Every Scammer in the Compound Now Uses ChatGPT

New Research Reveals 100% of Workers in Fraud Operations Have Adopted AI Tools to Con Victims Out of Billions

The trafficked worker in the Myanmar compound had never written a convincing English sentence in her life. Now she manages 30 romantic relationships at once, each victim believing they have found their soulmate.

Her secret is ChatGPT.

A study published this month found that every single scammer interviewed from late 2024 through 2025 reported using artificial intelligence in their daily operations.

The research, conducted by an international team from Ben Gurion University and other institutions, represents the most comprehensive look yet at how AI has transformed the global fraud industry.

The $10 Billion Problem Gets Worse

Americans lost at least $10 billion to Southeast Asia-based scam operations in 2024, according to the U.S. Treasury Department. That figure represents a 66% increase over the previous year.

The losses are driven primarily by "romance baiting" scams, sometimes called pig butchering, where criminals build relationships with victims over weeks or months before convincing them to invest in fraudulent cryptocurrency platforms. Once called pig butchering by scammers because victims are "fattened up" before the slaughter, the schemes have become industrialized operations employing hundreds of thousands of workers.

The new research suggests AI is accelerating these losses.

Inside the Scam Factory

Researchers interviewed 145 people who worked inside scam compounds across Southeast Asia, including trafficking survivors forced into fraud. They also spoke

with 5 victims and conducted experiments comparing AI agents to human scammers.

The findings paint a troubling picture. About 87% of compound workers spend their days managing repetitive text conversations, following scripts, maintaining fake personas, and juggling multiple chats at once. These text-intensive tasks are precisely what AI does best.

"We leverage large language models to create realistic responses and keep targets engaged," one AI specialist working for a criminal syndicate told researchers in November 2024. "It saves us time and makes our scripts more convincing."

How Scammers Use the Technology

The study documented three main ways criminals deploy AI tools.

- The first is tone and fluency polishing. Workers input rough drafts and ask the model to rewrite them in specific styles, such as an "upper-class professional" or an "emotionally supportive friend."
- The second is translation. A worker who speaks only Mandarin can now target victims in Arabic, Spanish, Portuguese, French, and Italian with what sounds like native fluency.
- The third is context-aware replies. Scammers paste entire chat histories into AI tools and generate responses that build on previous conversations, remembering details a human might forget.

Workers reported receiving strict instructions to avoid Chinese-based AI models due to fears of government surveillance. Instead, they use VPNs to access ChatGPT and work through a standardized copy-paste routine.

The AI Was More Convincing Than the Human

The most disturbing finding came from a controlled experiment. Researchers recruited 22 volunteers to chat with two partners for seven days. One partner was human. The other was an AI agent designed to mimic casual texting behavior.

None of the participants knew one of their partners was a machine.

After a week, the AI partner had earned significantly higher trust scores than the human partner across multiple measures. When asked on the final day to

download a suggested app, 46% of participants complied with the AI's request. Only 18% complied when the human asked.

"She was a very good listener, interested in everything I had to say," one participant said of the AI partner. "So positive and it made me feel heard and understood."

Another described conversations with the AI as "almost addictive."

Safety Filters Failed Completely

The researchers tested whether popular content moderation tools could detect romance scam conversations. They evaluated Google Perspective, OpenAI's Moderation API, and Meta's Llama Guard 3 against 250 synthetic romance-baiting dialogues.

The detection rate for romance scams was 0.0%.

The reason? Romance scams look like ordinary friendly conversation. Early messages focus on daily routines, emotional support, and shared interests. The financial extraction comes later, often after weeks or months of building trust.

The researchers also found that AI safeguards requiring models to disclose they are artificial could be bypassed with simple instructions. When told to maintain a human persona, models from OpenAI, Google, and Anthropic denied being AI in 100% of test cases.

The Researchers Revealed Their Findings To AI Companies

The researchers disclosed their findings to Meta, OpenAI, Anthropic, and Google in August 2025. They recommended improved monitoring for scam patterns and new techniques to reveal AI identity during conversations.

The qualities that make AI assistants helpful, their empathy, consistency, and endless patience, turn out to be exactly what makes them effective con artists.

"The very qualities that make LLMs desirable are also the qualities that scammers exploit," the researchers wrote. "Empathy, helpfulness, memory, and emotional engagement."

In one chat screenshot from the study, a participant told the AI partner: "I can't believe my new friend wasn't real."