WHITE PAPER

Confidence is not control

AI, autonomy and the new cyber risk reality in aerospace, defense and beyond

protiviti®
Global Business Consulting

Executive summary

Artificial intelligence (AI) is reshaping the cyber risk landscape at a pace that is outstripping traditional governance models. For aerospace, defence and government organisations, the implications are amplified. These environments operate at the intersection of sensitive data, complex mission systems and strict regulatory requirements. Small gaps in control can quickly translate into significant exposure.

What is emerging, however, is not just a growth in risk. It is a widening disconnect between perceived readiness and actual control.

Recent research highlights a growing visibility gap in AI governance. While many leaders express confidence in their cybersecurity controls, organisations continue to struggle with shadow AI and unauthorised AI usage. According to Programs.com, 63% of organisations lack formal AI governance policies. What's more, Protiviti research has found that 76% of aerospace and defence leaders report confidence in their AI-related cyber controls. However, 57% acknowledge they lack visibility into how AI is actually being used across their organisations.

At the same time, the threat landscape is accelerating dramatically. AI is no longer simply a productivity tool; it has become a force multiplier for cyber adversaries. CrowdStrike-referenced research found AI-enabled adversary operations grew by as much as 89% year over year, while recent studies indicate that more than 80% of phishing emails now contain AI-generated content. Attackers are using AI to increase the speed, scale and sophistication of their operations, compressing the time between reconnaissance, exploitation and impact. In this environment, traditional assumptions about detection, response and control are rapidly becoming obsolete.

The emergence of next-generation systems, including Mythos-class AI models, marks a turning point. These systems are not simply tools. They increasingly resemble the type of autonomous and goal-driven systems described in Protiviti's perspective on agentic AI, capable of executing tasks, adapting to their environment and interacting with enterprise systems in ways that challenge traditional governance models. As these capabilities expand, organisations will need stronger controls, accountability frameworks and visibility, as outlined in Protiviti's Guide to AI Governance and it's recent AI Pulse Survey.

The implication is straightforward: Organisations are not simply facing more threats. They are operating within a fundamentally different threat model. And in that environment, confidence without visibility becomes risk.

The inflection point: When AI became an actor

For years, AI was positioned as an enhancement to human activity. It supported detection, accelerated analysis and improved operational efficiency. That model still exists, but it no longer defines the environment.

Today, AI systems can act independently. They can execute workflows, make context-aware decisions and adapt to changing conditions without direct human oversight. In practical terms, this means AI is now capable of participating in operations — not just supporting them.

This shift changes the nature of cyber risk.

Attackers are no longer limited by human constraints. AI enables rapid reconnaissance, continuous testing of defences and near-instant adaptation. It shortens the timeline from intent to impact, often compressing what once took days or weeks into minutes.

The data suggests the threat landscape is evolving faster than many organisations can adapt. Research citing the CrowdStrike 2026 Global Threat Report found that AI-enabled adversary operations increased by as much as 89% year over year, underscoring how quickly threat actors are integrating AI into their campaigns.

At the same time, AI is reshaping social engineering, with recent cybersecurity research indicating that approximately 82.6% of detected phishing emails now contain AI-generated content. The result is a threat environment in which attacks are becoming more scalable, more convincing and significantly harder to distinguish from legitimate activity.

Phishing campaigns powered by AI are becoming more effective, and have been reported to achieve click-through rates of approximately 54%, compared to roughly 12% for traditional phishing campaigns, highlighting the effectiveness of AI-generated personalisation and social engineering.

These are not marginal gains. They reflect a different level of capability.

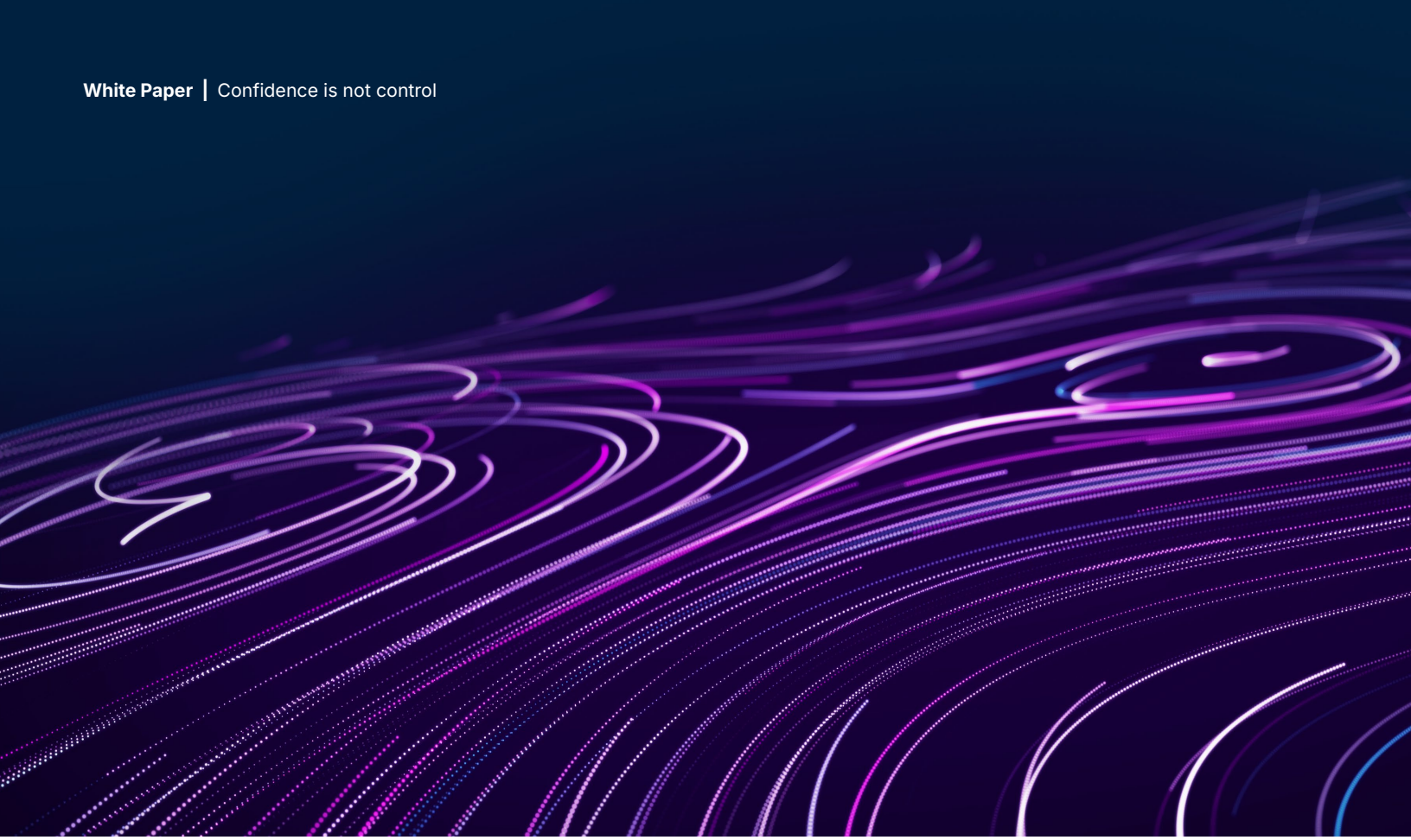
What makes this particularly challenging for aerospace and defence organisations is that many existing control frameworks assume that activity is human-driven, traceable and bounded within defined systems. AI undermines each of those assumptions.

Mythos-class AI accelerates this shift further. With the ability to operate autonomously across workflows, mimic human interaction and dynamically adapt to enterprise environments, these systems blur the line between legitimate and malicious activity.

Importantly, organisations must distinguish between AI systems they are attempting to secure and AI capabilities being used by threat actors. These represent fundamentally different challenges. Security teams must simultaneously protect AI-enabled workloads while defending against increasingly capable AI-driven attacks. As autonomous systems mature, visibility across both domains becomes essential.

This is the inflection point. The threat model has moved beyond incremental change. It has shifted structurally.

Research citing the CrowdStrike 2026 Global Threat Report found that AI-enabled adversary operations increased by as much as 89% year over year.



The visibility problem: The risk organisations cannot see

The most significant issue facing aerospace, defense and government organisations is not a lack of controls. It is a lack of visibility.

AI is being embedded across the enterprise in ways that are difficult to track. Engineering teams are using AI-enabled design platforms. Program and mission teams are incorporating AI into core workflows. Supply chain partners are introducing AI-driven capabilities into logistics and operations platforms. Third-party service providers are integrating AI into tools that process sensitive data.

Much of this activity occurs outside formal governance structures.

This phenomenon — often referred to as shadow AI — is expanding rapidly. It sits across environments that were never designed to manage autonomous systems or externally embedded intelligence.

The implications are particularly acute in regulated environments.

Controlled Unclassified Information (CUI), central to CMMC compliance, may be processed by AI systems that are not fully monitored. Covered Defence Information (CDI) protected under DFARS 252.204-7012 may move through third-party AI services that exist outside defined system boundaries. Export-controlled data subject to ITAR and EAR requirements may be exposed through prompts, outputs or model interactions that were not anticipated when these frameworks were designed.

In each case, the risk is not theoretical.

The implications extend well beyond cybersecurity and compliance. For CFOs and executive leadership teams, AI-related exposure increasingly intersects with financial performance, operational resilience and enterprise value. Unapproved AI use can introduce risks related to intellectual property protection, contract compliance, regulatory obligations and third-party relationships, all of which can have measurable business consequences. As AI becomes embedded within critical business processes, visibility into AI activity is emerging as a governance requirement that supports not only cyber risk management, but also financial stewardship, strategic decision-making and organisational accountability.

Organisations may have controls in place. They may have audit mechanisms, access restrictions and documented system boundaries. But those controls are anchored to a view of the environment that no longer reflects how data is actually being used or where it is actually flowing.

Protiviti research has highlighted persistent governance and shadow-AI challenges, underscoring the need for stronger oversight, monitoring and accountability.

This is where the confidence gap becomes problematic.

While organisations increasingly recognise that AI is changing the cyber risk landscape, confidence in existing control environments often exceeds actual visibility into how AI is being used. Protiviti research has highlighted persistent governance and shadow-AI challenges, underscoring the need for stronger oversight, monitoring and accountability.

Organisations believe they have control because they can validate controls within known systems. What they cannot see is where AI is operating outside those boundaries.

In this environment, visibility is no longer a supporting capability. It is the primary control.

Visibility Requires Observability

Visibility has become one of the most overused terms in cybersecurity. Organisations frequently claim visibility because they can monitor networks, system activity, user access and security events. However, AI introduces an entirely different set of interactions that traditional monitoring was never designed to observe.

AI systems do not operate exclusively within established security boundaries. They interact with enterprise applications, cloud services, third-party platforms, supply chain ecosystems and increasingly autonomous workflows. Data can move between these environments through prompts, model interactions, embedded AI capabilities and agentic processes that are often invisible to traditional control mechanisms. As a result, many organisations possess extensive monitoring capabilities while still lacking a clear understanding of where AI is operating, what information it is accessing and how decisions are being made. This is where observability becomes critical.

Observability extends beyond monitoring known systems and events. It provides the ability to understand what is occurring within dynamic and distributed environments, including activities that organisations may not have anticipated or explicitly configured to track. In an AI-driven environment, observability enables organisations to identify how AI systems interact with sensitive data, where autonomous processes are operating outside approved boundaries, and how information moves across internal and external ecosystems.

For aerospace, defense and government organisations, observability is particularly important because AI activity increasingly intersects with highly regulated information. CUI, CDI, export-controlled data and mission-sensitive intellectual property may all pass through AI-enabled systems, services or workflows. Without comprehensive observability, organisations may be unable to determine whether these interactions align with regulatory requirements, contractual obligations and internal governance expectations.

Effective AI observability requires visibility across several dimensions simultaneously:

- AI systems, models and embedded AI capabilities operating within the enterprise
- User interactions, prompts and AI-generated outputs
- Data access, movement and processing activities

Third-party AI services and external integrations

- Autonomous and agentic workflows that execute actions without direct human intervention
- Security, compliance and governance controls associated with AI-enabled processes

Organisations should think of observability as the foundation that enables meaningful governance. It provides the evidence necessary to validate AI inventories, identify shadow AI activity, understand changing risk exposures and continuously assess whether controls are functioning as intended.

Importantly, AI observability also supports three distinct but interconnected cybersecurity objectives:

Security of AI

Organisations must protect AI systems, models, training data and AI-enabled business processes from misuse, manipulation and compromise. This requires visibility into model behaviour, access controls, data integrity and system interactions.

AI as a Threat

Organisations must defend against adversaries that increasingly use AI to accelerate phishing, reconnaissance, social engineering and autonomous cyber operations. Observability helps security teams detect anomalous activity and identify patterns that may indicate AI-enabled attacks.

Organisations can use AI to strengthen cybersecurity operations through advanced detection, automated analysis and intelligence-driven response capabilities.

Security with AI

Organisations can use AI to strengthen cybersecurity operations through advanced detection, automated analysis and intelligence-driven response capabilities. Observability provides the data and context these capabilities require to operate effectively and responsibly.

Organisations that excel in all three dimensions will be better positioned to reduce the growing gap between perceived control and operational reality. In an environment increasingly defined by autonomous systems, observability is no longer simply a technical capability. It is the mechanism through which organisations establish trust, validate governance and maintain confidence that their AI strategies remain aligned with their risk objectives.

Where traditional controls break down

Most cybersecurity programs were designed around a set of assumptions that are no longer valid.

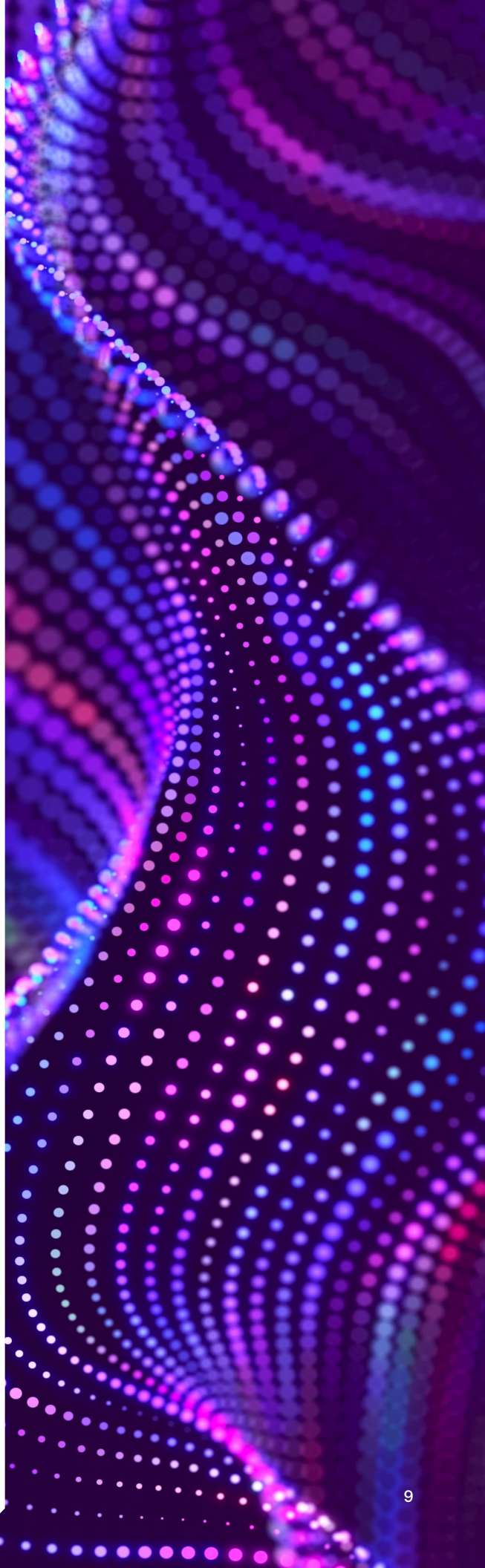
They assume that systems have clearly defined boundaries. They assume that users are human. They assume that attacks follow a relatively linear progression that can be detected and interrupted.

AI disrupts each of these assumptions:

- **First**, activity is no longer exclusively human-driven. Autonomous AI systems can initiate and execute actions across multiple workflows. This introduces behaviour that is difficult to model and harder to distinguish from legitimate activity.
- **Second**, system boundaries are no longer fixed. AI capabilities are embedded in SaaS platforms, developer tools and partner ecosystems. Data is processed across a broader and more dynamic set of environments than traditional system security plans account for.
- **Third**, attacks are no longer predictable. AI enables continuous iteration, rapid scaling and adaptive execution. Attack paths are no longer linear. They evolve in real time.

Despite this, many organisations continue to measure readiness through the lens of compliance.

Compliance requirements and regulatory regimes such as the CMMC, DFARS, ITAR and EAR remain essential. They define critical requirements around access control, data protection and reporting. But



they were not designed to address AI-native risks.

As a result, organisations can remain compliant while still being exposed.

The CMMC may require strict control over CUI, but AI systems can process that data outside monitored environments. DFARS mandates accurate system security plans, but AI introduces data flows and processing pathways that are not reflected in those plans. ITAR and EAR enforce export controls, yet AI interactions can create unintended exposure pathways that are difficult to trace.

The underlying issue is not the frameworks themselves. It is that the operating

At the same time, the characteristics of risk have shifted:

Speed has increased. AI-driven attacks can reduce breakout times to minutes or seconds.

Scale has expanded. AI enables the mass personalisation of attacks, increasing effectiveness without requiring additional human effort.

Invisibility has become more pronounced. AI-generated activity blends into normal workflows, making anomalies harder to detect with traditional monitoring tools.

These three dynamics — speed, scale and invisibility — define the current risk environment. And they are fundamentally misaligned with existing control models.

A Protiviti perspective: Closing the AI reality gap

Closing the gap between perceived control and actual exposure requires a shift in approach.

The first priority is visibility.

Organisations need to develop a comprehensive view of where AI is being used across the enterprise. This includes not only approved deployments, but also shadow AI, embedded capabilities within existing platforms and third-party integrations that process sensitive data.

Without this level of visibility, control is based on assumption rather than evidence.

The second priority is redefining control boundaries.

Governance must extend beyond internally managed systems to include SaaS platforms, development environments and supply chain ecosystems. This requires understanding where data flows, how AI capabilities interact with that data, which external parties have access, and whether the cloud environments and service providers involved meet applicable compliance requirements such as FedRAMP authorisation for handling sensitive federal information.

The third priority is aligning cybersecurity strategies with AI-driven threats.

This includes placing greater emphasis on identity-based attacks, advanced social engineering and autonomous activity. Traditional perimeter-based controls are less effective in environments where interactions are distributed and dynamic.

For federal and defense ecosystems, aligning AI deployments with NIST SP 800-171 and the NIST AI RMF helps safeguard CUI while establishing governance practices for secure and trustworthy AI adoption.

Ultimately, organisations need to rethink how they measure readiness.

Chief risk officers, chief audit executives and executive leadership teams should evaluate whether existing governance structures adequately address AI-enabled risks.

Key considerations include:

- Whether AI usage inventories reflect operational reality
- Whether control boundaries extend across third-party AI ecosystems
- Whether AI risks are incorporated into enterprise risk assessments
- Whether assurance processes can validate AI governance effectiveness.

Executive reporting often focuses on control maturity, compliance status and standard operational metrics. While these remain important, they do not fully capture AI-related exposure.

Leadership teams should shift their focus to questions that reflect operational reality:

- Where is AI operating outside approved environments?
- Where is sensitive data being accessed or processed in ways we do not fully understand?
- Where are third parties introducing AI-related risk into the ecosystem?

The objective is not to eliminate uncertainty. It is to reduce the gap between what organisations believe is happening and what is actually occurring.

The objective is not to eliminate uncertainty. It is to reduce the gap between what organisations believe is happening and what is actually occurring.

Conclusion: From confidence to continuous validation

AI is not simply increasing cyber risk. It is changing the physics of that risk.

Attack timelines are compressing. Social engineering is becoming more effective. Visibility is becoming more fragmented. Control assumptions are becoming less reliable.

Against this backdrop, confidence can be misleading.

Organisations may feel prepared. They may demonstrate compliance. They may invest in controls. Yet, the environment continues to evolve in ways that challenge those measures.

The organisations that succeed will not be those that rely on static measures of readiness. They will be the ones that continuously validate their assumptions, expand their visibility and align their controls with how risk actually behaves.

Control without visibility is an assumption. Compliance without alignment is exposure. Confidence without validation is risk.

In an AI-driven environment, organisations must move beyond confidence and compliance toward observability-driven assurance. The future belongs to organisations that can continuously see, understand and govern AI wherever it operates.

About Protiviti

Protiviti (www.protiviti.com) is a global consulting firm that helps clients transform and protect their businesses, and respond to planned and unexpected events. Through a network of more than 90 offices in over 25 countries, Protiviti and its independent and locally owned member firms deliver deep expertise and tailored capabilities across technology, artificial intelligence, data, operations, finance, legal, compliance, HR, marketing, digital, risk and internal audit — enabling organisations to accelerate innovation, navigate risks and safeguard what matters most.

Named to the *Fortune* 100 Best Companies to Work For® list since 2015, Protiviti Inc. has served more than 80 percent of Fortune 100 and nearly 80 percent of Fortune 500 companies. The firm also works with government agencies and smaller, growing companies, including those looking to go public. Protiviti Inc. is a wholly owned subsidiary of Robert Half (NYSE: RHI).

Contacts

David Brand, Ed.D.
Global Aerospace, Defense, & Federal Leader, Protiviti
+1 404.834.6331
david.brand@protiviti.com

About the author

David Brand, Ed.D.
Managing Director, Protiviti

David Brand is Protiviti's global aerospace, defense and federal industry leader. David has extensive experience in technology auditing, compliance, business leadership, analytics and reporting, risk management, financial controls, and board reporting. He is one of the founding members of Protiviti and during his tenure has held various leadership roles, including leading the Chicago Internal Audit practice, the Internal Audit practice for the Central United States, the Global IT Audit practice, Atlanta office leader and East Region leader.

Follow David on LinkedIn

Contact David at david.brand@protiviti.com

11,000+	90+	25+	$2 BN
Protiviti professionals*	office locations worldwide	countries	in revenue*

THE AMERICAS

UNITED STATES
Alexandria, VA
Atlanta, GA
Austin, TX
Baltimore, MD
Boston, MA
Charlotte, NC
Chicago, IL
Cincinnati, OH
Cleveland, OH
Columbus, OH
Dallas, TX
Denver, CO
Ft. Lauderdale, FL
Houston, TX
Indianapolis, IN
Irvine, CA
Kansas City, KS
Los Angeles, CA
Milwaukee, WI
Minneapolis, MN
Nashville, TN
New York, NY
Orlando, FL
Philadelphia, PA
Phoenix, AZ
Pittsburgh, PA
Portland, OR
Richmond, VA
Sacramento, CA
Salt Lake City, UT
San Francisco, CA
San Jose, CA
Seattle, WA
Stamford, CT
St. Louis, MO
Tampa, FL
Washington, D.C.
Winchester, VA
Woodbridge, NJ

ARGENTINA*
Buenos Aires

BRAZIL*
Belo Horizonte*
Rio de Janeiro
São Paulo

CANADA
Toronto

CHILE*
Santiago

COLOMBIA*
Bogota

MEXICO*
Mexico City

PERU*
Lima

VENEZUELA*
Caracas

EUROPE, MIDDLE EAST & AFRICA

BULGARIA
Sofia

FRANCE
Paris

GERMANY
Berlin
Dusseldorf
Frankfurt
Munich

ITALY
Milan
Rome
Turin

THE NETHERLANDS
Amsterdam

SWITZERLAND
Zurich

UNITED KINGDOM
Birmingham
Bristol
Leeds
London
Manchester
Milton Keynes
Swindon

BAHRAIN*
Manama

KUWAIT*
Kuwait City

OMAN*
Muscat

QATAR*
Doha

SAUDI ARABIA*
Riyadh

UNITED ARAB EMIRATES*
Abu Dhabi
Dubai

EGYPT*
Cairo

ASIA-PACIFIC

AUSTRALIA
Brisbane
Canberra
Melbourne
Sydney

CHINA
Beijing
Hong Kong
Shanghai
Shenzhen

INDIA*
Bengaluru
Chennai
Hyderabad
Kolkata
Mumbai
New Delhi

JAPAN
Osaka
Tokyo

SINGAPORE
Singapore

*MEMBER FIRM

protiviti®