DIGITAL TRANSFORMATIONS FOR HEALTH LAB

GOVERNING HEALTH FUTURES 2030

Navigating the digital playground

Youth health amid tech promises and practices

Erza Selmani

This report has been made possible by financial contribution from Fondation Botnar, Switzerland, to the Project IMG-22-005 Digital Transformations of Health Lab (DTH-Lab). Their commitment to advancing global health is deeply appreciated, and this project would not have been possible without their contribution. DTH-Lab is hosted by Université de Genève (UNIGE), Switzerland. DTH-Lab is committed to ensuring and enabling Global Access. The knowledge and information gained from the Project will be promptly and broadly disseminated and its 'Funded Developments' will be made available and accessible free of costs. The Global Access Commitments will survive the term of the Project.

Suggested citation: Selmani. E. (2025). Navigating the digital playground: Youth health amid tech promises and practices. Geneva: Digital Transformations for Health Lab.

Date of publication: December 2025

Contents

Abbreviations

AI	Artificial Intelligence
COPPA	Children's Online Privacy Protection Act
DTH-Lab	Digital Transformations for Health Lab
EU	European Union
FOMO	Fear of Missing Out
LGBTQ+	Lesbian, Gay, Bisexual, Transgender, Queer (or Questioning)
NGO	Non-Governmental Organization
UK	United Kingdom
WHO	World Health Organization

Executive summary

Youth interaction with digital technologies has evolved: from casual interactions to an all-inclusive digital way of life, present every day, from early on. This has brought countless opportunities for networking, education and self-expression, but at the same time has increased mental health related risks. Vulnerable groups, such as children from low-income households, LGBTQ+ youth and those from rural areas, are often exposed to harmful content and face hate speech, cyberbullying and algorithm enforced discrimination. The narratives promoted by the tech industry often highlight their commitment to user safety and mental health via the reduction of harmful content features and promotion of positive engagement. Nevertheless, evidence suggests that these measures are usually overshadowed by internal practices that prioritize engagement and corporate benefit over user well-being. Algorithms, for instance, are developed to maximize screen time, often by amplifying the content that triggers emotional responses, such as anxiety or insecurity.

This report examines the gap between public claims, made by social media platforms, on protecting young people online and the internal knowledge and reality of their practices. While social media platforms publicly claim they are continuously rolling out new mechanisms to protect young users through safety settings, well-being tools and content moderation, internal knowledge shows that financial revenue and increased user engagement continue to be the main drivers in internal decision-making. In practice, social media platforms' design prioritizes revenue generation over public health, resulting in harmful content amplification, insufficient content moderation and delayed response to known risks.

> *Putting public health at the core of digital policy is no longer an option, it is a duty. By replacing the emphasis on financial revenue with a focus on public interest, the global community can build a digital future that protects, empowers and respects youth.*

Furthermore, social media platforms often shift the burden of online safety to the users and their families, relying on self-regulation, parental control or reporting mechanisms as the primary safeguards. However, research tells us that this approach, putting responsibility on the user, is insufficient and has continuously failed to mitigate mental health harms related to overuse or harmful use of social media. Whistle-blower testimony, including that of Frances Haugen, former product manager at Facebook, reveals that companies like Meta are fully aware of the harm their platforms can cause to young users. Meta's own internal research has shown that Instagram worsens body perception for one in three girls, a finding that is validated by scientific literature which correlates excessive use of social media with decreased mental health and well-being. However, the measures undertaken to

combat this harm have been minimal and mainly motivated by public, regulatory or legal pressure and not by a true dedication to youth well-being and mental health.

The analysis draws on a combination of sources, including public communications and social media platforms' declarations, internal leaked documents, whistle-blower evidence, investigative reports and scientific literature. The latter provides scientific evidence on the documented impact of social media use on youth mental health, including increased anxiety, depression, negative body perception and exposure to health disinformation and cyberbullying.

This report describes a state of affairs which calls out for urgent and coordinated action. International examples including the EU's Digital Services Act, UK's Online Safety Act and the new Australian minimum age law, demonstrate a growing global consensus: we need stronger and mandatory regulations to protect young users. Public health agencies, healthcare professionals, educators and researchers should demand transparency, mandatory accountability frameworks, algorithm audits and regular public reports on the harm caused by digital technology. In addition, they need to promote digital literacy in schools, train healthcare professionals on online behaviour and expand access to confidential mental health services for adolescents.

Putting public health at the core of digital policy is no longer an option, it is a duty. By replacing the emphasis on financial revenue with a focus on public interest, the global community can build a digital future that protects, empowers and respects youth.

Section 1

Introduction

Over the past ten years, there has been a significant change in the interaction between youth and digital technologies (Bohnert & Gracia, 2021). The early social media platforms, including Facebook, have evolved, creating an all-encompassing algorithm-based digital existence which young people experience daily, for extended periods, starting at a young age. The COVID-19 pandemic created an environment where digital spaces became the main location for people to connect with others and learn, while developing their identities (Delogu et al., 2025). The digital transformation has delivered positive aspects, but it has simultaneously increased the amount of dangerous content and behaviours which can negatively affect mental health.

From the launch of the first social media platform to now, young people have gained new opportunities in connecting, networking, communicating and interacting with peers, families and communities all over the world. At the same time, these platforms have transformed into a self-expression method, where young people from all over the world are able to express their skills and talents, whether in arts, sports or gaming (Khalaf et al., n.d.). Different platforms, for instance YouTube, have provided the opportunity for students to take additional informal classes, simplified teaching content and tutorials. In certain situations, social media has offered a safe space for vulnerable groups such as LGBTQ+ youth (Prothero, 2024).

However, not all young users share the same experience with these platforms. Vulnerable groups of young people such as children from low-income households, LGBTQ+ youth and young people living in areas impacted by violence or insecurity, frequently face increased risks online. These include lacking digital literacy or social support, hate speech, cyberbullying and discrimination that is reinforced by algorithms. Existing disparities in offline

life are mirrored and often heightened by inequities in the digital world (Tumakuru & S, 2023; Khalaf et al., n.d.; Weir, 2023).

According to an article featured by the American Psychological Association, during adolescence the brain region associated with desire for attention, feedback and reinforcement becomes more sensitive, while the brain regions involved in self-control haven't fully matured. Features such as "like" buttons, notifications, unlimited videos and content take advantage of the brain immaturity stage and sensitivity, making young users over-engage, seek attention, feedback and validity (Weir, 2023).

This dual nature of social media is reshaping the developmental process of children and adolescents. While these platforms often promote the narratives of connection, empowerment and inclusivity, extensive use is correlated with an increased risk of anxiety, depression and sleep disorder (Office of the Surgeon General, 2023). This duality emphasizes the complexity of the relationship between technology and youth health, challenging simplified narratives that refer to digital technologies only in terms of them being benefits or risk factors (Office of the Surgeon General, 2023).

The narratives promoted by the tech industry often highlight its commitment to user safety and mental health. For example, Instagram and TikTok have presented features that aim to reduce harmful content and promote positive engagement (Magoven Creative Studio, 2024; Presser, 2025). Nevertheless, studies show that these measures are usually overshadowed by internal practices that prioritize engagement and benefit over user well-being. Algorithms, for instance, are developed to maximize screen time, often by amplifying content that triggers emotional responses, such as anxiety or insecurity. This model of "attention economy" is correlated with negative mental health results, including an increase in anxiety, depression and body image issues in adolescents. Furthermore, the increase of disinformation, toxic language and AI-generated content has further complicated the digital landscape (Costello et al., 2023).

1.1 Framing digital behaviour: From determinants to personal responsibility

Recent debates are framing the challenges related to social media through the lens of digital determinants of health (DDoH), which so far has been defined by World Health Organization (WHO) and the Organisation for Economic Co-operation and Development (OECD) as technological factors that shape health systems and healthcare, especially the divides in digital connectivity and digital literacy that limit the expansion of digital health solutions. However, it is argued that a more holistic interpretation of DDoH should include all characteristics, features, uses and governance of digital technologies and ecosystems that influence health and well-being (Holly et al., 2025). For young people growing up in digital environments, these determinants are exceptionally influential, shaping not only access to information and services but also their identities, social connections and mental well-being.

Despite growing awareness of the digital harms that young people can experience, a substantial amount of the public discourse is placing the responsibility for mitigating digital harms on the individual. Most of the online safety initiatives launched by social media platforms focus on limiting screen time, creating healthy digital habits or parental control, framing online harms as a result of bad choices or lack of self-control. While individual behaviour is extremely important, the way these platforms are built is a deeper force that should not be ignored. If we place the focus on screen time, we risk overlooking that social media platforms are designed to grab and hold users' attention due to commercial motivations. This perspective aligns with WHO's definition of health promotion, according to which, health promotion goes beyond individual efforts to broader social and environmental efforts. In this context, effective health promotion requires addressing structural and commercial drivers instead of just encouraging individuals to change their habits (WHO, n.d.).

1.2. Building safer digital spaces: International and national actions

Internationally, countries and organizations are taking action to address the negative impacts of digital determinants of health and well-being for children. In response to the trend, the WHO is working with Member States to address the needs of young people, placing a particular emphasis on digital determinants of mental health and well-being. The main focus remains on prioritizing young people's digital well-being, while applying evidence-based intersectoral public health strategies, developing clear guidance on tech use for health, holding industry and commercial interests accountable, reframing the regulative landscape, promoting healthy tech use and increasing research efforts to have a better understanding of the impact that social media has on young people's mental health (WHO, 2025).

The European Union obliges social media platforms to reduce any design feature that develops addiction, creates harmful content, enables unwanted contact or youth data processing, with fines for violation going up to millions of euros (The EU's Digital Services Act, 2022). An age-verification app is being tested across France, Greece, Italy, Spain and Denmark (Foo, 2025), further evidencing this growing concern. The United Kingdom imposes a legal obligation on digital platforms to protect children from unlawful and harmful content, under the Online Safety Act, 2023, implemented by Ofcom. Failure to comply can result in fines up to £18 million or 10 per cent of overall annual profit. NGOs and other organizations advocating for children's well-being highlight that these provisions are a necessary response to the limitations of self-regulation (Online Safety Act, 2023). In Australia, the Online Safety Amendment (Social Media Minimum Age) Bill 2024 prohibits social media platforms from allowing accounts for users under the age of 16. Fines will be enforced from December 2025 (Online Safety Amendment Bill (Australia), 2025).

These examples of regional and national policy responses reflect a global consensus that self-regulation has failed and we now have the chance to demand accountability and ensure that digital environments are built with young people's best interests at their core.

This report contributes to DTH-Lab's work on the digital determinants of health, particularly in understanding how tech industry practices impact youth health and well-being in digital environments. It lays the groundwork for acknowledging the importance of putting youth health and well-being at the centre of digital transformations and governance, while allowing young people to have an active role in shaping their digital futures.

As such, this report does not aim to demonize social media platforms, but rather to explore the dynamics between narratives and shed light on the gaps between what these platforms claim publicly and what actually happens internally.

Accordingly, this study aims to answer the following questions:

1. What strategies do social media platforms publicly promote for protecting youth health and well-being?
2. What do internal knowledge and whistle-blower reports suggest is known by social media platforms about their impact on youth health and well-being?
3. What are the discrepancies (if any) between their public narratives and internal knowledge?
4. What impact does this discrepancy have on youth health and well-being?

Section 2

Methods

This report draws on a mixed body of publicly available evidence to explore the gap between what major social media platforms (Facebook, Instagram, TikTok, Snapchat and YouTube) claim to do to protect young users and what internal practices they follow, as revealed through whistle-blower disclosures and investigative reporting.

Data sources include:

- Official statements, news and policy documents published by Meta, TikTok, Snapchat and YouTube
- Testimonies and leaks from high profile whistle-blowers
- Legal filings and lawsuits related to youth safety on social media platforms
- Peer-reviewed scientific studies and systematic reviews examining the health impact of social media on adolescents
- Reports from health organizations and NGOs

Sources were selected based on relevance to youth health outcomes and credibility, with an emphasis on materials that directly reveal internal platform knowledge or contradictions between public claims and internal decisions. The analysis was conducted thematically, focusing on four major platforms – Meta, TikTok, Snapchat and YouTube – and organizing the findings around their public narratives versus documented internal practices. The selection of these platforms was informed by previous research that revealed these platforms have the biggest young user base, they perform high data collection, their privacy policies remain ambiguous, they carry addiction risks, they demonstrate inconsistent age verification and poor security measures (Selmani, 2025).

This method enables a triangulated view of youth safety in digital environments, combining corporate messaging, insider disclosures, and external health evidence to better understand systemic risks and accountability gaps.

Section 3

Findings

3.1. Age restrictions and enforcement

Social networking sites frequently advertise age restrictions and parental control features as a critical aspect of preventing harm against young people. Meta platforms (Facebook, Instagram), TikTok, Snapchat and YouTube claim that the minimum age to join their platforms is 13 years old, with a basic age gate where users can self-declare their birthday. Meta and TikTok use AI tools to identify and delete accounts under the age of thirteen and other users can report underage accounts too. YouTube provides YouTube Kids for users aged under-13 and has introduced AI tools to estimate users' age, which is currently being tested in the US (Selmani, 2025; Duffy, 2025).

Meta (Facebook and Instagram), Snapchat and TikTok provide Family Center and Family Pairing features. All platforms provide "winddown" or "take a break" mechanisms. Instagram launched Instagram Teens Account in 2024, with default privacy settings, quiet times, break reminders and parental control (Meta, 2024). In 2025, Meta added new restrictions to Instagram Teens, by disabling users' ability to broadcast live or receive direct messages with unwanted content – now these features can only be enabled with parental consent. In April 2024, Meta introduced Teen Accounts to Facebook and Messenger, which is being rolled out in the US, UK, Australia and Canada (Meta, 2025a). In addition, Facebook and Instagram provide youth-specific features such as abuse detection as well as educational resources and, on Instagram, hidden like counts (Meta, 2025b). Recently, Meta introduced further mechanisms for youth protection online, such as more information on accounts, easier blocking and reporting, automatic strict safety settings for young users and action against inappropriate accounts (Kitson, 2025). YouTube has a special platform for users under the age of 13, called YouTube Kids, where parents

are fully in charge of selecting content. For users between 13 and 17 years of age, there are specific accounts called Supervised Experience, with limited options in content and commenting (YouTube, n.d. - a). As of 2025, YouTube has started to implement AI to estimate users' ages. If the model flags an underage account, it automatically personalizes adverts, turns on digital well-being tools, reminders to take breaks and installs safeguards (Duffy, 2025).

But insider witnesses, as well as independent reporting, indicate that such controls are mostly flimsy. Francis Haugen, a former product manager at Facebook, has alleged publicly that Meta knew that millions of children under 13 were on Instagram but did not take concrete steps to address this (Waterson & Milmo, 2021). While Instagram's Teen Accounts initiative has been welcomed by parents, critics say that it does not address all potential harms faced by teens on Instagram (Chedraoui, 2024). According to researchers, only eight out of forty seven tested mechanisms are effective for youth protection (Horwitz, 2025a). In addition, Fairplay for Kids (2025) found out that two-thirds of Meta's safety mechanisms for teens are ineffective, breaking their promises to parents for youth safety in their platforms and leaving teens vulnerable to online harm and abuse.

TikTok has also consistently failed to enforce its age restrictions, leaving youth accounts to slip through its filters and was fined by the UK data watchdog in 2023 (Vivarelli, 2023). In Snap's case, though it said it is secure, several lawsuits have alleged drug dealers and sexual predators have been able to connect with underage users on the platform (Miller, 2024). YouTube Kids has also been reported to contain videos and content that are not child-friendly. In addition, YouTube Kids can also be addictive for children due to its design and auto-play feature (Nwoko, 2025).

Independent research supports this concern. Children aged 10-12 are participating actively in social media platforms and have reported having higher levels of social comparison, screen addiction and lower emotional self-regulation (Tedjawidjaja & Christanti, 2022). According to Denmark's Digital Minister, online age verification is a must and should be considered the same as age verification for buying alcohol or entering a nightclub (The Local, 2025).

Table 1. Age restrictions and enforcement

Platform	Minimum age	Protection mechanisms	Real-world limitations
Facebook	13+	Age gate, Family Center for parental supervision, Teen Account, messaging limits, screen time and wellness tools, AI and abuse detection, educational resources	Weak verification, widespread underage use, ineffective enforcement
Instagram	13+	Age gate, Teen Account, AI age detection, Family Center for parental supervision, messaging limits, screen time and wellness tools, AI and abuse detection, hidden like counts, educational resources	Weak verification, widespread underage use, ineffective enforcement
TikTok	13+	Age gate, Family Pairing, screen time limits, wind down reminders at night	Easily bypassed controls, minimal ID verification, enforcement depends on parents
Snapchat	13+	Family Center, monitoring of contacts and interactions	No ID verification, ongoing lawsuits over safety failures, limited accountability
YouTube / YouTube Kids	13+ (YouTube), under 13 (YouTube Kids)	YouTube Kids app, supervised accounts, non-personalized ads, restricted features	Autoplay risks, harmful comment sections, anonymous or misreported age usage

3.2. AI and algorithms

Social media platforms regularly promote a positive narrative of their impact in community building, mental health support and creating safe spaces for self-expression. Instagram's Wellbeing Guides (Instagram, 2020) and TikTok's collaboration with mental health groups (Grace, 2024), are two examples of platforms initiating mental health awareness initiatives. However, this narrative becomes overshadowed by the algorithmic design which prioritizes engagement over well-being. Meta uses AI-based personalization for users' feeds and introduced the Meta AI assistant in 2023, which gained momentum after Meta launched a standalone Meta AI app (Robinson, 2025). TikTok, in addition to personalized content, provides infinite scrolling (Gilbert et al., 2025). Snapchat uses an internal scoring system and AI to recommend new friends via their quick add feature, but also claims to prioritize privacy and not recommend content to the public (van Essen & Van Ouytsel, 2023; Snap Inc, n.d.). According to their claims, YouTube uses a recommendation engine for personalized content and only exposes users to age-appropriate content (Southern, 2025). These sites claim that their algorithms are designed to provide users positive, personalized content: Meta has emphasized the use of AI to downrank harmful content (Meta, 2021), TikTok claimed that algorithms step in when users have been shown too much of a certain type of content (TikTok, n.d. – a, c) and YouTube stresses the importance of responsible AI innovation (O'Connor et al., 2023).

Internal documents and employee testimony tell a very different story. Frances Haugen discovered that Instagram served body image and weight loss content to teenage girls so aggressively that it significantly undermined their self-esteem (Clayton, 2021). A recent internal document leak from Meta, reviewed by Reuters, shows Meta's standards which allow their chatbots to engage in romantic and sexual interactions with children, generate false information and racist content. The document, titled GenAI: Content Risk Standards, was approved by the legal and ethics teams at Meta, defining what behaviour was acceptable by Meta AI and chatbots integrated in Facebook, WhatsApp and Instagram. After being questioned by Reuters, Meta removed some sections of the document and confirmed that it was wrong, claiming it did not align with internal company policies. However, the document shows that Meta allowed their chatbots to generate false, sexual and racist content, raising an important ethical and legal dilemma in AI utilization (Horwitz, 2025b).

TikTok's own internal document, published in 2023, concluded that even a single engagement with a video is sufficient to plunge a user into content on self-harm or eating disorders (5Rights, 2024). The site has also been accused of harmful algorithmic recommendations that direct young users towards inappropriate content. A report by the Molly Rose Foundation (2025) found that despite the new regulations, TikTok and Instagram still recommend massive volumes of harmful content to teens, concluding that both platforms remain unsafe by design, prioritizing engagement over well-being and leaving teens exposed to preventable risks. YouTube Kids has been reported many times by parents for having masked content with imitations of children's favourite animated characters, who promote inappropriate and harmful behaviour (BBC News,

2017; Bridle, 2018). A report by Statistics Canada in 2022, revealed that 22 per cent of adolescents were exposed to harmful online content, including hate speech and self-harm material, due to algorithmic amplification (Government of Canada, 2024).

Scientific studies confirm the risk. According to a study published in 2023 algorithms promoting content on personal appearance, violence and depressing themes have a direct impact on the rise of anxiety and depressive symptoms among youths who consume this content (Costello et al., 2023). Another study published in 2022 found that "Snapchat streaks" encourage problematic smartphone use and fear of missing out (FOMO) and the younger users are when they experience FOMO using Snapchat, the more likely they are to build and maintain streaks (van Essen & Van Ouytsel, 2022).

Table 2. AI and algorithms

Platform	Algorithmic feature	Official claim	Real-world limitations
Facebook	Feed ranking, AI-based personalization	Promotes positive content; uses AI to detect harmful material	AI chatbot promoting false, sexual and racist content to children
Instagram	Feed ranking, AI-based personalization	Promotes positive content; uses AI to detect harmful material	Internal documents show Instagram's algorithm amplifies body image issues; teens pushed into harmful content loops AI chatbot promoting false, sexual and racist content to children
TikTok	For You feed, Infinite scroll, Quick add	Personalized content to match user interests and safety for teens	Algorithm quickly pushes self-harm, eating disorders and suicide-related content to teens; creates "filter bubbles" and addiction
Snapchat	Streaks, Snap Score, Quick Add, AI chatbot	Prioritizes privacy, does not recommend public content by default	Gamified features drive compulsive use; Snap Score and streaks encourage unhealthy attachment and validation-seeking
YouTube	Recommendation engine	Promotes helpful, educational and age-appropriate content	Studies show amplification of emotionally negative or harmful videos (e.g. eating disorders, rage-bait) and masked content

3.3. Data security and privacy

Where privacy is concerned, the platforms make a show of standing up for the data privacy of minors. Meta says it restricts advertising to users under-18 and provides transparency tools (Meta, 2023). TikTok says in its Children's Privacy Policy that it gathers little data from children and follows privacy laws worldwide. However, internal evidence points to a far more aggressive strategy. Former Meta CEO Sarah Wynn-Williams has disclosed that Facebook and Instagram can determine when teens are feeling vulnerable or lonely, allowing the companies to tailor advertising accordingly. Internal documents explain how the platforms steer users towards user-generated or commercial content, depending on their mood (Bellens, 2025). TikTok has also been sued for its use of facial and voice recognition, even for underage accounts (BBC News, 2021). In 2021, TikTok was fined by the Dutch Data Protection Authority for violating children's privacy (European Data Protection Board, 2021). In 2023, TikTok was fined by the Irish regulators for the same reason (Gerken & McMahon, 2023). YouTube was also fined by the Federal Trade Commission and The New York Attorney General for violating children's online privacy (Federal Trade Commission, 2019). While Snapchat is clear that it does not allow targeted advertising for young users under the age of 13, it definitely uses targeted ads for audiences aged 13 and over. Snapchat uses its young audience as a selling point for attracting business (Snapchat For Business, n.d. - a). Recently, Snapchat has introduced Sponsored Snaps, delivered to users' chat inboxes (Snapchat for Business, n.d. - b). In 2019, YouTube was fined $170 million for violating children's privacy laws by the Federal Trade Commission (FTC). Commissioner Rohit Chopra argued that YouTube has baited children with videos featuring nursery rhymes and cartoons (BBC News, 2019).

Table 3. Data security and privacy

Platform	Data practices	Official claim	Revealed limitations
Facebook	Data collection and cross-app tracking	No targeted ads to users under-18; privacy-first approach	Whistle-blower: emotional data shared with advertisers to exploit teen insecurities
Instagram	Data collecting and cross-app tracking	No targeted ads to users under 18; privacy-first approach	
TikTok	Activity profiling, collects device/ behaviour data	Minimal data collection from minors; COPPA* compliant	Fined in 2019; continues tracking app-wide activity and user behaviour without transparency
Snapchat	Standard data collection, less public reporting	No direct targeted ads; private messaging default	Data collection practices are unclear
YouTube	Personal data collection, behaviour tracking	Compliant with COPPA*; non-personal ads on "kids" content	Fined $170 million; trackers, cookies still active, adverts interspersed into "made for kids" videos

*COPPA: Children's Online Privacy Protection Act

3.4. Harmful content reporting and accountability

Social media platforms often promote their robust content moderation systems as a safeguard against harmful content. In 2021, Meta set up internal teams specifically to address problematic content, following a three-step approach based around a desire to "remove, reduce and inform" (Meta, n.d.). In 2025, Meta announced the introduction of "community notes", where commenting on the accuracy of posts is left to users – initially rolled out in the United States of America with plans to expand across the world (McMahon, 2025). Snapchat promotes integrated reporting tools (automated systems and human moderators) to enforce community guidelines (Snapchat, 2025), similar to TikTok which in addition partners with mental health organizations (Grace, 2024). YouTube also uses AI-driven moderation tools in combination with human reviewers, claiming that their content moderation is not reactive but rather pro-active (Halprin et al., 2022).

However, the enforcement of these moderation policies is frequently inconsistent, leading to unlimited harmful content. Behind the scenes, there is a massive lack of response, according to insiders. Reports of sexual harassment, bullying and other inappropriate content shared by teenagers were frequently ignored, or not taken seriously, said Arturo Béjar, a former Facebook executive (Kerr, 2023). With the introduction of Community Notes, Meta is jettisoning independent fact-checkers in the U.S. and planning to do so globally, which can be considered

a reduced content moderation effort, as detailed in their corporate publication How fact-checking works (Meta, 2025c). Snapchat has been sued by families in cases alleging it was not doing enough to stop drug sales and predatory access to underage users (Kindelan, 2024). TikTok has also faced repeated complaints about inadequate moderation of harmful content, which has been left online for long stretches (Murphy Kelly, 2022). In 2025, YouTube loosened rules guiding content moderation, promoting "freedom of expression" (Grant & Francisco, 2025).

A study by the United States of America Surgeon General's Office (2023) shows that 64 per cent of adolescents have reported engagement with harmful content. The results of the same study show that there is a significant relationship between social media use and body image concerns and eating disorders. About 75 per cent of young people have reported that social media platforms are doing a fair to poor job at treating reports of bullying or harassment (Office of the Surgeon General, 2023). An experimental investigation by Amnesty International in 2023 found that the "For You" page on TikTok risks pushing young users towards harmful content (Amnesty International, 2023). In 2025, Amnesty asked TikTok whether they had implemented any new mechanisms since the investigation and TikTok listed the same well-being measures which do not address the "rabbit hole" problem (Amnesty International, 2025). Meta platforms have been accused of deliberately not taking action in response to the reports of young users, despite alarming findings from an internal study (Hatmaker, 2021; Common Sense Media, 2025).

Social media platforms remain non-transparent about their metrics and user engagement time with harmful content, despite legislative efforts in the U.S. and Europe (Lu et al., 2024; Cho & Zhu, 2025). A study found that social media platforms avoid dealing with how adolescents use their platforms, instead seeming to pretend they are not there and hiding inconvenient data in large reports (Owens & Lenhart, 2023). This approach, along with the lack of transparency, affects the possibilities for evidence-based policies and interventions that could address the negative effects of social media on adolescents.

Table 4. Harmful content reporting and accountability

Platform	Reported mechanisms	Official position	Real-world outcome / limitations
Facebook	AI and human moderation, fact-checking programmes in other countries (non-U.S.), Community Notes in U.S.	Safety-first transparent enforcement	Internal leaks show reports ignored, engagement prioritized over safety, reduced content moderation efforts
Instagram	AI and human moderation, fact-checking programmes in other countries (non U.S.), Community Notes in the U.S.	Safety-first, transparent enforcement	Internal leaks show reports ignored, engagement prioritized over safety, reduced content moderation efforts
TikTok	Automated filtering and human detection	Quick removal of flagged content	High rate of false negatives; user complaints of explicit content not being flagged; lawsuits over teen suicides
Snapchat	Automated system, human detection	Zero-tolerance for sexual exploitation	Sued by families for failing to prevent drug dealers using the platform
YouTube	Algorithm and human moderators, YouTube Kids flagged content, transparency reporting	Enforce policies, remove harmful content	High exposure rates among children; harmful content persists; radicalization pathways documented

3.5. Addiction and engagement-driven design

Social media companies often emphasize how committed they are to promoting meaningful engagement and healthy digital habits. Platforms like Instagram and TikTok provide features which remind users to take breaks, make suggestions to stop use and which monitor time spent using the platforms (McGrath, 2021; TikTok, n.d. - b). YouTube has launched a “take a break” feature to support digital self-regulation, Snapchat provides a similar feature, but it is a part of a feature which reminds users to send streaks to friends so they do not lose their streak score, as detailed in their support pages.

Social media companies use these mechanisms as a proof that they prioritize user well-being, especially that of young users. Despite the popularity of these tools, there are former employees at these companies who reveal a different reality. Frances Haugen, a former Meta employee, declared that internal research continuously found that engagement-based algorithms were designed with the purpose of having

people spend more time on the platform, even if it risks users' mental health (Hao, 2021). Multiple U.S States have sued TikTok due to its addictive features, such as the bite-sized content, infinite scrolling and beauty filters (Bhuiyan & Robins-Early, 2024). Instagram's internal studies show that for teens it is very difficult to stop using the platform, even if they want to, resembling a characteristic of addiction (Wakefield, 2021).

Former YouTube and Snapchat engineers have also admitted that features like auto-play and "streaks" were developed to increase daily user retention (Chapman, 2025). Young users can be more prone to becoming addicted to Snapchat due to their greater desire for validation and gratification and their feelings of urgency around presenting a certain self-image (Matta, n.d.). Personalized content, binge watching and short-term rewards are present in YouTube, just like in any other social media platform (Cherry, 2023).

This contradictory combination of characteristics, both promoting digital well-being while creating features that encourage addictive behaviours, combine powerfully to contribute to users' overuse, sleep disturbances, problems with concentration and loss of control. This is particularly problematic for adolescents, whose self-regulation and reward systems in the brain are still developing (Office of the Surgeon General, 2023; Weir, 2023).

While digital addiction has not been widely recognized until recently, there is now growing concern about the implications for brain neuroplasticity (Doidge, 2018). Adolescence is a critical period in which the brain is still wiring itself in areas linked to attention, self-regulation and decision-making (National Academies of Sciences, Engineering and Medicine, 2019). Social media platforms developed on algorithms and reward mechanisms – such as likes, comments, notifications and infinite scrolling – utilize this neuroplasticity to exploit adolescents' tendencies towards short attention spans and the desire to seek dopamine and instant gratification (Goldman, 2021). Over time, this can alter neural pathways, increasing tendencies towards impulsive behaviour, weakening the capacity for sustained attention. From a normative point of view, this raises deep questions about youth's autonomy and freedom of thought. If young people's cognitive development is being unconsciously influenced by commercially-motivated social media platforms, then concerns regarding their safety as users go beyond the effects on their mental health and privacy, extending to a need to protect the basis of their independent and critical thinking.

By designing platforms that aim for maximum engagement, while publicly promoting healthy habits, companies shift the responsibility to young users and continue to expose them to systems that maintain digital addiction and, in the long run, strip them of their ability to think.

Table 5. Addiction and engagement-driven design

Platform	Engagement feature	Official claim	Real-world outcome / limitations
Facebook	Infinite scroll, push notifications, likes/comments, algorithmic feed ranking	Prioritizes user well-being personalization targeted for safety; includes screen time limits and rest prompts; promotes entertainment, creativity, and community engagement	Internal leaks and research suggest algorithms prioritize engagement over well-being; users can develop addictive patterns, experience emotional fatigue, and encounter harmful content
Instagram	Infinite scroll, reels auto-play, Likes and comments notifications, algorithmic ranking,	Promotes privacy and ephemeral messaging, take-a-break; encourages communication, creativity and connection	Internal leaks show algorithms push for engagement despite harm; exacerbates body issues (one in three women); users become addicted and experience emotional fatigue, anxiety and depression
TikTok	For You page (FYP) infinite scroll, auto-play videos, "likes" and "shares" algorithm, trending content; notifications	Focused on helpful, appropriate content; Take a break; Learning and creating	Algorithm exploits dopamine loops, causes attention loss, mental health harm; Countries are suing for addiction effects; beauty filters; inappropriate content; sleep disruption
Snapchat	Streaks, Snap Score, Quick Add, Discover feed AR lenses, AI chatbot, notifications	Prioritizes user well-being; Personalization targeted for safety; includes screen time limits and rest prompts; Promotes entertainment, creativity and community engagement	Gamified features drive compulsive, repeated use, self-image, validation, reinforcing peer pressure and habit-formation; high notification frequency leads to distraction and stress
YouTube	Auto-play, recommendation engine, short infinite scrolls, notifications	Promotes privacy and ephemeral messaging; Take a break; Encourages communication, creativity, and connection	Binge watching and short-term rewarding; exposure to harmful and inappropriate content

Section 4

Conclusion

While digital technologies and social media platforms provide incredible opportunities for networking, education and self-expression, the risks to young users' mental health, especially those in vulnerable groups, remain high. Algorithms and internal corporate practices often prioritize engagement over well-being, thereby amplifying harmful content and challenging the existing safety and positive content promotion measures provided by social media platforms. The need to find a way to address the negative impacts of the digital determinants of health and well-being is urgent, especially for children and without risking their exclusion from information and connection.

> *By putting public interest ahead of tech companies' profits, we can build a digital future that is safer and fairer for everyone. Silence is no longer an option. Action is urgent. Accountability is a duty.*

4.1. Profits over public health

Platforms such as Meta's Facebook and Instagram are fully aware of the impact they have on youth mental health and well-being, as stated by the whistle-blower Francis Haugen, a former Product Manager for Facebook. Meta's internal research also shows that Instagram makes body issues worse for one in three girls (Wakefield, 2021). The results of scientific studies are in agreement that excessive social media use is linked to heightened anxiety, depression and sleep issues (Costello et al., 2023). However, the reaction of social media platforms has, so far, been minimal and mainly motivated by the need to respond to public, regulatory or legal pressure, and not because of any genuine

commitment to protecting young users. Self-regulation has already failed multiple times, while internal policies and initiatives are not bringing about the needed systematic change that puts youth well-being over profit.

4.2. Stronger regulation takes centre stage

Countries and international bodies are taking stronger measures to address digital determinants of health and well-being for children. The European Union, under the Digital Services Act (DSA), demands that social media platforms reduce design that encourages addiction, harmful content, unwanted contact and youth data processing, with breaches resulting in fines of up to millions of euros. Age verification systems are being tested in France, Greece, Spain, Italy and Denmark. The United Kingdom, following the Online Safety Act (2023), places a legal obligation on social media companies to ensure youth protection from harmful and unwanted content, with fines going up to £18 million, or 10 per cent of annual profit. In Australia, The Amendment for Online Safety (2024) prohibits social media platforms from allowing accounts for users under the age of 16, with fines being implemented from December 2025. According to Holly (2024) a complete ban of children from social media is not the ultimate fix and policymaker focus should be on ensuring that online environments are safe and health-promoting. However, targeted, context-specific measures such as smartphone restrictions in schools can reduce distractions and improve well-being during school hours. Such efforts can be more effective if combined with digital literacy classes, structured guidelines on healthy and balanced screen use and investments in alternatives to screen-based socialization and entertainment for young people. It is necessary to accept that digital devices are part of everyone's day-to-day lives, while creating spaces that promote social interaction, mental health and the ability to focus.

4.3. Recommendations for health agencies and professionals

Public health agencies, health professionals, educators, civil society organizations and researchers play a vital role in demanding accountability from social media platforms to uphold their responsibilities to young people's health, well-being and rights. All concerned stakeholders are recommended to actively demand and shape policies and regulations that ensure safe, fair and health-promoting digital environments.

It is recommended for public health agencies, health professionals and others to:

- **Advocate for binding regulations to ensure that the rights and safety of youth are respected by tech platforms**, including requirements to incorporate safety and well-being into platform design and to ensure transparency, algorithm audits and reporting obligations.

- **Demand a global accountability framework for digital platforms.** Insist on regular public reporting from social media platforms on potential and actual harms to youth health and well-being, and demand strong penalties when they fail to uphold their responsibilities.
- **Support comprehensive digital literacy and clinical support services for young people harmed by digital platforms.** Integrate critical education on social media in schools, train healthcare professionals to screen for health issues related to online behaviour and expand access to confidential mental health services for adolescents, because access to education and care is a fundamental right.
- **Implement WHO's recommendations for addressing the digital determinants of young people's mental health and well-being.** In a recent policy brief, WHO proposed eight priority actions (see text box) for governments and industry to design digital spaces that support, not harm, youth mental health and well-being (World Health Organization, 2025). The priority actions should not be seen only as guides, but rather be implemented with urgency to protect youth from online risks.

Banning youth from social media platforms is not a solution as it risks their exclusion from important information, opportunities and social connections. The global community – from health authorities and educators to families and young users themselves – has a clear responsibility to demand a digital environment that protects, empowers and respects youth.

By putting public interest ahead of tech companies' profits, we can build a digital future that is safer and fairer for everyone. Silence is no longer an option. Action is urgent. Accountability is a duty.

Box 1. WHO priorities for policy action on the digital determinants of youth mental health and well-being

1. Making young people's digital well-being a policy priority.
2. Applying proven, intersectoral public health strategies to improve digital well-being.
3. Developing clear guidance on digital well-being and healthy technology use.
4. Holding industry and commercial interests to account.
5. Supporting and initiating future laws and regulation to make digital environments healthy places for young people.
6. Bolstering the capacity of the health workforce to promote healthy technology use.
7. Increasing research into the impacts of social media on young people's mental health and well-being.
8. Investing in and promoting alternatives to screen-based parenting, play and entertainment.

Source: World Health Organization. (2025, May). Addressing the digital determinants of youth mental health and well-being: Policy brief.

References

Amnesty International. (2023, November 7). Global: TikTok's 'For You' feed risks pushing children and young people towards harmful mental health content. [Press release]. https://www.amnesty.org/en/latest/news/2023/11/tiktok-risks-pushing-children-towards-harmful-content/

Amnesty International. (2025, May 12). TikTok fails to address risks to children and young people's mental health despite past warnings. [Public statement]. https://www.amnesty.org/en/documents/act10/9354/2025/en/

The Australian Government Department of Infrastructure, Transport, Regional Development, Communications, Sport and the Arts. (2025, July 30). Online Safety Amendment (Social Media Minimum Age) Bill 2024 – Fact sheet. https://www.infrastructure.gov.au/department/media/publications/online-safety-amendment-social-media-minimum-age-bill-2024-fact-sheet

BBC News. (2017). The disturbing YouTube videos that are tricking children. BBC News. Retrieved from https://www.bbc.com/news/blogs-trending-39381889

BBC News. (2019). YouTube fined $170m in US over children's privacy violation. BBC News. https://www.bbc.com/news/technology-49578971

BBC News. (2021). TikTok agrees legal payout over facial recognition. https://www.bbc.com/news/technology-56210052

Bellens, E. (2025, April 9). Meta allegedly targeted ads at teens based on their emotional state. Business & Human Rights Resource Centre. https://www.business-humanrights.org/en/latest-news/meta-allegedly-targeted-ads-at-teens-based-on-their-emotional-state/

Bhuiyan, J., & Robins-Early, N. (2024, October 8). US states sue TikTok, claiming its addictive features harm youth mental health. The Guardian. https://www.theguardian.com/technology/2024/oct/08/us-states-tiktok-lawsuit-mental-health

Bohnert, M., & Gracia, P. (2021). Emerging digital generations? Impacts of child digital use on mental and socioemotional well-being across two cohorts in Ireland, 2007–2018. Child Indicators Research, 14(2), 629–659. https://doi.org/10.1007/s12187-020-09767-z

Bridle, J. (2018, June 17). How Peppa Pig became a video nightmare for children. The Guardian. https://www.theguardian.com/technology/2018/jun/17/peppa-pig-youtube-weird-algorithms-automated-content

Chapman, P. (2025). Advancing Platform Accountability: The Promise and Perils of DSA Risk Assessments.Tech Policy Press. https://techpolicy.press/advancing-platform-accountability-the-promise-and-perils-of-dsa-risk-assessments

Chedraoui, K. (2024). Parents welcome Instagram Teen Accounts amid online safety worries. CNET. https://www.cnet.com/news/social-media/parents-welcome-instagram-teen-accounts-amid-online-safety-worries/

Cherry, K. (2023, November 30). What is YouTube Addiction? Verywell Mind. https://www.verywellmind.com/youtube-addiction-signs-causes-effects-and-treatment-5222236

Cho, C.Y., & Zhu, L. (2025, March 20). Social Media: Content Dissemination and Moderation Practices. (CRS Report No. R46662). Congressional Research Service (CRS) for the United States Government. Retrieved July 16, 2025, from https://www.congress.gov/crs-product/R46662

Common Sense Media. (2025, January 7). Common Sense Media condemns Meta's choice to end fact-checking and censorship of harmful content.https://www.commonsensemedia.org/press-releases/common-sense-media-condemns-metas-choice-to-end-fact-checking-and-censorship-of-harmful-content

Costello, N., Sutton, R., Jones, M., Almassian, M., Raffoul, A., Ojumu, O., Salvia, M., Santoso, M., Kavanaugh, J. R., & Austin, S. B. (2023). Algorithms, addiction, and adolescent mental health: An interdisciplinary study to inform state-level policy action to protect youth from the dangers of social media. American Journal of Law & Medicine, 49(2–3), 135–172. https://doi.org/10.1017/amj.2023.25

Delogu, F., Morgan-Paisley, K., Del Gatto, C., Indraccolo, A., Pedale, T., & Brunetti, R. (2025). The impact of the COVID-19 pandemic on the use of social media: A cross-national comparison. Acta Psychologica, 255, 104888. https://doi.org/10.1016/j.actpsy.2025.104888

Department for Science Innovation and Technology. (2025, April 24). Online Safety Act: Explainer. GOV.UK. Retrieved August 14, 2025, from https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer

Doidge, N. (2018). Screen time, the brain, privacy and mental health. In J. Aaronson & R. Medhora (Eds.), Data governance in the digital age: A CIGI essay series (pp. 75–82). Centre for International Governance Innovation. https://www.cigionline.org/articles/screen-time-brain-privacy-and-mental-health/

Duffy, C. (2025, August 13). YouTube will start using AI to guess your age. If it's wrong, you'll have to prove it. CNN Business. https://edition.cnn.com/2025/08/13/tech/youtube-ai-age-verification

European Data Protection Board. (2021, July 22.). Dutch DPA: TikTok fined for violating children's privacy. https://www.edpb.europa.eu/news/national-news/2021/dutch-dpa-tiktok-fined-violating-childrens-privacy_en

Fairplay for Kids. (2025). Teen accounts, broken promises: How Instagram is failing to protect minors. https://fairplayforkids.org/wp-content/uploads/2025/09/Teen-Accounts-Broken-Promises-How-Instagram-is-failing-to-protect-minors.pdf

Federal Trade Commission. (2019, September 4). Google and YouTube will pay record $170 million for alleged violations of children's privacy law. United States Government. https://www.ftc.gov/news-events/news/press-releases/2019/09/google-youtube-will-pay-record-170-million-alleged-violations-childrens-privacy-law

5rights. (2024, October 18). TikTok knows it is harming children. https://5rightsfoundation.com/tiktok-knows-it-is-harming-children/

Foo, Y.C., (2025, July 14) Five EU states to test age verification app to protect children. Reuters. https://www.reuters.com/sustainability/boards-policy-regulation/five-eu-states-test-age-verification-app-protect-children-2025-07-14/

Gerken, T. & McMahon, L. (2023, September 15). TikTok fined €345m over children's data privacy. BBC News. https://www.bbc.com/news/technology-66819174

Goldman, B. (2021, October 29). Addictive potential of social media, explained. Stanford Medicine New Center. https://med.stanford.edu/news/insights/2021/10/addictive-potential-of-social-media-explained.html

Grant, N., & Francisco, T. M. from S. (2025). YouTube loosens rules guiding the moderation of videos. The New York Times. https://www.nytimes.com/2025/06/09/technology/youtube-videos-content-moderation.html

Grace, A. (2024). TikTok teams up with WHO to promote professional mental health content. The Independent. https://www.independent.co.uk/news/uk/tiktok-who-jeremy-farrar-oxford-university-france-b2619470.html

Halprin, M., & O'Connor, J.F., (2022, December 1). Inside YouTube: On policy development at YouTube. YouTube Official Blog. https://blog.youtube/inside-youtube/policy-development-at-youtube/

Hatmaker, T. (2021, September 16). Facebook knows Instagram harms teens. Now, its plan to open the app to kids looks worse than ever. TechCrunch. https://techcrunch.com/2021/09/16/facebook-instagram-for-kids-mosseri-wsj-teen-girls/

Hao, K. (2021, October 5). The Facebook whistleblower says its algorithms are dangerous. Here's why. MIT Technology Review. https://www.technologyreview.com/2021/10/05/1036519/facebook-whistleblower-frances-haugen-algorithms/

Holly, L. (2024, November 22). Tackling digital harms: Why simply banning children from social media won't protect them. The BMJ, 387, q2617. https://doi.org/10.1136/bmj.q2617

Holly, L., Naing, S. Y., Pitt, H., Thomas, S., & Kickbusch, I. (2025, March 26). Health promotion and the digital determinants of health. Health Promotion International, 40(2), daaf033. https://doi.org/10.1093/heapro/daaf033

Horwitz, J. (2025a). Instagram's teen safety features are flawed, researchers say. Reuters. https://www.reuters.com/business/instagrams-teen-safety-features-are-flawed-researchers-say-2025-09-25/

Horwitz, J. (2025b). Meta's AI rules have let bots hold 'sensual' chats with kids, offer false medical info. Reuters. https://www.reuters.com/investigates/special-report/meta-ai-chatbot-guidelines

Instagram. (2020, May 18). Supporting well-being with guides on Instagram. Retrieved July 15, 2025, from https://about.instagram.com/blog/announcements/supporting-well-being-with-instagram-guides

Kerr, D (2023, November 7). Meta failed to address harm to teens, whistleblower testifies as senators vow action. NPR. https://www.npr.org/2023/11/07/1211339737/meta-failed-to-address-harm-to-teens-whistleblower-testifies-as-senators-vow-act

Khalaf, A. M., Alubied, A. A., Khalaf, A. M., & Rifaey, A. A. (2023, August 5). The impact of social media on the mental health of adolescents and young adults: A systematic review. Cureus, 15(8), e42990. https://doi.org/10.7759/cureus.42990

Kindelan, K. (2024, January 4). Families' lawsuit against Snapchat alleging the platform enables drug dealers allowed to move forward. ABC News. https://abcnews.go.com/GMA/News/families-lawsuit-snapchat-alleging-platform-enables-drug-dealers/story?id=106076434

Kitson, N. (2025, July 25). Meta launches new security features on Instagram and Facebook to protect teens. TechCentral. https://www.techcentral.ie/meta-launches-new-security-features-on-instagram-and-facebook-to-protect-teens/

Lu, M., Lamond, M., & Fry, D. (2024). A content analysis of metrics on online child sexual exploitation and abuse used by online content-sharing services. Child Abuse & Neglect, 157, 107046. https://doi.org/10.1016/j.chiabu.2024.107046

Magoven Creative Studio. (2024). Instagram Features and Safety Guidelines. https://magoven.io/instagram-features-and-safety-guidelines/

Matta, N.. (n.d.). Snapchat Addiction. Addiction Center. Retrieved July 16, 2025, from https://www.addictioncenter.com/behavioral-addictions/social-media-addiction/snapchat-addiction/

McGrath, L. (2021, November 21). Instagram Tests "Take a Break" Feature to let users self regulate. Bloomberg. https://www.bloomberg.com/news/articles/2021-11-10/instagram-tests-take-a-break-tool-to-let-users-self-regulate

McMahon, L., Kleinman, Z., & Subramanin, C. (2025, January 7). Facebook and Instagram get rid of fact checkers. BBC News. https://www.bbc.com/news/articles/cly74mpy8klo

Meta. (2025a, April 8). Facebook & Messenger Teen Accounts: How built-in protections help keep your teen safe. Meta Family Center. Retrieved from https://familycenter.meta.com/resources/teen-accounts-facebook-messenger/

Meta. (2025b, April 8). We're Introducing New Built-In Restrictions for Instagram Teen Accounts, and Expanding to Facebook and Messenger. https://about.fb.com/news/2025/04/introducing-new-built-in-restrictions-instagram-teen-accounts-expanding-facebook-messenger/

Meta Transparency Center. (2025c). How fact-checking works. Retrieved July 21, 2025, from https://transparency.meta.com/features/how-fact-checking-works

Meta. (2023, January 10). Continuing to Create Age-Appropriate Ad Experiences for Teens. https://about.fb.com/news/2023/01/age-appropriate-ads-for-teens/

Meta. (2021, December 8). Our new AI system to help tackle harmful content. https://about.fb.com/news/2021/12/metas-new-ai-system-tackles-harmful-content/

Meta. (n.d.). How Meta enforces its policies. Meta. Retrieved October, 13, 2025, from https://transparency.meta.com/enforcement/

Miller, G. (2024, September 5). Snapchat Is A "Breeding Ground" For Child Predators, According To New Mexico Lawsuit. Tech Policy Press. https://techpolicy.press/snapchat-is-a-breeding-ground-for-child-predators-according-to-new-mexico-lawsuit

Murphy Kelly, S. (2022, December 15). TikTok may push potentially harmful content to teens within minutes, study finds. CNN Business. https://www.cnn.com/2022/12/15/tech/tiktok-teens-study-trnd

National Academies of Sciences, Engineering, and Medicine. (2019). The Promise of Adolescence: Realizing Opportunity for All Youth. (E.P. Backes & R.J. Bonnie, Eds.).Washington, DC: The National Academies Press. https://doi.org/10.17226/25388

Nwoko, H. (2025 March 8). Is YouTube Kids Safe? This Doctor Is Sending Parents a Warning. Parents. Retrieved July 15, 2025, from https://www.parents.com/doctor-sends-warning-to-parents-about-youtube-kids-11695578

O'Connor, J.F., & Moxley, E. (2023, November 14). Inside YouTube: Our approach to responsible AI innovation. YouTube Official Blog. https://blog.youtube/inside-youtube/our-approach-to-responsible-ai-innovation/

Office of the Surgeon General (OSG). (2023). Social media and youth mental health: The U.S. Surgeon General's Advisory. US Department of Health and Human Services. https://www.ncbi.nlm.nih.gov/books/NBK594763/

Owens, K., & Lenhart, A. (2025). The unseen teen: Social platforms and accountability for addressing adolescent well-being. Anals of the New York Academy of Sciences, 1549(1), 160–170. https://doi.org/10.1111/nyas.15375

Presser, A. (2025, July 30). New trust and safety tools for creators, families, and the TikTok Community. TikTok Newsroom. https://newsroom.tiktok.com/en-us/tiktok-announces-a-suite-product-features

Prothero, A. (2024, February 16). How social media may benefit teens' mental health. Education Week. https://www.edweek.org/leadership/how-social-media-may-benefit-teens-mental-health/2024/02

Robison, K. (2025, June 12). The Meta AI app lets you 'Discover' people's bizarrely personal chats. WIRED. https://www.wired.com/story/meta-artificial-intelligence-chatbot-conversations/

Selmani, E. (2025, January). Analysing major tech companies' policies and strategies for the health and safety of young users. Digital Transformations for Health Lab. https://dthlab.org/wp-content/uploads/2025/01/Erza-Selmani-Analysing-Major-Tech-Companies-Policies-and-Strategies-for-the-Health-and-Safety-of-Young-Users.pdf

Snapchat For Business. (n.d. - a). : Reach an untapped audience on Snapchat.. Snapchat. Retrieved July 15, 2025, from https://forbusiness.snapchat.com/

Snapchat For Business. (n.d. - b). Introducing sponsored snaps. Snapchat. Retrieved September 30, 2025, from https://forbusiness.snapchat.com/advertising/sponsored-snaps

Snapchat (2025, March). Snapchat Moderation, Enforcement, and Appeals: Community Guidelines Explainer Series. Snapchat. Retrieved July 16, 2025, from https://values.snap.com/privacy/transparency/community-guidelines/moderation

Southern, M. G. (2025, January 27). How YouTube's Recommendation System Works In 2025. Search Engine Journal. https://www.searchenginejournal.com/how-youtubes-recommendation-system-works-in-2025/538379/

Statistics Canada. Government of Canada, S. C. (2024, February 27). Young people and exposure to harmful online content in 2022. Government of Canada. https://www150.statcan.gc.ca/n1/pub/11-627-m/11-627-m2024005-eng.htm

Tedjawidjaja, D., & Christanti, D. (2022). Effect of social comparison in social media on psychological distress in adolescents: Role of emotion regulation as moderator. Journal of Educational, Health and Community Psychology, 11(4), 836. https://doi.org/10.12928/jehcp.v11i4.25056

The European Commission. (2022, October 27). The Digital Services Act. The European Union. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en

TikTok. (n.d. - a). How TikTok recommends content. Retrieved July 15, 2025, from https://support.tiktok.com/en/using-tiktok/exploring-videos/how-tiktok-recommends-content

TikTok. (n.d. - b). Screen time. Retrieved July 16, 2025, from https://support.tiktok.com/en/account-and-privacy/account-information/screen-time

TikTok Transparency Center. (n.d. - c). Our approach to content moderation. TikTok. Retrieved July 16, 2025, from https://www.tiktok.com/transparency/en/content-moderation/

Tumakuru, & S, S. (2023). Impact of social media on youth: Comprehensive analysis. Shodh Sari 2(4), 286–301. https://doi.org/10.59231/sari7640

Van Essen, C. M., & Van Ouytsel, J. (2023, September). Snapchat streaks—How are these forms of gamified interactions associated with problematic smartphone use and fear of missing out among early adolescents? Telematics and Informatics Reports, 11, 100087. https://doi.org/10.1016/j.teler.2023.100087

Vivarelli, N. (2023, April 4). TikTok Slapped With Fine by Britain's Data Watchdog for Failing to Enforce Age Limits on Kids. Variety. https://variety.com/2023/digital/global/tiktok-fine-by-britain-data-watchdog-age-limits-kids-1235572933/

Wakefield, J. (2021, September 15). Facebook under fire over secret teen research. BBC News. https://www.bbc.co.uk/news/technology-58570353

Waterson, J., & Milmo, D. (2021, October 25). Facebook whistleblower Frances Haugen calls for urgent external regulation. The Guardian. https://www.theguardian.com/technology/2021/oct/25/facebook-whistleblower-frances-haugen-calls-for-urgent-external-regulation

Weir, Kirsten, (2023). Social media brings benefits and risks to teens. Psychology can help identify a path forward. Monitor on Psychology, 54(6), 46. https://www.apa.org/monitor/2023/09/protecting-teens-on-social-media

World Health Organization. (2025, May). Addressing the digital determinants of youth mental health and well-being: Policy brief. https://www.who.int/europe/publications/i/item/WHO-EURO-2025-12187-51959-79685

World Health Organization. (n.d.). Health promotion. https://www.who.int/westernpacific/about/how-we-work/programmes/health-promotion

YouTube. (n.d.). Terms of service. Retrieved July 15, 2025, from https://kids.youtube.com/t/terms

About DTH-Lab

DTH-Lab is a global consortium of partners working to drive implementation of The Lancet and Financial Times Commission on Governing Health Futures 2030's recommendations for value-based digital transformations for health co-created with young people. DTH-Lab operates through a distributive governance model, led by three core partners: Ashoka University (India), DTH-Lab (hosted by the University of Geneva, Switzerland) and PharmAccess (Nigeria).

Leadership Team

Aferdita Bytyqi, DTH-Lab Executive Director and Founding Member.

Ilona Kickbusch, DTH-Lab Director and Founding Member.

Anurag Agrawal, DTH-Lab Founding Member. Dean of Biosciences and Health Research, Ashoka University.

Rohinton Medhora, DTH-Lab Founding Member. Professor of Practice, McGill University's Institute for the Study of International Development.

Njide Ndili, DTH-Lab Founding Member. Country Director for PharmAccess Nigeria.

DIGITAL TRANSFORMATIONS FOR HEALTH LAB

GOVERNING HEALTH FUTURES 2030

Digital Transformations for Health Lab (DTH-Lab)
Hosted by: The University of Geneva
Campus Biotech, Chemin des Mines 9
1202 Geneva, Switzerland

www.DTHLab.org