

# TechScape: ‘Lives are ruined in an afternoon’ – social media and the Huw Edwards story

Publication Date: 2023-07-18

Author: Chris Stokel-Walker

Section: Technology

Tags: Technology, TechScape newsletter, X, Internet, Blogging, Digital media, Huw Edwards, Privacy, features

Article URL: <https://www.theguardian.com/technology/2023/jul/18/lives-are-ruined-in-an-afternoon-how-social-media-shaped-the-huw-edwards-story>



Social media imploded and the BBC practically ate itself last week as the scandal over Huw Edwards allegedly paying for explicit images from an unnamed young person unspooled. But what you knew, and when, depended largely on where you looked. Consume only traditional media – television, radio and newspapers and news websites like the Guardian – and you would not have had much inkling of who was in the frame until Edwards’s wife named the BBC News presenter as the one at the centre of the storm. Sniff around social media, however, and you likely knew who was involved days before – and probably also thought a lot less of other names bandied about in connection with the concern. How Twitter’s trending tab works One former member of Twitter’s curation team, who asked not to be named, believes the failure on Twitter’s part was down to a combination of short-staffing and tech changes since Elon Musk took over. This former staff member, who was let go as part of Musk’s mass layoffs after he bought the company in October 2022, said that when they worked at Twitter, the platform maintained a blocklist of terms that would not be allowed on to the trending page. The blocklist was designed for issues like the one faced last week with Edwards: without it, potentially libellous mentions could end up in Twitter’s trending tab, as happened (thanks, Elon). Twitter’s trending tab used to work by highlighting “canonical terms” – or the main term people are talking about. Early in the week, that would have been “BBC presenter”. Curation team members could then append related terms to that canonical one, shown as smaller text beneath the main one, though Twitter’s algorithm will also automatically do that when it experiences high traffic around a certain set of terms. Curation staff can also ensure names or terms don’t get attached by attaching it to a blocklist. “Obviously mistakes happen and the tech fucked up sometimes, and you’d have a word slip through the blocklist, but when that happened, because we had a team of people, we could manually quickly get to it,” they say. “It was never left up there too long.” Back in November, Musk defended mass layoffs at Twitter, including those involved in content moderation, tweeting that the company had fired 15% of its Trust & Safety staff, “with our front-line moderation staff experiencing the least impact”, and adding: “Again, to be crystal clear, Twitter’s strong commitment to content moderation remains absolutely unchanged.” Last month, Musk said on French TV that Twitter will follow the EU’s content moderation rules and claimed that hate speech on the platform had decreased 30% during his tenure, despite being presented with statistics that suggested the opposite during the interview, Politico reported. A load of ‘nonce’-ense – and the rise of other inappropriate terms One of the more unpleasant facets of social media is that it reflects our society like a funhouse mirror, amplifying all the bad bits. Thus, “nonce” trending on Twitter last week. First: it’s worth noting again, for the record, that the Sun appears to have walked back its core allegation that Edwards paid for images from

an underage contact, and there is no evidence he has done anything illegal. The police have concluded the News at Ten presenter has no criminal case to answer. Second: the word “nonce” is not only deeply unpleasant but also inaccurate. It shouldn’t have trended, the former Twitter employee suggests, because it was on a blocklist when they were at the company – not that long ago. “It shouldn’t be there, which tells me something has degraded in the backend,” they say. Even if a less contentious phrase was trending, its connection with Edwards could be manually broken, the former Twitter employee told me. That it wasn’t suggests how thinly stretched staffing is at Twitter now. Employees at the social media platform also had the ability to reduce visibility on terms if they were legally contentious for eight hours at a time. “It shouldn’t happen,” says the former Twitter employee. “What we were doing before meant we were taking action. Even if there were cause for people to try to sue us, we could probably prove we had gone to a lot of steps to stop that from happening – which could sway in favour of us overall.” “I don’t know what their excuse is [now],” says the former staffer. Twitter did not respond to a request for comment when I put these allegations to them last week for a separate story on the matter. A collective failure In some respects, singling out Twitter is unfair: it was a collective failure of social media. People were able to name Edwards as the BBC presenter with impunity in social media comment sections. TikTok suggested Edwards and other BBC presenters’ names as “hot” search terms, appending the fire emoji to their names. Google showed news stories and videos about the then-unnamed BBC presenter to people who searched for Huw Edwards’ name, connecting him to the scandal. “Several tech platforms had already, based on the ‘wisdom of the crowd’, named Huw Edwards in their own ways,” says Dan Barker, an SEO expert and tech consultant, who monitors the pulse of public opinion through digital means. “Google and Twitter seem to have put a lot of work in over the years to try and stop their algorithms from inadvertently naming people in situations like this, but even so, a Google search for ‘BBC presenter’ resulted in the top suggested ‘related search’ of ‘Huw Edwards’, as well as suggesting people search for ‘who is the bbc presenter reddit’, and other similar phrases,” he says. (TikTok declined to comment on the claims; Google did not respond.) Edwards’s own actions were also potentially revealing when he deactivated his Instagram account, tipping off those watching that something was up. Solving the problem The last week was far from edifying for anyone. As the scandal and speculation roiled on social media, I spoke to David Auerbach, a former Google and Microsoft engineer, who has written a book called Meganets. The conversation was for a podcast I’ll soon be hosting for the freedom of expression group Article 19. (The Observer also spoke to Auerbach in March about the problems with big tech.) Auerbach and I didn’t directly address the Huw Edwards situation: I figured a scandal about a UK news presenter being dragged through the mud was a little too parochial for the US-based software engineer to have picked up on. But what he said struck me as emblematic of the issue: social media was designed by maths and engineering graduates who didn’t account for what happens when you let humans on to the apps they build. I also thought of another conversation I had that week, with David Banks, a media law trainer and consultant – who appeared near-constantly on TV to carefully outline the legal story as it evolved. “At the moment,” he told me, “we’ve got a situation where privacy law in this particular type of situation – where you’ve got someone really famous, really well known – just doesn’t work.” We spoke after Edwards had been officially identified, and the chaos began to subside somewhat. Banks was gloomy about this cycle happening again. “Social media moves so fast and the mob has seemingly so little care for the consequences of what it does, that lives are ruined in the space of an afternoon,” he said. “If we think there should be a privacy law as a society, then why are we sitting here pretending social media isn’t happening?” The week in AI Elon Musk claimed his xAI development would be more trustworthy than OpenAI and Google. What AI researchers think of that can be found here (£). The AI revolution puts skilled jobs at highest risk, the OECD says. Meanwhile, capitalism is unlikely to end with a robot rebellion, argues Jeff Sparrow. Google Bard is facing questions about its overreliance on Daily Mail stories to check its large language model is working. The US Federal Trade Commission is investigating OpenAI about, among other things, where its training data comes from. In the UK, financial firms must boost protections against AI scams, the regulator is to warn. People are turning to chatbot impersonations of lost loved ones to help them grieve. This long read asks: will AI help us live after we’re dead? Fake reviews: can we trust what we read online as use of AI explodes? A ChatGPT rival has launched a chatbot that can summarise a novel. If you want to read the complete version of the newsletter please subscribe to receive TechScape in your inbox every Tuesday.