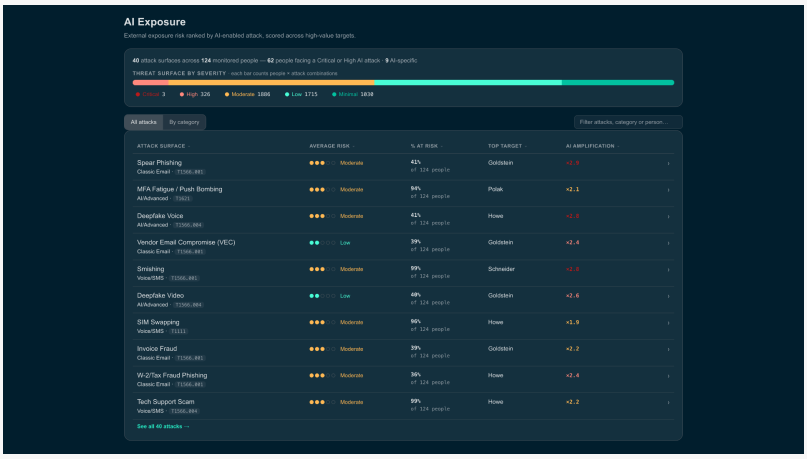
VanishID Previews AI Exploitability Management, Measuring What Frontier AI Can Weaponize From Public Data

New AI Exploitability Management capability measures AI risk to organizations, including deepfakes, impersonation, and spear phishing.



VanishID's AI Exploitability Management dashboard ranks external attack risk by AI-enabled threat. It maps 40 attack types and monitors employee AI-driven attack risk levels.

LAS VEGAS, NV, UNITED STATES, August 4, 2026 /EINPresswire.com/ -- VanishID, the external identity security company, today unveiled AI Exploitability Management, previewing how organizations can measure what frontier AI can build from their people's public exposure, before an attacker builds it. VanishID is demonstrating the capability publicly for the first time this week at Black Hat USA 2026, where attendees can sign up for a complimentary private scan of their own AI exploitability. A personal findings report will follow after the show.

> We built AI Exploitability Management to make it measurable, attack by attack and person by person, from the attacker's side of the fence."
>
> *Matt Polak, CEO of VanishID*

AI Exploitability Management in Brief

- Measures AI exploitability risk by evaluating more than 40 AI-powered attack scenarios, including executive impersonation, real-time deepfakes, voice cloning, spear phishing, business email compromise, and other AI-driven social engineering techniques.

- Calculates an AI Exploitability Score for every individual assessed, helping security teams understand what publicly available information creates the greatest AI-driven risk and prioritize mitigation based on the attacks most likely to succeed.

- Runs entirely from the outside, the same vantage point an attacker has, with nothing installed, no integrations, and no credentials.

“Attacks on people used to be artisanal. One credential broker, patiently piecing together what they could find. AI industrialized that. Now it just needs the breadcrumbs we've all left scattered across the web to easily enable a compromise,” said Jason Barnett, Vice President and Chief Information Security Officer of Oracle Health & Global Industries.

“Most security teams can quote their patch rate to a decimal point. Almost none can tell you what AI could assemble from their CFO’s public exposure this afternoon, and that blind spot is where the attacks now start,” said Matt Polak, CEO of VanishID. “We built AI Exploitability Management to make it measurable, attack by attack and person by person, from the attacker’s side of the fence. You can’t patch a face photo, but once you can measure what it enables, remove what’s possible to remove, and neutralize residual risk with existing security operations and tooling instead of leaving it unowned. Measurement is where managing AI risk begins.”

How AI Is Weaponizing The Human Attack Surface

Frontier AI attacks are assembled entirely outside the perimeter. A real-time deepfake video call needs a face photo, a voice clone needs seconds of audio, and a convincing spear phish needs little more than a writing sample and an org-chart position. Every ingredient is harvested from public exposure, then compiled by AI into an attack that arrives through a trusted channel, a video call, a phone call, or an email that sounds exactly right. That is how attackers extracted $25.6 million from engineering firm Arup in 2024. A finance employee wired the funds after a video conference where every other participant, including the CFO, was a real-time deepfake. There was no malware and no exploit, and nothing fired an alert. Security teams can measure patch levels, misconfigurations, and phishing click rates. Until now, they had no way to measure what AI can do with their people.

How AI Exploitability Management Works

VanishID’s AI Exploitability Management makes that risk measurable, and it starts from a simple observation: every AI-enabled attack has a recipe. The capability breaks more than 40 attacks down into the specific ingredients each one requires, then classifies those ingredients as required, high-value, useful, or contextual. The attack types range from deepfake video and voice cloning to browser-in-the-browser phishing, whaling, and nation-state insider recruitment. Deepfake video, for example, requires a face photo and is amplified by a voice sample, an org-chart position, and a writing style. The platform then shows which ingredients are already public,

for which people, so teams see not just abstract exposure but the exact attacks it enables and who they land on.

The assessment works entirely from the outside, the same vantage point an attacker has. VanishID's AI examines only what is publicly reachable. It requires nothing installed on devices, no integrations, and no credentials, never touches internal systems, and asks nothing of the security team to run. It is AI measuring the risk AI poses. And because exposure grows and attacker capabilities keep improving, the assessment runs continuously rather than as a one-time scan.

Experience AI Exploitability Management at Black Hat USA 2026

At Black Hat USA 2026 attendees can see AI Exploitability Management live at the VanishID Booth #5813, August 4–6, and can sign up on site, with just a name and email, for a complimentary assessment of their personal AI exploitability. This will show what an attacker's AI could build from their own public exposure, scored attack by attack, with a findings report delivered after the conference.

About VanishID

VanishID is the external identity security company. Its agentic AI platform puts autonomous agents to work finding and neutralizing the exposed personal information attackers use against executives, employees, and their families. The agents hunt exposure the way an adversary would, at machine speed, across data brokers, breach dumps, the dark web, and public records. They remove what they can and mitigate residual risk. Onboarding takes a name and a corporate email, with no lift for the security team. The platform never trains on personal data and is SOC 2 Type II certified. More than 100 organizations rely on VanishID, from Fortune 500 enterprises and financial institutions to health systems and public sector agencies. Founded in 2019 and headquartered in Bethesda, Maryland, VanishID is backed by Dell Technologies Capital, Crosslink Capital, Rally Ventures, Energy Impact Partners, Bright Pixel, LockStep Ventures, and strategic angels including former Palo Alto Networks CEO Mark McLaughlin and former Mandiant CPO Christopher Key.

Learn more at vanishid.com.

Marketing at VanishID
VanishID
email us here
Visit us on social media:
LinkedIn
YouTube
X