



PHOTOGRAPH: MOISES SAMAN/MAGNUM

ADRIAN CHEN BACKCHANNEL OCT 23, 2014 6:30 AM

The Laborers Who Keep Dick Pics and Beheadings Out of Your Facebook Feed

Inside the soul-crushing world of content moderation, where low-wage workers soak up the worst of humanity and keep it off your screens.



[The AI Database →](#)

SECTOR: SOCIAL MEDIA

THE CAMPUSES OF the tech industry are famous for their lavish cafeterias, cushy shuttles, and on-site laundry services. But on a muggy February afternoon, some of these companies' most important work is being done 7,000 miles away, on the second floor of a former elementary school at the end of a row of auto mechanics' stalls in Bacoor, a gritty Filipino town 13 miles southwest of Manila. When I climb the building's narrow stairwell, I need to press against the wall to slide by workers heading down for a smoke break. Up one flight, a drowsy security guard staffs what passes for a front desk: a wooden table in a dark hallway overflowing with file folders.

Past the guard, in a large room packed with workers manning PCs on long tables, I meet Michael Baybayan, an enthusiastic 21-year-old with a jaunty pouf of reddish-brown hair. If the space does not resemble a typical startup's office, the image on Baybayan's screen does not resemble typical startup work: It appears to show a super-close-up photo of a two-pronged dildo wedged in a vagina. I say *appears* because I can barely begin to make sense of the image, a baseball-card-sized abstraction of flesh and translucent pink plastic, before he disappears it with a casual flick of his mouse.

The road leading to the former office of TaskUs in Bacoor, a Manila suburb. PHOTOGRAPH: MOISES SAMAN/MAGNUM

Baybayan is part of a massive labor force that handles “content moderation”—the removal of offensive material—for US social-networking sites. As social media connects more people more intimately than ever before, companies have been confronted with the Grandma Problem: Now that grandparents routinely use services like Facebook to connect with their kids and grandkids, they are potentially exposed to the internet's panoply of jerks, racists, creeps, criminals, and bullies. They won't continue to log on if they find their family photos sandwiched between a gruesome Russian highway accident and a hardcore porn video. Social media's growth into a multibillion-dollar industry, and its lasting mainstream appeal, has depended in large part on companies' ability to police the borders of their user-generated content—to ensure that Grandma never has to see images like the one Baybayan just nuked.

“Everybody hits the wall. You just think, ‘Holy shit, what am I spending my day doing?’”

So companies like Facebook and Twitter rely on an army of workers employed to soak up the worst of humanity in order to protect the rest of us. And there are legions of them—a vast, invisible pool of human labor. Hemanshu Nigam, the former chief security officer of MySpace who now runs online safety consultancy SSP Blue, estimates that the number of content moderators scrubbing the world's social media sites, mobile apps, and cloud storage services runs to “well over 100,000”—that is, about twice the total head count of Google and nearly 14 times that of Facebook.

This work is increasingly done in the Philippines. A former US colony, the Philippines has maintained close cultural ties to the United States, which content moderation companies say helps Filipinos determine what Americans find offensive. And moderators in the Philippines can be hired for a fraction of American wages. Ryan Cardeno, a former contractor for Microsoft in the Philippines, told me that he made \$500 per month by the end of his three-and-a-half-year tenure with outsourcing firm Sykes. Last year, Cardeno was offered \$312 per month by another firm to moderate content for Facebook, paltry even by industry standards.

A suicidal message posted by a Whisper user and flagged for deletion by TaskUs employee. PHOTOGRAPH: MOISES SAMAN/MAGNUM

Here in the former elementary school, Baybayan and his coworkers are screening content for Whisper, an LA-based mobile startup—recently valued at \$200 million by its VCs—that lets users post photos and share secrets anonymously. They work for a US-based outsourcing firm called TaskUs. It's something of a surprise that Whisper would let a reporter in to see this process. When I asked Microsoft, Google, and Facebook for information about how they moderate their services, they offered vague statements about protecting users but declined to discuss specifics. Many tech companies make their moderators sign strict nondisclosure agreements, barring them from talking even to other employees of the same outsourcing firm about their work.

“I think if there's not an explicit campaign to hide it, there's certainly a tacit one,” says Sarah Roberts, a media studies scholar at the University of Western Ontario and one of the few academics who study commercial content moderation. Companies would prefer not to acknowledge the hands-on effort required to curate our social media experiences, Roberts says. “It goes to our misunderstandings about the internet and our view of technology as being somehow magically not human.”

I was given a look at the Whisper moderation process because Michael Heyward,

Whisper's CEO, sees moderation as an integral feature and a key selling point of his app. Whisper practices “active moderation,” an especially labor-intensive process in which every single post is screened in real time; many other companies moderate content only if it's been flagged as objectionable by users, which is known as reactive moderating. “The type of space we're trying to create with anonymity is one where we're asking users to put themselves out there and feel vulnerable,” he tells me. “Once the toothpaste is out of the tube, it's tough to put it back in.”

Watching Baybayan's work makes terrifyingly clear the amount of labor that goes into keeping Whisper's toothpaste in the tube. (After my visit, Baybayan left his job and the Bacoor office of TaskUs was raided by the Philippine version of the FBI for allegedly using pirated software on its computers. The company has since moved its content moderation operations to a new facility in Manila.) He begins with a grid of posts, each of which is a rectangular photo, many with bold text overlays—the same rough format as old-school internet memes. In its freewheeling anonymity, Whisper functions for its users as a sort of externalized id, an outlet for confessions, rants, and secret desires that might be too sensitive (or too boring) for Facebook or Twitter. Moderators here view a raw feed of Whisper posts in real time. Shorn from context, the posts read like the collected tics of a Tourette's sufferer. *Any bisexual women in NYC wanna chat? Or: I hate Irish accents! Or: I fucked my stepdad then blackmailed him into buying me a car.*

A list of categories, scrawled on a whiteboard, reminds the workers of what they're hunting for: pornography, gore, minors, sexual solicitation, sexual body parts/images, racism. When Baybayan sees a potential violation, he drills in on it to confirm, then sends it away—erasing it from the user's account and the service altogether—and moves back to the grid. Within 25 minutes, Baybayan has eliminated an impressive variety of dick pics, thong shots, exotic objects inserted into bodies, hateful taunts, and requests for oral sex.

More difficult is a post that features a stock image of a man's chiseled torso, overlaid with the text “I want to have a gay experience, M18 here.” Is this the confession of a hidden desire (allowed) or a hookup request (forbidden)?

Baybayan—who, like most employees of TaskUs, has a college degree—spoke thoughtfully about how to judge this distinction.

“What is the intention?” Baybayan says. “You have to determine the difference between thought and solicitation.” He has only a few seconds to decide. New posts are appearing constantly at the top of the screen, pushing the others down. He judges the post to be sexual solicitation and deletes it; somewhere, a horny teen's hopes are dashed. Baybayan scrolls back to the top of the screen and begins scanning again.

TaskUs contractors review content on their computers at their office in the Taguig district of Manila, on August 28, 2014.

PHOTOGRAPH: MOISES SAMAN/MAGNUM

Employees of Open Access BPO review content at the company's offices in Manila, on August 28, 2014. PHOTOGRAPH : MOISES SAMAN/MAGNUM

EIGHT YEARS AFTER the fact, Jake Swearingen can still recall the video that made him quit. He was 24 years old and between jobs in the Bay Area when he got a gig as a moderator for a then-new startup called VideoEgg. Three days in, a video of an apparent beheading came across his queue.

“Oh fuck! I've got a beheading!” he blurted out. A slightly older colleague in a black hoodie casually turned around in his chair. “Oh,” he said, “which one?” At that moment Swearingen decided he did not want to become a connoisseur of beheading videos. “I didn't want to look back and say I became so blasé to watching people have these really horrible things happen to them that I'm ironic or jokey about it,” says Swearingen, now the social media editor at Atlantic Media. (Swearingen was also an intern at WIRED in 2007.)

While a large amount of content moderation takes place overseas, much is still

done in the US, often by young college graduates like Swearingen was. Many companies employ a two-tiered moderation system, where the most basic moderation is outsourced abroad while more complex screening, which requires greater cultural familiarity, is done domestically. US-based moderators are much better compensated than their overseas counterparts: A brand-new American moderator for a large tech company in the US can make more in an hour than a veteran Filipino moderator makes in a day. But then a career in the outsourcing industry is something many young Filipinos aspire to, whereas American moderators often fall into the job as a last resort, and burnout is common.

“Everybody hits the wall, generally between three and five months,” says a former YouTube content moderator I’ll call Rob. “You just think, ‘Holy shit, what am I spending my day doing? This is awful.’”

Rob became a content moderator in 2010. He’d graduated from college and followed his girlfriend to the Bay Area, where he found his history degree had approximately the same effect on employers as a face tattoo. Months went by, and Rob grew increasingly desperate. Then came the cold call from CDI, a contracting firm. The recruiter wanted him to interview for a position with Google, moderating videos on YouTube. *Google!* Sure, he would just be a contractor, but he was told there was a chance of turning the job into a real career there. The pay, at roughly \$20 an hour, was far superior to a fast-food salary. He interviewed and was given a one-year contract. “I was pretty stoked,” Rob said. “It paid well, and I figured YouTube would look good on a résumé.”

For the first few months, Rob didn’t mind his job moderating videos at YouTube’s headquarters in San Bruno. His coworkers were mostly new graduates like himself, many of them liberal arts majors just happy to have found employment that didn’t require a hairnet. His supervisor was great, and there were even a few perks, like free lunch at the cafeteria. During his eight-hour shifts, Rob sat at a desk in YouTube’s open office with two monitors. On one he flicked through batches of 10 videos at a time. On the other monitor, he could do whatever he wanted. He watched the entire *Battlestar Galactica* series with one eye while nuking torture videos and hate speech with the other. He also got a fascinating glimpse into the inner workings of YouTube. For instance, in late 2010, Google’s

legal team gave moderators the urgent task of deleting the violent sermons of American radical Islamist preacher Anwar al-Awlaki, after a British woman said she was inspired by them to stab a politician.

But as months dragged on, the rough stuff began to take a toll. The worst was the gore: brutal street fights, animal torture, suicide bombings, decapitations, and horrific traffic accidents. The Arab Spring was in full swing, and activists were using YouTube to show the world the government crackdowns that resulted. Moderators were instructed to leave such “newsworthy” videos up with a warning, even if they violated the content guidelines. But the close-ups of protesters' corpses and street battles were tough for Rob and his coworkers to handle. So were the videos that documented misery just for the sick thrill of it.

“If someone was uploading animal abuse, a lot of the time it was the person who did it. He was *proud* of that,” Rob says. “And seeing it from the eyes of someone who was proud to *do* the fucked-up thing, rather than news reporting on the fucked-up thing—it just hurts you so much harder, for some reason. It just gives you a much darker view of humanity.”

Rob began to dwell on the videos outside of work. He became withdrawn and testy. YouTube employs counselors whom moderators can theoretically talk to, but Rob had no idea how to access them. He didn't know anyone who had. Instead, he self-medicated. He began drinking more and gained weight.

It became clear to Rob that he would likely never become a real Google employee. A few months into his contract, he applied for a job with Google but says he was turned down for an interview because his GPA didn't meet the requirement. (Google denies that GPA alone would be a deciding factor in its hiring.) Even if it had, Rob says, he's heard of only a few contractors who ended up with staff positions at Google.

A couple of months before the end of his contract, he found another job and quit. When Rob's last shift ended at 7 pm, he left feeling elated. He jumped into his car, drove to his parents' house in Orange County, and slept for three days straight.

Several content moderation companies have their Manila offices on Ayala Avenue, in the main business district. PHOTOGRAPH : MOISES SAMAN/MAGNUM

GIVEN THAT CONTENT moderators might very well comprise as much as half the total workforce for social media sites, it's worth pondering just what the long-term psychological toll of this work can be. Jane Stevenson was head of the occupational health and welfare department for Britain's National Crime Squad—the UK equivalent of the FBI—in the early 2000s, when the first wave of international anti-child-pornography operations was launched. She saw investigators become overwhelmed by the images; even after she left her post, agencies and private organizations continued to ask for her help dealing with the fallout, so she started an occupational health consultancy, Workplace Wellbeing, focused on high-pressure industries. She has since advised social media companies in the UK and found that the challenges facing their content moderators echo those of child-pornography and anti-terrorism investigators in law enforcement.

“From the moment you see the first image, you will change for good,” Stevenson says. But where law enforcement has developed specialized programs and hires experienced mental health professionals, Stevenson says that many technology companies have yet to grasp the seriousness of the problem.

“There's the thought that it's just the same as bereavement, or bullying at work, and the same people can deal with it,” Stevenson says. “All of us will go through a bereavement, almost all of us will be distressed by somebody saying something we don't like. All of these things are normal things. But is having sex with a 2-year-old normal? Is cutting somebody's head off—quite slowly, mind you; I don't mean to traumatize you but beheadings don't happen quickly—is that normal behavior? Is that something you expect?”

In Manila, I meet Denise (not her real name), a psychologist who consults for two content-moderation firms in the Philippines. “It's like PTSD,” she tells me as we sit in her office above one of the city's perpetually snarled freeways. “There is a memory trace in their mind.” Denise and her team set up extensive monitoring systems for their clients. Employees are given a battery of psychological tests to determine their mental baseline, then interviewed and counseled regularly to minimize the effect of disturbing images. But even with the best counseling, staring into the heart of human darkness exacts a toll. Workers quit because they feel desensitized by the hours of pornography they watch each day and no longer want to be intimate with their spouses. Others report a supercharged sex drive. “How would you feel watching pornography for eight hours a day, every day?” Denise says. “How long can you take that?”

An employee at the Manila offices of Open Access BPO, another company that provides content moderation services. Moises Saman/Magnum

Nearby, in a shopping mall, I meet a young woman who I'll call Maria. She's on her lunch break from an outsourcing firm, where she works on a team that moderates photos and videos for the cloud storage service of a major US technology company. Maria is a quality-assurance representative, which means her duties include double-checking the work of the dozens of agents on her team to make sure they catch everything. This requires her to view many videos that

have been flagged by moderators.

“I get really affected by bestiality with children,” she says. “I have to stop. I have to stop for a moment and loosen up, maybe go to Starbucks and have a coffee.” She laughs at the absurd juxtaposition of a horrific sex crime and an overpriced latte.

See What's Next in Tech With the Fast Forward Newsletter

From artificial intelligence and self-driving cars to transformed cities and new startups, sign up for the latest news.

Your email

Enter your email

By signing up you agree to our [User Agreement](#) (including the [class action waiver and arbitration provisions](#)), our [Privacy Policy](#) & [Cookie Statement](#) and to receive marketing and account-related emails from WIRED. You can unsubscribe at any time.

Constant exposure to videos like this has turned some of Maria's coworkers intensely paranoid. Every day they see proof of the infinite variety of human depravity. They begin to suspect the worst of people they meet in real life, wondering what secrets their hard drives might hold. Two of Maria's female coworkers have become so suspicious that they no longer leave their children with babysitters. They sometimes miss work because they can't find someone they trust to take care of their kids.

Maria is especially haunted by one video that came across her queue soon after she started the job. “There's this lady,” she says, dropping her voice. “Probably in the age of 15 to 18, I don't know. She looks like a minor. There's this bald guy putting his head to the lady's vagina. The lady is blindfolded, handcuffed, screaming and crying.”




The video was more than a half hour long. After watching just over a minute, Maria began to tremble with sadness and rage. Who would do something so cruel to another person? She examined the man on the screen. He was bald and

appeared to be of Middle Eastern descent but was otherwise completely unremarkable. The face of evil was someone you might pass by in the mall without a second glance.

After two and a half years on the cloud storage moderation team, Maria plans to quit later this year and go to medical school. But she expects that video of the blindfolded girl to stick with her long after she's gone. "I don't know if I can forget it," she says. "I watched that a long time ago, but it's like I just watched it yesterday."

UPDATE: 24/10/2014 11:44 PT: Several captions in this story were updated to accurately reflect the name of the company Open Access BPO.

More Great WIRED Stories

-  The latest on tech, science, and more: [Get our newsletters!](#)
- Unmasking [Pedro Pascal](#), the new face of sci-fi
- The flight tracker behind [@ElonJet](#) took a left turn
- I spent a week using only [TikTok](#) for search
- You may survive a [nuclear blast](#)—with the right shelter
- Alphabet's [layoffs](#) aren't very Googley
-  Explore AI like never before with [our new database](#)
-  Upgrade your work game with our Gear team's [favorite laptops](#), [keyboards](#), [typing alternatives](#), and [noise-canceling headphones](#)

TOPICS MAGAZINE-22.11 CONTENT MODERATION SOCIAL MEDIA FACEBOOK TWITTER GOOGLE
WIRED CLASSIC LONGREADS
