LONDON POLITICA

Risks of An Increasingly Digitalised World.

Date: 12th March 2022

Digital risks report covering cryptocurrency and social media-related issues and implications.

Written by: *Lina Gabel, Yann Guillaume, Issy Ronald, Tara Sahgal, Marina Tovar, Marko Cem Zerunyan (Research Analysts)*

Edited By: *Marina Kutumova-Sidwell (Research Director)*

Table of Contents

Executive Summary

Marina Kutumova-Sidwell

The COVID-19 pandemic accelerated the digitalisation of various activities around the world resulting in several changes. Firstly, it led to a sharp decrease of cash circulation thereby paving the way for cashless payments and the use of digital currencies such as cryptocurrencies. Overall, since the start of 2020 cryptocurrency market capitalisation has seen a relatively stable growth reaching the peak value of $3 trillion dollars in September 2021. In addition to the coronavirus pandemic, experts also predict that the war in Ukraine will lead to further growth in the cryptocurrency sector following a whopping total of $106 million in cryptocurrency donations to help alleviate Ukrainian humanitarian crisis.

However, it is important to note that despite such a widespread use of cryptocurrencies around the world, multiple risks and concerns remain, particularly in association with the decentralised nature of these digital currencies. These issues include regulatory problems stemming from market manipulation, procyclicality and high volatility of crypto markets, money laundering activities and fraud schemes developments. Chapter one of this report tackles these matters in an in-depth analysis and offers possible mitigation tactics.

Secondly, the COVID-19 pandemic also resulted in an increased use of social media. It is estimated that as of February 2022, there are approximately 4.2 billion people that are actively engaging in social media activities worldwide making social media one of the most widely used sources of information.

To illustrate, the Reuters Institute reports that both the traditional TV news coverage and newspaper readership are declining, while social media is experiencing a major surge of use as a news source. As this public reliance on social media goes up, journalists have less control over the reported information leading people to be exposed to misleading information or false facts thereby leading to a range of potential risks. These include political manipulation, escalation of violence, intensification of hate speech, and global security risks. The second chapter of this report examines these issues and provides potential risk mitigation strategies.

Chapter One. Cryptocurrency's Role and Place

Market Manipulation

Marko Cem Zerunyan

In 2021, the cryptocurrency market made major headway in terms of market size to assert itself as an established field in the financial sector. Thus, varied reports that up to 90% of trading volume in the cryptocurrency market could be exposed to manipulation is a serious concern. Inevitably, the issue of market manipulation and abuse brings up the question whether the cryptocurrency industry requires further regulation. But with the decentralised nature of cryptocurrency being one of its key appeals, the prospect of regulatory intervention in the cryptocurrency market is controversial. Hence, the steps taken to address crypto market manipulation pose a formidable challenge, especially entering into 2022.

Although it is difficult to gauge the precise scale and extent of price manipulation across the cryptocurrency market, it has been subject to numerous historical criticisms. Indeed, some financial analysts suggest that Bitcoin's bull run to $1000 in 2013 was induced by artificial bot trading directed by the infamous Mt. Gox exchange. Similarly, some have argued that the rally to $20,000 in 2017 was instigated by the issuers of the Tether dollar-pegged stablecoin. It is self-evident that the elimination of market manipulation is paramount for the creation of a reliable and efficient market. This section of the report will mention eight schemes or ways through which the cryptocurrency sector finds itself vulnerable to market abuse, manipulation, and fraud. The elimination of these manipulation schemes is vital for cryptocurrencies' future as a major technology in the world economy.

Surveillance challenges

Before proceeding to the manipulation methods, it serves to contrast transparency in cryptocurrency markets and traditional public markets. Indeed, public markets are currently well-equipped with strictly regulated centralised exchanges and other investigatory tools to counteract manipulative behaviour. In the crypto industry, the broad anonymity of transactions combined with the existence of multiple centralised exchanges like Binance and decentralised exchanges like Uniswap makes it difficult to achieve the same level of surveillance. Another caveat to regard in terms of surveillance is that quantitative measurements on the impact of market manipulation are inevitably speculative. Especially with the widespread increase in cryptocurrency adoption in 2021, it is true that many statistics may not be indicative of the current scale of unjust profiteering through manipulation.

Relevant market players

There are four key market players in the cryptocurrency market: *retail investors, institutional investors, exchanges*, and *cryptocurrency organisations*. Different forms of market manipulation are propagated by each of these players, but it is also the case that they are all uniquely damaged by the different schemes of market manipulation. To specify, retail investors refers to nonprofessional and non-high wealth individual investors; institutional investors refer to professional and incorporated venture capital, private equity, and hedge fund investors; exchanges refer to marketplaces brokering the buying and selling of cryptocurrencies; and cryptocurrency organisations refers to the foundations and companies responsible for engineering and issuing cryptocurrency protocols like Ripple for XRP.

Forms of market manipulation

i. Pump and Dump

Pump and dump refers to schemes whereby particular investor groups consolidate a large amount of cryptocurrency assets and subsequently try to inflate the price in order to sell it off at a high-profit margin. The traditional method for these schemes is the promotion of predominantly misleading or false information to create artificial demand in a digital asset. In 2021, infamous technology entrepreneur John McAfee and his associates were indicted for conspiracy to commit commodities and securities fraud and conspiracy to commit securities and touting fraud among other charges for pump and dump schemes. They had employed his popular Twitter account to support niche cryptocurrencies or advertise initial coin offerings (ICOS) without disclosing where he stood to profit through investment gains and promotional fees. Critics have singled out certain popular Twitter accounts in 2021, including Elon Musk's, for also propagating pump and dump schemes. In regards to Elon Musk, it has been suggested that he has inflated prices of certain 'memecoins' such as Dogecoin through indirect tweets before selling them at high prices and leaving retail investors at major losses.

Another form of pump and dump activity that has emerged in recent years has been 'pump groups', who use social media platforms like Discord, Reddit, Telegram, and Twitter to anonymously arrange coordinated purchases. Through these coordinated purchases, these groups hope to drive up prices and attract further speculative investment from outside investors who also pick up on increasing prices. In August 2018, when cryptocurrency was still far smaller relative to its scale in 2022, it was reported that almost $825 million in trading volume over a period of 6 months was connected to pump and dump trade groups. As cryptocurrency remains largely unregulated, the ability for vulnerable retail investors to seek legal remedies against pump and dump losses remains limited.

ii. Derivatives

A large number of regulators worldwide ban the existence of crypto-derivatives. Derivatives do not entail real ownership of a cryptocurrency but instead involve ownership of a securitized contract whose value is dependent upon an underlying speculation on the future price of a cryptocurrency. In an already speculative and risky cryptocurrency market, derivatives involve

even more speculation, volatility, and high-leverage. Hence, regulators in places like the EU and UK have protected investors against their lack of understanding associated with this risk- and manipulation-exposed product.

In connection with pump and dump schemes, retail investors in the crypto-derivatives market can be hurt in the long term because institutional investors are able to place market-moving bets on future cryptocurrency prices. Because the details of derivative contracts are fully available to the public, it has not been uncommon for pump and dump groups to place major buy orders on the spot market ahead of a contract's settlement date. This pushes up the price of the cryptocurrency and engenders a serious profit for institutional investors on futures positions. Separately, it is also argued that crypto-derivatives almost exclusively serve the interests of the cryptocurrency exchanges, as it allows them to hedge their risk exposures that arise from spot market volatility.

iii. Wash Trading

Wash trading is a technique that creates a false sense of market liquidity and distorts prices by the execution of artificial trading volume. This is particularly difficult to inhibit in the cryptocurrency market as decentralised exchanges are anonymous, meaning malicious traders can easily place sell orders and buy the same order without exposing their identity. One report indicates that 70% of decentralised trade activity is suspicious and probably fake. One aforesaid study proposes that the Mt. Gox exchange had been manipulated by two trading bots producing artificial trading volume.

iv. Insider Trading and Frontrunning

Although insider trading broadly refers to the use of private information to facilitate a knowledge advantage in the market, it is a legal term technically confined to common stock. Thus, it is contested whether bonds, commercial real estate, commodities, and cryptocurrency alike should be subject to insider trading law. Given their open-source and decentralised nature, it is argued that private information in cryptocurrency is largely immaterial and implausible. Nevertheless, the dangers of insider trading have been highlighted and even prosecuted in the US in the *Berk v Coinbase* case. That case involved the executives and employees of Coinbase purchasing Bitcoin Cash prior to announcing on Twitter that it would support listing the asset shortly after repeatedly rejecting it would list the asset. Accordingly, exchanges and cryptocurrency organisations do have considerable control over 'private information' such as listings and impending ICOs, resulting in a similar dynamic to insider trading. In terms of ICOs, although many founders self-impose schemes such as lock-up periods, this does not bind early institutional investors like venture capitalists. These venture capitalists, in fact, usually have direct relationships with the founders and can take advantage of undisclosed information acquired from interactions with these founders.

A similar phenomenon to insider trading is frontrunning, whereby certain market players capitalise upon information asymmetry, essentially obtaining a head start on public information. Cryptocurrency frontrunning mainly entails miners accessing transactional information on the blockchain prior to other participants and using this information to pre-determine the direction of digital asset prices.

v. Spoofing and Quote Stuffing

Spoofing is another scheme involving the distortion of market dynamics by the placement of buy and sell orders without the intention of eventually executing them. This creates artificial buying and selling pressure and is chiefly facilitated through the use of bots and algorithms which automate the placement of these spoof orders at 5-10 second intervals. The analogous method of quote stuffing involves a steadier form of market pressure by placing a vast number of cancelled high and low orders, engendering the creation of an artificial average price.

vi. Distributed Denial of Service (DDoS) and Trading Freezes

A DDoS attack involves attempts to disable a website or network through an overload of server requests. These attacks can force a blockchain into becoming inoperable for the duration of the attack, completely halting transactions and market action. These attacks thus damage crypto-exchange performance and can even temporarily render them unavailable, subjecting them to freeze trading. A trading freeze can be exploited in both directions of the market in that it can inhibit mass selloffs or mass purchasing. During the crypto crash in May 2021, many exchanges were 'forced' into freezing trading temporarily. It has been suggested that these freezes were not actually DDoS situations but instead consolidated efforts by exchanges to contain and manipulate the price of cryptocurrencies. The nascent cryptocurrency Solana experienced DDoS attacks in late 2021 leaving the protocol offline for about 17 hours on one occassion.

vii. Stablecoins

The manipulative capability of stablecoins is disputed, but it is suggested that stablecoin groups like Tether make 'grants' to print new coins and make purchases of cryptocurrency when their prices are falling to create a false sense of market stability. Many studies have rejected the existence of manipulation by stablecoin 'printing', but it is a concern that stablecoin organisations may be printing new tokens without having the reserves to back each of these coins.

viii. Centralised Consolidation

As noted above, institutional investors including venture capitalists and hedge funds have benefits such as insider information and the ability to take advantage of complex market instruments like crypto derivatives. With the significant level of capital disposable to institutional investors, it is possible that they can increase their stakes in particular cryptocurrencies to the extent that they can single handedly move valuations. Recently, in December 2021, the consolidation of institutional investors in cryptocurrency and more specifically Web3 projects sparked controversy with Jack Dorsey claiming, "You don't own Web3. The VCs and their LPs do. It will never escape their incentives. It's ultimately a centralised entity with a different label. Know what you're getting into." Thus, it is feared that retail investors will fall victim to institutional centralisation of cryptocurrencies.

Possible mitigation

Conclusion

Although progressive economies such as the EU and Canada have accepted the issuance of ground-breaking crypto ETNs and ETFs, manipulation concerns have been at the heart of the SEC and FCA's rejections. Moreover, the volatility of the crypto-derivative market sparks fear of possible risk over-leveraging as was recently seen in the Archegos scandal. Considering the regulatory response in previous instances of manipulation like the LIBOR scandal in 2008, the Forex market from 2008 to 2013, and the Gold-fixing scandal in 2004, cryptocurrency supporters will seek to tackle the issue of market manipulation as effectively and quickly as possible to ensure the credibility of the market.

High Leverage Leading to Procyclicality and High Volatility

Yann Guillaume

On January 24th 2022, the value of Bitcoin fell below $34 000, which was the lowest point it had reached since July 2021. This massive drop that had been ongoing since November 2021, when the Bitcoin had skyrocketed to an all-time high superior to $68 000 after several months of upward fluctuations, practically eliminated all of the gains that had been made in the Bitcoin's price in 2021. The value of various other crypto assets, including Ethereum, followed a similar trajectory. This slump in prices was intensified after the Federal Reserve's Chairman Jerome Powell announced that he will increase interest rates in March in order to reduce inflation, which prompted many investors to attempt to remove risk from their portfolios.

Those impressive fluctuations are symptomatic of the procyclicality and very high volatility that characterise crypto markets. The latter phenomena are largely due to the high leverage that is recurring in decentralised finance's (DeFi) lending and trading platforms, and constitute a serious risk for investors, especially as the size of crypto markets is set to pursue its expansion in 2022.

Nature of the risk

One of the most prominent characteristics of DeFi is the high leverage that tends to emerge from the new forms of lending and trading platforms it proposes. Although lending in DeFi tends to be over collateralised, it is by no means a guarantee of stability for crypto markets since borrowers are allowed to re-use the funds they have borrowed in one instance in the form of collateral in other transactions. This possibility incentivises investments during risk-on periods when investors' optimism about the prospects for the economy is high as it permits them to increase their exposure for a given amount of collateral. Similarly, the trading of derivatives on decentralised exchanges (DEXs) induces leverage for the payments agreed as part of automated market-maker (AMM) protocols are set to be made in the future. Furthermore, the highest margin allowed in DEXs is superior to that permitted in exchanges encompassed in the traditional financial system, and the leverage permitted in crypto centralised exchanges (CEXs) is even greater.

This high leverage in crypto markets substantially heightens the risk of procyclicality. Leverage allows a higher number of assets to be acquired for a given amount of initial capital invested. However, this situation becomes problematic during risk-off periods, when investors become risk-averse because the economic outlook is perceived as being uncertain. In fact, in such periods investors are keen to remove risky investments from their portfolio by fear of incurring losses. In crypto markets, this trend translates into investors seeking to reduce their debt, which leads them to dispose of their crypto assets, often as a result of already existing investment losses and depreciating collateral values. This sherd movement in turn intensifies the downward pressure on collateral prices, which causes margins to rise and sustains procyclicality in crypto markets.

Potential implications

Thus, a serious risk for financial stability in 2022 would be a brutal and complete shift to a risk-off environment, as investors' reluctance to make and keep risky investments would induce a continued collapse of cryptocurrencies' values. In addition to cause substantial investment losses, those spirals of downward prices could spread to the rest of the financial system given the importance of the cryptocurrency market' size. As a matter of fact, the major stable coins in circulation cumulated a value of approximately $120 billion by late 2021, and the overall market value of cryptocurrencies was roughly $2.3 trillion in December 2021 after having grown by $1.5 trillion during the course of the year. This appreciation is expected to endure in 2022, which therefore increases the potential impact that a crash of collateral prices could have on the traditional financial system as linkages between the latter and DeFi will probably increase.

Possible mitigation

The possibilities of mitigation for this risk are inherently limited because of the very nature of DeFi. In fact, financial intermediation on this platform solely relies on private backstops in the form of collateral to facilitate transactions and mitigate risks, which means that it does not include shock absorbers to protect investors during risk-off periods. Meanwhile, the traditional financial system includes such safety nets since banks possess the ability to extend their balance sheets through the issuance of bank deposits, which relies on their access to central bank balance sheets. In concrete terms, it means that banks can extend loans or acquire depreciating assets during risk-off periods in order to bring stability and reassure investors, which tends to attenuate the risk of procyclicality and volatility.

The decentralisation of DeFi substantially complicates public authorities' task to enforce policies aiming to provide similar protections for investment made in crypto markets. But, a possible venue for regulators is to gain access to the groups of stakeholders who adopt the main decisions related to the management of those new platforms. Accessing the latter groups could allow policymakers implement the regulatory safeguards ensuring that DeFi participants internalise the adverse consequences tied to the procyclicality and volatility caused by high leverage.

Money Laundering and Fraud Schemes

Marina Tovar

Criminals and threat actors laundered approximately $8.6 billion of cryptocurrency in the year 2021, increasing up to 30% compared to 2020. In the same line, researchers estimated criminals gathered $14 billion in cryptocurrencies in 2021. Those who used cryptocurrencies for Anti-Money Laundering (AML) very likely range from cybercriminals, ransomware groups, human traffickers and malware operators to terrorist groups. Because cryptocurrencies operate in cyberspace, the risk is occuring everywhere. Therefore, all governments should prioritise combating AML and developing effective anti-money laundering practises to combat the use of cryptocurrencies for this purpose. Threat actors take advantage of the poor AML and fraud practises the crypto exchange market is characterised by.

Cryptocurrency allows for a high level of anonymity and decentralisation of assets for groups, making it very likely that more organisations will continue to further migrate to cryptocurrency platforms, to diversify their assets, enable money laundering, and expand their available financing methods. The risk is persistent and appealing to criminals, but not all criminals will consider cryptocurrencies as an appealing option as the crypto exchange market has a high fluctuation and a difficult conversion to currencies in accepted banking systems.

Cryptocurrencies can allow for the creation of fraud schemes, one of the most cryptocurrency-related crimes, with fraudsters raking in over $2.6 billion in 2020 alone. Trafficking is another potential risk derived from the use of cryptocurrencies as it facilitates the sale of hacking tools, drugs and stolen data, among some. Cryptocurrencies and their use for terrorist financing is also a potential risk to gather unanimity and diversity the sources of their funds and to buy supplies like weapons. Cyber extortion is a potential threat as ransomware and malware groups are more frequently demanding the ransom payment in cryptocurrencies, where Bitcoin accounts for nearly 98% of ransomware payments. The use of cryptocurrencies for AML purposes directly threatens the political and economic stability of a country. All of the benefits derived from fraud schemes, trafficking or cyber extortion can be laundered with cryptocurrencies.

The use of cryptocurrencies for AML purposes will likely continue to be used by cybercriminals to expand their profits and lure from the anonymity cryptocurrencies provide. Law enforcement should use advanced blockchain analytics solutions to fight threat actors who launder money. Furthermore, risk assessments and due diligence investigations would help tracking and detecting the origins and possible threat actors.

Is regulation the answer?

Policymakers on the elaboration of legislation and public policies to tackle the challenges cryptocurrencies pose might encounter internal challenges. The first challenge that arises lies on the multiplicity of actors involved in the process of policymaking and their diverging interests. A clear example is the Russian one, where the Central Bank advocates for the prohibition of cryptocurrency mining while Putin aims to take the contrary direction.

Therefore, the complex institutional complex that shapes policy processes is one of the most relevant internal challenges. To overcome that, an increase in awareness is needed and a framework where all actors are involved and actively heard will be essential to provide a common, joint, and accepted policy proposal. Following that, in the process of policy-making questions like "what is the optimal policy choice" or "is a band potentially more fruitful than regulation" will be present and diverge the opinions of the actors involved.

The regulation of cryptocurrencies is also likely to experience external challenges, like the cross-border and decentralised nature of the cryptocurrency market. Due to the cross-border nature of cryptocurrency networks, the question of who or which institutions will be in charge of overseeing the cryptocurrency market and the infrastructures associated that interact with crypto assets in payments and other related activities. Furthermore, the decentralised nature of the crypto market can lead to the risk of fragmented solutions and inconsistent interpretive guidance. A lack of common agreement between international stakeholders could end up in the creation of fragmented legislation that would deeply harm consumers and investors in the long run. To combat that, financial institutions and potential stakeholders should understand national and local regulatory considerations when establishing a business or offering crypto-enabled services, like know-your-customer (KYC) licensing requirements or the possible anti-money laundering (AML) obligations.

Chapter Two. The Rise Of Social Media and Its Risks

Regulation of Social Media

Lina Gabel

The rapid rise of social media over the last decade has provided individuals with new opportunities to access information, express opinions and participate in democratic processes. However, social media platforms have become increasingly scrutinised and criticised for a few key reasons including, the blatant disregard for people's rights to data protection and privacy, spread of false information, aiding in the process of political manipulation and undermining freedom of expression. This risk has existed for some time, however, it has reached unprecedented heights with rapid technology developments.

Several social media aspects, such as surveillance, personalisation, and disinformation, combine to create a web of interrelated political risks on social media. Today's largest platforms provide efficient means of monitoring people's online presence, which can be used by governments to target politically active citizens and in turn silence opposition. The immense collection of data by social media and technology companies creates privacy risks for users which may affect their ability to form and express original opinions leading to a loss of privacy and autonomy. The subsequent attention capture model, enabled through the collection of vast amounts of data, seeks to exploit human biases to increase engagement but simultaneously undermines personal freedom and autonomy. Furthermore, the promotion of personalised content on social media may lock citizens in informational bubbles thus contributing to the narrowing of world views. Additionally, the expansion of false information on social media can distort the views and preferences of individuals, which can in turn be used to undermine the integrity of elections.

The real-world implications of increasing social media risks have been underway and openly observed for several years. For example, in the lead up to the 2016 U.S. general election, between 110 and 130 million adult American population saw fake news. Russia's plans to influence the U.S. election began in April 2014 with the creation of troll farms that could spread false messages on social media. The Russians studied political groups in the US and developed a large network of fake accounts allowing them to post and spread divisive content on key issues such as black lives matter, immigration laws and gun control. Experts find that this was incredibly successful and that Russian influence most likely swayed the outcome of the 2016 election.

China-linked networks of social media bots and trolls first appeared on the global disinformation radar in 2019. Recently, several in-depth investigations shed light on the expansion of Chinese disinformation campaigns which indicate that significant amounts of

both human and financial resources are being devoted to the increasing disinformation effort. Additionally, the overall sophistication and impact have been elevated, with linkages between fake accounts and official government accounts growing more evident thus limiting deniability opportunities by the Chinese government. The reports also point to several instances in which journalists and traditional media outlets in different countries have unknowingly shared disinformation on their own social media accounts, news websites or television broadcasts. This enhances the credibility of the content and ensures that it reaches a far wider audience.

The increase of risks associated with widening influence of social media platforms demands carefully planned regulation, especially as the harms associated with moderation of social media have further implications for free speech and democracy. For example, efforts by social media platforms to tackle disinformation and deception may very well threaten individuals' freedom of expression and enable forms of political censorship. What if, following attempts at stricter regulation regarding data collection and disinformation, social media transforms from an open and public space in which to exchange ideas and opinions into a harsh and guard space where only a select few are able to express their views? Who will be prioritised and perhaps more importantly who chooses which groups to prioritise?

In recent months, social media networks have embarked on large scale clean ups. For example, Facebook's removal of hate speech has increased tenfold in the past two years. It disables over 17 million fake accounts every day which is more than twice the number two years ago. YouTube removed 11.4 million videos and 2.1 billion user comments in less than three months. With regards to false information surrounding the Covid-19 pandemic, YouTube identified over 200,000 either 'dangerous or misleading' videos. These social media platforms set out to be neutral platforms on which users provide the content thus keeping the companies off any editorial decisions. However, as the platforms implement more advanced algorithms that rank content and simultaneously moderate undesirable uploads, they veer further away from free-flowing ideas, instead moving towards a more curated and edited selection of content.

The imagery of Russian troops invading Ukraine, with frequent missile strikes raining down across both Ukrainian military infrastructure and civilian residential areas in several regional capitals has once again turned the world's attention towards Russian military tactics. This brings back memories of the Russian annexation of Crimea in 2014 during which Russia used disinformation as a tactic to sow confusion into the overarching conflict strategy. They created a campaign to sway and gather the support of ethnic Russians residing in Crimea. State media and social media accounts linked to Russia have started to spread allegations that the West have been manipulating protests and promoting tales of crimes committed by Ukrainian soldiers.

This time around, in 2022, the US states that Russia is using new disinformation campaigns to portray Ukrainian leaders as aggressors and to persuade both Ukrainian and Russian citizens to support military action. European Union officials have reported that Russian outlets have increasingly promoted content that justifies conflict in a similar manner to 2014. Disinformation experts state that they have observed a forceful effort from Russian leaders and state sponsored media to spread a false narrative around the reasons for the invasion of Ukraine.

Furthermore, experts expect this effort to multiply in size as both international and domestic resistance towards war grows. Following every major attack in Ukraine there is a flood of new propaganda and disinformation, which can take several different forms. It can try to take certain videos out of context thus claiming them to be something else, for example showing Russian attacks to be more powerful than they actually were. It can also help to build the illusion of a Ukraine that is not standing strong and fighting back when in reality it most certainly is.

Thus, with democracy at stake, it is apparent that increasing regulation is fundamental to mitigate the harmful and global consequences of disinformation campaigns. However, signs are emerging that the Kremlin is aiming at a full monopoly on how Russians view the invasion of Ukraine by censoring independent media outlets and by banning social media platforms such as Facebook and Twitter. Even though regulation of social media is successfully implemented in large parts of the world, disinformation will continue to polarise world views since several countries, such as Russia and China, who are increasingly creating their own domestic online bubbles will still be able to fabricate and spread exceptionally harmful disinformation within their own spheres of influence.

The United Kingdom has moved closer to large-scale regulation of social media, when a parliamentary committee recommended major changes to the country's online safety bill with the aim of holding internet services providers accountable for the material published on their platforms. Furthermore, the European commission has proposed two legislative initiatives to upgrade rules regarding digital services in the EU with the goal of creating a safer digital space in which the fundamental rights of all users are protected and to establish a level playing field in order to foster innovation and competitiveness. As expanded regulation looks set to arrive in the near future, it will be crucial to ensure that the approach is balanced and mitigates any risks of infringements on individuals' freedom of speech rights.

Impact Of Social Media On Conflict

Issy Ronald

The risks inherent in the widespread use of social media intensify when within a conflict setting, in both intrastate and interstate conflicts. Social media's tendency to amplify hate speech and disseminate disinformation carries with it many risks such as escalating violence and providing cover for an authoritarian crackdown.

In Myanmar, Facebook's algorithms amplified hate speech, potentially contributing to the genocide as years of deliberate disinformation by Buddhist nationalists helping to build resentment against the Muslim Rohingyas. Similarly, in Iraq, the ISIS media network has disseminated online videos of Shia or Kurdish militias carrying out human rights abuses against Sunni communities. Both these examples suggest that countries with existing ethnic and sectarian tensions are more vulnerable to these types of disinformation campaigns on social media that can intensify, or trigger conflict; a pattern that is likely to continue in 2022.

This disinformation does not remain in the online world but penetrates offline communities too. In countries where conflict is either simmering or has already broken out, long-standing suspicions of the central government foster an appetite for alternative sources of information. Reports from Ethiopia, for example, describe conspiracy theories developed online as circulating among SMS chains.

Already, Facebook is taking a more active approach towards mitigating these risks. After the coup in Myanmar, it banned accounts of the military junta, directly contrasting with its refusal to regulate hate speech during the genocidal violence against Myanmar's Rohingya population. Facebook whistle-blower Frances Haugen has alleged that 87% of the spending on combatting disinformation at Facebook is spent on English content, while only 9% of users are English speakers. This suggests that non-English speaking countries are at greater risk from the spread of disinformation.

Yet, the removal of disinformation to prevent conflict is itself fraught with risk. Initiatives currently arise from social media networks themselves or national governments, limiting the actors able to mitigate this risk. Often, national governments have adopted punitive approaches criminalising disinformation which are used to censor the media or arrest journalists and opposition activists. The Iraqi parliament, for example, has introduced a cybercrime law delineating "harming the reputation of the country online" as a crime which carries a life sentence.

To reduce these risks and their platform's volatility, in March 2021 Facebook adopted a human rights policy following the UN Guiding Principles on Business and Human Rights. Facebook's Guiding Principles on Business and Human Rights. Under this new policy, Facebook must publish an annual report on its impact on human rights and establish a fund for human rights defenders, but the platform's current community standards, privacy policies and code of conduct will remain unchanged.

In the early 2010s, when social media was still in its infancy, Western commentators praised its democratic potential. While this early optimism seemed largely misplaced, social media can still facilitate grassroots movement, potentially creating instability. In Myanmar, social media has played a crucial role in creating and sustaining the Civil Disobedience Movement (CDM) that opposes the newly formed military junta. Social media allows the distribution of resources such as what to do if someone gets arrested, how to access VPNs, and crowdfunding options. In this way, social media can facilitate bottom-up activism, as well as top-down repression.

Social media can also be used to disseminate information to an international audience, particularly in conflicts where there is an information vacuum due to restricted access and internet blackouts. In the Tigray conflict, for example, the government and opposition forces have both sought to present its own version of events to English-speaking audiences, using social media to do so. Stand with Tigray, one of the most prominent pro-Tigrayan groups, has accumulated more 36,000 followers on Twitter, drawing attention to the humanitarian crises in the region. Meanwhile, pro-government groups such as Ethiopia Current Issues Fact Check (ECIFC) used official-sounding directives and statements that often condemned international coverage. These conflicting campaigns hinder the capture of accurate, impartial information, and so create further uncertainty within conflict settings. This gives rise to a greater risk for humanitarian and international organisations seeking to distribute aid and facilitate conflict resolution.

External actors must also be wary of using social media to engage in intrastate conflicts for it is increasingly clear that even well-meaning global social media campaigns can interfere with conflict dynamics. In Nigeria, for example, the #BringBackOurGirls campaign hindered rescue attempts, and many have encouraged Boko Haram's growing reliance on gender violence and kidnapping for international attention.

In light of the ongoing war in Ukraine, this ability to exercise greater control over social media networks has become a valuable tool to police the dissemination of information. Indeed, ten days into the conflict, Russia completely blocked access to Facebook and greatly restricted access to Twitter within its borders, in tandem with its efforts to control more traditional media sources too.

Simultaneously, Ukrainians' use of social media has shaped much of the war's coverage. Videos posted on TikTok, pictures on Instagram and first-hand reports on Twitter have driven much of the global outrage that has prompted drastic changes in the foreign policies of Germany, Sweden, Finland and Switzerland, among others. President Zelensky, in many ways, encapsulates this, with his addresses to the nation often going viral on social media, helping to mobilise global support for his cause.

Morality vs Security: Social Media Algorithms & Terror Networks

Tara Sahgal

The 21st century has witnessed an unprecedented rise in the use of Information and Communication Technologies (ICT) across the globe. At the end of 2019, the International Telecommunications Union estimated that 4 billion people – slightly over 51% of the total global population – were using the internet, of whom 2.89 billion were also active on Facebook (20 times higher than in 2008). This increased digitisation has brought with it immense potential as well as immense risk, with the most significant being the manipulation of social media platforms and their algorithms by extremists. This is especially important because of the global reach of terrorist groups (both through internet-enabled informal networks and on-ground sleeper cells), which creates global rather than country-specific security risks.

Studies have found that terrorist groups actively use social media platforms to gather intelligence and recruit users from around the world. This is exacerbated by the large amounts of radicalistic material – including literature, audio tapes and video clips – that have become increasingly accessible to the global youth. A simple Facebook search can draw up hundreds of pieces of propaganda, often surpassing the platform's airtight moderation algorithm: the Fuouaris Upload Network on Facebook, uncovered by Institute of Strategic Dialogue (ISD) researchers in early 2020, is a prime example of this.

ISD's three-month long investigation found 288 accounts that supported the Islamic State of Iraq and the Levant (ISIS) within the same network, named Fuouaris Upload (coming from the Arabic term *Furūsiyya* to invoke visions of knightsmanship in the virtual world). Many of these accounts were controlled by the same individuals aiming to flood the internet with propaganda, with most being 'stolen' from other users and referred to as *mughtanim*, or war spoils. One of the primary accounts went by the name Luqmen Ben Tachafin and was discovered to be the centre of the network, controlling several other subsidiary accounts under the same name and a third of all the accounts within the network. All accounts featured the same statement: 'Luqmen Ben Tachafin. I shake your throne and destroy your dreams. Never tired, never bored, until the Judgement Day.' A total of 50 pieces of terrorist material were circulated across these accounts, including videos of beheadings, audiotapes from *Al-Bayan*, the Islamic State's radio channel, speeches by ideologues and clips from ISIS operations in the Middle East. Interestingly, these pages also spread a Covid-19 narrative celebrating the virus and resultant *kuffar* death toll (or 'kill count'), highlighting their ability to adapt to and exploit adverse global circumstances.

Overall, the Fuouaris Upload case study points to major security risks stemming from social media algorithms. It is nowhere near a lone case: the live-tweeting of the 2013 terror attack by Al Shabaab in Kenya, ISIS' use of social media to recruit and radicalise students in Sudan, and Al Qaeda in the Arabian Peninsula's (AQAP) English-language digital magazine, which inspired the 2013 Boston terror attack, are all instances of terrorist organisations exploiting social media platforms for their benefit. Usually, this successful exploitation is the result of

unique variations employed by these groups when uploading propaganda onto social media, which enables them to amass thousands of views and followers before being flagged by authorities. These include coordinated raids on popular pages, such as the United States Department of Defence; hijacking popular hashtags, such as #BlackLivesMatter; masking content by adding innocent clips from documentaries and films at the beginning of propaganda videos; and gaming text analysis through the use of broken text.

These methods not only surpass Facebook's algorithms but are also at times *reinforced* by them – which poses the most imminent risk for countries trying to control radicalisation without infringing on individual rights. For example, a key part of Facebook's algorithmic process is the creation of filter bubbles and auto-generated pages: it is common knowledge that the platform is known for its heavy personalisation, achieved by micro-targeting the interests and interactions of its users. While this may be seen as a positive from an individual point of view, personalisation to this extent can create a filter bubble where users are only exposed to similar beliefs and ideas, thereby reinforcing existing biases and polarising discourse. So, while Facebook may be committed to de-radicalising its platform, its own algorithms are responsible for creating spaces that enable users to repeatedly view radical content, masked by the methods mentioned previously, before they can be flagged. Similarly, the platform has also been criticised for its tendency to auto-generate pages tailored to users' interests. This feature can easily be exploited by extremists, who can list specific terrorist organisations under their occupation, confusing the algorithm into creating interest pages for their 'business' – as was the case in 2019, when it was found that Facebook unintentionally produced almost 200 pages for ISIS and Al-Qaeda.

The issues posed by auto-generation, filter bubbles and the varied methods of surpassing moderation are furthered by the lack of transparency surrounding the algorithms used by major social media platforms. For example, how do we know what content gets removed from Facebook? How does it get removed, and when? What tools are employed? These questions are crucial to understanding and thereby mitigating cyber terror risk. Unfortunately, their answers are largely inaccessible to the public due to the opacity of virtual social structures.

It is thus clear that there are significant international security risks associated with opaque and sovereign social media platforms; however, the mitigation of these risks has proven to be difficult. The key obstacle is formed by the blurry line between state interference and national and international security. When is it acceptable for a national government to access users' details or activity history? Should social media platforms be held accountable for content published on them? How can international security be preserved without infringing on individual rights?

Public discourse across nations needs to focus on these questions in the digital age in order to effectively curb security risks in the long run. In the short run, perhaps the best way forward is to identify individual attributes and demographics that point to an increased vulnerability to radicalisation (such as teenagers and those with socio-economic struggles) and facilitate national efforts to confront this in countries where it is feasible.

Secondly, states can continue to work with social media platforms on a case-by-case basis to monitor propaganda on already established terrorist channels, although the ethical concerns surrounding this need to be acknowledged and addressed. Finally, states can potentially

collaborate with other parts of the private sector, as was previously highlighted by the US Government in a 2016 report. This involves the amplification of credible voices through a continuous online presence of anti-terror groups (formed by private sector companies) in order to challenge radicalistic material with facts and alternative interpretations. That being said, while these efforts can help alleviate security risks in the short run, long-run mitigation remains the biggest concern and should be the priority for all countries.


LONDON POLITICA