge for Public Discourse' (Algorithm Watch, 26 May 2020), https://algorithmwatch.org/en/wp-content/uploads/2020/05/Governing-Platforms-communications-study-Stark-May-2020-AlgorithmWatch.pdf.
[60] Michael Patrick Lynch, 'The Value of Truth', Text, Boston Review, 26 February 2021, http://bostonreview.net/philosophy-religion/michael-patrick-lynch-value-truth.
[61] Sabina Mihelj and César Jiménez-Martínez, 'Digital Nationalism: Understanding the Role of Digital Media in the Rise of "New" Nationalism', *Nations and Nationalism* 27, no. 2 (April 2021): 331–46, https://doi.org/10.1111/nana.12685.
[62] 'Market Failures, Public Goods, and Externalities', Econlib, n.d., https://www.econlib.org/library/Topics/College/marketfailures.html.
[63] Lina M. Khan, 'Sources of Tech Platform Power', *Georgetown Law Technology Review*, 21 July 2018, https://georgetownlawtechreview.org/sources-of-tech-platform-power/GLTR-07-2018/.
[64] Ram Ahuja, *Social Problems in India* (Jaipur: Rawat Publications, 2014).
[65] Gordon Pennycook and David G. Rand, 'The Psychology of Fake News', *Trends in Cognitive Sciences* 25, no. 5 (1 May 2021): 388–402, https://doi.org/10.1016/j.tics.2021.02.007.

VIII. Appendix

Social Problems

These descriptions have been adapted from common definitions of the terms (where available).

- Addiction: Inability to stop engaging in a behaviour even though it causes psychological or physical harm.
- Child Abuse: Act resulting in death, serious physical or emotional harm, sexual abuse or exploitation of children.
- Civic Integrity: Interference with elections or other civic processes to address matters of common concern through public participation[1].
- Crime: Violation of laws in force.
- Harassment: Behaviour intended to trouble, annoy, threaten and cause emotional or mental suffering.
- Inequity: Actions leading to unequal outcomes for different groups.
- Information Integrity: Actions/Behaviours leading to adverse effects on the information ecosystem.
- Misaligned Incentives: Situations where one set of participants benefits from actions that lead to adverse outcomes for multiple individuals/groups.
- Privacy Impact: Actions/Behaviours leading to individual, collective or both types of privacy harms.
- Radicalisation: Process through with groups/individuals adopt or accept radical or extreme positions[2].
- Sexual Abuse/Harassment: Unwanted sexual activity/advances where perpetrators use force or threats against victims.[3]
- Social Division: Divisions in society based on membership in particular social groups in terms of advantages, disadvantages, inequalities and differences.[4]

[1] Twitter Inc., 'Civic Integrity Policy', January 2021, https://help.twitter.com/en/rules-and-policies/election-integrity-policy.

[2] 'Counter-Terrorism Module 2 Key Issues: Radicalization & Violent Extremism', n.d., https://www.unodc.org/e4j/en/terrorism/module-2/key-issues/radicalization-violent-extremism.html.

[3] 'Sexual Abuse', https://www.apa.org, n.d., https://www.apa.org/topics/sexual-assault-harassment.

[4] Jane McCarthy and Rosalind Edwards, 'Social Divisions', in *Key Concepts in Family Studies* (London: SAGE Publications Ltd, 2011), 180–83, https://doi.org/10.4135/9781446250990.

- Suppression: Preventing groups/individuals from acting on or expressing views.

Cognitive Biases

Cognitive biases adapted from 'The Evolution of Cognitive Bias'[5]

- Anchoring Bias: Decision making is tied to (anchored to) a previously or frequently-provided piece of information, making it difficult for a person to act beyond the boundaries of the arbitrary set reference point.
- Availability Cascade: Increasing belief in a concept due to its rising availability in public discourse.
- Bandwagon Effect: The chances of a person believing the veracity of information/belief being presented is more if a greater number of people support it.
- Blind-spot Bias: The inability to self-identify flawed decision making due to cognitive biases of our own.
- Confirmation bias: The tendency to ignore parts of new data that don't fit pre-conceived mental schemas, and using only the data that fits.
- Conservatism Bias: The reluctance to accept new ideas. Similar to an availability bias, but used in the context of general ideas and not discrete bits of information.
- Information Bias: The tendency to think more information leads to better decision making.
- Recency Bias: Tendency to rely more on recent bits of information compared to older information when both sets of information are received with little temporal delay.
- Stereotyping: Attribution of characters to an idea or people as a whole without having access to information of the units forming an idea or people.

[5] Martie G. Haselton, Daniel Nettle, and Damian R. Murray, 'The Evolution of Cognitive Bias', in *The Handbook of Evolutionary Psychology* (American Cancer Society, 2015), 1–20, https://doi.org/10.1002/9781119125563.evpsych241.