

Optimizing Machine Learning Models Through Data Labeling

For A B2B Software Company

A leading B2B software company urgently needed to combat phishing and other scams that were targeting its CRM platform. By implementing a streamlined review process, Vaco successfully labeled policy violations in 13+ languages, achieving a 98% confidence level with minimal false positives.

Client
Multinational Software Products and Services Company

Services
Managed Services

Area of Expertise
Digital Transformation: Trust & Safety

Industry
Software Products and Services

Challenge

Facing a surge in phishing and other scams targeting its CRM platform, a multinational software company urgently needed to address its backlog of policy violations. Vaco and three other firms engaged in a competitive data "bake-off" to showcase their expertise in data labeling, vying for the chance to become the company's outsourced business partner. Vaco prevailed and was selected as the partner to deliver high-quality data labeling results while minimizing the onboarding of resources and upfront time investment.

The average global cost of a data breach reached an all-time reporting high of \$4.45 million in 2023, with 1,339 brands targeted by a phishing attack in Q4 2023 alone. (Sources: [IBM](#) & [Techopedia](#))

Our Solution

Vaco quickly engaged a dedicated team, including quality assurance specialists to establish quality calibration layers. Initial training sessions and ongoing syncs were led by Vaco to ensure alignment on policy guidelines, documentation requirements and quality expectations. The team dedicated 55 hours to the review process and 12 hours to quality assurance efforts. Equipped with proficiency in five languages, the review team handled nearly all reviews. For unsupported languages, Google Translate was leveraged.

Our Impact

Within four days, the team successfully labeled 2,000 items in 13 native languages and various formats. Vaco delivered all reviewed items in a ZIP file as images and developed a new review process and interface. Furthermore, 432 items underwent rigorous quality assurance testing, achieving a 98% confidence level with a margin of error of 5%. Less than nine false positives were identified. This success enabled the review team to align their training with active reviews. Vaco provided the organization with a comprehensive recap, including relevant metrics, actionable insights and suggested policy clarifications, enabling the enhancement of machine learning modeling for the organization.