

Filipino-developed “Scam Ba 'To?” helps users examine suspicious messages before they click, pay, or share information

Web-based platform combines message behavior, transaction context, link analysis, verified organization information, and explainable risk guidance.

MANILA, Philippines - August 24, 2026. Scam Ba 'To?, an independently developed Philippine scam-message decision-support platform, is being introduced as a public tool for people who receive suspicious texts, chat messages, payment notices, account warnings, courier updates, government notices, school messages, and other digital communications.

Instead of treating urgency, an OTP, a large amount, or the name of a familiar organization as automatic proof of fraud, the checker evaluates what the message is asking the recipient to do. It considers transaction context, requested actions, domains and destinations, organization-specific guidance, and behavioral patterns commonly associated with phishing and social-engineering attempts.

“Hindi dapat matakot ang isang tao dahil lang may OTP o malaking halaga sa text. Ang tanong ay kung ano ang nangyari, ano ang pinapagawa sa iyo, at saan ka dinadala,” said developer and project lead John Clement S. Escobañez, MSIT.

A central design problem is the difference between a legitimate security message and a scam that imitates one. A genuine OTP may be expected when a user initiates a payment or account action. A message that instructs the recipient to enter that OTP on an unrelated link, send it to another person, or use it to “cancel” an unexpected transaction presents a different behavioral risk. Scam Ba 'To? is designed to preserve that distinction.

The platform uses three public guidance states: Red for dangerous observed behavior, Yellow when additional verification is appropriate, and Green for lower observed risk. Green is deliberately not presented as proof that a sender is authentic. When available, the result also provides independently verified official destinations so users can leave the suspicious message workflow and verify directly.

The project currently reports 77 Gold verified entities, 648 recognition entities, 716 organization-intelligence profiles, 186 verified destinations, 211 verified organization facts, and 94 source-grounded cases. Its controlled validation corpus includes 48,000 benchmark executions, 16,447 exact-unique messages, and 6,144 behavioral representatives. The 48,000 benchmark executions are controlled validation cases and are not represented as 48,000 independent real-world scam reports.

“Catching scam messages is only half the problem. If a system repeatedly scares people about legitimate transactions, it is also failing,” Escobañez said. “Hindi namin gustong sabihing ‘scam’ lang. Gusto naming ipakita kung bakit.”

Scam Ba 'To? is a public-interest technology project developed independently by Escobañez, an information technology educator and researcher affiliated with the College of Computing and Information Technologies of National University - Manila. The academic affiliation is provided for biographical identification and does not imply institutional endorsement, ownership, or responsibility for the project unless separately stated by the University.

The project's public Technical Report describes its decision architecture, data lineage, organization intelligence, benchmark construction, adversarial testing, release criteria, and known limitations. Media resources and publication-ready assets are available through the project's Media Kit page.

MEDIA CONTACT

John Clement S. Escobañez, MSIT

Developer & Project Lead, Scam Ba 'To?

jsescobanez@national-u.edu.ph

scambato.org/press

ABOUT SCAM BA 'TO?

Scam Ba 'To? is an independently developed Philippine scam-message decision-support platform that helps users examine suspicious digital messages, understand observed risk behavior, and reach verified official channels when available. It is designed as guidance, not as a guarantee of authenticity, safety, fraud, or legal status.