



AI or Reality?

www.euawareness.org

#KnowledgeAgainstManipulation

AI or Reality? #5

Tea Parulava

An Institute Built to Be Cited

How a Russian operation used AI to fake a think tank that never existed

The opinions and conclusions expressed are those of the author and do not necessarily reflect the views of the EU Awareness Centre.

© [2026] EU Awareness Centre. All rights reserved

14 October 2026

An institute built to be cited

Every disinformation case we've looked at to date in our "AI or Reality?" series shares one thing: a fake message at its heart. A fake newspaper article. A fake video. An image which looks like it's from a real moment and isn't. The manipulation lies in the content, and as soon as the content is revealed, the manipulation ends.

A case OpenAI disclosed on 25 August 2026 works differently, and it deserves attention precisely because of that difference. It did not fake a message. It faked the messages' source.

Anatomy of a fake think tank

OpenAI's investigation began, by its own account, with something narrow: a look into a cluster of AI-generated social media posts. What it found underneath was a fully constructed research institute.

It called itself the International Burke Institute (IBI) and claimed to be an independent expert community based in Israel, complete with a street address and a public-facing website registered in February 2025. Its materials claimed links to well-known international relations figures, including Francis Fukuyama and Noam Chomsky. The site made these claims without any confirmation from the scholars themselves.

The institute's flagship product was the Burke Sovereignty Index (a rather slick-sounding index of national sovereignty). An index has a different kind of authority than an opinion piece. It's based on method, data and comparison rather than argument. The Sovereignty Index used that method to reach conclusions that consistently favoured Russia and opposed the West. The reports would be full of attacks on "Macronism" and the French economy, calls for Germany to pursue "real military sovereignty rather than dependence on American bases", and descriptions of EU member states as the "Hostage Countries of the 'Big' Union". France, Germany and the EU were not just in the crosshairs. They were the point.

Built on stolen work

Institutes need publications, and IBI had dozens of articles carrying the trappings of academic output. OpenAI sampled 36 of them, covering September 2025 through May 2026, and found that 34 had been copied from elsewhere. Some traced back to Cambridge University Press. One was Migration Policy Institute research, reattributed in IBI's version to a real Australian food chemistry professor who had no part in it.

This is the part worth sitting with. Manufacturing the appearance of expertise this way is inexpensive: take real scholarship, strip its real authorship, attach a false byline, and publish it under a fabricated institutional letterhead. The intellectual labour is real. Only the source is not.

The tell hidden in a language model's own output

OpenAI don't allow access to their models from Russia, so the operators used VPNs to get around that. They then prompted the models in Russian, asking them to generate English output and mask any trace of Russian. It was a deliberate attempt at origin laundering.

It didn't work in full. One of the surviving phrases referred to Germany's former 'traffic light' coalition as the 'Svetofofor coalition'. Svetofofor is the Russian word for 'traffic light' (a direct translation of the informal German term for that coalition). No native English or German speaker would write that phrase. It is a clear fingerprint of machine translation from Russian, a rare and concrete piece of forensic evidence in an influence operation. On the totality of the evidence, OpenAI concluded the accounts very likely originated in Russia, though it didn't name a specific actor.

The output itself did not stay contained to one website. Promotional material spread across X, LinkedIn, Facebook, Substack, and Telegram. The operation's Telegram channels, aimed at audiences in Germany, the United States, France, Poland, and Türkiye, each drew 10,000 to 20,000 followers.

By the [Brookings Institution's Breakout Scale](#), a standard measure of how far an influence operation has spread beyond its own manufactured audience into authentic ones, OpenAI rated this operation at the lower end of Category Three: multiple platforms, with only some indication of reaching real audiences. In other words, this was not (yet) a large operation in terms of measurable reach. Its importance lies elsewhere.

Why manufacturing a source is more dangerous than manufacturing a message

A fake video can be debunked in an hour and then completely forgotten within a week. A fake institute, on the other hand, is built to survive exposure, since its *raison d'être* was never to be believed at the time. It was to be cited later by someone who wouldn't check the source.

Once a footnote points to the Burke Sovereignty Index, or a policy brief cites an "IBI expert," the fabrication has done its work regardless of whether IBI's own website is still online. Each citation launders it a step further from its origin. This is what separates the case from the disinformation this Centre has documented before, from cloned newspaper domains to fabricated video statements: those operations counterfeited things that already existed. This one manufactured a thing that had never existed at all, purpose-built for other people's future research to lean on without knowing it.

A note on our own position

EU Awareness Centre is itself a small independent research organisation, precisely the profile IBI tried to counterfeit. The difference between a real institute and a manufactured one isn't size or output, but whether the address, authors, and sources are real and checkable. Naming sources, linking to primary material, and correcting the record make an institute verifiable, and they would have exposed IBI to anyone who looked.

How to tell a real institute from a manufactured one

For anyone encountering an unfamiliar "institute," "index," or "expert community" online, a few checks go a long way:

- Can you independently verify its physical address — through public registration records, an actual office, or a domain history that predates its claims?
- Do the experts it names confirm the affiliation themselves, in their own words, rather than only appearing on the institute's own site?

- Is its flagship methodology or index disclosed and reproducible, or only asserted as a conclusion?
- Do its citations trace back to their actual original authors and publications?
- Is it referenced by any independent, verifiable directory of research institutions or does it only reference itself?

None of these checks requires special tools. They require the same five minutes of scepticism this Centre asks of every claim it publishes. In a case built entirely on the appearance of authority, that is usually all it takes to see through it.

Source: OpenAI, "Disrupting a new covert influence campaign from Russia," 25 August 2026.

Tea Parulava

14 October 2026