Bell Integration

Ambitious Transformation



Trustworthy AI: Why Enterprise AI demands a new model of trust

As AI gains greater access, authority and autonomy, how do organisations ensure it remains within approved boundaries?

Meet the author

Alison Hodivala is the Director of AI & Data Portfolio Strategy at Bell Integration, where she focuses on ensuring the company's AI and data services reflect market needs and are supported by delivery models that scale reliably. She works across product, delivery and commercial teams to develop offerings that are practical, relevant and capable of delivering business impact.



Her career spans more than 25 years in start-ups and global organisations, covering consultative sales and delivery, enablement and change management. Drawing on a background in organisational psychology, Alison applies an evidence-based approach to AI adoption, aligning people, processes and technology to achieve effective and responsible outcomes.

Business outcomes now depend increasingly on the **quality of customer interactions** across the full lifecycle.

Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

©Bell Integration 2026

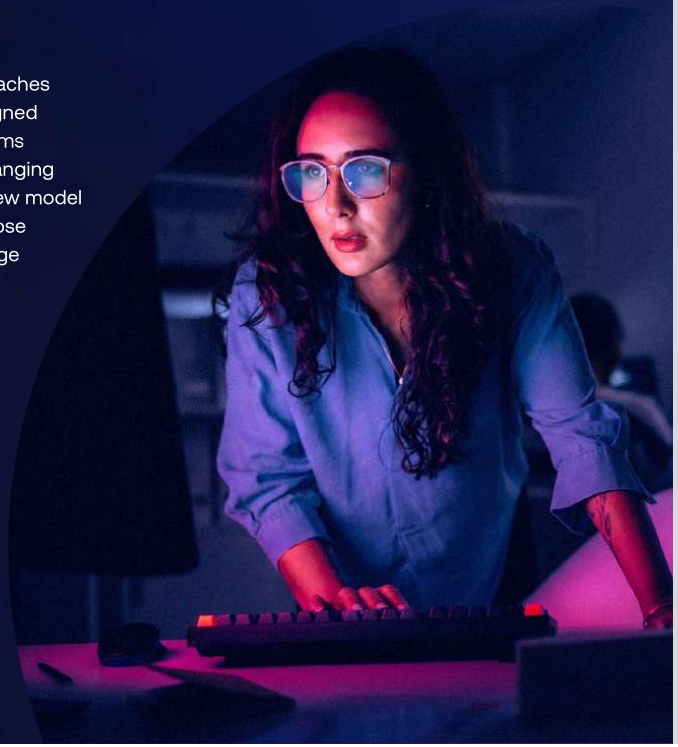
Introduction

Enterprise AI is moving beyond answering questions and generating content. Organisations are connecting it to enterprise knowledge, embedding it in business processes and beginning to trust it with work that has traditionally required human decision-making.

As AI gains greater access, authority and autonomy, the question is no longer simply whether it creates value. It is whether organisations can maintain confidence in the decisions and actions AI takes on their behalf.

Many organisations recognise that challenge, but fewer have addressed it. The IBM Institute for Business Value found that 82%[1] of C-suite executives believe secure and trustworthy AI is essential to their business, yet only 24%[1] of current generative AI initiatives are being secured. That gap may become more significant as AI adoption expands. Gartner predicts that by 2027, "40%[2] of enterprises will demote or decommission autonomous AI agents due to governance gaps identified only after production incidents occur".

The issue is not that organisations are adopting AI too quickly. Existing approaches to governance and security were designed for deterministic software, not AI systems that interpret context and adapt to changing conditions. Enterprise AI demands a new model of trust, one designed for systems whose behaviour and operating context change throughout their operational life.



©Bell Integration 2026

Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

Enterprise AI changes the assumption

For decades, enterprise software operated largely through deterministic rules and processes. It accepted an input, followed a defined set of instructions and produced an expected result. Governance focused on the people making decisions, while security focused on protecting the systems that supported them. The software executed defined instructions; judgement remained with people.



©Bell Integration 2026

Today, AI interprets information, reasons over context and produces outcomes that were never explicitly programmed. As organisations embed AI into business processes and connect it to enterprise knowledge, they are trusting it to perform tasks that once required human judgement.

A claims assistant that once highlighted exceptions now resolves most of them. A coding assistant that once suggested changes now prepares them for approval.

This is more than another step in automation. Organisations are beginning to delegate judgement as well as work.

That changes what it means to trust a system. Governance and security remain fundamental, but neither provides, on its own, ongoing evidence that AI continues to operate within organisational expectations.

Answering that question is what makes AI trustworthy.

Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

What makes AI trustworthy?

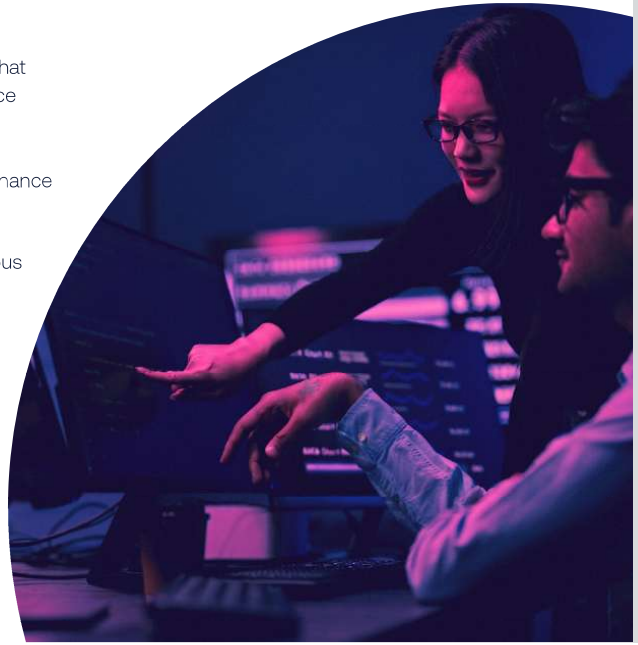
Trustworthy AI is confidence that AI continues to operate within defined expectations and boundaries. That confidence does not come from a point-in-time assessment, but from an organisation's ability to demonstrate that those expectations and boundaries continue to hold as models, data, integrations and authority evolve. Governance and security establish that confidence, supported by continuous assurance.

- Governance defines what AI is permitted to do, what data it can use, what authority it can exercise and where people remain accountable for decisions.
- Security protects the models, data, identities and integrations that AI depends on from manipulation, compromise and misuse.
- Continuous assurance provides ongoing evidence that governance and security remain effective as AI systems change.

None of these works in isolation.

Governance without security leaves policies that can be bypassed. Security without governance protects systems that may still make poor decisions. Without continuous assurance, organisations cannot be confident that governance and security remain effective as AI changes.

Together, governance, security and continuous assurance enable organisations to rely on AI as it evolves.



Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

©Bell Integration 2026

Two kinds of AI failure

The same AI failure can have very different causes.

Consider two organisations using the same procurement assistant. In both cases, the AI recommends a supplier with a poor record of performance. The outcome is identical. The cause is not.

In the first organisation, the recommendation reflects weaknesses in governance. Recent supplier performance was missing from the training data, and there was no policy requiring a human review for high-value contracts. The AI behaved exactly as it had been designed to. The decision was still the wrong one.

In the second organisation, the recommendation was influenced by manipulated information retrieved from an external source, or by a prompt injection that altered the AI's behaviour.

This time, the problem wasn't the policy or the model. The AI had been manipulated through the information or instructions it received.

Both organisations arrive at the same outcome: a poor supplier is selected. But they face very different risks.

One is a governance failure. The other is a security failure.

The response depends on the cause. Better governance will not prevent prompt injection. Stronger security will not correct incomplete policies or poor-quality training data. Unless organisations distinguish governance failures from security failures, they risk strengthening controls that cannot address the underlying problem.

Recognising the difference is the first step towards addressing the right problem. Governance and security solve different problems. Organisations need both, but each has a distinct role in ensuring AI operates within defined expectations and boundaries.

Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

©Bell Integration 2026

AI Governance + AI Security = Trustworthy AI

Distinct controls. Shared outcome.



Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

©Bell Integration 2026

When trusted AI becomes trusted risk

For decades, organisations have managed insider risk by controlling access to sensitive information, business processes and critical systems. Trusted access is necessary for people to perform their roles, but it also creates risk when that trust is misused, compromised or exercised outside established boundaries.

AI introduces a new dimension to that risk. As AI becomes embedded in business processes, it is being given access to sensitive information, connected to enterprise systems and increasingly granted authority to influence decisions and initiate actions.

This places AI within the insider risk landscape. The defining issue is not whether AI is a person, but whether it holds trusted access and delegated authority that can create material consequences for the organisation if that trust is compromised, manipulated or poorly governed.

Those consequences can affect the same areas organisations already consider when assessing enterprise risk: revenue, reputation, regulation and the ability to run operations. AI may expose confidential information, act outside approved boundaries, make or influence inappropriate decisions, or be manipulated into taking unauthorised actions.

Managing this risk therefore requires organisations to look beyond who has access. They must also understand what access and authority have been granted to AI, what decisions and actions it can influence, and how that authority is governed, secured and continuously assured.

As organisations delegate more responsibility to AI, this becomes increasingly important. Trustworthy AI is not about preventing AI from taking on greater responsibility. It is about giving organisations the confidence to extend that responsibility within understood and controlled boundaries.

Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

©Bell Integration 2026

Governance defines the boundaries. Security protects them.

Governance and security are complementary, but they solve different problems. Confusing their roles risks applying the right control to the wrong problem.

Governance defines the boundaries within which AI can operate. It establishes what AI is permitted to do, what data it can use, what authority it can exercise, where human oversight is required and who remains accountable for the decisions it supports. Governance asks: **what should AI be allowed to do?**

Security protects those boundaries. It protects the models, data, identities and integrations AI depends on from compromise, manipulation and misuse. Security asks: **how do we prevent AI from being manipulated or compromised in ways that take it outside those approved boundaries?**



Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

©Bell Integration 2026

Neither discipline replaces the other.

Even a well-governed AI system can still be undermined by a compromised data source, an insecure integration or a successful prompt injection. Equally, strong security controls cannot compensate for poor governance, incomplete policies or inadequate training data. Trustworthy AI depends on both disciplines because they address different sources of failure.

Most organisations do not need new governance and security functions. They need those functions to operate against a shared understanding of AI risk. Without that alignment, governance may approve behaviour that security cannot protect, while security may protect systems whose decisions remain poorly governed.

But governance and security alone are not enough.

Unlike traditional software, AI does not operate in a static environment. Models evolve, data changes, integrations expand and patterns of use shift over time. Organisations therefore need confidence not only that governance and security are in place, but that they remain effective as AI systems, data and business processes change.

Even a well-governed AI system can still be undermined by a compromised data source, an insecure integration or a successful prompt injection.



Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

©Bell Integration 2026

Maintaining trust in Enterprise AI

Governance and security establish the foundations of trustworthy AI. The challenge is ensuring they remain effective as AI evolves.

It is tempting to assume that once an AI solution has been tested and approved, it can be trusted indefinitely. AI does not work that way. An AI solution may perform exactly as expected during testing yet behave differently six months later. Models are updated, new information becomes available, business processes evolve and integrations expand. AI systems interact with different users, operate in changing environments and may work alongside other AI systems. The conditions in which AI operates rarely stay the same, even when the model itself does not.

Traditional software assurance focused on validating functionality before release. Once deployed, security controls helped ensure applications remained available, protected and resilient. Enterprise AI requires a broader approach. Organisations need confidence not only that AI is functioning correctly, but that it continues to operate within approved policies and organisational expectations as circumstances change.

Point-in-time approval cannot provide enduring confidence when the data, integrations, users, access and authority surrounding an AI system can continue to change after deployment. A model may remain unchanged while its data, integrations, users, authority and business context evolve around it.

Continuous assurance is not another governance activity. It is how trust is maintained after deployment.

That means asking questions such as:

- Are AI's decisions still appropriate?
- Has the information it relies on changed?
- Are users interacting with it in unexpected ways?
- Has its access or authority expanded since it was approved?
- Is it still operating within its approved boundaries?

These are not questions that can be answered before deployment alone. They require continuous assurance, giving organisations ongoing confidence that governance and security continue to achieve their intended outcomes as AI evolves.

Enterprise AI changes how organisations must maintain confidence in the systems they rely on. Trust is no longer established once and assumed to persist. It is maintained through continuous assurance, giving organisations the confidence to scale AI, automate more complex work and rely on AI in more valuable parts of the business.

©Bell Integration 2026

Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

The confidence to scale AI

Enterprise AI is changing more than the technology organisations use. It is changing how work is delegated, how decisions are supported and what organisations need in order to trust the systems they rely on.

The question is no longer whether organisations should adopt AI. For many, that decision has already been made. The challenge now is ensuring AI continues to operate within approved boundaries as its role within the business expands.

Governance and security provide the foundation for that confidence. Continuous assurance sustains it as AI evolves. Together, they enable organisations to maintain confidence that AI's access, authority and actions remain appropriate as its operating context changes.

Organisations that can demonstrate that their AI remains within approved boundaries will be able to automate more valuable work, delegate greater authority and scale AI with confidence. Those that cannot will be forced to limit autonomy, retain manual controls or withdraw systems after incidents occur.

Trust therefore becomes an enabler of scale rather than a barrier to adoption.

Ultimately, the future of enterprise AI will not be determined by what AI is capable of doing. It will be determined by what organisations have the confidence to trust it to do.



Meet the author
Introduction
Enterprise AI changes the assumption
What makes AI trustworthy?
Two kinds of AI failure
When trusted AI becomes trusted risk
Governance defines the boundaries. Security protects them.
Maintaining trust in Enterprise AI
The confidence to scale AI
References

©Bell Integration 2026

References

1. IBM 2024, “Unlock trustworthy AI with integrated governance and security.” Unlock trustworthy AI with integrated governance and security | IBM

2. Gartner, 2 March 2026. “Avoid Governance Mismatch: Classify AI Agents by Autonomy Level.” ID G00846081. By Shiva Varma. Avoid Governance Mismatch: Classify AI Agents by Autonomy Level

Further reading:
AI Governance – IBM Documentation



Meet the author

Introduction

Enterprise AI changes the assumption

What makes AI trustworthy?

Two kinds of AI failure

When trusted AI becomes trusted risk

Governance defines the boundaries. Security protects them.

Maintaining trust in Enterprise AI

The confidence to scale AI

References

©Bell Integration 2026

About Bell Integration

Bell Integration is a global technology leader that partners with customers to create lasting, ambitious digital transformation.

We are our customer's accelerator – orchestrating the right partners, systems, and ideas to create measurable impact.

Bell Integration brings over 30 years of trust, expertise, and innovation to the table, evolving from an independent UK-based business to a global powerhouse.

With talented teams operating across continents, we support the full IT life cycle from strategic consulting to project design and delivery.

See how Bell Integration can help your business succeed. Please contact us on marketing@bell-integration.com or visit www.bell-integration.com.

www.bell-integration.com

©Bell Integration 2026