PANORAMIC

ONLINE SAFETY REGULATION

USA

LEXOLOGY

Online Safety Regulation

Contributing Editors

Jenna Rennie, Rory Hishon and Alexander Beaton

White & Case LLP

Generated on: March 31, 2026

Contents

Online Safety Regulation

Contributors

USA

White & Case LLP

WHITE & CASE

Hope Anderson

hope.anderson@whitecase.com

LEGAL FRAMEWORK

Legal regime

Does your jurisdiction have a legal regime governing or addressing online safety? If so, how does it operate?

The United States (US) does not have a comprehensive, cross-sector legal regime that governs online safety. Instead, online safety obligations are derived from a patchwork of federal laws, state laws, and agency regulations, which are typically targeted to address specific harms – for example, child sexual abuse material (CSAM), sex trafficking and privacy violations.

Accordingly, in practice, the US online safety framework operates less as a unified legal regime and more as a decentralised system of targeted enforcement that relies on, ***inter alia***, consumer protection standards, criminal law requirements, online privacy laws and emerging state-level online youth safety laws.

Law stated - 16 December 2025

Online harms covered

Which online harms are covered under the relevant legislation and how are these harms defined?

US law does not define 'online harms' within a single statutory framework. Instead, specific harms are addressed through targeted federal and state laws, each using its own terminology. At the federal level, the most clearly articulated harms relate to criminal content, including but not limited to CSAM, sex trafficking, terrorism, credible threats and stalking. CSAM is defined broadly under 18 USC section 2256 to include any visual depiction of sexually explicit conduct involving a minor, and under 18 USC section 2258A, online providers must report apparent violations to the National Center for Missing & Exploited Children (NCMEC). Federal sex-trafficking provisions including the Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA) and the Stop Enabling Sex Traffickers Act (SESTA) prohibit knowingly assisting or facilitating trafficking activities. Terrorism-related harms are addressed through material-support statutes covering dissemination of content intended to aid designated foreign terrorist organisations.

Beyond criminal content, the primary federal mechanism for addressing online harms is the Federal Trade Commission (FTC)'s authority over 'unfair or deceptive' acts or practices under section 5 of the FTC Act. Although not expressly framed as combating 'online harms,' FTC enforcement has increasingly focused on website and app design choices – for example, dark patterns, manipulative interfaces, misleading safety claims and inadequate protections for children – that create or heighten online safety risks. The Children's Online Privacy Protection Act (COPPA) also treats certain data processing and profiling practices as harmful when directed at users under 13.

Emerging state-level legislation is attempting to expand regulation of online harms, particularly in the context of youth safety. For example, recently enacted age-appropriate design code laws in states such as California, Maryland, Nebraska and Vermont have introduced relatively comprehensive youth safeguards to minimise minors' exposure

to certain types of harmful or inappropriate content, protect minors from certain targeted advertising practices and regulate platform features perceived as 'addictive' or psychologically harmful. Various states have also passed more targeted legislation aimed at addressing risks related to, ***inter alia***, 'addictive' platform features or engagement mechanics, exposure to dangerous or unwanted contacts, and profiling of minors. Scoping and definitions vary, but these laws commonly target perceived online harms arising from, for example, algorithmic recommendations, social media engagement mechanics, exposure to high-risk categories of content and other practices that may impair a minor's well-being.

In addition to passing or considering laws targeted at youth-specific risks, several states are beginning to address broader risks that fall within the online safety landscape, including the spread of misleading AI-generated or manipulated content (including deepfake sexual imagery). A small but growing subset of state legislation also targets harms associated with AI chatbots, particularly where minors or consumers may be misled or manipulated. Several of these state laws seek to address general consumer deception risks (for example, failing to inform users that they are interacting with an automated system rather than a live human), while others are narrowly tailored to address more specific harms (eg, chatbots providing mental health services and chatbots interacting with minors). While these laws are still nascent and vary significantly in scope, they reflect an emerging regulatory trend that treats certain AI-generated interactions – especially those that obscure a bot's automated nature or that could encourage risky conduct – as an independent category of online harm.

However, many of these emerging state-level youth online safety regimes are facing significant First Amendment and/or federal preemption challenges. Courts have already enjoined or partially blocked several laws on grounds that their content-based restrictions, parental-consent requirements and age-verification mandates impermissibly burden access to lawful speech. As a result, the constitutional durability of this emerging generation of online safety statutes remains uncertain and ongoing litigation is likely to continue shaping – and potentially narrowing – the scope of ongoing state-level attempts to define and regulate online harms.

Law stated - 16 December 2025

Online services covered

Which online services are covered under the law and how are these services defined?

The US does not take a unified approach to defining 'online services' for purposes of online safety regulation. Instead, different federal and state laws apply to specific types of online services, often using their own definitional and scoping terminology. Accordingly, coverage is context-specific.

At the federal level, the broadest and most influential definition appears in section 230 of the Communications Decency Act, which applies to 'interactive computer services' – defined as 'any information service, system or access software provider that provides or enables computer access by multiple users to a computer server.' 47 USC section 230(f)(2). This term has been interpreted expansively to cover virtually any online platform that enables users to generate, share or access content. Accordingly, 'interactive computer services' as used in section 230 includes social media services, messaging platforms, forums, search

engines, cloud hosts, and many other intermediaries. Federal criminal statutes governing CSAM, trafficking, and terrorism apply broadly to 'electronic service providers' and 'remote computing services,' terms drawn from the Stored Communications Act and construed to include most entities offering online communications or storage capabilities. COPPA, by contrast, applies to operators of commercial websites and online services 'directed to children' under 13 or that knowingly collect data from such users, sweeping in a wide range of apps, platforms, games, and ad networks.

Because the US lacks a comprehensive online-safety regime, federal law does not impose platform-wide systemic duties on 'social media services' or other narrower classes of online intermediaries. Instead, obligations arise in a piecemeal manner depending on the service's functions and public representations – through, ***inter alia***, content-reporting mandates, privacy rules or consumer-protection theories.

State-level legislation is beginning to define online services more narrowly and prescriptively – especially in the context of youth safety. Recent laws in Utah, Arkansas, Texas and Louisiana specifically regulate 'social media platforms,' often defined by reference to user accounts, content-sharing functionality and platform size thresholds (eg, reaching a specified number of account holders). The emerging age-appropriate design code laws apply to 'covered businesses' providing online services that are 'reasonably likely to be accessed' by minors in the state – a deliberately expansive formulation that can encompass general-purpose platforms, gaming services, streaming media and even educational technology. These definitions vary somewhat across states and are subject to ongoing constitutional litigation, leaving their long-term applicability uncertain.

Law stated - 16 December 2025

Territorial scope

What is the territorial scope of the relevant law?

In practice, the territorial scope of US online safety obligations is broad and user-centric. All online services that make their products available to US residents (or state residents) are generally expected to comply with relevant privacy, consumer protection, youth safety and criminal laws.

US federal laws related to online safety generally apply with broad extraterritorial reach, though the exact scope varies by statute. Federal criminal laws governing CSAM, trafficking, terrorism and related offences apply to conduct with a sufficient US nexus, including activity involving US persons, services operated from or accessible within the US, or data stored in the US. COPPA applies to operators of websites and online services that are 'directed to' children in the US or that knowingly collect personal information from US-based users under 13, regardless of where the operator is physically established. The FTC's authority under Section 5 of the FTC Act also extends to foreign entities where the relevant conduct causes, or is likely to cause, substantial injury to US consumers.

Emerging state-level youth safety laws typically assert jurisdiction over online service providers that offer their services to residents of the regulating state. These laws generally apply based on the location of the user, not the operator, and require covered services to implement, ***inter alia***, age verification, parental consent or design safeguards when providing access to minors within the state. Many of these statutes include explicit extraterritorial

commitments – such as applying to any platform with users in the state – but some of these emerging laws are facing substantial constitutional scrutiny, including Dormant Commerce Clause challenges. As a result, their ultimate territorial reach remains uncertain.

Law stated - 16 December 2025

Codes of practice

Are there any codes of practice or other non-binding guidelines or recommendations relating to online safety in your jurisdiction?

The US does not maintain formal, statutory 'codes of practice' relating to online safety comparable to those issued under, for example, the UK Online Safety Act or Australia's Online Safety Act. Instead, online safety expectations derive from a patchwork of federal and state laws, non-binding federal guidance, agency policy statements and voluntary industry frameworks.

The FTC is the most significant source of such guidance. Although the FTC does not issue online safety codes, it regularly publishes policy statements, enforcement policy guides, business guidance documents, and COPPA compliance materials that outline expectations regarding dark patterns, teen protections, data minimisation, product design choices and safety-related representations. These materials often shape industry practice with respect to online safety, because deviations from them can inform FTC investigations and enforcement actions for 'unfair' or 'deceptive' conduct.

Other federal agencies provide targeted guidance relevant to specific harms. The Department of Justice (DOJ) and NCMEC publish best practices for reporting and mitigating CSAM. The Department of Homeland Security, FBI, and Global Internet Forum to Counter Terrorism (GIFCT) collaborate with industry partners to develop voluntary frameworks for detecting and removing terrorist and extremist content. Similarly, the Cybersecurity and Infrastructure Security Agency (CISA) issues voluntary cybersecurity and digital resilience guidance that can intersect with online safety.

In the absence of a centralised federal regime, industry self-regulation also plays a meaningful role. Major platforms often participate in voluntary standards-setting initiatives focused on youth safety, transparency reporting, crisis protocols, non-consensual intimate imagery and coordinated inauthentic behaviour. These frameworks are not legally enforceable but often serve as benchmarks evaluated by regulators, courts and civil society organisations.

Law stated - 16 December 2025

Harmful versus illegal content

How does the law in your jurisdiction distinguish between harmful and illegal content?

US law largely regulates illegal content (eg, CSAM, copyright violations, credible threats and trafficking). The First Amendment severely limits governmental regulation of content that is legal but may be considered harmful (eg, bullying and hate speech).

Illegal content is defined narrowly and primarily through federal and state criminal statutes. This includes, for example, CSAM, sex-trafficking facilitation, material support for terrorism, true threats, extortion, and certain forms of stalking or harassment. These categories are tightly defined – for example, CSAM is extensively codified in 18 USC sections 2256–2258A, and platforms must report 'apparent' violations to NCMEC. Similarly, FOSTA-SESTA and related provisions impose liability for knowingly facilitating trafficking, and terrorism laws prohibit distributing material intended to provide support to designated foreign terrorist organisations. Because these forms of content fall outside First Amendment protection, platforms may face mandatory reporting duties, along with potential criminal penalties or loss of safe-harbour protections when they fail to act on them.

By contrast, harmful content – including, for example, graphic violence, hate speech, sexual content or content otherwise considered inappropriate for minors – is generally lawful to host or access. The First Amendment typically prohibits the government from imposing broad, content-based restrictions on content that may be offensive, harmful or developmentally inappropriate for children. As a result, regulation of harmful-but-legal content typically arises indirectly – for example, through the FTC's authority to police unfair or deceptive business practices (eg, misleading safety representations and harmful design features) or through emerging state youth safety statutes that attempt to limit minors' exposure to certain categories of content (eg, pornography). Notably, many of these emerging state laws are facing constitutional challenges because they still seek to regulate access to lawful speech.

Accordingly, with certain exceptions, regulation of harmful content depends largely on voluntary platform governance, consumer-protection enforcement and emerging state-level youth safety obligations.

Law stated - 16 December 2025

Extremist and terrorism-related content

How does your jurisdiction regulate the dissemination of extremist and terrorism-related content online?

The US regulates extremist and terrorism-related content primarily through federal criminal law. The core statutory framework consists of the 'material support' provisions of 18 USC sections 2339A-2339B, which prohibit providing material support or resources to designated foreign terrorist organisations (FTOs). Section 2339A defines 'material support' to encompass 'any property, tangible or intangible, or service.' Based on this broad definition, US courts have interpreted 'material support' broadly to include online platform functionality when knowingly provided to advance terrorist activity. Thus, although mere content hosting is not itself criminalised, platforms can incur liability if they knowingly facilitate or coordinate with an FTO. For example, knowingly aiding in the dissemination of propaganda, recruitment materials, or operational communications intended to aid an FTO can fall within the scope of these provisions.

Importantly, section 230 of the Communications Decency Act generally protects intermediaries from civil liability for third-party terrorist content, but this protection does not extend to federal criminal prosecutions or to claims that the platform itself knowingly provided material support. Litigation following major terrorist attacks has tested the

boundaries of this doctrine, and the US Supreme Court has recently reaffirmed that routine algorithmic recommendation of third-party content – absent evidence of knowing, purposeful assistance – does not constitute material support.

Beyond statutory requirements, the federal government relies heavily on voluntary industry cooperation. The Department of Homeland Security, FBI and State Department collaborate with private companies through information-sharing groups, and many major platforms participate in the Global Internet Forum to Counter Terrorism (GIFCT), which maintains shared hash databases and best-practice frameworks for detecting, removing and reporting terrorist content. These voluntary measures are not legally mandated but significantly shape practical expectations for online platforms operating in the US.

Law stated - 16 December 2025

Disinformation versus misinformation

How, if at all, does the law in your jurisdiction distinguish between misinformation and disinformation online? Does it include malinformation?

US law does not provide formal statutory definitions of 'misinformation,' 'disinformation' or 'malinformation,' nor does it impose direct legal obligations on online platforms to identify or remove such content. The First Amendment heavily limits government regulation of false or misleading speech unless it falls within specific, historically recognised exceptions such as fraud, defamation or deceptive commercial practices. As a result, US law addresses harmful falsehoods through narrow subject-matter statutes and general consumer protection principles, rather than through any dedicated online safety regime.

In practice, misinformation (ie, false content shared without intent to deceive) is generally protected by the First Amendment. Even factually incorrect statements about politics, health, or social issues typically cannot be restricted by government actors. Disinformation, meaning false information shared with intent to mislead, is also largely protected unless it meets the elements of fraud, defamation, election-related deception targeted at voters, or other specifically prohibited conduct. Several federal and state statutes criminalise narrowly tailored forms of deceptive interference with voting, impersonation of government officials, or intentional dissemination of false information used for financial or other tangible harm, but these laws do not establish a broad regulatory category for online disinformation.

The related concept of malinformation (ie, accurate information used in a misleading or harmful context) similarly has no legal recognition under US law. Because it typically involves truthful speech, it receives First Amendment protection absent a separate, actionable harm (for example, doxxing that violates a privacy statute, non-consensual intimate imagery or targeted harassment that meets criminal thresholds).

Given these constitutional constraints, federal agencies such as the FTC and Food and Drug Administration (FDA) regulate misinformation primarily through consumer protection and product liability frameworks, focusing on deceptive commercial assertions (eg, false medical claims and misleading advertising). Beyond the domain of commercial speech, the US relies largely on voluntary industry cooperation, election-security collaborations and civil-society initiatives to address misinformation, disinformation and malinformation online.

In summary, the US legal system treats false or misleading online content as lawful speech by default, with legal and regulatory intervention limited to well-established exceptions. The distinction among misinformation, disinformation and malinformation, therefore, remains more a matter of platform policy and academic terminology, rather than a defined or regulated legal category.

Law stated - 16 December 2025

OBLIGATIONS FOR ONLINE SERVICE PROVIDERS

General obligations

What general legal obligations relating to safety are imposed on providers of online services, including providers of online intermediary services?

The US does not impose a broad duty of care on online service providers to promote or ensure online safety. Instead, safety-related obligations arise through a combination of federal criminal law, consumer-protection enforcement, children's privacy rules and emerging state-level youth safety statutes. As a result, providers of online services face a fragmented set of expectations rather than a single, comprehensive safety regime. Below is a non-comprehensive list of examples.

At the federal level, the clearest obligations relate to certain categories of illegal content. Online service providers that become aware of apparent child sexual abuse material (CSAM) must report it to the National Center for Missing & Exploited Children (NCMEC) under 18 USC section 2258A. Federal law also prohibits knowingly facilitating sex trafficking, terrorism-related activities or certain forms of threats, extortion and stalking. While section 230 generally shields intermediaries from civil liability for user content, it does not immunise such platforms from federal criminal liability for knowing facilitation of illegal acts or excuse them from complying with relevant reporting responsibilities.

A second, and increasingly significant, source of obligations arises under consumer protection law. the Federal Trade Commission (FTC) enforces section 5 of the FTC Act, which prohibits 'unfair or deceptive' acts or practices. Though not framed in online safety terminology, the FTC has applied this authority to a wide range of platform practices, including misleading safety representations, inadequate safeguards for minors, harmful or manipulative design features, and failures to reasonably secure user data. Operators of services directed to children under 13 also face specific obligations under the Children's Online Privacy Protection Act (COPPA), including data minimisation, verifiable parental consent and restrictions on tracking and profiling.

Outside these federal requirements, general content monitoring obligations are expressly limited. Section 230 protects providers' discretionary moderation choices and prevents the imposition of broad duties to proactively police user content. As a result, safety obligations are largely tied to statutory triggers (such as encountering illegal content) or to representations that platforms make about their own products.

States are beginning to fill perceived gaps with more prescriptive youth-safety and age-assurance laws. Recent statutes in, for example, California, Utah, Arkansas, Texas and Louisiana impose duties related to age verification, parental consent and protection of minors from harmful or addictive features. However, many of these statutes face ongoing

constitutional challenges, leaving the durability and enforceability of these obligations uncertain.

To summarise, US law generally imposes targeted, reactive safety obligations focused on illegal content and prohibits broad, proactive content moderation mandates. The result is a system in which safety expectations are shaped by federal enforcement, product-design scrutiny, and evolving (and legally contested) state-level youth safety initiatives.

Law stated - 16 December 2025

Risk assessments and mitigation

Are there any specific legal obligations for online service providers to conduct risk assessments and mitigate risks to safety?

At the federal level, there is no legal obligation to conduct online safety risk assessments. The closest analogue to a federal online safety risk assessment obligation comes from the FTC's consumer protection authority. Through enforcement actions and consent orders, the FTC has required companies to build and maintain reasonably designed safeguards, conduct periodic security and privacy assessments, and implement design-level controls when products pose foreseeable risks – particularly where minors are involved. While these obligations are not framed as 'online safety' risk assessments, they function similarly in practice, especially in cases involving dark patterns, manipulative interfaces or misleading claims about teen safety or content moderation.

In the privacy context, the COPPA compliance framework requires operators to conduct annual risk assessments to identify internal and external risks to the confidentiality, security and integrity of personal information collected from children under 13, and to document and justify the safeguards adopted to mitigate such risks. While not specific to online safety, COPPA's risk-assessment obligations effectively function as an early US analogue to more formalised DPIA-style requirements in other jurisdictions, though they remain limited to privacy and security harms rather than broader risks associated with design features, algorithmic exposure pathways and potentially harmful content.

A more explicit move toward risk assessment and mitigation duties designed to address a broader spectrum of online safety issues is emerging at the state level, particularly in youth safety legislation. For example, California's currently enjoined Age-Appropriate Design Code (AADC) would require businesses offering services likely to be accessed by children to conduct data protection impact assessments (DPIAs) examining potential harms to minors arising from product design, profiling, or content exposure. Similarly, the Maryland AADC (currently subject to legal challenge) would require covered entities to prepare a DPIA for any online services reasonably likely to be accessed by children, assessing (among other issues) whether the product is designed consistent with the best interests of children in line with criteria specified in the law.

Several other states have adopted narrower but related online safety laws that effectively create risk evaluation obligations in practice, even in the absence of an express mandate for written risk assessments. For example, new laws in Arkansas, Colorado, Florida, New York, Louisiana and Texas impose varying duties related to age assurance, parental controls, 'addictive' engagement features, exposure to certain high-risk content categories, and interactions between minors and unknown adults. As a practical matter, compliance with

these laws may effectively require platforms to assess whether particular features or content-delivery mechanics could reasonably increase the likelihood of specified harms and to adjust their design choices accordingly. As with the AADC statutes, many of these laws remain subject to active legal challenges, leaving their long-term enforceability uncertain; however, they illustrate a broader trend toward embedding forward-looking risk-assessment concepts into US online safety law.

Law stated - 16 December 2025

Protection of minors and age verification

Are there any specific legal obligations to protect minors online? If so, what measures are required or advised, such as age verification?

US law imposes several targeted obligations to protect minors online, but there is no comprehensive federal youth safety regime. Federal requirements focus primarily on children's privacy and illegal content, while states are beginning to introduce more prescriptive (but often constitutionally contested) youth safety obligations.

At the federal level, the principal statute governing minors is COPPA, which is a privacy statute that applies to operators of websites and online services directed to children under 13 or that knowingly collect personal information from such users. COPPA requires businesses to implement verifiable parental consent before collecting data from children, to provide clear privacy notices, to limit data collection to what is 'reasonably necessary' and to maintain reasonable security for children's data. COPPA does not impose broader youth safety or age verification duties, but it functions as the closest federal analogue to a child-protection code of practice.

Federal criminal statutes impose additional responsibilities related to CSAM. Providers that become aware of apparent CSAM must report it to NCMEC, and failure to do so can result in criminal penalties. These laws do not create a general duty to monitor, but they require action once providers obtain knowledge of illegal material. Beyond these narrow federal requirements, there are no general obligations to age-verify users, restrict minors' access to harmful content, or undertake preventative safety measures. Such requirements would likely raise significant First Amendment concerns if mandated at the federal level.

At the state level, legislatures have begun to adopt more prescriptive youth safety laws, often directed at social media platforms. States such as Arkansas, California, Florida, Georgia, Maryland, Nebraska, Utah and Vermont have passed laws that either require or strongly encourage age assurance, parental consent and/or enhanced protections for minors, including limits on nighttime access, default privacy settings, restrictions on 'addictive' features and measures intended to reduce exposure to harmful or inappropriate content. Some of these statutes focus narrowly on specific risks (eg, age assurance for social media accounts or limitations on perceived 'addictive' features), while others adopt a more comprehensive 'design-code' approach that requires platforms to assess and mitigate a variety of foreseeable risks to minors arising from data practices and product architecture. Certain state laws also mandate that covered platforms provide high-privacy defaults for minors, assess design-related risks, and limit profiling of minors. However, as noted above, many of these new laws face ongoing constitutional challenges, and their long-term enforceability remains uncertain.

Law stated - 16 December 2025

Civil and human rights

Are there any obligations for online service providers to balance civil and human rights, such as privacy rights and freedom of expression, with safety regulations? If so, what measures are required or advised?

US law does not impose an express statutory obligation on online service providers to 'balance' civil or human rights with safety considerations. Instead, the balancing is inherent in the structure of US constitutional and regulatory law – most notably the First Amendment, privacy and consumer protection frameworks, and anti-discrimination principles – which together shape the permissible bounds of online safety regulation and influence platform practices in this area.

The most significant constraint is the First Amendment, which sharply limits government attempts to regulate lawful but harmful content, mandate viewpoint-based moderation or require platforms to remove or restrict certain categories of speech. Courts routinely invalidate or narrow state-level online safety statutes for restricting minors' and/or adults' access to lawful speech, making constitutional balancing an unavoidable part of the legislative landscape around online safety.

Privacy considerations also influence online safety obligations. COPPA incorporates baseline privacy protections for children under 13, while the FTC uses its consumer protection authority to police deceptive or unfair data practices, dark patterns and design choices that may undermine user autonomy or safety. FTC consent orders frequently require companies to implement privacy-by-design and data minimisation measures that indirectly balance safety against privacy and expression-related concerns by requiring transparency, choice, and safeguards against over-collection or misuse of data.

Emerging state privacy laws such as the California Privacy Rights Act and similar statutes in Virginia, Colorado and Connecticut (among other states) impose additional obligations and restrictions related to data processing, which may effectively compel companies to evaluate how safety features interact with user privacy and autonomy.

Separately, federal and state anti-discrimination laws may limit practices that could result in discriminatory access, targeting or suspension decisions. While these laws do not mandate any formal human rights assessment, they may influence how online platforms design safety and content moderation systems.

Law stated - 16 December 2025

Disinformation and misinformation

Are there any specific legal obligations to combat disinformation and misinformation online? If so, what measures are required or advised?

No. US law does not impose any generally-applicable legal obligations on online service providers to detect, remove, or otherwise combat misinformation or disinformation. As detailed above, the First Amendment sharply limits government regulation of false or

misleading speech, and as such, most efforts to address misinformation are voluntary, industry-led, or arise indirectly through consumer protection and election-integrity laws.

In practice, federal regulation of misinformation occurs only in narrow domains. For example, the FTC can take action when false statements constitute deceptive commercial conduct, such as misleading health or product claims. The FDA has regulatory oversight over false medical and pharmaceutical claims. Election-related statutes prohibit certain forms of intentional voter deception, such as misrepresenting voting locations, dates, or eligibility rules, but these laws are narrowly tailored and do not establish a broad prohibition on political misinformation.

There are also no federal obligations for platforms to verify factual accuracy, label contested information, or implement specific mitigation measures to address false statements. The federal government may encourage voluntary cooperation for certain issues related to, ***inter alia***, national security, electoral integrity or public health, but it cannot mandate content moderation without raising significant constitutional concerns.

At the state level, several legislatures have explored or enacted laws targeting deceptive election practices or harmful falsehoods, but many such statutes face First Amendment challenges when they stray beyond narrow categories of unlawful deception. Broad obligations to restrict or monitor misinformation are especially vulnerable to constitutional scrutiny. As a result, most measures to combat misinformation and disinformation – such as fact-checking programmes, labeling, downranking, or coordinated takedown efforts – are the product of platform policies, industry initiatives and civil-society partnerships, rather than legally mandated requirements. Where regulation exists, it is typically narrowly focused on fraudulent or deceptive commercial conduct, medical claims or election-related interference, reflecting the strong constitutional protection afforded to false speech.

Law stated - 16 December 2025

Notice and takedown

Is there a legislative 'notice and takedown' mechanism or similar in your jurisdiction? If so, how does it operate?

Yes. The United States has several sector-specific notice-and-takedown regimes, but no general statutory obligation to remove all unlawful or harmful online content upon notice. The most well-established notice-and-takedown regime is outlined in the Digital Millennium Copyright Act (DMCA), which provides safe harbour to online service providers that follow the notice-and-takedown procedures in 17 USC section 512. A second, narrower mechanism applies to CSAM – under 18 USC section 2258A, providers that obtain knowledge of apparent CSAM must report it promptly to NCMEC. Although this is not a comprehensive notice-and-takedown framework, in practice, online service providers typically remove CSAM immediately upon notice to avoid criminal exposure and to comply with industry standards. Beyond these categories, federal law imposes no statutory notice-and-takedown process for defamation, harassment, hate speech, extremist content, misinformation or other harmful-but-lawful content. Section 230 of the Communications Decency Act reinforces this by shielding intermediaries from liability for most user-generated content and by preventing courts from imposing generalised monitoring or takedown obligations. As a result, decisions

about responding to user notices in these areas are typically left to platform discretion, guided by terms of service, community standards and voluntary industry programs.

Notably, the US is seeing the emergence of targeted notice-and-takedown regimes for certain other types of content – most prominently, non-consensual intimate images (NCII). At the federal level, the recently passed Take It Down Act requires covered platforms to establish a process for individuals to notify the platform of non-consensual intimate images and to request their removal. Upon receiving a valid request, platforms are required to remove the image within 48 hours and make 'reasonable efforts' to identify and remove any identical copies of such depiction. Emerging state laws have begun to mirror or complement this approach. For example, Texas and Florida recently enacted NCII and sexual-deepfake statutes that impose platform-facing removal obligations once a victim provides notice.

Law stated - 16 December 2025

ENFORCEMENT AND PENALTIES

Enforcement
How is the online safety regime enforced in your jurisdiction?

Because the US lacks a unified online safety framework, enforcement is dispersed across multiple federal and state actors, each with authority over a narrow slice of conduct. As a result, the US enforcement model is sectoral, enforcement-driven and primarily reactive, rather than a centralised supervisory regime.

At the federal level, the DOJ and the FBI enforce criminal statutes related to child sexual abuse material (CSAM), sex trafficking, terrorism, extortion and other unlawful online conduct. These laws operate through traditional criminal investigations, subpoenas, search warrants and prosecutions. Providers that fail to comply with CSAM reporting obligations or that knowingly facilitate criminal activity may face significant penalties.

The Federal Trade Commission (FTC) serves as the closest analogue to a general online safety regulator. Using its authority under section 5 of the FTC Act, the agency investigates and prosecutes companies for 'unfair' or 'deceptive' practices, including misleading safety representations, inadequate protections for minors, manipulative design features and data practices that elevate safety risks. Enforcement may result in consent orders, mandatory design changes, third-party assessments, data- and model-deletion requirements, and monetary penalties.

The Children's Online Privacy Protection Act (COPPA) is enforced exclusively by the FTC and state attorneys general. Penalties can include substantial civil fines, limits on data collection, ongoing monitoring and mandated privacy safeguards for services directed to or knowingly used by children under 13.

At the state level, enforcement of emerging youth-safety and age-verification laws is carried out by state attorneys general, consumer protection divisions or dedicated regulatory offices. Remedies may include civil penalties, injunctive relief and restrictions on platform operations within the state. Several of these laws, however, are actively litigated on First Amendment and Dormant Commerce Clause grounds, which has paused or limited enforcement in some jurisdictions.

Private litigation also plays a role, although section 230 significantly limits civil liability for user-generated content. Plaintiffs are increasingly pursuing alternative theories such as product liability, negligent design or deceptive practices claims, but courts often scrutinise such claims closely in light of the robust constitutional and statutory protections afforded to online intermediaries.

Law stated - 16 December 2025

Authorities

Which authorities are responsible for enforcement? What is the basis, nature and extent of their enforcement powers?

Enforcement authority is distributed across several federal and state bodies, each operating under distinct statutory mandates. The FTC is the primary civil regulator for consumer-facing online services. Acting under section 5 of the FTC Act, the FTC enforces prohibitions on unfair or deceptive acts or practices and increasingly applies this authority to online safety issues, including dark patterns, manipulative design, misleading safety or privacy claims, and failures to provide adequate protections for children. The FTC may investigate companies, issue civil investigative demands, obtain injunctive relief, enter into long-term consent orders, and seek civil penalties where statutory penalty hooks exist (eg, COPPA violations). COPPA also grants the FTC explicit authority to enforce children's privacy protections and requires operators to implement reasonable security and data-handling practices for users under 13.

The DOJ – often with assistance from the FBI, US Attorneys' Offices, and other specialised units – handles criminal enforcement of online conduct involving CSAM, trafficking and exploitation, threats and harassment, terrorism-related offences, computer fraud, and other federal crimes. The DOJ's enforcement powers derive from federal criminal statutes, including 18 USC sections 2251–2258A (CSAM), section 1591 (sex trafficking), and sections 2339A–2339B (material support for terrorism). These authorities permit subpoenas, search warrants, arrests, prosecutions, asset seizures and extraterritorial enforcement where a sufficient US nexus exists.

At the state level, state AGs play a central role in enforcing online safety obligations through, ***inter alia***, state consumer protection laws, privacy statutes and emerging online safety legislation. State AGs may investigate deceptive practices, privacy violations, youth-related harms and other online safety issues. Their enforcement powers typically include civil penalties, injunctive relief, and restitution. Relevant regulatory authorities at the state level also have rulemaking and/or enforcement discretion under certain online safety regimes.

Additional authorities may be involved in specific circumstances, including but not limited to the Federal Communications Commission (FCC) for communications-related issues, NCMEC for CSAM reporting intake, state child-protection agencies for harm to a child or consumer-affairs agencies for consumer deception issues. However, the bulk of online-safety enforcement is carried out by the FTC, DOJ and state AGs, each exercising powers tailored to their statutory domains.

Law stated - 16 December 2025

Penalties and liability

What are the potential fines or penalties for non-compliance? Are there risks of liability for employees or directors of online service providers?

The penalties for non-compliance with online safety obligations vary significantly depending on the underlying statutory regime, but they can include civil fines, injunctive relief, long-term compliance obligations, and (in limited circumstances) criminal exposure. For example, violations of COPPA and the TAKE IT DOWN Act are treated as 'unfair or deceptive acts or practices' in violation of section 5 of the FTC Act. Accordingly, the FTC may seek injunctive relief, civil monetary penalties, and other forms of redress for consumers as authorised by the FTC Act. For 2025, the maximum statutory fine per violation is US$53,088 (this number is adjusted annually to account for inflation). In certain circumstances, the FTC may also impose comprehensive conduct and reporting obligations through consent orders.

Criminal statutes may present more serious consequences. For example, FOSTA-SESTA created new criminal and civil liability for covered online platforms. Under 18 USC section 2421A, penalties for owning, managing, or operating a covered online platform with the intent to promote or facilitate prostitution include fines and up to 10 years of imprisonment. Certain 'aggravated' violations for owning, managing, or operating a covered platform with the intent to promote or facilitate the prostitution of five or more persons or acting in reckless disregard of the fact that such conduct contributed to sex trafficking carry fines and penalties of up to 25 years of imprisonment. Offences involving CSAM under 18 USC sections 2251–2258A also carry substantial criminal penalties, including lengthy imprisonment.

At the state level, penalties under youth safety, privacy and consumer protection laws typically take the form of civil fines, which may vary significantly. Some statutes provide for higher maximum penalties where minors are involved or where violations are deemed wilful. State attorneys general may also seek injunctive relief, restitution, or audits of a provider's business practices.

Individual liability for employees or directors is uncommon but possible. In civil matters (eg, FTC enforcement), liability typically attaches to the corporate entity unless an individual directly participated in, controlled, or had knowledge of the unlawful conduct. Criminal statutes also pose limited risk – in particular, executives may face liability if they knowingly facilitate criminal offences (eg, CSAM distribution, failure to report apparent violations and sex-trafficking facilitation). Some emerging state youth safety laws also include provisions that could, in theory, reach individuals who knowingly violate statutory obligations. In practice, however, enforcement actions almost always focus on the organisation unless there is clear evidence of knowing, intentional or wilfully negligent misconduct by a responsible corporate officer.

Law stated - 16 December 2025

DISPUTES

Claims

What claims relating to online safety are available and most common in your jurisdiction?

Because the United States does not have a unified online safety framework, claims in this area may arise through a combination of criminal provisions, consumer protection laws, privacy statutes, and common-law causes of action. As discussed above, civil liability for harmful user-generated content is significantly constrained by section 230 of the Communications Decency Act, which generally bars suits seeking to hold platforms responsible for third-party content. As a result, the claims that do arise typically focus not on the content itself, but on the platform's own conduct, such as allegations of deceptive practices, data misuse, defective or unsafe design, or failures to implement reasonable safeguards.

Consumer protection and deceptive practices claims from the FTC, state attorneys general and private plaintiffs remain some of the most common avenues for online safety enforcement. These actions frequently allege misrepresentations regarding, ***inter alia***, safety features, privacy protections, content moderation practices, parental controls, or the risks posed to minors. Privacy-based claims under the Children's Online Privacy Protection Act (COPPA) or state privacy statutes are also increasingly common.

In parallel, a growing body of teen mental health (TMH) litigation has emerged, with school districts, parents and state attorneys general asserting that social media platforms' product-design choices – such as infinite scroll, algorithmic amplification of harmful content, engagement-driving reward loops or inadequate safety mechanisms – create foreseeable risks to minors' mental health. Plaintiffs are generally framing these suits as product liability, negligence or public nuisance actions in attempts to avoid section 230 by characterising online harms as stemming from defective or unreasonably dangerous design rather than third-party speech. Although certain claims have survived early motions to dismiss, these cases continue to face substantial First Amendment, section 230 and causation hurdles, and no consistent judicial framework has yet emerged. Collectively, they reflect an attempt to use traditional tort principles to fill perceived gaps in the US online safety regime, and may shape future debates about online platforms' duty of care, platform design expectations and broader legislative reform around online safety.

Finally, various tort claims (eg, defamation, online harassment and non-consensual intimate images (NCII)-related claims) may be brought under state law, although platforms generally retain strong immunity under section 230 for content posted by users. As a result, such claims are more commonly pursued against the individual posters rather than the online service provider, further underscoring how US litigation tends to focus on platform conduct and design choices as opposed to specific instances of harmful content.

Law stated - 16 December 2025

Procedure

What is the procedure for claimants to bring actions relating to online safety in your jurisdiction?

The procedure for bringing claims depends on the type of legal theory asserted and the forum in which the action is brought. Most online safety-related actions proceed through ordinary civil or administrative procedures, while criminal matters follow traditional prosecutorial processes. Section 230 may significantly shape these procedures by limiting many direct claims against intermediaries.

Law stated - 16 December 2025

Remedies

What interim and substantive remedies may be imposed in relation to online safety claims?

Interim (or provisional) remedies that are available through ordinary US civil and administrative mechanisms may include:

- Preliminary injunctions: courts may order a company to halt specific practices – such as misleading safety representations, harmful design features, or unlawful data collection involving minors – while litigation is pending. Preliminary injunctions require showing likelihood of success on the merits and irreparable harm.

- Temporary restraining orders (TROs): in urgent cases, courts may issue short-term orders restricting conduct before a hearing; for example, this might occur in FTC enforcement or state AG actions involving alleged deceptive or harmful practices directed at minors.

- Preservation and production orders: courts may require platforms to preserve or produce data, content or internal documents relevant to ongoing litigation or investigations (eg, in CSAM or trafficking matters).

- Administrative interim measures: the FTC may impose temporary compliance obligations during investigations through stipulated orders or through federal court filings seeking emergency relief.

Substantive remedies are more varied and depend on whether the case involves regulatory enforcement, civil litigation, or criminal prosecution. They may include injunctions (which are usually narrowly tailored due to First Amendment limitations), damages (eg, statutory damages under DMCA/COPPA, compensatory damages in tort), consent decrees (imposed by the FTC or state AGs) or civil penalties under state and federal statutes.

Law stated - 16 December 2025

Defences and exemptions

Does your jurisdiction provide any defences or exemptions from liability for online safety claims? If so, how do they operate and which online services providers may avail of them?

US law offers some of the strongest liability protections in the world for online service providers and these defences shape virtually every aspect of online safety litigation. The most powerful of these is section 230 of the Communications Decency Act, which broadly

shields providers of interactive computer services from civil liability for third-party content. Section 230 immunity generally extends to hosting, distributing, ranking and recommending user-generated content, making it one of the most consequential legal protections available to online platforms. In practice, section 230 may also prevent plaintiffs from reframing harmful-content claims as negligence, product design defects, emotional distress theories or failures to protect minors; various courts have dismissed such claims at the pleading stage where they ultimately hinge on the content of third-party speech. Nonetheless, this remains an evolving and contested area, and some recent decisions have shown an openness to design or conduct-based theories that purport to seek liability independent of user content.

Section 230 does contain statutory exceptions – immunity generally will not apply to suits brought under federal criminal law, intellectual property law, any state law 'consistent' with section 230, certain privacy laws applicable to electronic communications or certain federal and state laws relating to sex trafficking.

Layered on top of section 230 is the First Amendment, which functions as a structural constraint against many regulatory attempts to impose safety-driven content restrictions. Because the First Amendment protects both users' speech and platforms' editorial discretion, courts have struck down or enjoined state laws that require age verification, access restrictions or the removal of lawful but harmful content. These constitutional constraints mean that authorities are limited in their ability to compel online platforms to moderate lawful content, redesign their platforms or adopt particular safety standards. As a result, many of the more ambitious online youth safety statutes remain tied up in litigation, and some have already been narrowed or fully blocked.

Certain content-specific safe harbours also provide important protections. The DMCA shields platforms from copyright liability when they implement a compliant notice-and-takedown system, while COPPA offers a safe-harbour mechanism for operators that participate in FTC-approved self-regulatory programmes for children's privacy. Federal child safety law also grants immunity to providers that report suspected child sexual abuse material to NCMEC in good faith, ensuring that reporting does not expose them to civil liability. Section 230 itself includes an additional safe harbour for the good-faith removal or restriction of objectionable material, reinforcing that most platform safety interventions in the US are discretionary rather than legally compelled.

Taken together, these defences create a liability environment that is far more protective of intermediaries than that found in many other jurisdictions. Platforms can generally avoid civil exposure for harmful third-party content, enjoy strong constitutional protection against compelled moderation and benefit from targeted safe harbours.

Law stated - 16 December 2025

UPDATE AND TRENDS

Key trends and future developments

What are the most noteworthy recent trends and developments in online safety regulation in your jurisdiction? What developments are expected in the coming year?

The current US online safety landscape is evolving quickly, but the long-term stability of emerging laws and regulations remains uncertain. Courts are drawing First Amendment boundaries around what states may require of online services, resulting in a wave of constitutional challenges and injunctions to newly enacted youth safety statutes. Despite this judicial pushback, state legislatures continue to advance ambitious online safety frameworks, including age assurance requirements, design restrictions, parental consent requirements, transparency duties and heightened protections for minors. Federal regulators are moving in parallel; for example, the Federal Trade Commission (FTC) has intensified scrutiny of engagement-driving mechanics, recommender systems and other youth-related design practices, signalling an increasing willingness to treat perceived manipulative or unsafe design choices as potentially unfair or deceptive practices under section 5 of the FTC Act.

At the same time, Congress continues to debate whether a national online safety regime is viable. Although proposals such as the Kids Online Safety Act (KOSA) and various AI and privacy bills remain under discussion, little concrete progress has been made toward enacting a comprehensive federal framework on these issues. In the absence of such legislation, states are expanding further into the online safety arena, including via laws targeting the proliferation of non-consensual intimate images (including AI-generated deepfake), design choices that encourage addictive or compulsive use by minors and harmful content disseminated by AI chatbots.

Layered on top of this increasingly complex patchwork is section 230 of the Communications Decency Act, which continues to shape the outer limits of platform liability by shielding services from most civil claims arising from third-party content. Although section 230 continues to provide broad immunity for many such claims, its legal contours are becoming more contested. Certain courts have shown a willingness to entertain theories that frame harms as stemming from platforms' design choices (including, for example, algorithmic recommendations or engagement mechanics) rather than platforms' hosting of third-party content, narrowing the space in which section 230 clearly applies. In parallel, various legislators have suggested there is a need to amend or carve out more exceptions to section 230. While no major reforms have yet been enacted, the combined effect of these judicial developments and legislative interest suggests that the scope of platform immunity under section 230 may continue to evolve, introducing additional uncertainty into the already complex and fragmented US online safety landscape.

Law stated - 16 December 2025