



AI or Reality?

www.euawareness.org

#KnowledgeAgainstManipulation

AI or Reality? #2

How a Fake Pope Image Fooled the Internet

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

13 September 2026

How a Fake Pope Image Fooled the Internet

In March 2023, an unusual photograph of Pope Francis went viral. It showed the Pope wearing a large white puffer jacket, looking more like a fashion model than the head of the Catholic Church. The image looked convincing, even though it was completely fake.

No photographer took it, and it depicts no real event. It was generated using artificial intelligence, with Midjourney.

What's interesting is that the image wasn't created as a deliberate attempt to deceive people. It appears to have started as a joke, a playful experiment with a new image generation technology. The creator was not necessarily trying to convince the world that Pope Francis had actually been photographed wearing a designer puffer jacket. The problem came later.

Once the image escaped its original context and began circulating across social media, many people encountered it simply as an image. The original joke and the circumstances of its creation became less visible, while the photograph itself travelled much further. In other words, the way the image was shared and consumed online amplified its misleading effect. And that distinction matters. Not every misleading image is created with the intention to mislead. But once synthetic content is detached from its original context, it can still be misleading.

The original image was posted on Reddit under the title "The Pope Drip". From there, it spread far beyond its original context, including other social media platforms. Many people who encountered it had no immediate reason to suspect that it was artificial. The person looked like Pope Francis, the white papal clothing looked plausible, and the composition resembled an ordinary candid photograph. So, nothing immediately screamed that the image was FAKE.

Our brains do not normally analyse every photograph we encounter online. We make rapid judgements based on familiarity, context and likelihood. If an image looks like something that could have happened, it is easy to treat it as something that did actually happen.

How was it made?

The image was generated with Midjourney, an AI image-generation system. Generative AI can create entirely new images from written descriptions. Instead of photographing a real person in a real location, the system generates an image based on patterns learned from vast amounts of visual data, and the result can look remarkably photorealistic.

But in the early stages of development, AI image generators had noticeable weaknesses. Since the Pope image was created in 2023, a closer inspection reveals several inconsistencies. The hand around the cup doesn't quite make sense; the glasses and eye area appear distorted and unnatural; the chain holding the crucifix appears incomplete, missing one part. This reflects some of the weaknesses that early image generators often had in anatomy, objects, and physical reality.

AI image generation has developed rapidly since then. Some of the clues that could help identify an AI-generated image a few years ago are becoming much less reliable. Hands and fingers can look realistic; reflections can be convincing; lighting and shadows can appear physically consistent, and faces can look completely natural.

This means that simply looking for strange fingers, distorted objects, or unnatural faces is no longer enough. The question is increasingly becoming not only "Can I spot something wrong with this image?", but also "Can I verify that this image is real?"

When you encounter a surprising or emotionally powerful image online, the first step is to stop and not share it immediately. Look more closely and examine the details: anatomy, objects, shadows, reflections, and the relationships between different elements in the image. But even if you cannot find anything obviously wrong, don't assume the image is authentic. Try to check its origin. Where did the image come from? Who posted it,

what was the original context and is there an original source somewhere? Also, don't forget to check whether reliable, independent sources confirm that the event actually happened.

In the case of the Pope image, fact-checkers found it was AI-generated, and no real photographic event behind it.

Technology is not the only thing changing. The European Union has also introduced new transparency requirements for AI-generated content under the AI Act. Since 2 August 2026, relevant transparency rules have applied. In certain circumstances, AI-generated content must be marked in a way that allows it to be detected, while deepfakes must be clearly disclosed. Such measures can make it easier for people to understand what they are looking at.

But regulation cannot replace our own judgement. Not everything we encounter online will necessarily be labelled. Content can be copied, altered, reposted or removed from its original context. A piece of content that began as a joke can acquire a completely different meaning once it reaches a much larger audience. The Pope image illustrates this problem well.

The story also shows why intention and impact should not always be treated as the same thing. The person who generated the image may have intended it as a joke or simply as an experiment to hone their AI skills. But that does not mean that everyone who subsequently encountered it understood the joke. Once an AI-generated image is separated from its original context, people may interpret it differently. Some may share it because they believe it is authentic. Others may share it without knowing whether it is authentic. Algorithms can then amplify the content further, reaching audiences who are unaware of its original source.

This does not necessarily make the original creator a propagandist, nor does every instance of accidental misinformation amount to a coordinated manipulation campaign. But the outcome can still be misleading.

That is one of the challenges of the generative-AI era: content does not need to be deliberately created as disinformation in order to contribute to a misleading information environment.

The real lesson

The Pope image was not remarkable because it was a perfect fake. In fact, it contained several imperfections that could be detected on close inspection. What made it significant was that it was plausible enough to pass as a real photograph and that its original context could easily disappear as it travelled online.

Today, when the technology is considerably more sophisticated, you do not need to become an expert in artificial intelligence to respond to this. You need to become a little harder to fool.

When an image makes you think, "Really?", pause before accepting it or sharing it. Look closely, check where it came from and consider its original context. Don't forget to look for independent confirmation as well. Because in the age of generative AI, the most important question may no longer be whether something looks real. It is whether you can verify that it is real.

Think before you trust.



David Dondua

Ambassador David Dondua is a diplomat and expert in international security, conflict resolution, and European integration. During his diplomatic career in the Georgian foreign service (1993–2022), he held key positions, including Ambassador to Austria, Greece, Serbia, the OSCE and NATO. He serves as Chairman of the Board of the EU Awareness Centre.