China Launches Special Campaign to "Clear Up and Rectify the Abuse of AI"

Article By:

Aaron Wininger

On April 30, 2025, China's Cyberspace Administration (CAC) launched a 3-month campaign to "clear up and rectify the abuse of AI technology" including using information that infringes on others' intellectual property rights, privacy rights and other rights. Per the Cyberspace Administration, "the first phase will strengthen the source governance of AI technology, clean up and rectify illegal AI applications, strengthen AI generation and synthesis technology and content identification management, and promote website platforms to improve their detection and identification capabilities. The second phase will focus on the abuse of AI technology to create and publish rumors, false information, pornographic and vulgar content, impersonate others, engage in online water army [paid posters] activities and other prominent issues, and concentrate on cleaning up related illegal and negative information, and deal with and punish illegal accounts, multi-channel networks (MCNs) and

中央网信办部署开展"清朗·整治AI技术滥用"专项行动

2025年04月30日 17:00　　来源： 中国网信网　　　　　　　　　　　　　　　　　　　　【打印】【纠错】　　　　　　　　　　　　A⁻ A⁺

Per the CAC, in the first phase, the focus is on rectifying six prominent problems:

- First, illegal AI products by failing to perform large model filing or registration procedures. Providing "one-click undressing" and other functions that violate laws and ethics. Cloning and editing other people's voices, faces and other biometric information without authorization and consent, infringing on other people's privacy.
- Second, teaching and selling illegal AI product tutorials and products. Teaching tutorial information on how to use illegal AI products to forge face-changing videos, voice-changing audio, etc. Selling illegal "speech synthesizers" and "face-changing tools" and other product information. Marketing, hyping, and promoting illegal AI product information.
- Third, lax management of training corpus. Using information that infringes on others' intellectual property rights, privacy rights and other rights. Using false, invalid, and untrue content crawled from the Internet. Using data from illegal sources. Failure to establish a training corpus management mechanism, and failure to regularly check and clean up illegal corpus.
- Fourth, weak security management measures. Failure to establish content review, intent recognition and other security measures that are commensurate with the scale of business. Failure to establish an effective illegal account management mechanism. Failure to conduct regular security self-assessments. Social platforms are unclear about the AI automatic reply and other services accessed through API interfaces, and do not strictly control them.
- Fifth, the content identification requirements have not been implemented. The service provider has not added implicit or explicit content identification to deep synthetic content, and has not provided or prompted explicit content identification functions to users. The content dissemination platform has not carried out monitoring and identification of generated synthetic content, resulting in false information misleading the public.
- Sixth, there are security risks in key areas. AI products that have been registered to provide question-and-answer services in key areas such as medical care, finance, and for minors have not set up targeted industry security audits and control measures, resulting in problems such as "AI prescribing", "inducing investment", and "AI hallucinations", misleading students and patients and disrupting the order of the financial market.

The second phase focuses on rectifying seven prominent problems:

- First, using AI to create and publish rumors. Fabricating all kinds of rumors and information involving current politics, public policies, social livelihood, international relations, emergencies, etc., or making arbitrary guesses and malicious interpretations of major policies. Fabricating and fabricating causes, progress, details, etc. by taking advantage of emergencies, disasters, etc. Impersonating official press conferences or news reports to publish rumors. Using content generated by AI cognitive bias to maliciously guide.
- Second, using AI to create and publish false information. Splicing and editing irrelevant pictures, texts, and videos to generate mixed, half-true and half-false information. Blurring and modifying the time, place, and people of the incident, and rehashing old news. Creating and publishing exaggerated, pseudo-scientific and other false content involving professional fields such as finance, education, justice, and medical care. Using AI fortune-telling and AI divination to mislead and deceive netizens and spread superstitious ideas.
- Third, using AI to create and publish pornographic and vulgar content. Using AI stripping, AI drawing and other functions to generate synthetic pornographic content or indecent pictures and videos of others, soft pornographic, two-dimensional borderline images such as revealing clothes and coquettish poses, or ugly and other negative content. Produce and publish bloody and violent scenes, distorted human bodies, surreal monsters and other terrifying and bizarre

images. Generate synthetic "pornographic texts" and "dirty jokes" and other novels, posts and notes with obvious sexual implications.

- Fourth, use AI to impersonate others to commit infringement and illegal acts. Through deep fake technologies such as AI face-changing and voice cloning, impersonate experts, entrepreneurs, celebrities and other public figures to deceive netizens and even market for profit. Use AI to spoof, smear, distort and alienate public figures or historical figures. Use AI to impersonate relatives and friends and engage in illegal activities such as online fraud. Improper use of AI to "resurrect the dead" and abuse the information of the dead.
- Fifth, use AI to engage in online water army [paid posting] activities. Use AI technology to "raise accounts" and simulate real people to register and operate social accounts in batches. Use AI content farms or AI to wash manuscripts to batch generate and publish low-quality homogeneous writing to gain traffic. Use AI group control software and social robots to like, post and comment in batches, control the volume and comments, and create hot topics to be listed.
- Sixth, AI products, services and applications violate regulations. Create and disseminate counterfeit and shell AI websites and applications. AI applications provide illegal functional services, such as creative tools that provide functions such as "expanding hot searches and hot lists into texts", and AI social and chat software that provide vulgar and soft pornographic dialogue services. Provide illegal AI applications, generate synthetic services or sell courses, promote and divert traffic, etc.
- Seventh, infringe on the rights and interests of minors. AI applications induce minors to become addicted, and there is content that affects the physical and mental health of minors in the minor mode.

The original text is available [here](#) (Chinese only).

Source URL:https://natlawreview.com/article/china-launches-special-campaign-clear-and-rectify-abuse-ai