



## A HOW-TO GUIDE FOR CONTENT MODERATION IN 2022



*A Conectys made Whitepaper*



## WHERE WE STAND NOW

Content moderation became a huge topic in the early days of 2022, especially around social platforms (i.e. Twitter) and their banning of certain public figures (i.e. then-U.S. President Trump). Entire episodes of podcasts were dedicated to content moderation, executives from some platform companies [were called out](#), and some even stepped up to claim that “content moderation does not matter.”

As a company that [works with many global, hyper-scale brands on content moderation](#), we can unequivocally say that content moderation matters, and brands should take it very seriously. It’s a complicated issue of speed, scale, and security -- you want your users to feel like their posts appear immediately (speed), but you also need them to feel safe in the community (security), and you need the entire ecosystem to scale with little to no warning.

It’s not easy to execute upon, and brands pivot and move around on their strategy periodically. We’ve worked with brands and helped them co-create new best practices, new coverage hours, new languages they need to moderate, etc. It’s a consistently moving target.

## SOME OF THE CORE ISSUES AROUND CONTENT MODERATION NOW

First is the issue of tech vs. human. When social platform Parler was effectively taken offline because their host (Amazon Web Services) declined to host for them anymore, Parler’s CEO vowed they would return -- with [content moderation done by algorithm](#). This is a common concept these days with current regulatory environments; to maintain scale and cost effectiveness, brands try to turn more of their content moderation efforts over to algorithm and Artificial Intelligence (AI).

Our view has always been that the best moderation work is performed with a mix of technology aiding humans. Tech helps with cost and scale, and it reduces the tedium on humans, which can have mental health consequences. It also reduces the amount of extreme or highly-negative content that moderators might be seeing. Those are positives, of course.

But humans need to be part of the equation, because they have language and societal nuance that today’s AI lack -- and when humans know what the core moderation issues are, they can tell executives back at brand HQ about what’s going on in their communities. That insight can be valuable for future brand decision-making. Periodically, it even informs new product decisions.

The other big issue hanging over current content moderation is training. Moderators need to be vigorously trained and re-trained. They need to understand what’s harmful, what’s not, where the brand is going, the guidelines, the best practices, the business model of the moment, and more.

## CONTENT MODERATION TRENDS THAT NEED ADDRESSING

- A mix of tech and human in moderation
- A balance of speed, security, and scale
- An eye towards the evolving regulatory environment
- Screening, training, re-training, and more re-training
- A constant focus on moderator mental health

We're now going to walk through these topics and indicate what you need to do for your brand in 2022 (and beyond). Let's go.

### THE MIX OF TECH AND HUMAN

The thinking is often: Get more tech to put towards these challenges of user-generated content (UGC). That definitely helps, but it's not the full picture. Remember: at its best, tech is a force multiplier. Tech is necessary within UGC, but it's not sufficient enough to moderate at scale.

Consider the role of AI in moderation. We've consistently thought AI will be the primary guiding force in UGC at scale, and AI is certainly helpful as a tool.

But -- there's a good deal of cultural nuance when moderating, so much so that it's been referred to as "the impossible job." (The good news is, it's not impossible.)

Consider this distinction: in the United Kingdom, irony can be the default style of humor. In the U.S., it's not. Hence, a comment such as "This is rubbish, we should do away with the poster" can take on a very different tone coming from someone in each country. In the US, a comment like that might be read as a direct threat; in the UK it might be dismissed as "ribbing" the poster.

AI is a strong tool in the moderation arsenal, but it lacks a deeper nuance right now. In many ways, [it's still very nascent](#).

Basing your scaled moderation strategy almost completely on these types of technology won't work right now.

If the answer is not "Throw more technology at it" for companies facing a massive scale of UGC moderation needs, what is the answer right now?

For many, it's outsourcing -- simply because companies lack the resources to moderate at speed and scale internally. And to be blunt, when moderation is done poorly, there are massive legal and brand repercussions, so it's often an area of the business best outsourced to experts in UGC and moderation.

**What to do now:** Find an outsourcing partner that works in moderation (we do, in 35+ languages with 24/7 global coverage from 12 locations + WfH). Then look over some of the [necessities of a moderation partner](#), and make sure the partners you're considering have experience, scale, past clients they can speak to, and use a mix of technology and human agent. Remember also: human agents dealing with moderation needs all day (with some of the tedium replaced by technology) will understand who and what is being posted on your platform, and that insight can become conversations that inform product decisions,

marketing campaigns, and sales approaches. Tech alone cannot get you there.

## SPEED, SECURITY, AND SCALE

As a user base grows, the scale of moderation can become immense. As the scale becomes that massive, you confront a major problem: how do you keep the community safe and in keeping with your community guidelines without sacrificing speed? If someone posts a comment or photo to your community, they want to see it show up immediately. This give-and-take between security and speed for large-scale communities can be very challenging to navigate.

What you need to know on scale is: Can the potential partner match the growth of your company and complexity of content? Can they be a part of your entry into new markets and geographies?

On the quality and speed front: How capable is the partner at each? The BPO space -- broadly -- has a bad reputation for over-promising and under-delivering. You want to make sure you're getting the right mix of quality and speed that you were promised during the negotiations.

**What to do now:** Ask the right questions. What security measures does the partner have? What metrics around security are tracked? What clients and in what industries are they currently working with? Can they provide global scale? Flexible staffing and pricing models? What certifications do they have? What you're ultimately looking for is someone who can take you as global as you desire to be, but can do so in a way that keeps your communities safe. Ask about different approaches to moderation as well: pre-moderation, live post moderation, platform tech support, and more. What is the full ecosystem that can be offered?

## REGULATIONS

This one is tricky and it changes seemingly every month in different countries. The EU has one context for the moderation of content, and the USA and South America and Asia are very different. While all governments have the same big goals in mind -- make sure that people are not harassed or subjected to negativity online, and make sure the framework of conversation is civil (ideally don't drag down basic human interaction principles in the process) -- the legal and regulatory approach can vary. The world is fairly diverse in its approach to content moderation.

**What to do now:** First, evaluate a potential moderation partner for localization. What languages and areas has the outsourcing partner worked in? If a potential partner can only work in a few languages and it's well below the breadth of your content reach, the partnership likely won't work, or you'll need to bring in an additional partner to cover those additional geographies. In addition to that, however, the more nations and languages that a provider has worked in, the more they tend to understand those specific regulatory environments.

Next: can the partner be a consultant? This is an area Conectys prides itself on; we don't consider ourselves a vendor so much as a strategic partner for the brands we have the pleasure of serving. You ideally want to find a partner in UGC and community moderation who is more than simply transactional -- you want advice and best practices from this business relationship too. If you're working with an expert, you should expect both strategy and execution. If we know the specific environment -- moderation best practices in Turkey vs. Poland vs. the USA, for example -- we will offer that insight to you and co-execute a recruitment plan, overall strategy, training process, re-training process, and more.

Our goal is to give you everything you need to be successful across multiple regulatory environments.

## TRAINING AND RE-TRAINING

For agents to be successful, they need to be recruited well, trained well, and then re-trained repeatedly, especially if anything around your business model or delivery methods have changed. If one type of speech or one comment (see “rubbish,” above) was permissible and then becomes not permissible based on new executive thinking or co-strategizing, the agents need to know that. So, like with every aspect of work, it all boils down to clear communication and effective training.

**What to do now:** Look for a partner with an end-to-end performance management approach that takes into account successful hiring, on-boarding, training, re-training, and more.

Here's an overview of ours, for example: *Ultui cordi iur, Cupiocris cor aus hoctatume ia*

## End-2-End Performance-Focused Approach




You need a rigorous commitment to training because business models and approaches pivot so often right now, and moderation needs + regulations can as well. So your people need to be the lifeblood of the partnership, and if you do that right, you tend to get the best moderation -- and the best future strategy -- out of it.

## MODERATOR MENTAL HEALTH

One of the big trends we've seen in the last two years has been a greater understanding of the psychological implications of content moderation.

This all begs the question: What is the responsibility of content moderation providers, what is the responsibility of the platform hosting the content, and how can both together to ensure the health and well-being of moderators?

**What to do now:** Look for partners that can offer preventive action. Our managers, for example, have extensive training on how to identify possible signs of mental health issues and how to offer support in such cases.

Protecting employees has been a paramount concern of executives we've spoken with -- a great development! -- so this is definitely an area to be asking potential partners about.

## THE BOTTOM LINE

The [CAGR on content moderation services is 10.3%](#), 2019-2026. It's a growing space -- and quickly. There's a good chance you are already moderating at this point, and might have been for several years. But now you need to think about the next steps: post-COVID, new business models, pivots, etc. What are you doing to set yourself up for moderation success and moderation strategy in the next, say, three years?

For example, region-specific moderation is on the rise and will increase as brands look to new opportunities (markets) post-COVID.

["Hidden digital labor"](#) is another increasing concern, and refers to the tedium and emotional labor of multi-hour moderation, and content moderation partners (like us) and brands are increasingly trying to protect brand agents against it.

As you keep thinking about content moderation trends and what your brand needs to do with content moderation, check out some of our moderation services.

We'd love to talk more, as well.





Contact us at [sales@conectys.com](mailto:sales@conectys.com)

Or learn more from our [Blog](#).

