CEO insights: beyond the AI model card

BY STIJN JANS · MAY 18, 2026

As part of our AI series, I recently released a blog on the topic of keeping the human layer sacred in the AI era. There, I shared my thoughts on where human intelligence fits, the decisions I believe companies will face in the next 3 to 5 years, and explored what I deem to be the beating heart of the Bug Bounty industry.

Considering that discussion, I want to continue the conversation regarding transparency in the age of AI. Transparency only works if it's continuous, visible, and contestable. In a time of rapid technology development, it's an ongoing process and something to challenge and continuously refine. Which is why our AI Model Card is just the start.

In this blog, I highlight our commitments to transparency and giving back to our hacking community, and the subsequent results.

Continuous mechanisms for radical transparency

To put it bluntly, if a researcher needs to read between the lines to understand what we're doing with their work, we've already failed.

So here are a few mechanisms we're committed to.

- **Public versioning:** This is when every adjustment to how we use researcher data is visible through a public changelog. The result means no silent updates and no buried changes. If the policy moves, the community sees the difference.
- **Internal governance:** Before we take any meaningful step with AI that touches researcher data or the submission flow, we have an open dialogue about the use of technology. This is something all layers of the organization are engaged in.
- **Voice of the researcher:** No AI decision is made in isolation. Our Head of Hackers, Head of Triage, and other domain leads are in the room for every meaningful conversation. They represent the people most impacted by what we ship. The people who live closest to the researcher community and the triage workflow have a seat at the table and a real voice in every outcome.
- **Human in the loop:** We are strong believers in keeping humans in the loop on decisions. Not as a fallback when AI fails, but as a structural requirement of how the platform works. AI surfaces, recommends, accelerates, but humans decide.
- **Incident transparency.** If something goes wrong, we publish what happened, what we changed, and what we learned. Again, this all falls to continuous transparency.

What giving back to the hacking community really means

Giving back only means something if the researcher's life and economics actually improve. So that's the bar for every AI feature we ship.

We ask ourselves four main things, does it:

1. Help them create more impact?
2. Pay researchers faster?
3. Help them earn more?
4. Save them time?

If yes, we build it. If no, it's not a giving back feature; it's something else.

The advantage of consistent reasoning

Alongside continuous transparency, there is another element that doesn't get enough attention, and that is consistent reasoning. What I mean here is that when AI is involved in supporting decisions, the same logic applies across submissions. This means less bias, less variance between analysts, and less chance that a researcher gets a different outcome on the same finding depending on who happens to pick up the ticket.

It is fairness at scale.

Here are a few concrete examples:

- **Pre-submission strengthening:** Tooling that helps researchers sharpen a report before they hit submit. Better reports earn higher bounties. The researcher captures that value.
- **AI skill matching:** We detect the technologies running on customer assets and match them against the track record of hackers who've successfully found vulnerabilities in those same stacks. Researchers get pointed at work where they're statistically most likely to succeed and earn. So not only do the right researchers see the right programs, but customers get better coverage.
- **Attack Surface insights:** With customer permission, we surface where a program has had the least attention. This means fewer collisions on well-tested endpoints and opens up more opportunities for fresh eyes to find something impactful.
- **Fairness tooling:** AI-assisted review of bounty decisions to flag potential inconsistencies across regions, languages, and reporting styles makes sure that the platform treats every researcher the same.

Next steps

Our Bug Bounty experts are at the forefront of AI insights, trends, and discussions.

Interested in the “Vulnpocalypse” and how AI is reshaping vulnerability discovery? Explore insights from our COO, Ed Parsons, on what this shift means for security teams. If you’re looking to separate AI hype from reality, our Head of Product, Greg Jenkins, breaks down common AI misconceptions and how they can affect your team. And for a closer look at what AI means for the future of Bug Bounty, Program Leader Chris Holt shares his perspective on where AI can support programs, and where it may create new challenges.

Or read my previous blog on ‘Holding the human layer sacred in the AI era’.

There are many more discussions in the pipeline. So, watch this space, sign up to our newsletter, and connect with us on socials to stay in the loop on all AI discussions.

AUTHOR

Stijn Jans

Stijn Jans is the Founder and CEO of Intigriti, the leading Bug Bounty and ethical hacking platform that connects organizations with a global community of security researchers. Under his leadership, Intigriti has grown into a trusted cybersecurity partner for enterprises worldwide, championing transparency and continuous vulnerability disclosure to help companies stay ahead of emerging threats.