GetReal



SECURING THE HUMAN CAPITAL SUPPLY CHAIN:

Extending Zero Trust to the Human Layer

> In the GenAI era, assuming the person on the other side of the screen is who they claim to be is no longer a safe assumption.

For centuries, security has focused on safeguarding assets, first physical and then digital. Each era adapted its defenses as technology reshaped how value was created, accessed, and stolen. Over the last several decades, enterprises correctly invested in protecting networks, endpoints, and identities as work moved online. But something fundamental has shifted again.

Today, trust itself has become the attack surface. As remote work, digital collaboration, and generative AI converge, adversaries no longer need to breach systems to compromise organizations. They can simply appear as people operating deceptively inside workflows never designed for adversarial conditions. Hiring pipelines, help desks, and collaboration platforms now sit squarely on the front lines of enterprise risk.

This paper introduces the human capital supply chain as a previously unrecognized attack surface now actively exploited by adversaries. It argues that Zero Trust must evolve beyond devices and credentials to include the human layer where identity is inferred through face, voice, and behavior. In the GenAI era, assuming the person on the other side of the screen is who they claim to be is no longer a safe assumption.



Matthew Moynahan
Chief Executive Officer
GetReal Security

Executive Summary

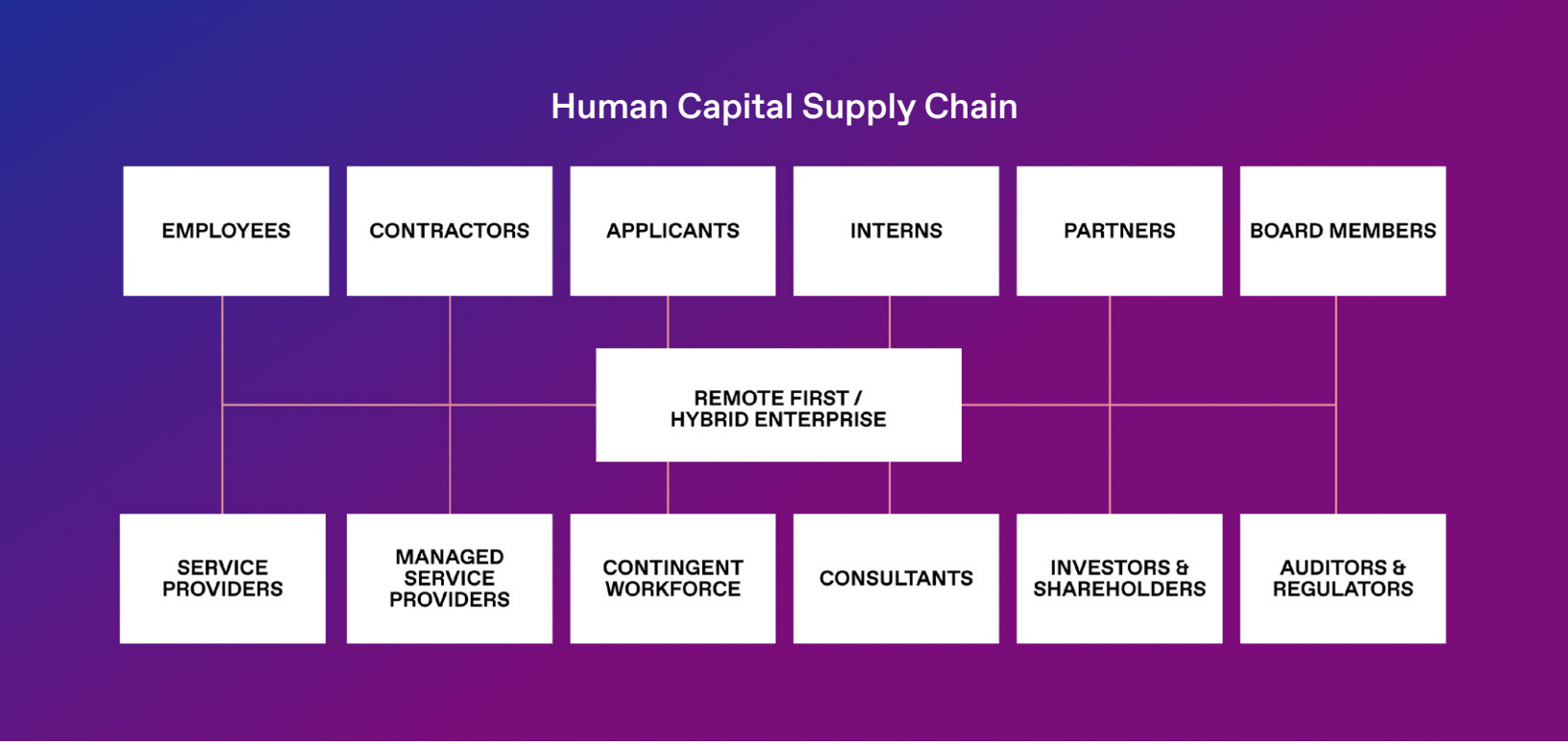
Digital transformation and generative AI have fundamentally altered how modern enterprises establish trust. Work now occurs primarily through remote, digital interactions, where identity is inferred through audio, video, and images rather than physical presence.

This shift has created a new category of enterprise risk: the **human capital supply chain threat**. Unlike traditional supply chain attacks that target software or infrastructure, these attacks exploit people themselves using synthetic, stolen, or manipulated human identities to gain access, influence decisions, or embed insider threats.

Recent incidents, including DPRK's phony IT worker and Scattered Spider attacks, demonstrate that adversaries no longer need malware or zero-day exploits to infiltrate an enterprise. They exploit trust-based workflows in hiring, identity recovery, and remote collaboration, operating in a security blind spot where traditional controls do not apply.

To address this risk, enterprises must recognize and secure a previously undefined attack surface: the human layer, where human identity is presented and interpreted through its digital representation in a digital media format. Extending Zero Trust principles to this layer requires treating human identity and behavior as continuously verifiable signals, not assumptions.

Enterprises that evolve their trust models accordingly will be better positioned to detect and disrupt human-based supply chain attacks. Those that do not will experience increased risk and fraud through new methods of compromise not grounded in malicious code but from people who were never who they claimed to be.



HUMAN CAPITAL SUPPLY CHAIN

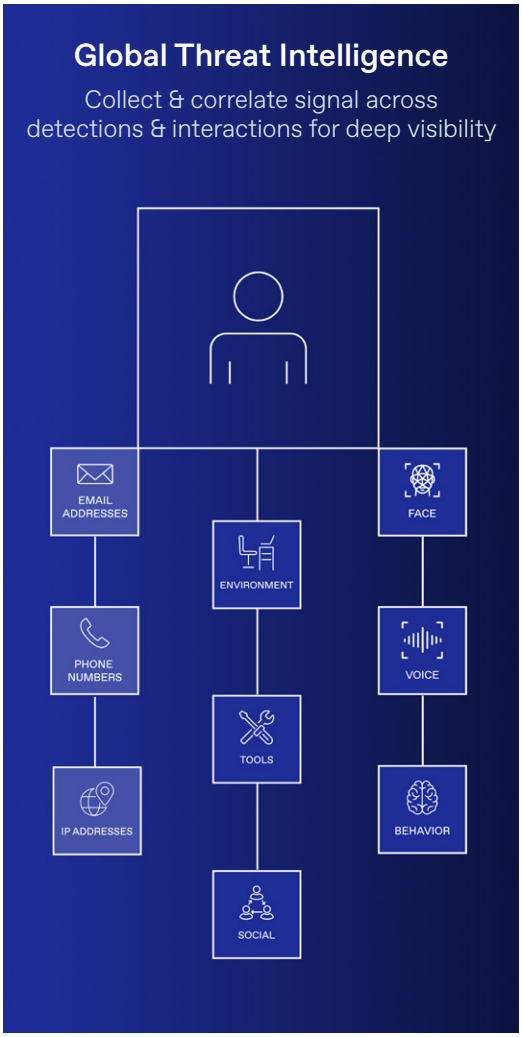
The network of people who enter, move through, and interact with an enterprise, often outside traditional security controls. This includes employees, contractors, job candidates, recruiters, and service providers.

Why the Human Capital Supply Chain Is Now a Security Problem

This paper introduces a new category of enterprise risk, the human capital supply chain threat, created by the collision of remote work, digital identity, and generative AI. It explains why traditional security architectures fail to detect or prevent adversaries operating at what we call the human layer, and outlines how Zero Trust principles need to evolve to defend the human layer of the enterprise.

The objective of the paper is to provide CEOs, CISOs, CIOs, and their teams:

- ➜ **A clear understanding** of how GenAI enables human-based supply chain attacks
- ➜ **Real-world evidence** that adversaries are already exploiting this risk
- ➜ **A new mental model** ("the human layer") for evaluating identity, trust, and access
- ➜ **Strategic guidance** for extending Zero Trust into the human layer and human capital supply chain



How Digital Transformation Created the Human Capital Supply Chain Risk

Enterprise work has changed dramatically over the past decade. Pursued as a strategic imperative, digital transformation ushered in a new era of work. Enterprises moved en masse from primarily on-premises infrastructure to digital-first operating models built on outsourced cloud infrastructure and application stacks.

This same capability enabled enterprises to manage through the Covid-19 pandemic when work-from-home mandates forced employees out of their physical offices almost overnight. This shift also normalized remote-first work and intensified global competition for talent, particularly for younger, technologically savvy workers who prioritize flexibility and choice over rigid enterprise policies.

Together, these trends have reshaped how work is conducted. Today, the majority of internal and external interactions among employees, job candidates, contractors, partners and customers occur through digital audio- and video-based communications across both managed (e.g., corporate laptops) and unmanaged environments (e.g., personal devices and consumer messaging platforms).

Advances in AI have significantly improved the quality of audio and video transmission, making modern collaboration platforms stable enough to sustain the remote, digital-first interaction model of today's extended enterprises.

Generative AI (GenAI) is now intersecting with this transformed enterprise environment, creating a new and largely underappreciated class of risk: the human capital supply chain threat. This threat does not directly target code repositories, hardware vendors, or software dependencies. Instead, it targets people themselves — their identities, skills, trust relationships, and decision-making authority — at machine speed and global scale.

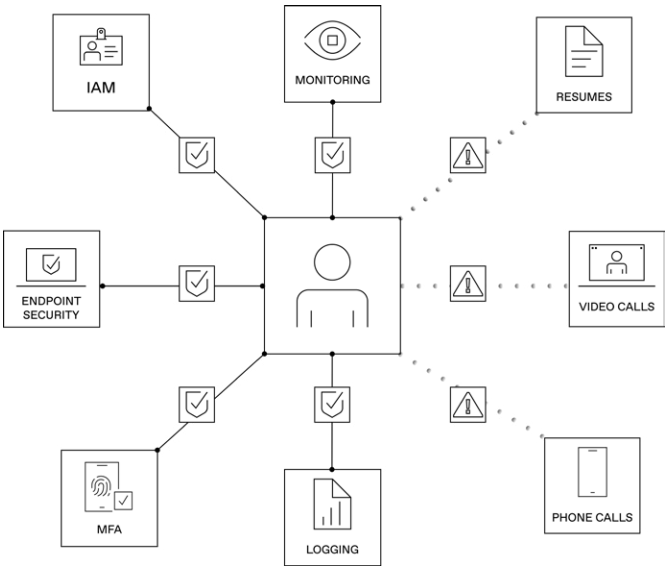
While security marketers have emphasized that "identity is the new perimeter" when referring to the importance of Identity and Access Management (IAM) infrastructure – this threat is considerably different.

For the first time, cybersecurity architectures are challenged by the fact that a person's physical characteristics in digital form, such as their face and voice, now constitute an attack surface. These attributes are being weaponized as techniques within adversary TTPs (Tactics, Techniques, and Procedures) used for social engineering and identity abuse. These techniques exploit human trust to gain access, influence decisions, or bypass traditional controls. While identity has been a battleground for decades, never before have we had to question the authenticity of a human sitting before us and speaking to us.

Now those same digital human attributes can also function as real-time IOCs (Indicators of Compromise), revealing that a trusted entity has been synthetically manipulated or impersonated. Digital human characteristics now function as both an attack vector and a detection signal in identity-based attacks, and most enterprises have not yet fully accounted for this shift.

> Digital human characteristics now function as both an attack vector and a detection signal in identity-based attacks, and most enterprises have not yet fully accounted for this shift.

Zero-Trust Blind Spots



> While meaningful progress has been made, managing the risk across hundreds or thousands of interconnected people and vendors remains difficult.

People Have Become a Supply Chain: From Technical Dependencies to Human Ones

Traditional supply chain risk management focuses on third parties that provide software, infrastructure, or services. Enterprises have responded with vendor risk management programs, software bills of materials (SBOMs), and zero trust architectures to mitigate these risks. While meaningful progress has been made, managing the risk across hundreds or thousands of interconnected people and vendors remains difficult.

Modern enterprises also depend on a human capital supply chain: employees, contractors, consultants, offshore teams, recruiters, and job candidates moving through hiring pipelines and global talent marketplaces. It is critical to recognize that this human supply chain operates almost entirely outside traditional security controls associated with employees. Employees are protected by IAM, MFA, endpoint security, logging, and monitoring. Interactions with job candidates and other third parties, however, exist in a security blind spot.

There is no zero trust model for resumes. No MFA for interviews. No continuous monitoring for people who have not yet been granted accounts—but who may soon receive deep access. And no enterprise policy or control exists to prevent employees from unintentionally inviting an adversary to join a company Zoom or Teams call. Yet in this complicated, globally inter-connected web of 8 billion people, enterprises still rely on antiquated background checks and back-channel references that are too light and too late in the process. They also fail to actually verify whether the person appearing on the other side of a screen, or phone call, is who they claim to be.

And, GenAI dramatically increases the security risk created by this gap.

Real-World Proof: Human Capital Supply Chain Attacks in Practice

The DPRK Phony IT Worker Campaign

The threat is not theoretical. One of the clearest examples is the DPRK (North Korea) phony IT worker campaign, and its success infiltrating the ranks of the F1000, typically the organizations with the most advanced cybersecurity programs.

In this operation, DPRK-linked threat actors pose as legitimate candidates applying for remote IT roles using fake, stolen, or borrowed identities. In many cases they use GenAI to augment the persona with fabricated artifacts such as phony resumes, LinkedIn profiles, GitHub accounts, and other supporting details. In other words, nearly everything about the candidate is fabricated, except the physical person showing up for the interview.

These candidates successfully pass resume screenings, technical interviews, and onboarding. They gain employment at Western enterprises—often in sensitive engineering and IT roles. In many cases, the same physical person is applying for dozens of jobs across functions and geographic regions with only slight modifications to their persona to evade detection.

This campaign has been effective largely because traditional identity and security controls are mostly absent from enterprise hiring processes and those of third-party service providers, such as contractors or outsourced recruiting firms doing the majority of their job placements. The DPRK threat has exposed many weaknesses in the digitized workflows in the Human Resource function, including:

- **Background checks** that validate documents rather than in-person authenticity
- **Identity verification** is point-in-time rather than continuous
- **Recruiting workflows** are optimized for speed rather than adversarial resistance
- **Security teams** rarely monitor or influence talent acquisition workflows and systems

GenAI significantly amplifies this model. AI-generated resumes, GitHub histories, certifications, technical answers, and real-time interview assistance make synthetic candidates nearly indistinguishable from legitimate ones. In many cases, meaningful security controls are not applied until after the attacker has been hired and becomes an insider threat.

By that point, the human capital supply chain has already been compromised.

Scattered Spider Attacks on IT Help Desks

Another example of social engineering at the human layer is the set of techniques used by organized criminal groups such as Scattered Spider, which has successfully demonstrated how effective human-centered attacks can be without exploiting software vulnerabilities. The MGM incident (September 2023) is a defining case study in modern cyber risk because:

- **No software** vulnerability was exploited
- **No malware** was initially required
- **No zero-day** was involved

The breach succeeded almost entirely through human social engineering, exploiting identity recovery and help desk workflows— in this case, impersonating an MGM employee to manipulate a help desk staffer, compromising the identity recovery process workflow, and provisioning new devices and credentials providing network access to the attacker.

> Digital human characteristics now function as both an attack vector and a detection signal in identity-based attacks, and most enterprises have not yet fully accounted for this shift.

These attacks succeed because security controls are unevenly applied to remote interactions with humans we have grown to trust or believe we should trust. Help desk workers often act in good faith and are incentivized to resolve issues quickly and maintain caller satisfaction, which places speed of identity and credential recovery at the center of performance evaluations and creates an exploitable attack angle. In addition, contractors and non-employees ironically receive weaker scrutiny than full-time employees. As a result, attackers know most enterprise IT workflows often operate outside of modern security threat models. This is a textbook example of a human capital supply chain attack.

In remote work environments, exploiting the trust-based processes and systems enterprises rely on every day is proving to be the weakest link in enterprise security architectures. Essentially, where trust exists in human interactions, the humans involved become the “last control point” for the enterprise. Social engineering, now coupled with GenAI, dramatically increases the success rate of these attacks through highly personalized pretexting, contextual awareness, and voice or text impersonation. What appears to be a “user or employee problem” is, in reality, a systemic trust exploitation problem.

GenAI and the Collapse of Trust-Based Enterprise Assumptions

Digital transformation and the COVID-19 pandemic forced enterprises to optimize for speed, scale, and flexibility. It also normalized the notion of mediated trust or trusting the person you are speaking to on the phone or a Zoom or Microsoft Teams call. As phishing emails have illustrated, humans have a cognitive bias to trust. Numerous psychological studies show that human trust increases as you move from written word to spoken word (audio) to audio and visual confirmation (streaming voice + video). When you step back, you can see this implicit trust model baked into the fabric of the modern digital enterprise:

- **Remote and hybrid work**, where trust is established digitally
- **Identity-first architectures**, where credentials assume legitimacy
- **SaaS sprawl**, which fragments sensitive workflows
- **Agile talent models**, that emphasize contractors and short-term access
- **Self-service HR and IT**, designed to reduce friction

These systems implicitly assume that humans are real, honest, and verifiable. North Korea and Scattered Spider have shown this cannot and should not be assumed in any case, even when a physical human is showing up and directly engaging with the enterprise. GenAI will only break this assumption down even further as it will soon be indiscernible as to whether a person's face and voice is human or fake in real-time calls or video. Recent advances in generative AI (e.g., OpenAI's Sora and Google's Veo) and agentic AI illustrate the pace of progress. Combined with the fact that hiring, onboarding, and access recovery are optimized for efficiency and disconnected from security oversight, the result is deception that becomes global, scalable and affordable.

The Four Categories of Threat Actors



The Human Control Point and the Rise of Synthetic Threat Actors

A new risk archetype has emerged: the synthetic threat actor. Threat actors have historically operated in the seams of their target's IT infrastructure, preferring to hide their digital footprints as they move laterally on the network and never surface unless detected. In digital and remote-first work environments, GenAI has enabled and emboldened these actors to come out of the shadows and operate in the daylight, presenting themselves as legitimate people through fully synthetic or modified appearances. In doing so, they weaponize human trust and violate the inherent trust model of the enterprise through its people and systems.

These threat actors typically fall into four categories:

1. **A physical threat actor cloaked in a fake identity** (e.g. DPRK phony IT worker candidate);
2. **A physical threat actor impersonating a legitimate identity** (e.g. Scattered Spider);
3. **A synthetic threat actor using a legitimate identity** (e.g. executive impersonation deepfake)
4. **A different individual than the originally vetted person**, introduced through "bait-and-switch" schemes, to embed an insider threat or engage in wage/labor arbitrage

Regardless of the form, these actors gain legitimate access using fake or AI-assisted identities. Synthetic attackers become synthetic insiders by exploiting a fundamental ordering flaw: trust is granted before security is applied. They need only a short window of legitimate access to exfiltrate sensitive data, divert payroll or payments, insert backdoors, poison internal knowledge and AI systems, or manipulate business decisions.

This is a supply chain attack executed through people rather than code. Even Zero Trust architectures struggle when adversaries are indistinguishable from legitimate users. GenAI accelerates the challenge by breaking identity models that do not operate at the human layer. They verify credentials and devices but not whether the face, voice, or persona presented in a digital interaction corresponds to the physical individual the enterprise believes them to be.

Security awareness training was designed for a world in which humans are the targets of deception, not the deception itself. It teaches employees to verify requests, recognize social engineering patterns, and remain skeptical of suspicious interactions. But it also assumes the existence of a real individual on the other end.

GenAI breaks this assumption. Adversaries can manufacture, swap, or modify human personas. As a result, enterprises need threat intelligence focused not just on phishing techniques, social engineering techniques, or malware families, but on human adversaries themselves – the real individuals or groups themselves operating on borrowed, stolen, or synthetic identities. These actors reuse personas, reappear across organizations, and exploit trust at the human layer. They should be identified, correlated, and responded to in the same way as malware authors in earlier decades.

A New Class of Human and Identity-Centric Threat Intelligence

To address this gap, enterprises need a new category of security threat intelligence—one focused not on malware hashes or IP addresses, but on real human beings operating under fake or shifting personas in the digital realm.

These threat actors reuse identities, resumes, and playbooks across organizations, appearing repeatedly in hiring pipelines, contractor networks, and social engineering incidents. The sophisticated criminal organizations will coordinate across campaigns, companies, and geographies and commit repeatable, and therefore trackable, attacks. Just as malware authors in the 1980s and 1990s were eventually identified, cataloged, and hunted, today's human adversaries must be treated as first-class threat actors, given the damage they can cause.

This new category of threat intelligence needs to include:

- **Known threat actors** operating under synthetic or stolen identities
- **Suspicious personas** tied to prior incidents crowdsourced from enterprises
- **Behavioral and activity-based fingerprints** observed across interviews, communications, and enterprise workflows over time
- **Cross-company and cross-industry** correlation of recurring human-based attack patterns



These risks are concrete and applied repeatedly consisting of real people executing real attacks under shifting digital disguises across multiple enterprises.

Rethinking Security to Defend the Human Control Point

Defending against this threat requires expanding supply chain security to include human trust pathways and human-focused threat intelligence.

This responsibility extends beyond the enterprise to any party operating in the human capital supply chain. This includes outsourced recruiting and staffing agencies, BPO service providers, social platforms like LinkedIn, candidate sites such as Indeed, and technology service providers supplying the underlying technology infrastructure. Each of these entities should treat hiring and recruiting as security-critical attack surfaces that could place the enterprise – or a supplier's relationship with their enterprise customer – at risk. Identity assurance must be applied in every interaction to ensure biometric consistency between the represented identity and participating individual before the candidate is placed at an enterprise or any access is granted.

In addition, the following should be done:

- **Design HR and help desk workflows** for resilience to adversarial conditions
- **Integrate security, HR, legal, and procurement** into shared threat models
- **Track and respond to human-layer threats** with the same seriousness once reserved for malware
- **Ensure continuous verification** that the human represented in the digital content reflects the actual individual participating in the corresponding human-layer interaction

This is not simply a cybersecurity problem; it is a governance and organizational design problem that requires ownership.

Applying Zero Trust Principles to the Human Layer

Traditional ZTA evaluates identity, device security posture, and contextual signals to determine access to enterprise resources. These principles remain foundational, but they primarily assess trust through accounts, credentials, and devices rather than how identities are represented and interpreted in remote human interactions. As enterprises increasingly rely on remote and digital collaboration, participants make trust decisions based on how an identity is represented through face, voice, and behavior in digital content such as profile images, recorded media, and live audio and video calls. This expands the attack surface into the human layer itself.

The table below illustrates how Zero Trust extended to the human layer expands traditional trust evaluations.

Traditional Zero Trust Assessments: *Source: "Zero Trust Architecture" NIST SP 800-207*	**Human-Layer Zero Trust Adds:**
➜ What is the level of confidence in the subject's identity for this unique request?	➜ Is the declared identity represented in the digital interaction known or unknown? Is it human or sanctioned synthetic (AI agent)?
➜ Is access to the resource allowable given the level of confidence in the subject's identity?	➜ Based on context, what level of trust or influence should be extended to the represented identity in this interaction?
➜ Does the device used for the request have the proper security posture?	➜ What is the trust status of the represented identity (is it who or what participants believe it to be, and has it recently been seen or impersonated in other digital interactions)?
➜ Are there other factors that should be considered and that affect confidence (e.g., time, location of subject, subject security posture)?	➜ Are the attributes presented in the human layer (face, voice, behavior and environment context) consistent across interactions over time?

The Strategic Imperative Ahead

The question is no longer whether digital transformation or GenAI adoption will occur, but enterprises respond to their convergence. Their collision has created a new category of enterprise risk—one already exploited by nation-states and organized criminal groups. The DPRK phony IT worker and Scattered Spider threats attacks are not isolated anomalies. They are early indicators of a future where people themselves become the attack vector. GenAI only accelerates this trend at scale.

Enterprises that evolve their trust models—and their threat intelligence—to account for human adversaries will be better positioned to defend themselves. Those that do not may find their next breach, fraud event, or strategic failure does not originate from malicious code—but from streaming digital content that convincingly displays a human interaction that was never legitimate to begin with. In the GenAI era, the most dangerous supply chain vulnerability may not be software at all—it may be a person pretending to be someone they are not.

A zero-trust architecture at the human layer means treating human identity and behavior with the same skepticism we already apply to networks and devices. In short: No human is trusted by default. Every person must continuously prove they are real, authorized, and behaving legitimately.

This goes beyond IAM and MFA. It requires assuming accounts ≠ humans and credentials ≠ intent.

About GetReal Security

GetReal Security is the cybersecurity leader specializing in the detection and mitigation of threats posed by malicious generative AI content including deepfakes and impersonation attacks. Its technology serves multinational corporations, financial institutions, media organizations, government agencies, and social media companies.

The company was incubated by Ballistic Ventures, the venture capital firm dedicated exclusively to funding and incubating entrepreneurs and innovations in cybersecurity, and Dr. Hany Farid, the preeminent expert in media forensics.

getrealsecurity.com