AI slop: truth is losing the algorithm

Wendy Sloane

What happens when we are unable to distinguish between what's fake and what's real?

I admit my online reading habits are a little unconventional. I'm a longtime admirer of conjoined twins Abby and Brittany Hensel, one of whom recently got married, and I'm fascinated by the resilience and determination of 26-year-old Loren Shauers, whose body was cut in half following a grisly forklift accident in Montana. After years writing "triumph over adversity" articles for women's magazines, I'm often drawn to bizarre stories that don't captivate others.

That may explain why my social media feeds are packed with an endless stream of extraordinary human-interest stories, accompanied by striking videos and images. Increasingly, though, it is impossible to tell which are genuine and which are AI-generated. According to Ofcom, more than half the population now gets its news from social media, rising to around 75 per cent among 16 to 24-year-olds. At the same time, those platforms - TikTok, Instagram and YouTube - are being flooded with a new kind of content: "AI slop", fake images, videos and posts designed to look authentic, provoke strong emotional reactions to maximise – and monetise - online engagement, regardless of their authenticity.

This flood of fabricated AI content is produced "without judgement, care, authorship, responsibility or any meaningful relation to truth," writes David Webster, Professor of Philosophy and Pedagogy at Buckinghamshire New University. Much of it exists for one purpose: to generate clicks, shares and advertising revenue. But, he argues, the problem extends beyond "uncanny images of a many-fingered Jesus that someone's grandparent shares on Facebook". It also includes SEO articles, fake reviews and plausible-sounding blog posts that quietly reshape how information is consumed.

"The danger is not only that AI-generated material might be wrong. The danger is that it may help produce an environment in which the distinction between thought, content, answer, summary, position and argument becomes increasingly blurred," he writes. "It is a nihilistic attitude to actual truth."

Webster's point is that AI slop does not simply spread falsehoods; it changes the environment in which we encounter information. When journalism, satire, advertising, propaganda and AI-generated fiction all appear side-by-side in the same endless feed, the boundaries between truth and fabrication begin to blur. The result is a paradox: we may become less able to recognise falsehoods while also becoming less willing to believe anything at all - even the truth.

The collapse of trust

My own social media feeds offer a revealing case study of how quickly these systems learn what captures our attention - and how readily AI-generated content exploits it. Knowing my fascination with eclectic human-interest stories, platforms now regularly serve me up a smorgasbord of fabricated photos of married couples with Down's syndrome with (improbably) five or six children, and surreal reels featuring glamorous conjoined twins with three or four heads, who (naturally) resemble Playboy models from the neck down. Then there are the surprisingly realistic videos of beautiful female amputees beckoning viewers to take a closer look, alongside cruel AI-generated clips of extremely overweight people breaking chairs, capsizing boats or cracking swimming pools beneath their weight while onlookers laugh in amusement and disgust.

Some AI slop is obviously absurd. But its success doesn't depend on being consistently believable. It depends on keeping us scrolling, reacting and sharing. The more convincing it becomes, the harder it

is to separate authentic reporting from synthetic fiction. Even when users suspect the content is AI-generated - or it is clearly labelled as such - millions still engage with it, either treating it as God's honest truth or sharing it as entertainment that, perhaps unsurprisingly, confirms their own bias.

The stakes become considerably higher when AI slop moves beyond bizarre entertainment and begins to shape public life. Political deepfakes are among the most powerful examples: the widely circulated image of Trump depicted as Jesus on Truth Social, shared not long after he posted a video portraying the Obamas as apes; a fabricated video of Volodymyr Zelenskyy surrendering to Putin, and fake images showing a York Labour politician handing out cash to men in balaclavas. Other manipulated photos showed Boris Johnson and Nigel Farage in hospital beds - separately, I hasten to add - alongside printed signs urging citizens to pray for their recovery.

Every convincing deepfake makes authentic photographs easier to dismiss, credible reporting easier to question and legitimate evidence easier to undermine. Making matters more complicated, AI's large language models, or LLMs, are becoming a major source of "truth", although they are trained on an information environment that is already distorted. In a world where anything can be fabricated, everything can be denied - threatening not only politics and journalism, but also victims of crime, public trust and even our collective memory of major events.

The consequences of AI slop are no longer confined to politics or internet hoaxes. The same technology used to generate harmless fantasy and political misinformation is increasingly being weaponised against individuals - even children. Police recently warned UK parents to share images of their children only within "close friends" groups after criminals used photographs of fully clothed children to create explicit deepfake child pornography. This included the case of a 15-year-old girl who told Childline that someone had created fake nude images of her after stealing photos from her Instagram account, the BBC reported.

The threat extends beyond deliberate abuse: even professionals can be deceived by increasingly convincing AI-generated content. "I recently saw a video of a monkey stealing someone's umbrella, and I had no idea it was AI until the monkey flew away on it," Conrad Pope, a 25-year-old freelance digital content strategist, told me. "AI is getting so good now that I get why people are tricked by it."

For Pope, who said his generation are "guinea pigs" for social media, the problem is compounded by the way the online environment is increasingly designed to reward attention over accuracy. "So much news posted to social media isn't cited as coming from anywhere. It's just a big headline – and big headlines trend, because that's how the social media algorithm works."

Those algorithms reward certainty and emotion, meaning people who project confidence - regardless of whether their claims are accurate - can rise quickly, in some cases becoming mouthpieces for a whole generation. Figures such as Andrew Tate and Joe Rogan have built large audiences in part by presenting themselves as authoritative voices across a wide range of subjects, through sheer algorithmic determination.

"People are becoming fonts of knowledge without having done any research," Pope says. "Charisma travels. If you say something with enough confidence or say something that is shocking, the algorithm will promote it because engagement is higher."

Beyond misinformation

Experts in media literacy, psychology and AI ethics warn that the impact of AI slop extends beyond misinformation: for a generation raised on algorithmically generated spectacle, it is changing the way they process information. Layla Marwat, a 21-year-old third-year International Relations and Politics student at Oxford Brookes University, says she doesn't always know what is real or what is fake on

the internet. Often, she says, authentic creators who don't rely on AI are paying the price of the public's confusion.

"There are so many influencer accounts that if I didn't know who they were I would just automatically think they were AI now as they're just too perfect," she said in an interview. "People are now AI-generating their bodies like crazy and making them look completely unreal. Some of them I know are real people, only because I saw them on YouTube when I was 13."

If social media already shapes how young people understand the world, AI-generated content raises a deeper question: what happens when our feeds become increasingly detached from reality? For some experts, this is where "brain rot" begins. Monica Black, a life coach and clinical hypnotherapist specialising in depression and trauma, believes AI slop is "frying" young people's brains, and argues that constant herd exposure to low-quality, emotionally charged content is reducing their ability to think critically and engage in meaningful discussion.

"They get all their information from social media and spend their time just sitting on their phones, exchanging videos," she told me. "Because everyone is swapping the same stuff, they believe the fake stuff is real." The proliferation of fake images is desensitising young people to genuine issues, not only normalising cruelty but encouraging and amplifying it, while dulling our capacity for empathy, she adds.

"Youngsters see these huge people with rolls of fat pouring off them, which is definitely AI, and then in a blink of an eye they shed it. You see them on the train watching these things and giggling. We should pull these kids up on this behaviour, but we don't," she says. "We are also becoming desensitised to their reactions, when we should be pointing it out to them."

Regulation or re-education?

The UK's planned social media restrictions for under-16s could help curb the spread of AI-generated misinformation. But young people are not the only ones being influenced by AI slop. Dr Pawel Gasiorowski, a lecturer in Computer Science and Applied Computing at London Metropolitan University, warns that the flood of synthetic content is fuelling increasingly extreme worldviews across all age groups by allowing by tailored propaganda to spread faster, evade regulation and escape scrutiny.

"Slopaganda" uses highly targeted memes, deepfakes and other AI manipulations to trigger outrage and rally political support. "Young people are being seduced by the suggestion of fakes," Dr Gasiorowski told me. "They don't have many people to look up to, so they look up to people online. That's how extremists come about." A 2021 analysis of 6,000 individuals in the Middle East found those who used the internet were more likely to support ISIS than those who did not, while a 2024 study found that those who already were ISIS supporters use the internet to enforce their views.

What are the tech giants doing about it? Not much, it appears. Meta CEO Mark Zuckerberg has announced plans to make it even easier for creators to publish AI content, while YouTube CEO Neal Mohan revealed in December that more than one million of his channels were using AI to make content, the BBC reported. Elon Musk's Grok chatbot, for its part, has been marketed as less restricted than many rival systems – a design choice that has fuelled repeated controversy.

Media literacy experts argue that the best defence may not be better AI detectors, but better digital habits: pausing before sharing emotionally charged posts, checking whether stories appear in multiple reputable outlets, and tracing claims back to their original source. Yet these habits run directly against the incentives of social media, where speed, outrage and novelty are rewarded far more than caution.

The burden therefore falls not only on users but also on the platforms that profit from engagement, whether fuelled by truth or fiction. Technology will continue to improve, and regulation will struggle to keep pace.

My feed is still full of children bravely battling terminal illnesses and people overcoming facial differences. But now I find myself looking twice - sometimes three times - to work out whether the images and interviews are genuine. A recent video of a two-month-old baby with a tumour covering one side of his head looked authentic at first glance. Only on closer inspection did it become clear that it was generated by AI: the blurred, oversized growth had been digitally grafted onto footage of a healthy child.

The greatest danger of AI slop, however, is not that it fools us once. It's that it trains us to stop believing anything at all. In an internet where everything can be faked, nothing has to be proved. And when engagement matters more than evidence, truth doesn't simply lose the argument. It loses the algorithm.

Dr Wendy Sloane is deputy dean of the School of Computing and Digital Media at London Metropolitan University, an associate professor, and the journalism course leader. A former Moscow correspondent, she also was deputy editor of Woman's Own and Eva magazines, and deputy features editor at Marie Claire.
@wendyutah