



HUD CASE STUDY

Cutting Fakes & Embracing Feels

How the team behind HUD dating app work with WebPurify to keep users safe and their platform growing

CUTTING FAKES & EMBRACING FEELS

Since it arrived on the dating app scene in 2015, HUD has attracted an impressive 16.5 million users to its platform.

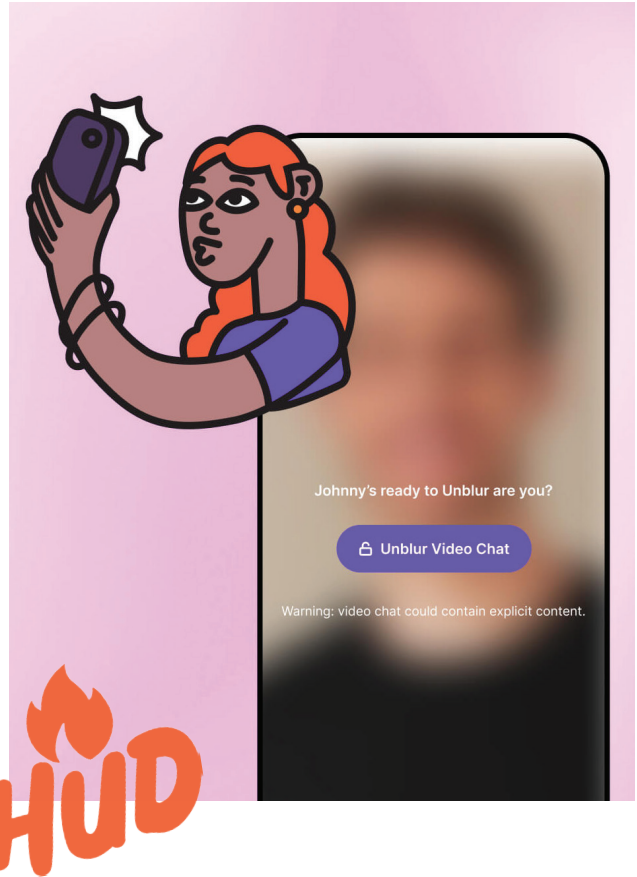
The HUD App provides a space for individuals seeking short-term relationships with no strings attached. But while users might be keen to keep things casual, the team behind HUD App are committed to keeping their community safe 24/7, bolstered by a large human moderation team.

The chief benefit of using human moderation is that it's more accurate, suggests Amanda Kuhns, Community Experience Team Manager at HUD App. *"We find more, because we can really look into context more than a lot of AI can."*

There is a price to employing human moderators, but it's nothing compared to the cost of failing to moderate effectively.

For example, it's reported that almost 40% of adults who have used a dating app or website have received unwanted sexually explicit messages or images on the platforms, and 30% have had someone contact them after they specified they weren't interested.

Scammers continue to be a problem as well. It's not just hearts that are broken by romance fraudsters—bank accounts can be devastated too. In 2022, reported losses of romance scamming hit a staggering \$1.3 billion, according to data from the Federal Trade Commission.



Amanda explains that as well as helping users to explore new options with dating and hookups, keeping users safe and educating them about their interactions and potential scams is a prime focus for HUD App.

"We focus on the more casual side of dating, and there's a gray area when it comes to facilitating people to have the kinds of conversations you might not easily have in real life. But we try to teach our users that consent is still integral to every interaction, intimate or not. We want them to have a great time and meet somebody, but be safe while doing it."

THE STATUS QUO

What are HUD's biggest content moderation challenges?

Screenshot from the HUD dating app.



01

Fake profiles

When it comes to dealing with the scammer side of dating apps, it's very important for HUD App to keep fake profiles out. "That's why we rely on many tools, including photo verification," explains Amanda, pictured below.

It's not just the initial profile photograph submitted during the signup process that gets verified by human moderators, but every photo that's subsequently uploaded to an account. "It means that users can't make a profile with a real photo, then come back with fake ones and try to pass them off as real," Amanda says.

02

AI-generated or enhanced Images

HUD App's moderators are vigilant about fake profiles created with AI-generated imagery too.

"We're starting to see our users attempt to implement AI into their profiles more, albeit in a very basic way."

"For example, if our moderators see multiple AI-generated photos of the same person, we flag that, and if anyone attempts to create a profile using one or more of those photos, it's immediately shut down."



WHAT ARE HUD'S BIGGEST CONTENT MODERATION CHALLENGES?

Photos that users have taken and enhanced to the point where the image no longer resembles reality may also require moderation.

"AI-powered moderation will struggle with images that have been heavily filtered or enhanced with effects, as it can't distinguish if those photos are of a real person," says Amanda. "HUD App's human moderation team are expert at catching scammer photos and fake profiles, as trained human eyes are much more accurate at seeing what's real and what's not."

It's another example where human moderation gives HUD the advantage, despite the resource implications of prioritizing human moderation.



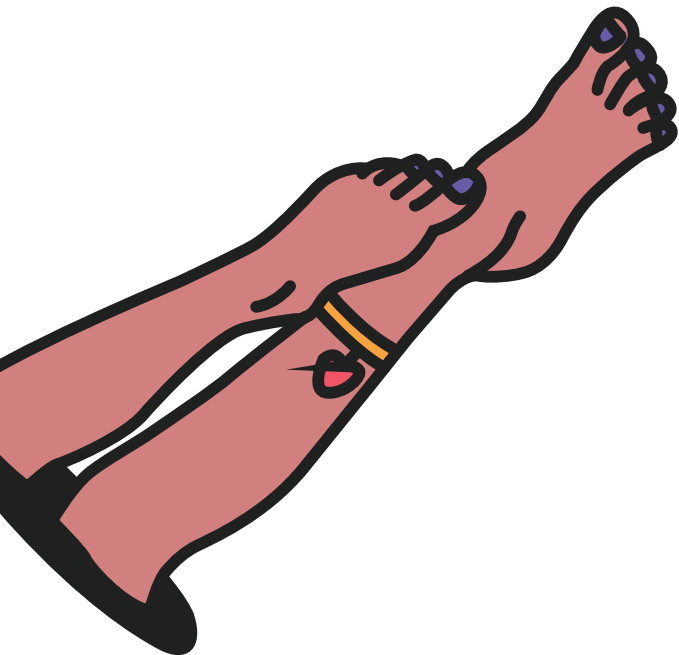
03

Scale & speed

Amanda says the biggest challenge is time.

"A human moderator reviews every report we get, which does take time, but it's worth it," says Amanda. "It gives our users peace of mind and confidence that we're listening to them, taking their concerns seriously, and reviewing and investigating anything that they come to us with."

When she was first hired by the company four years ago, Amanda was the only person doing verification. Additional moderators were recruited as the HUD App user base amped up, and it soon became clear that the in-house team's time would be better served investigating more involved customer service emails and user reports.



THE SOLUTION

WebPurify's Speed & Scale

Screenshots from the HUD dating app.



THE SOLUTION: WEBPURIFY'S SPEED & SCALE

An external vendor that could help lighten the load for routine verification and moderation tasks was the solution.

Amanda wanted to collaborate with a team that offered flexibility and had experience in multiple forms of moderation—not only images, but also text and locations too—because scammers use all sorts of tricks to obscure who and where they really are.

“We were also looking to work with a team that offered the right time zone for us in order to fulfill a 24/7 human moderation service, with what we felt was a very good support system from their managers,” she says.

“WebPurify checked all the boxes. They were easy to work, attentive and took direction really well.”

“You could tell this is what they do and they’re very good at it, because they picked it up so fast and the work was accurate. Explaining it to someone totally new is hard, so finding a team of moderators who were so quick with everything was great.”

WebPurify’s initial brief was to focus on the verification of user photos, profile text and location.

WebPurify x HUD Services

Profile verification

Providing human 24/7, fast, accurate verification of users’ profiles via bespoke tools including photo verification, ensuring all images users submit are real and authentic. While the exact process is proprietary, it is human labour-intensive and ensures users aren’t attempting to sign up with copyrighted or AI-generated images.

Profile text

Ensuring submitted descriptions comply with the HUD community guidelines.

Scams

Checking IP addresses against geo location to detect scammers.

THE SOLUTION: WEBPURIFY'S SPEED & SCALE

HUD App already had a gesture photo verification process in place, which WebPurify's moderators were able to use to check that people weren't signing up using old photos.

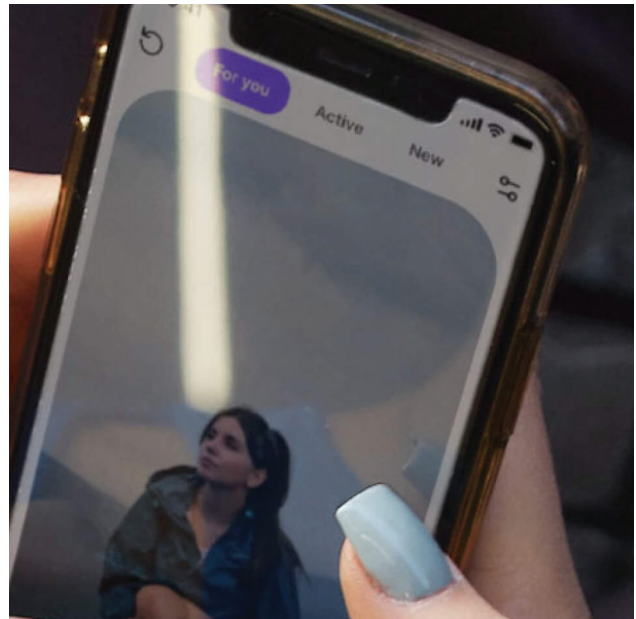
Amanda explains that gesture photo verification is a process that requires a human moderator. "In order for a user to sign up, they must provide an authentic photo of themselves, and our gesture verification process includes WebPurify ensuring these images are accurate and authentic. WebPurify also checks to ensure that the rest of their profile is within community guidelines before the user is verified."

"They've been carrying out these verifications for every signup, as well as for users who upload additional photos—making sure they're real photos that fall within the guidelines and show nothing inappropriate.

"And if they see anything in the user's intro text which goes against our guidelines or suggests a scam, WebPurify is able to flag that profile for us so we can suspend it."

HUD App doesn't allow anyone to sell services of any kind on their platform, but unfortunately, some users still attempt to advertise subscription-based or other paid user interactions, contravening HUD App's terms and conditions.

The team is starting to look at introducing more in the way of AI moderation in order to provide a double layer of protection.



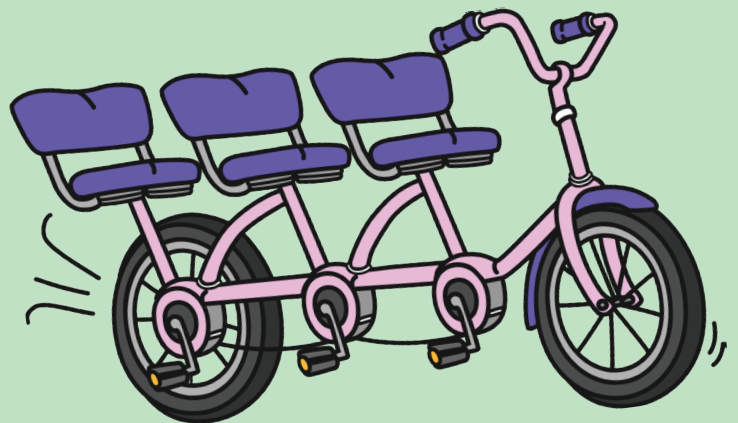
"We use it to flag profile text and conversations that contain certain keywords," Amanda says. "If someone keeps saying the word 'cash', for example, we'll get a report and then we'll go in and see if they're actually asking for cash, which is against our rules—in which case we'll take action—or if they're just talking about working."

Despite the speed advantage of AI moderation, she adds, human moderation continues to be most effective.

"Our AI will flag other words, such as 'sugar', which can be indicative of a sugar daddy or sugar mommy scam. But the users may just be talking about having a sweet tooth, in which case a human moderator can see the distinction. However, our AI will continue to flag those words, and others which might point to something far more sinister. It's important for our users that we're using all the tools we can to protect them."

THE RESULTS

Verifying daily dating app sign-ups



The Results

4k-6k daily sign-ups
to the HUD app

The HUD app receives 4,000–6,000 daily sign-ups, which means that WebPurify has been required to complete a high volume of verifications each day. “The way that they have multiple moderators who can overlap with each other means we never get behind and the queue is always cleared quickly, meaning any potential scammers will be stopped quickly, too,” Amanda says.

“You want to be fast and accurate, but then also be aware of yourself—you don’t want to burn out.”

Why HUD chose WebPurify

24-7

dedicated human moderators

Mental health

support for content moderators

Communication

Speed & scale

It isn’t just about speed of verification, she adds:

Working with an ethical partner that had a proper mental health provision was an important consideration when HUD selected WebPurify. “Having such good leaders in place meant that I knew their moderators were well taken care of. We look after our admins as well, because we’re sometimes exposed to things that we don’t want to see.

"Our in-house moderation team is entirely female, and sadly, a majority of the reports we receive are from female users reporting male users for harassment," Amanda says. "Safety is an important responsibility and priority for us, whether you're a user of HUD App or you're a moderator working to keep users safe."

As well as checking the new users, WebPurify reviews any edits made to existing user profile text and new photo uploads. "Every time users update their intro, whether it's just adding a space or an emoji, that comes through to us so we can check it," Amanda explains. "WebPurify is helping with that now, to make sure no one's adding content they shouldn't be."

So what of user privacy? Should account holders be concerned that a person rather than anonymous AI is verifying and moderating personal content on a dating app?

"Our users' privacy is critical, and we adhere strictly to all privacy and personal information legislation," Amanda says. "However, if a user is reported for harassment, it is our responsibility and policy to check their profile, chats, and images closely. If they are behaving inappropriately and we need to take action, we do this decisively, and we support our moderators in this."

The moderation teams at HUD and WebPurify are making an impact though. More than 200 user accounts are suspended on average each day, for reasons such as violating policies or being reported by another user. Around 160 of those daily suspensions are because the account has been identified as a scammer.

On average,

160

scammers are
identified each day

“When we find a fraudulent profile, we have a clear set of steps to take,” Amanda says. “Our goal is always to support users to feel safe and secure when using our app, and to eliminate any scammers or ill-intentioned users.”

She explains that HUD App and WebPurify are beginning to see fewer fraudulent profiles on the app. “We are skilled at catching fake profiles and scammers using image verification and location confirmation methods,” Amanda says. “User safety, though, can be more time-consuming and detail-oriented, as our moderators need to look for context and supporting evidence when it comes to issues like harassment or inappropriate behavior. This is one of the most important parts of the moderation role, which AI simply cannot fulfill.”



She adds that working with WebPurify’s moderators has brought a different perspective to the process too: “They would pose questions that maybe we’d never asked ourselves because we were used to getting through it as fast and as accurately as possible. But they would point out potential issues before they became a larger problem.

“Likewise, we could message them to be on the lookout for a fake photo that we’d seen multiple times, for example, or to reinforce how important it is to flag anyone who looks under 18. WebPurify is very good at responding and leaving notes, and they have helped a lot in catching people who shouldn’t be on the app.”

Content moderation and verification are central to the quality of people’s experiences in online dating. And yet, if those tasks are carried out successfully they will go largely unnoticed by the users.

“We’re very invested and we work hard,” summarizes Amanda. “We really try to find as much as we can. We were concerned about whether a third-party team was going to be as invested as we are, but Ravi, WebPurify’s Head of Training, and Josh, WebPurify’s Head of Client Services, have been very attentive. They know what they’re doing and they communicate well.

With moderation, you need that, because everyone needs to be on the same page.”