

## TOO MANY USER REPORTS, NOT ENOUGH TIME

For years, gaming and social platforms have relied on users to report abuse, hate speech, and other NSFW content. Reports are sent to moderation teams who review each abuse report individually. Many platforms receive thousands of reports daily, most of which can be closed without taking action.

Meanwhile, reports containing time-sensitive content – suicide threats, violence, terrorism, and child abuse – risk going unseen.

## NOT ANYMORE

With Predictive Moderation, platforms can train a custom AI model on their moderation team's consistent decisions, automating the most time-consuming part of the moderation process by closing false reports, taking action on the obviously abusive reports, and triaging reports that require human eyes for priority review.

Predictive moderation is all about efficiency. We want moderation teams to streamline their workflows and focus on the work that matters – reports that require urgent review, and retention and engagement-boosting activities with the community.

## CHANGING REGULATIONS

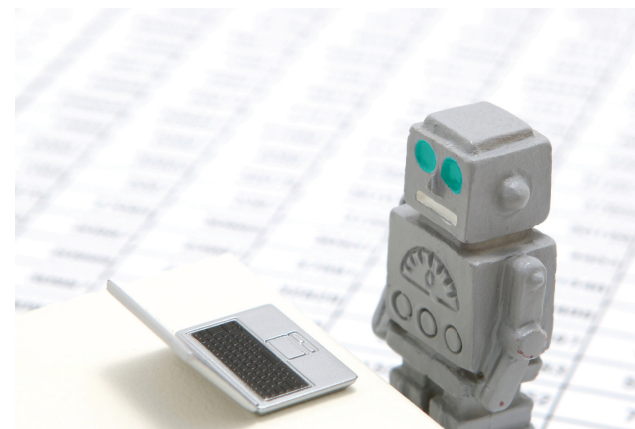
The German law known as NetzDG says that platforms must **remove reported hate speech and illegal content within 24 hours – or face fines of up to 50 million euros**. Similar laws concerning reported content are being introduced in France, Australia, the UK, and across the globe.



## IS PREDICTIVE MODERATION A GOOD FIT FOR YOU?

If you struggle with any of the following, Predictive Moderation might be the solution for you:

- An inefficient ticketing system to review and action on reports
- All reports placed into one queue without prioritization
- Not enough moderators or time to review all reports
- Missing critical reports due to workload
- Angry users due to reports not being actioned
- Concerned about regulatory compliance



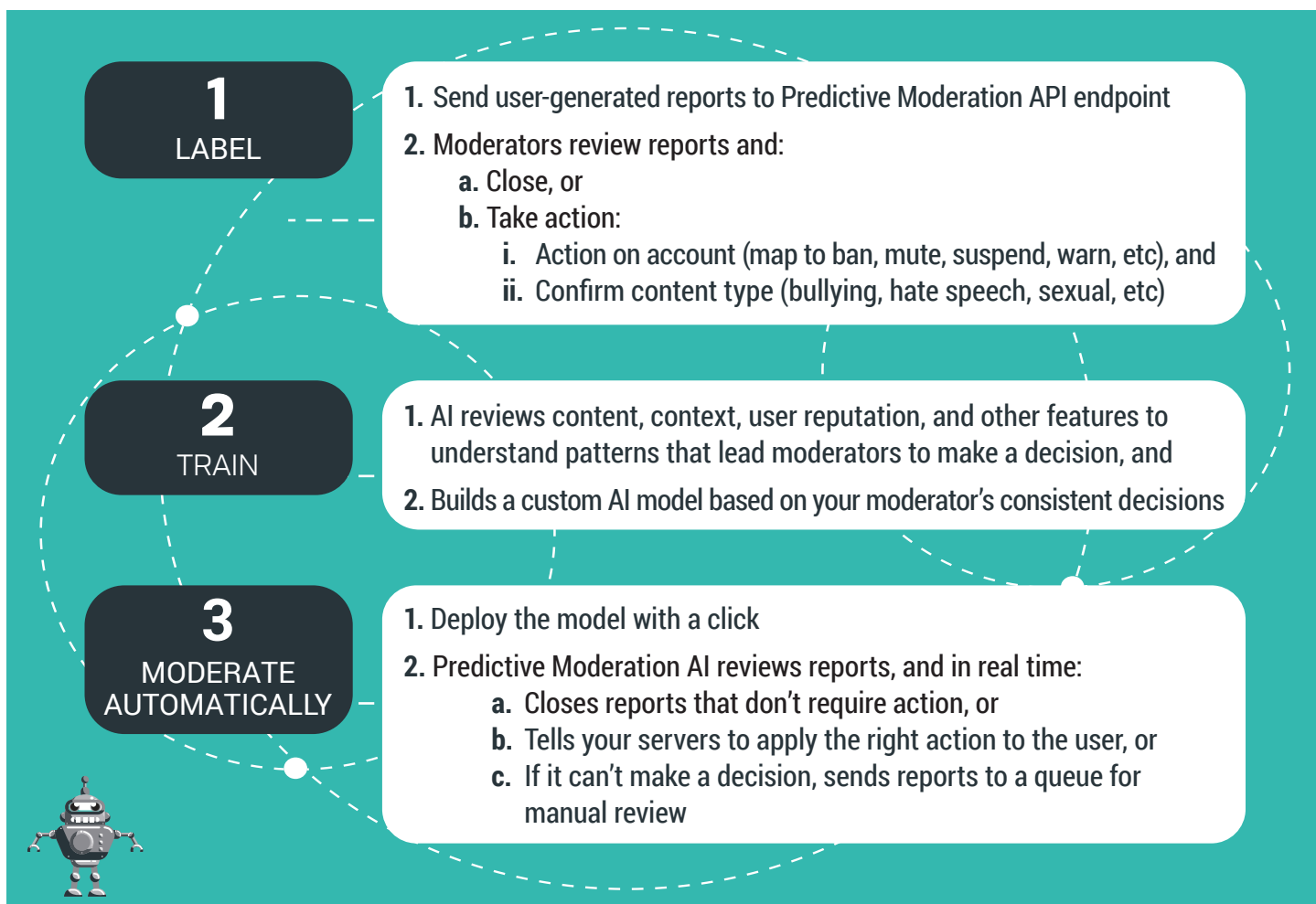
## BUILT FOR PRIVACY

Predictive Moderation was built in accordance with GDPR (see Article 89). All appropriate safeguards and measures are in place to ensure a high level of privacy.

# HOW IT WORKS



- Reduce moderation workload by 50-70%
- Free up moderators to engage the community in retention-boosting activities
- Custom AI built on your unique data
- Comply with global reporting regulations
- Protect moderators from exposure to high-risk content
- Increase user loyalty and trust by taking action on reports in real-time



Not ready to implement Predictive Moderation? No problem. Just send your user-generated reports to our report API. While moderators follow their normal workflow, Predictive Moderation stores their decisions in the background.

When you're ready to implement, we're ready to train your custom AI model.



For more information about Two Hat's other products, including our chat, image, and video content moderation solution Community Sift, visit our site.

We can be reached at [hello@twohat.com](mailto:hello@twohat.com) or 1.855.407.0889.