


Whitepaper

# NAVIGATING RISKS AND OPPORTUNITIES IN METAVERSES: EVERYTHING BUSINESS LEADERS NEED TO KNOW



**Users won't engage in an environment that feels unsafe. Content moderation will set the benchmark for brand success in metaverses – and here's how you can maximize it with a human-first approach.**



## Metaverses: Worlds of Data

There's a lot of hype surrounding metaverses, and for good reason. They'll unlock more ways to engage customers, elevate brand presence, and make internal business processes such as employee onboarding more efficient. And we'll all be exchanging a lot of data in the process.

What's the catch? More data often brings more problems.

Namely: more vulnerabilities that could expose your business and its customers to risks like identity fraud, asset theft, and transaction hijacking. At the same time, there'll be new ways for misinformation to spread, harmful or disturbing content to proliferate, and harassment to escalate. Keeping your users safe will be critical to keeping them coming back to your brand.

Given the volume of information being uploaded, exchanged, and promoted within metaverses, you'll need next-level safety and security technology to moderate identity and content at this scale. This will be key to protecting users and earning their trust. But the machines can't do it alone.

If this sounds counterintuitive, consider AI's limitations against the new complexity that metaverse content will bring. The best of human empathy and judgment is needed to mitigate bias and truly protect users.

And that's why your content moderation strategy must be human-first. By applying human intelligence to automation and using this to protect both your users and your content moderators, you'll set your business up for success in metaverses.

Read on to discover how best to go about it.





# Table of Contents

*/ The Online World Is Changing*

*/ 1. Navigating the Risks*

*/ 2. Turning Challenges Into Opportunities*

*/ 3. Content Moderation in Metaverses*

*/ Learn, Learn, and Learn Again*

*/ The Sutherland Approach*







# The Online World Is Changing

Do you remember 1995?

This was the year that saw a breakthrough in the internet's future, with 10 million users accessing an early web browser<sup>1</sup> named Netscape Navigator. After this, the world as we knew it went through significant change.

Well, it's happening again. Today, instead of dialing up on our modems, we're starting to grab our AR-enabled mobile phones, connect over fiber optics on PCs, and maybe pop on a VR headset for full immersion. And in doing so, we're gaining access to metaverses – rich, immersive virtual environments where you can interact, shop, game, learn, create, and work.

By 2026, an estimated 25% of the global population<sup>2</sup> will log onto a metaverse for at least an hour a day. And with this comes an issue we've seen before: exponential levels of users, resulting in exponential levels of data and complexity.

As of April 2022, users' online activities generated 1.7 million Facebook posts and 1 million hours of streaming content every minute.<sup>3</sup> That's already a lot of content and user information for platforms to handle. But by spending just 20 minutes in a VR simulation, a user gives away two million unique recordings of their body language alone.<sup>4</sup>

The levels of data creation and collection about users in a metaverse environment will be significantly larger than what we're all used to. You'll need to adapt quickly if you want to demonstrate responsible data practices to your users.

<sup>1</sup> Source: <https://www.scienceandmediamuseum.org.uk/objects-and-stories/short-history-internet>

<sup>2</sup> Source: <https://www.gartner.com/en/articles/what-is-a-metaverse>

<sup>3</sup> Source: <https://www.statista.com/statistics/195140/new-user-generated-content-uploaded-by-users-per-minute/>

<sup>4</sup> Source: <https://www.devex.com/news/sponsored/opinion-the-challenges-of-protecting-data-and-rights-in-the-metaverse-103026>



## The Biggest Currency in Metaverses (Hint: It's Not Crypto)

Consumers expect the brands they engage with to protect them. In fact, business relationships with customers bank on trust.

This comes naturally face-to-face. In a bricks-and-mortar store, for example, you expect staff to make you feel welcome and safe by providing a positive customer experience and maintaining a level of security – both physically and digitally.

These factors are what earn your trust. But how do you recreate this virtually?

Businesses will need new tools to connect with users and help them feel comfortable in this new environment. And early adopters will have the chance to help create and shape them. But at the same time, bad actors will use these same tools and the opportunities created by metaverses to wreak havoc on businesses and users.

Left unchecked, this will lead customers to distrust any interaction they have in this environment – including interactions with your business.

Trust, therefore, is the currency businesses in the metaverse space must bank on, just like in the physical world. But to earn this trust, they must first find a way to address the risks that are already concerning users.

## Content Moderation Strategies are Fundamental

In metaverses, it's not just the increased danger brought on by bad actors that we need to worry about. It's also the fact that in an immersive space, adverse incidents have the potential to become much more believable to users.

To address both factors, security systems and content moderation strategies are not just needed – they'll set the benchmark for your business' success in the metaverse.





# 1 / **Navigate the Risks**

The adoption of metaverses is set to increase the online threat landscape exponentially. Bad actors will discover new ways to:

- ▶ Spread false information.
- ▶ Steal identities, funds, or assets.
- ▶ Harass or intimidate users.

To protect both your business and its customers, you'll need to understand how risk is evolving. This will unlock new methods to mitigate danger and help your users feel safe.

## **Data: The Ultimate Double-Edged Sword**

In the online world, data offers both the most rewards and the biggest risks. Customer data, trends data, economic data – it's what drives customer experience and propels business forward.

At the same time, data is what exposes both brands and users to risks - ones that are becoming harder and harder to manage.

Which is why consumers are already worried about how their identity and data will be protected in metaverses. In fact, over one in two U.S. adults have cited tracking and misuse of their personal data as a major concern.<sup>5</sup> And with the amount of data they're giving away via VR holding the potential to reveal mannerisms, physical features, and even the layout of their homes, they've got a right to be worried.

---

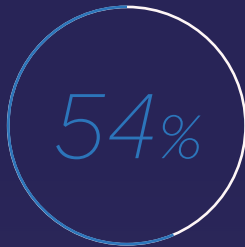
<sup>5</sup> Source: <https://morningconsult.com/2022/04/11/metaverse-public-concerns-survey/>



# ANONYMITY HAS A NEW FACE

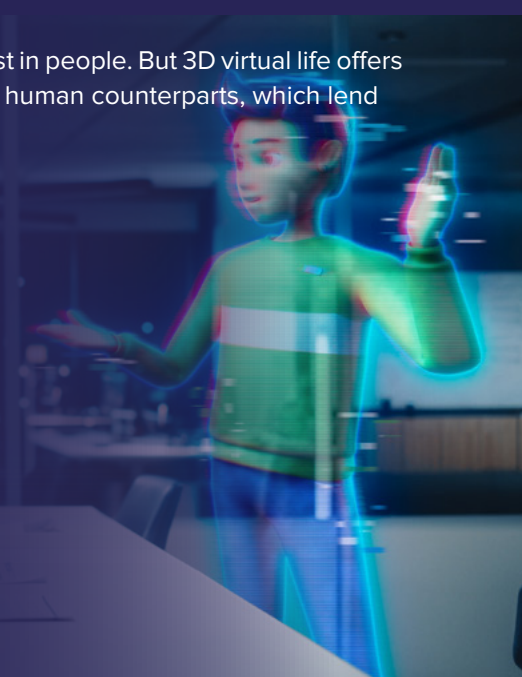
Facelessness has a heady influence on us. Example: would you phrase an in-person complaint in the same way that you would in an email?

We all know from the internet that anonymity sometimes brings out the worst in people. But 3D virtual life offers new risks. Avatars have the potential to evolve into mirror images of their human counterparts, which lend themselves well to impersonation and identity fraud.



*On the other hand, 54% have admitted that they would create an avatar in a the metaverse that does not look like them in reality.<sup>6</sup>*

This presents its own problems when it comes to shaping virtual identities and opens questions about how we'll create meaningful relationships.



## The Internet, but Not as We Know It

Businesses shaping a metaverse presence must think carefully about how they'll regulate identity and content creation. And how they will do so in a way that doesn't limit creative expression or the user's experience.

But this doesn't have to hinder the new ways you can engage with your customers. With the right solutions to back you, new risks can turn into new opportunities.

<sup>6</sup> Source: <https://www.tidio.com/blog/metaverse/>







## 2 / **Turning Challenges** Into Opportunities

It all boils down to getting safety and security right. If you nail this, you get to claim the position of an industry leader and maximize the opportunities that metaverses have to offer.

Let's look at some of the opportunities you'll have to evolve your content moderation and customer experience strategies in metaverses.

### **Game-Changing Content Moderation**

Content moderation already plays a critical role in protecting today's online spaces. Millions of social posts are assessed daily to identify spam, harassment, and harmful content.

Taking this to the next level will unlock the trust you need to attract and retain customers in the metaverse space. But this becomes even more challenging with the amount of data involved.





Knowledge management is a crucial element here. In metaverses, businesses need to establish robust content moderation policies from the get-go. But manually remembering all the policies that apply to even just one topic will be an impossible task. And when those policies change overnight in response to global events or trends, this could create a great deal of stress.

Leveraging proprietary AI tools can help enhance and streamline processes, from enabling content translation to establishing and managing moderation policies at scale.



### Taking CX to the Next Level

Users want to know that they're being looked out for, given that they're worried about the dangers of metaverses. You can make leaps towards nurturing their relationship with your brand by engaging the best new tools metaverses have to offer.

From “face-to-face” customer interactions to security avatars manned by real people, metaverses give us new ways to recreate the systems that make us feel safe in the real world.



When incidents do occur, the best customer communication tools are needed to manage the relationship and help users feel heard. Proprietary AI tools can also lend a hand in managing customer sentiment to continually enhance the communication process.





## 3 / Content Moderation in Metaverses

# 3 Best Practices

With the levels of information involved in a metaverse, technology will be critical to keeping it running. But the machines can't do it alone. AI falls down when it comes to complex decision-making, allowing human content moderators to take the stage. But how does this combination operate in practice?

Here are three best practices to consider.

### 1. Mitigate Bias

AI and automation can process vast volumes of information. It's unbelievably capable when it comes to quick content moderation at scale.

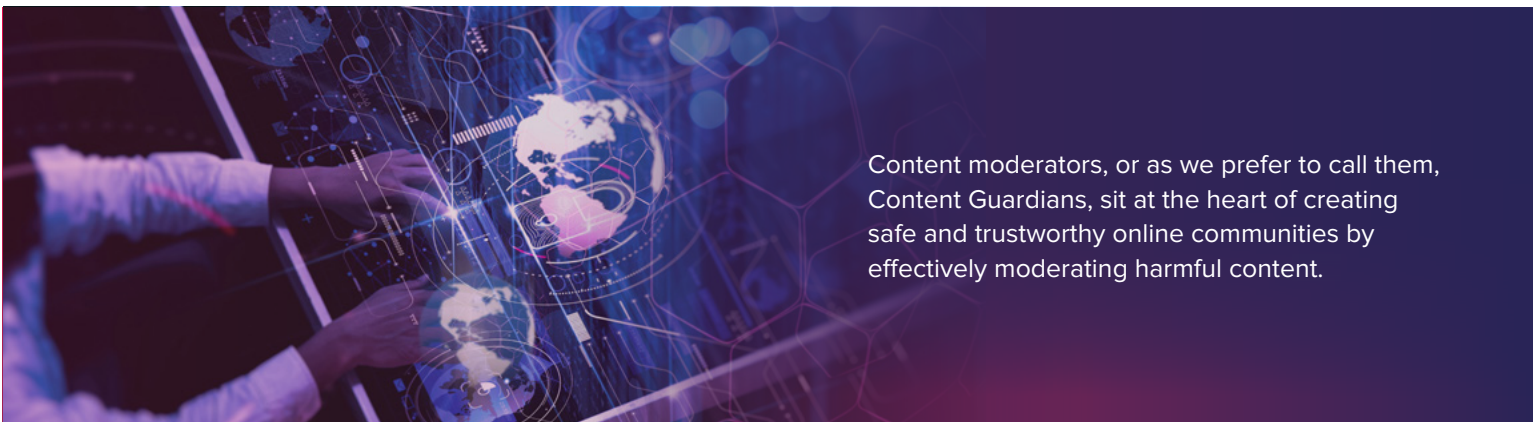
But where AI struggles is when complex situations arise. After all, it's only as good as the data sets it relies on. But these systems can't identify inappropriate or harmful content in some cases – in other words, they can't understand context and tone.

For instance, avatars bring new ways to express appearances and, with this, a whole host of new social cues and questions. Should virtual events have dress codes? Can an avatar's apparel be considered inappropriate or offensive? One thing is for sure: knowing the difference will be down to context, not rules.

Similarly, on the internet, new 'keywords' arise to take the place of prohibited terms. They can be used to perpetuate abuse but are often reclaimed by oppressed groups. Knowing the difference is hard enough, but in the metaverse, the diversity of content will make this even harder. It's a judgment call that the machines simply can't make.



When you add to the mix the new factor of what “personal space” means in a metaverse, there are some complex cases on the horizon. Making these nuanced decisions has to be done by humans. You need a diverse, multidisciplinary team that can handle the hard stuff, working hand-in-virtual-hand with AI.



Content moderators, or as we prefer to call them, Content Guardians, sit at the heart of creating safe and trustworthy online communities by effectively moderating harmful content.

## 2. Prioritize Your Protectors

The best of humans and machines must be called in to combat complex content moderation and keep everyone safe from harm. But Content Guardians are susceptible to harm themselves when faced with these decisions.

Making judgment calls is challenging when the line is blurred. And with content that has the potential to become even more graphic or disturbing, combined with the real-time pressure that requires them to act quickly, you're left with a human team that's tired, stressed, and even mentally unwell.

Short-term, this has a productivity and process impact. Long-term, it could ultimately damage their relationship with you and lead to higher turnover.

Protecting the wellbeing of your protectors is vital to ensure no damage to their mental health ensues. This means they can continue doing their best work - which makes sense from a business perspective - and because it's simply the right thing to do.

Your Content Guardians will benefit from a holistic approach that takes on their physical, social, and psychological wellbeing. You need internal processes that are proactive rather than reactive – flagging when someone feels unwell or unhappy and taking action immediately.

A strong wellbeing and resilience framework is key. Using tried and tested indices, you can monitor wellbeing on a day-to-day basis, anticipate overall performance, and deliver access to telemedicine and psychologists when needed. But how can you provide the best wellbeing to everyone on your team at the same time?





### 3. Bring in Automation To Protect Your Content Guardians

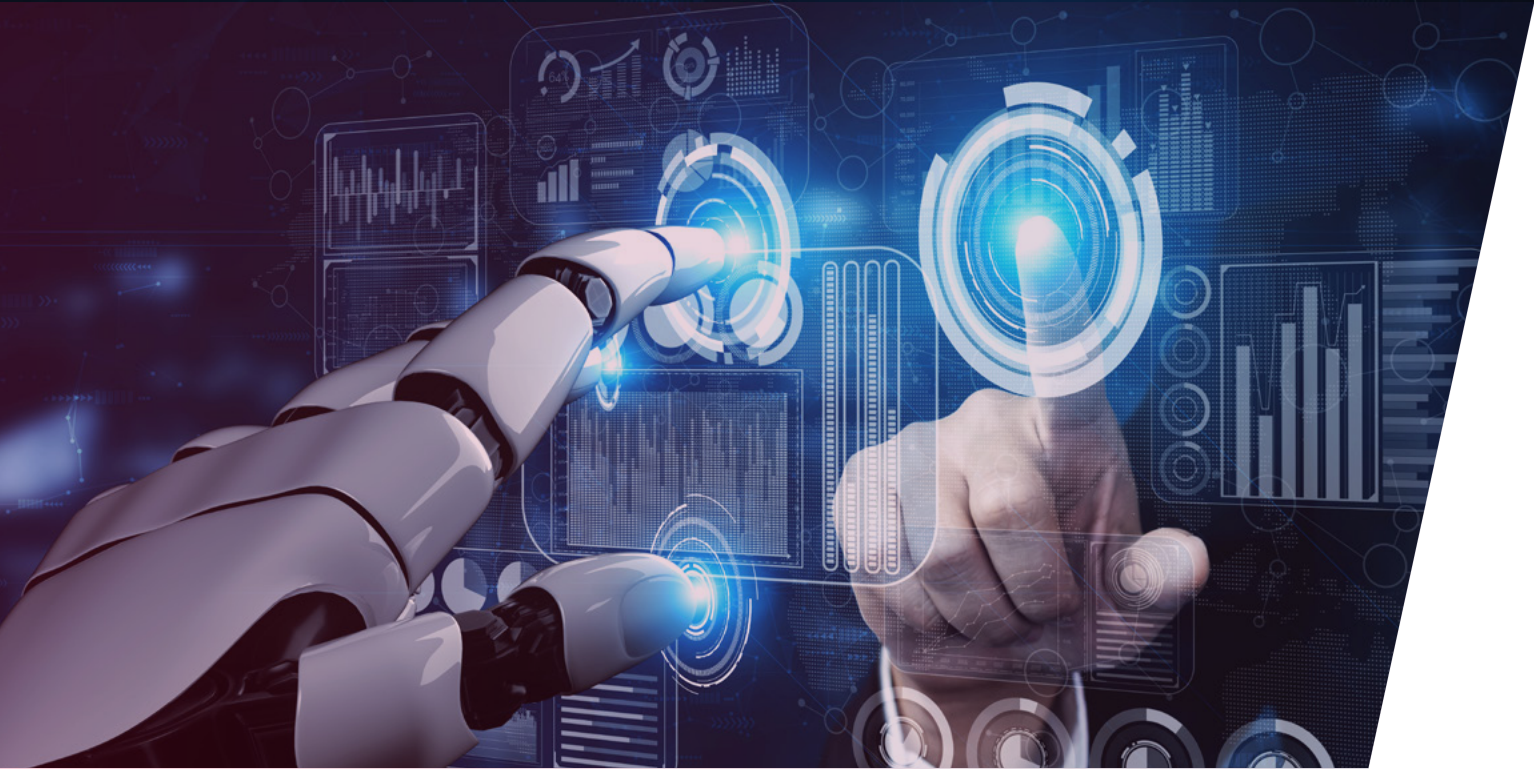
Human-centered automated processes do not just apply to protecting customers. They can also be brought under the hood to serve the protectors.

For example, empowering your Content Guardians with the systems to support their own health will ensure wellbeing is regularly maintained. Advanced technology will provide a big lift, from app-centered wellness activities to regular online quiz check-ins.

Automation can also support internal employee processes. This includes facilitating work allocation to stop individuals from becoming overburdened and helping to harmonize the human/machine dream team. It means you can streamline your digital hiring processes, bringing in AI-driven psychometric and stress tests to anticipate success. But intelligent technology can also support change management processes – which will be critical for your transition into the metaverse.

Your Content Guardians are the first line of defense in a challenging battle. Ultimately, they should feel proud of the vital work that they do to protect people from harm. And this starts with mitigating stress, promoting wellbeing, and harmonizing processes to invoke the best of human and artificial intelligence.





## Learn, Learn, and Learn Again

*No version of the internet can exist without human-oriented safety measures.*

The future of metaverses doesn't need to be an intimidating world fraught with unforeseen and complex risks. In fact, if we continue to apply our learnings from challenges we've already faced in the online world, we'll have a stable foundation to stand on as we explore the exciting opportunities and experiences that metaverses bring.

The bottom line? AI just can't do it alone, and neither can humans. The right proportion of both will fuel the best metaverse content moderation strategies – tackling scale with AI and automation while navigating ever-growing complexity with empathetic Guardians.

Transitioning safety and security measures to a metaverse environment will involve a huge digital transformation effort and a great deal of research and expertise. This will partly be down to learning, but you'll also need to fall back on tried and tested methods advised by seasoned experts.



# The Sutherland Approach

*Users, customers, employees – they won't engage in an unsafe or unwelcoming environment.*

When it comes to safety and security, Sutherland takes a human-centric approach that serves the needs of both end users and the Content Guardians that keep users safe.

Our innovative content moderation and trust and safety solutions stand out from the crowd, recognized by industry and market organizations. We support your metaverse transition by bringing:

- / Strategic support and setup from experience design and VR to content curation.
- / Domain-specific expertise, technology consulting, and implementation across industries: gaming, eCommerce, and more.
- / Design of proactive safety processes and policies, such as transaction screening, identity verification, and customer risk profiling.
- / Next level customer engagement and support solutions that deliver enhanced CX.
- / Labeling, moderating, and filtering inappropriate or harmful content in real-time, while strengthening AI measures.
- / A strong resilience and wellbeing framework to empower Content Guardians.

**Contact us now to discover how we can support your content moderation efforts in metaverses.**

We make digital  
**human™**

[sutherlandglobal.com](https://sutherlandglobal.com)  
[sales@sutherlandglobal.com](mailto:sales@sutherlandglobal.com)  
1.585.498.2042



Sutherland is an experience-led digital transformation company. Our mission is to deliver exceptionally designed and engineered experiences for customers and employees. For over 35 years, we have cared for our client's customers, delivering measurable results and accelerating growth. Our proprietary, AI-based products and platforms are built using robust IP and automation. We are a team of global professionals, operationally effective, culturally meshed, and committed to our clients and to one another. We call it One Sutherland.

