The Indian EXPRESS

That ‘AI caricature using everything about me’ trend could expose you to digital fraud

A social media trend asking AI tools to create caricatures using the prompt ‘everything you know about me’ is raising cybersecurity concerns. In this week’s The Safe Side, we will talk about the safety concerns about this viral trend.

"Create a caricature with everything you know about me"
THE SAFE SIDE
OUR WEEKLY GUIDE TO DIGITAL SAFETY

A user tries the viral AI prompt “create a caricature using everything you know about me,” a trend cybersecurity experts warn could expose personal data to scammers. (AI-generated Image: DALLE)

A new trend has gone viral on social media, where users share a personal photo and ask Artificial Intelligence (AI) tools to create a caricature or illustration based on their life, their job, and “everything the AI tool knows” about them. The result showing animated versions of the person at the office, with their family, or representing their profession has become popular content on Instagram, LinkedIn, and even X.

While the trend may seem creative and entertaining, cybersecurity specialists warn that this practice can expose personal information and enable the creation of personalised, large-scale fraudulent messages, a very real threat today.

According to Adrian Hia, Managing Director for Asia Pacific at Kaspersky, this type of request does not work like a simple visual filter. “To achieve more accurate images, people allow AI tools to access all the information associated with their profiles without restrictions, since the instruction itself is embedded in the command ‘create a caricature about me and my job based on everything you know about me'. In addition to the reference photo, extra data such as company name, corporate logos, job title, city, daily routines, hobbies, and other family details are often included and used to create the trend," he said.

What kind of personal data is exposed?

"This is the question most people do not think to ask—because the output looks like a cartoon, not a dossier. But pause and consider what goes in. A photo alone reveals your face, approximate age, ethnicity, and often a background that hints at your home or workplace. If you add a prompt like 'everything you know about me,' and the AI draws from whatever you have shared in that session or linked account—your name, job title, employer, family references, city, interests, even your communication style. What comes out the other side is not just a caricature. It is a synthesised personal profile, assembled in seconds, that you willingly handed over," Dr Sanjay Katkar, Joint Managing Director at Quick Heal Technologies Limited, told *indianexpress.com*.

How can sharing information make scams more convincing?

"Combining a person's photograph with contextual details significantly increases the risk of phishing or impersonation scams because it allows scammers to create highly believable identities. If scammers obtain a person's photograph, they can use it to create fake social media or professional profiles across different platforms. The image can also be used to generate voice or video deepfakes, or to clone company or professional profile pages where the victim's photo is used as the display image, making the profile appear authentic," says cybersecurity expert Anurag Mathur.

He adds, "If job or workplace details are available along with the photograph, scammers can exploit this information to craft convincing business emails or messages. By knowing the person's job title, position in the organisational hierarchy, and the company they work for, fraudsters can impersonate the individual and send fake instructions or requests to colleagues or employees within the organisation."

"Similarly, if information about family members is known, scammers may attempt emotional manipulation. They can send urgent messages or calls to relatives claiming that a family member has met with an accident or is in trouble, and demand immediate financial help or ask them to rush to a particular location," Mathur notes.

Hia mentions, "Each of these data points is a key piece in building a detailed digital profile. By combining image, text, and context, habits, relationships, frequently visited places, and professional responsibilities are revealed—information that cybercriminals can exploit to craft more sophisticated scams. As a result, a fraud attempt that mentions where someone works, their job title, or even a family member becomes far more convincing and increases the likelihood that the victim will trust it and share sensitive information or money."

How useful is an AI-generated personal profile to scammers?

"Extremely useful. Cybercriminals usually spend considerable time conducting open-source intelligence gathering (OSINT), piecing together fragments of a person's digital footprint from platforms like LinkedIn, Instagram, news articles, or data breaches. Viral AI caricature trends can effectively do that work for them in one consolidated output. Instead of scraping information from multiple sources, scammers get a ready-made profile containing a face, personality cues, professional context, and personal details," Katkar opines.

Amit Relan, CEO and co-founder of mFilterIt, says, "What makes this trend uniquely dangerous as a fraud channel is the aggregation effect. No single piece of information a user shares is harmful in isolation—but the AI synthesises it all into a consolidated identity profile: your workplace, your anxieties, your relationships, your habits. That synthesis is exactly what a fraudster needs to craft a phishing attack you will believe, an impersonation you will not question. A scam tailored precisely to your blind spots."

Hia says this risk is particularly acute in the APAC region. Despite a high AI adoption rate, with 78 per cent of professionals using AI weekly (surpassing the global average of 72 per cent), many users still struggle with basic technical literacy, leaving them vulnerable to social engineering and phishing.

In addition, when interacting with these platforms, users are not only sharing the final image. Depending on the service and its privacy policies, the original photo, the user's text or instructions, usage history, and certain technical data such as IP address, device, or interaction patterns may also be stored. Part of this information may be retained to operate the service, improve performance, or train AI models, meaning the content does not necessarily disappear after the caricature is generated and may remain longer than users expect.

"This viral trend of caricature creation of our lives may seem like harmless fun, but it is effectively a voluntary briefing for cybercriminals," says Hia. "In a region where AI adoption is leading the world but technical literacy is still catching up, these digital portraits are becoming dangerous maps," adds Hia.

Here's what you can do

While these tools can be a fun way to experiment with digital creativity, experts *indianexpress.com* spoke to recommend adopting more cautious habits:

📍Avoid entering identifiable data in prompts such as your full name, job title, company, city, address, schedules, or daily routines, even if it seems like it is only for personalising the image.

📍 Do not upload photos that show logos, credentials, documents, licence plates, screens, building façades, or any other element that could help identify your location or link you to an organisation.

📍 Avoid uploading photos that clearly show your face along with identifiable details such as children, ID cards, or workplace badges.

📍 Do not share images or information about minors, and avoid revealing family details that could be used to impersonate relatives or create emotional scams.

📍 Do not use prompts such as "everything you know about me." Instead, specify: "Create using only the details in this message and do not use past chats or history."

📍 Crop or remove obvious identifiers in photos before uploading them, such as street signs, certificates, building names, or society name boards.

📍 Avoid including precise geographic information in prompts, for example, "me as a Bangalore engineer working at XYZ Tech Park."

📍 Do not upload photos of other people without their consent.

📍 Disable chat history or data-saving features if the platform provides that option and you do not want your data reused.

📍 Review the platform's privacy policy and permissions before using it, especially regarding how long content is stored and whether it is used for model training or service improvements.

📍 Consider using privacy modes such as 'Temporary Chat', where prompts, photos, and outputs are not retained beyond a limited time or used for training AI models.

📍 If you want to try the trend, use an older or edited image instead of a high-resolution current photograph and check whether the platform allows you to request deletion of your data afterwards.

📍 As an added layer of protection, security tools can help reduce risks from malicious links, dangerous downloads, and phishing attempts associated with viral trends.

📍 Most importantly, before participating in any viral AI trend, ask yourself: What personal information am I actually sharing? Avoid uploading high-resolution photos or identifiable personal details if you have not reviewed the platform’s data retention policies.