



The Leadership Guide to

Automated Data Governance.

Build a tech-powered data
governance programme that:



- Extends the reach and value of MS Copilot, Purview, Databricks
- Automates governance, 24/7/365
- Saves millions through efficient data storage
- Mitigates data risk with always-on compliance
- Aligns AI models, data, and use cases
- Powers safe, impactful AI at scale.

[+ more](#)

Table of Contents

The Challenge: Elevating Data Governance to meet the AI agenda.	3
Opportunity One	
From enterprise chaos to clarity.	4
Trapped knowledge: a hidden tax on the enterprise.	5
Why classification matters.	6
Case study: Enabling AI search with 330% ROI for manufacturer.	7
Leadership perspective: Govern data like your AI depends on it.	8
Opportunity Two	
Closing the AI trust gap.	9
Connecting structured and unstructured data to make progress with AI.	10
Extend the value of tools like Databricks and Microsoft Copilot.	10
Protecting sensitive data when using AI assistants.	11
Case study: Building trusted data foundations for AI at KPMG UK.	12
Opportunity Three	
Govern your data, reduce your risk.	13
Better together: Extending the reach of Microsoft Purview.	14
Case Study: Delivering business value through classification and enrichment.	15
The future of data governance technology.	16
Case study: Saving £1m on data storage, ready for AI success.	17
Meet the platform that powers value-driven data governance.	18
Get in touch with us.	19

The Challenge

Elevating Data Governance to meet the AI agenda.

The real value of data governance has always been in improving the quality, completeness, and relevance of the data that feeds your most critical decisions. Now, AI promises to accelerate workplace decisions and actions like never before – but AI needs the right data to deliver on these big promises.

When we speak with business leaders about their data challenges, they fall into one or more of the following categories:



They're struggling to find and use information that exists in the enterprise.



They're worried about not being compliant.



They're desperate to unlock valuable insights that are buried in their data.



They're under pressure to deploy AI to improve decision-making and increase productivity.

For leaders responsible for data governance at scale, now is the time to prove the value of getting governance right – and lead the way with AI adoption.

AI has added fuel to the data governance fire; in fact, it's pretty much burned the house down. The rules of data governance – how to start, where to start, and what you can achieve with good data – have totally changed.

The challenge can be broken down into the three areas addressed in this guide:

- 1 Enterprise chaos and trapped knowledge
- 2 The AI trust gap
- 3 Governing data risk

Leaders have the chance to create a whole new set of founding principles for modern data governance, where AI is at the centre. And, importantly, the business case for change is strong.

Opportunity One

Opportunity Two

Opportunity Three

From enterprise chaos to clarity.

Enterprise data is growing exponentially and is largely ungoverned. Scattered across thousands of systems, more than half is redundant, obsolete, trivial, or duplicate data – making it impossible to find ‘golden records’, the single sources of truth that you can trust.

At the same time, most human knowledge lives in unstructured data, like documents, emails, web, and messages. Many companies don’t delete anything – and never have. And the worlds of structured content and unstructured data remain disconnected – leaving valuable lessons learned and corporate intel buried and effectively unusable.

Finding a needle in a haystack is hard, but it’s even harder when you don’t know which haystack to look in – or if what you’re looking for actually exists.



Enterprise Chaos.

Data is growing, ungoverned, and scattered across 1000s of systems.

1,100

Corporate systems in the average Fortune 500 company.

30%

of records are estimated to be duplicates, in organisations not actively working on data quality.

73%

of information is only ever used once – proving that users systematically recreate what they cannot find.



Trapped Knowledge.

Lessons learned and corporate intel are buried and unusable.

44%

of people have made a wrong decision because they were unaware of information that could’ve helped.

1%

On average, most employees can only see 1% of information that exists.

LEADERS’ QUESTIONS, ANSWERED.



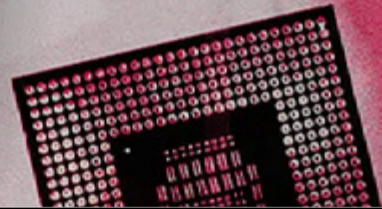
“Our data is spaghetti. We know we should clean it up – but it could be a big job. How do I prove the value of classifying and governing data at scale?”

Automation is the only way to tackle Data Governance at scale. The best approaches use AI to discover, label, and automatically govern data 24/7/365 – and this all happens in situ, without disruption or lengthy data cleansing programmes.

Once you understand what data you have, where it lives, and who owns it, you can put that data to work and start releasing measurable value – whether that’s to explain an irregular business trend, support a new digital experience, power a data science project, fix a compliance issue, or drive an AI use case.

Trapped knowledge: a hidden tax on the enterprise.

Large global organisations lose 100s of millions of dollars a year as a result of trapped knowledge – through time wasted, duplicated work, delays, errors, poor decisions, increased risk, and missed opportunities for innovation.



Access control.

Corporate systems are designed to support specific roles, but strict access controls often lock out everyone else, making knowledge hard to share. Overly complex permissions keep information hidden from people who need it.

WITH AUTOMATED DATA GOVERNANCE

An 'open by default' approach makes most information accessible to everyone, and automatically restricts only truly sensitive data – opening up knowledge while keeping the business protected.



Invisible insights.

When employees move roles or leave, valuable data and insights are left behind and become invisible. Valuable lessons from past jobs, incidents, and projects are lost.

WITH AUTOMATED DATA GOVERNANCE

By automatically discovering and surfacing these hidden insights and golden records, you can make them available to AI and people alike, to inform better answers and decision-making.



Missing data, hidden in plain sight.

Corporate systems often suffer from poor data quality and missing information, much of which is hidden in unstructured content. Insight teams are limited to what's in their specific systems.

WITH AUTOMATED DATA GOVERNANCE

When you find and fix missing or invalid data in unstructured sources, you unlock new insights and improve data quality simultaneously.



OneDrive.

People share information directly from their OneDrive, because saving it to the right SharePoint library is too difficult. This just creates more hidden duplicates that no one else can see.

WITH AUTOMATED DATA GOVERNANCE

Automated governance ensures information is stored in the right locations with the correct permissions, making it easy for everyone to find and reuse what they need.

Why classification matters.

Where manual classification once relied on users to select from dropdowns and checkboxes – leading to inconsistencies and errors – AI now combines natural language processing and machine learning to classify and enrich data automatically. This means permissions, labels, and even metadata can be managed at scale, with AI learning and improving as it processes more information. The result is a governance process that's not only faster and more accurate, but also less dependent on user participation – AI does the heavy lifting, while humans oversee and validate.



Automated classification cuts through data sprawl, surfacing what's valuable and compliant.



AI Agents handle routine governance tasks at scale, freeing up your team for higher-value work.



Dynamic, evergreen governance means your data estate stays healthy and AI-ready, without endless manual intervention.

How does Aiimi's Workplace AI platform identify golden records at scale?



Unified Data Layer

Connects all enterprise data (structured and unstructured) in a single virtual layer, eliminating silos and enabling a complete view.



AI-Driven Classification

Identifies exact and near-duplicate records, supporting data minimisation and helping determine the most accurate, up-to-date version.



Duplicate Detection

Identifies exact and near-duplicate records, supporting data minimisation and helping determine the most accurate, up-to-date version.



AI Agents

Intelligent agents continuously scan, compare, and apply rules to pinpoint and label golden records automatically and on an ongoing basis.



Governance and Trust

Golden records are labelled, governed, and protected with permissions and policies, making them discoverable, usable, and trusted.



Ongoing Monitoring

Automated governance runs 24/7/365, with AI learning and feedback loops to enable continuous improvement and ongoing reliability.

Data Governance that looks after itself, and your business.

When Data Governance is a success, it's happening constantly but users never have to think about it. Both structured data and unstructured content are tagged and contextualised; sensitive data is automatically protected; and accurate permissions mean users are able to find what they need. Previously hard-to-reach data is served to the SMEs who need it, increasing productivity, accuracy, and efficiency.

CASE STUDY

Enabling AI search with 330% ROI for a major manufacturer.

This global manufacturer wanted the ability to navigate hard-to-reach historical data, in order to optimise processes, unify its information, and, ultimately, save time and money.

We implemented Aiimi Workplace AI to index 400 million documents and drawings across multiple sites, creating a comprehensive knowledge network. This has enabled users to quickly find relevant data, identify subject matter experts, and evidence design decisions, significantly improving efficiency. Searches that previously took up to 90 minutes now take as little as five. Users report 80% efficiency gains through better access to data and are freed up for value-add activities. With a 75% first-search success rate and over 330% annual ROI, the solution has become the go-to platform for Product Engineering, driving operational, quality, and manufacturing improvements, reducing risk, and unlocking value from historic data.



**Read more about
this case study:**



Leadership Perspective

Govern data like your AI depends on it.

“We’ve reached a tipping point where the use of data determines who leads and who falls behind. The rise of AI – and the ambitions that come with it – has only amplified this.

Information is exploding in volume and complexity. And it’s not just structured data – it’s video, audio, chat, images, and more, across countless platforms, formats, and sources. All this, against a backdrop of increasingly demanding customers and unprecedented regulatory scrutiny.

So, organisations face more information, complexity, obligations, and expectations than ever before. Is AI the answer?

In part, yes. AI is a powerful tool – but just pointing AI at the problem doesn’t solve it. In fact, it exposes the underlying problems, like weak data foundations, broken governance, and missing context. AI doesn’t paper over the cracks; it magnifies them. It’s why disillusionment in AI is rising, and why organisations are spending more than ever, on platforms, people, compute

power – just to stay standing, not to get ahead.

Right now, most organisations are still chasing their data, not controlling it. We need to change that – to get control over data and make it work for you.

This means visibility and governance across your entire data estate: knowing where your data is, what it means, and what it’s worth. It means building an environment where always-on compliance becomes a competitive advantage, not a burden. Then you can really start pulling levers to benefit from your data: joining the dots, uncovering insight, building trust, and turning information into action. From here, you can operationalise AI with confidence; applying the right data, with the right model, to solve the right problem – making your data and its governance your ultimate point of difference.”



Steve Salvin
Aiimi, CEO

Opportunity One

Opportunity Two

Opportunity Three

Closing the AI Trust Gap.

Businesses are under increased pressure from the board to implement AI – and as they embrace AI assistants, a critical challenge has emerged: AI can only be as effective – and as trustworthy – as the data it can access. This is the heart of the AI Trust Gap.

Most AI assistants, like Microsoft Copilot and Claude, are often limited to the information stored in specific systems, such as M365. Even with native connectors, the information they can access isn't classified and well-governed, leaving powerful AI with an incomplete view. As a result, their information retrieval capability (or search) can bring back data that is not the most relevant to the prompt – such as old versions or data considered to be redundant, obsolete or trivial (ROT), which in turn leads to errors and a lack of confidence in AI outputs.

To be trusted, AI needs the full data picture – and that data picture needs to be safe and well-governed in order to protect your business and sensitive data from exposure. Identifying 'golden records' is crucial, as it ensures AI has access to trusted, high-quality data – the 'single source of truth' that its answers should be based on.



AI Trust Gap.

When AI copilots can't access well-governed data, their outputs can't be trusted.

74%

of companies say AI is now a top-three strategic priority.

60%

of AI projects that are started are abandoned due to poor or incomplete data.

LEADERS' QUESTIONS, ANSWERED.



“Help! I don't really trust our AI outputs – they're missing the full context.”

AI assistants are only as good as the information they can access, so their answers often miss the full context and may not be reliable. Often, AI assistants either don't provide citation or, worse still, they hallucinate citations, so their outputs are not explainable or trustworthy. But they have huge potential value for users: they're familiar, they work in the apps they use every day, and they're part of your enterprise ecosystem. Fortunately, they can deliver significantly more value when underpinned by good data governance and smart tech integrations.

Using an automated Data Governance platform that can classify all data and integrate securely with AI copilots extends their reach across the enterprise – giving them the additional context needed to deliver highly relevant responses, while ensuring sensitive data remains protected. With built-in explainability and citation that extends to all sources, every AI-generated answer is transparent and traceable – so you can trust and verify the information behind every response.

Connecting structured and unstructured data to make progress with AI.

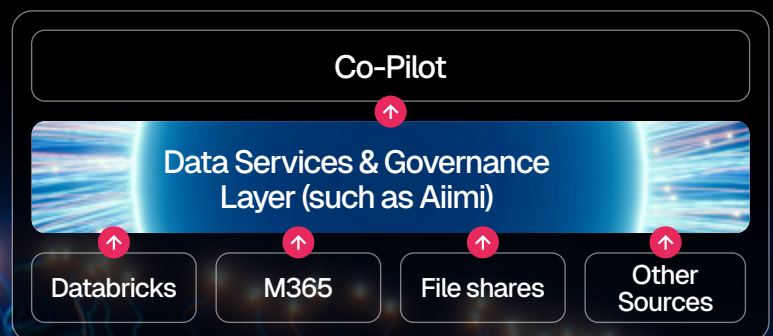
Organisations often focus heavily on structured data – neatly organised information in databases and spreadsheets – because it's easy to analyse and report on. But this approach only captures a small portion of the intelligence within a business. The majority of enterprise data exists in unstructured formats such as emails, documents, messages, images, and PDFs.

The real value lies in bringing structured and unstructured data together. By applying AI technologies such as natural language processing and machine learning, businesses can extract meaning, context, and relationships from unstructured content and combine it with structured datasets. This integrated approach results in richer insights, more accurate decision-making, and greater opportunities for AI and automation to add value.

Extend the value of tools like Databricks and Microsoft Copilot.

Aiimi's platform acts as a data services and governance layer, securely connecting structured data in platforms like Databricks with unstructured content, to give you more complete, accurate answers to power decisions.

Aiimi gives AI copilots a unified view of well-governed data across all systems, with consistent labelling, protection for sensitive data, and kitemarked high-quality data – safely extending their reach, to maximise the value of your investment.



Industry Impact

Transforming asset data outcomes in the water sector.

Historical challenges with data governance have resulted in low data quality and incomplete records in parts of the water sector. With increasing regulatory scrutiny, the industry is under pressure to demonstrate improved governance of data.

Previously, one of our water customer's teams spent over 100 hours manually searching through unstructured historical data and resurveying assets to fill gaps in their records. Aiimi's platform will enable them to connect structured and unstructured data for a unified view of asset information. This will significantly reduce the time taken to respond to asset data requests, minimise the number of systems and people involved, and increase the percentage of assets with complete document sets and key metadata.

Ultimately, this approach is designed to improve the success rate in finding the information needed, supporting both operational efficiency and regulatory compliance.

Protecting sensitive data when using AI assistants.

Enterprise-scale AI assistants boost productivity precisely because they can uncover and bring together data that a user may never have found. But it's a double-edged sword; unrestricted access to all enterprise data introduces new risks. And if trust is broken when AI copilots are rolled out, it can be costly to rebuild it both inside and outside your organisation.



Sensitive Personnel Information & PII

It can be human nature to seek out information we know isn't ours to see – and AI makes it easy to ask the question. Personally Identifiable Information (PII) that's not well governed can be inadvertently surfaced, leading to severe privacy breaches and harm to individual wellbeing.



Customer & User Data Risks

Generative AI tools often bring together data from multiple sources to provide an answer, meaning disparate sensitive data with poor permissions may be combined to expose significant insight into vulnerable individuals or protected groups.



Client Confidentiality & Ethical Walls

Professional services organisations, like legal and audit firms, must restrict the exchange of confidential information between teams to avoid conflicts of interest, maintain regulatory compliance, and keep client trust. Unchecked corporate AI tools accessing data across systems present a significant risk to this.



“Data Governance is vital to protecting an organisation’s (and its customers’) information when deploying AI. Information might need to be withheld from AI services because it contains sensitive information such as personal data, or because it is protected by client contractual conditions, for example. Conversely, effective data classification can drive the right information to an AI process to improve the accuracy and relevance of LLM outputs.”

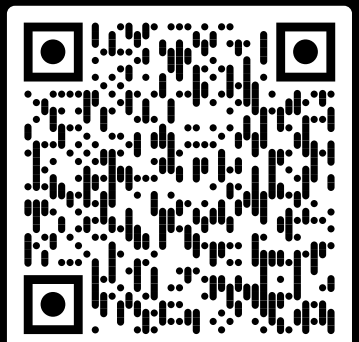
Matt Eustace

Aiimi Data Protection Officer and
Head of Solution Engineering



Businesses need the benefits of daily AI-powered productivity, without taking a gamble on leaking their most sensitive data. To safely scale these tools, organisations must implement robust data governance and automated sensitive data classification.

Read more in our blog: →





CASE STUDY

Building trusted data foundations for AI at KPMG UK.

Data and information governance have never been more critical for safe, secure, and valuable AI outputs. KPMG UK chose the Aiimi platform to intelligently classify unstructured data in situ, so they could better understand what data they hold across their complex estate.

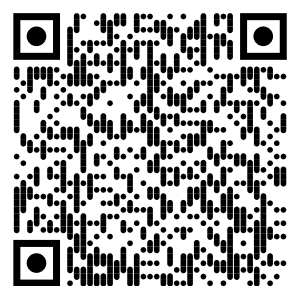
The firm can identify data that has limited value and doesn't need to be retained for regulatory compliance, so they can delete it. They can label their most valuable data and make it available to downstream AI applications in a way that respects retention rules, access controls, and the contractual obligations they have to their clients and audited entities.



“To have trust in AI outputs you must first have trust in your underlying data. The Aiimi platform gives us that trust.”

Chris Allen
Chief Data Officer at KPMG UK

Read more about
this case study:



Opportunity One

Opportunity Two

Opportunity Three

Govern your data, reduce your risk.

We know that enterprise data chaos leaves valuable knowledge buried and unusable. But the flipside of people not being able to find what they need is that they may well be able to see what they shouldn't. The danger here isn't just inefficiency or lost productivity; it's the risk of accidental exposure, misuse, and compliance failures.

Unchecked data risk is one of the greatest threats facing the enterprise. When data isn't properly classified and governed, it becomes impossible to know what information exists, where it lives, its risk profile, and who can access it. This lack of control leaves sensitive personal information, confidential records, and valuable IP exposed and vulnerable.

Even a single vulnerability, slow detection, or missed regulatory requirement can have severe consequences – from direct financial losses and legal penalties to reputational damage and even operational shutdown.

This risk is amplified when businesses start using AI. Data that was previously impossible to find – protected by “security through obscurity” – can suddenly become easy to find by people who shouldn't have access and, even worse, left visible to AI.

The risks of shadow AI are clear to see: when users adopt third-party AI tools that aren't sanctioned by their organisation, they're effectively putting corporate data into the public domain.



Ungoverned Risk.

Personal information, sensitive records, and IP are left exposed.

70%

of enterprise data is over-retained, with no current business, legal, or regulatory value.

87%

of organisations have sensitive data that is accessible to all employees.

€20m

penalty for breach of data protection regulations - or 4% of global annual turnover, whichever is higher.

72%

of customer contracts now include some uncapped liability for data privacy breaches.

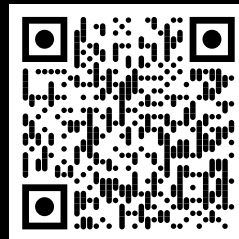
Better together: Extending the reach of Microsoft Purview.

Classification and governance must span all data sources and be ongoing, rather than a static programme that fails to keep pace with real-world risk.

Aiimi's Workplace AI platform works together with Microsoft Purview, extending its capabilities to a broader range of sources beyond the Microsoft estate. By providing deeper classification abilities, Aiimi enables a stronger foundation for data management, risk reduction, and compliance.

For organisations without Purview, the Aiimi platform can provide a comprehensive solution for data classification that supports risk and compliance management and provides high-quality data to the business for AI use cases.

Find out more:



“Privacy and data management will be key market differentiators, and organisations that lead in these areas will set new standards.”

Steve Salvin
Aiimi, CEO

Complete control and compliance through automation.

True enterprise compliance requires always-on governance – not periodic checks or complete overhauls that take years in the making. Automated tools deliver scalable classification, risk-scoring, and remediation. As new data types and regulations emerge, continuous, adaptive data governance ensures they are automatically identified and managed as part of everyday operations.

By focusing on classification and governance, organisations can reduce ungoverned risk, maintain compliance, and protect valuable information assets.

Legal compliance

Confident handling of legal holds, disclosures, and GDPR subject access requests.

Statutory compliance

Timely responses to Freedom of Information and Environmental requests.

Regulatory compliance

Accurate reporting to regulators and open data platforms.

Corporate compliance

Assurance that internal policies and rules are consistently followed.

NOW:
Pensions

CASE STUDY

Delivering business value through classification and enrichment at NOW: Pensions.

With over 1TB of documents to migrate, NOW: Pensions was rethinking its approach to data – prioritising discoverability, governance, and multiple regulatory responsibilities. Aiimi's solution delivered three core areas of value – all while migrating the data to a new shared Cloud drive.



Compliance

With data enriched by intelligent AI-powered technology that applies access policies and satisfies the requirements of the ICO and the Pensions Regulator, including GDPR. Using Aiimi Workplace AI, NOW: Pensions was able to identify over three million spreadsheets that needed classifying under new internal and external data policies, with the appropriate sensitivity labels and storage applied.



Cost

With legacy drives decommissioned, maintenance costs lowered, and resources freed up for internal IT. Surplus documents have been removed and the remainder housed within the company SharePoint, NOW: Pensions has been able to decommission its legacy in-house drives, saving on maintenance costs, freeing up data storage, and providing more operational bandwidth for their IT team.



Efficiency

With redundant documents removed and clear data structures within SharePoint that make it easier for employees to find the files they need. More than 1TB of data has been successfully migrated to SharePoint and 4 million documents that were no longer within the acceptable retention policy have been identified and deleted.



“With Aiimi, I’ve added three key areas of business value: compliance, cost, and efficiencies.”

Adam Stancliffe
Head of Data Governance, NOW: Pensions

Read more about this case study:



The future of data governance technology.

As enterprise data grows in scale and complexity, traditional governance methods are quickly becoming obsolete. The future lies in intelligent automation – where AI-powered classification and always-on AI Agents transform data management from a manual, error-prone process into a strategic advantage. Automated classification is now the essential first step, enabling organisations to quickly identify, label, and control their information, reduce storage costs, and strengthen compliance.

But the real game-changer is the rise of AI Agents. These digital workers operate 24/7/365, not only classifying data but also migrating it, managing permissions, and continuously improving data quality. By creating real-time collections of 'golden records' and adapting dynamically as new data types emerge, AI Agents make governance evergreen – ensuring your information landscape is ready for whatever comes next.



Matt Eustace

Aiiimi, Head of Solution Engineering

What's next?

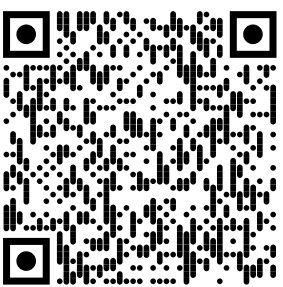
As data landscapes evolve, the next wave of governance technologies will be defined by their adaptability. AI won't just automate classification; it will actively monitor for emerging document types, business processes, or regulatory changes, updating classification schemes and governance frameworks in real time. This means organisations can stay ahead of compliance requirements and business needs, with AI surfacing new patterns and prompting human review where necessary. The result is a living, responsive governance environment that grows and adapts alongside the enterprise, ensuring data remains organised, accessible, and ready for whatever the future brings.

CASE STUDY

Saving £1m on data storage – and setting the stage for AI success.

For one professional services firm, Aiimi's automated data governance has saved £1m in data storage costs within just six months, by removing obsolete data. It's enabled the firm to remediate 200m files, delete 150TB of data, and remove thousands of data leakage risks, while also reducing the CO₂ emissions associated with data storage. The programme has so far reached more than 25,000 employees, and is now helping to lay the foundations for the organisation's internal Data Service, curating gold-standard data for around 30,000 employees.

Read more about
this case study:





Meet the platform that powers value-driven Data Governance.

Our AI platform turns enterprise data into a trusted, governed, and business-ready asset that drives better decisions, reduces exposure to risk, and enables AI. Valuable data is made usable for your teams and your AI tools, and sensitive data is kept safe.



Classify and govern data, 24/7/365

Aiimi connects to all enterprise data across all sources and systems, for automated data governance and always-on compliance.

[Find out more](#) →



Activate unstructured data for AI

Transform messy, inconsistent enterprise data into a trusted asset, ready to power AI. Aiimi uses AI and ML to automate the classification and management of data.

[Find out more](#) →



Extend the reach and value of copilots.

Aiimi provides a governed data layer to AI copilots, maximising the value of your investment – and providing accuracy, explainability, and confidence in AI outputs.

[Find out more](#) →

To talk to us about enterprise
data governance, visit:



[https://aiimi.com/platform/
enterprise-data-governance](https://aiimi.com/platform/enterprise-data-governance)

or get in touch via:



enquiries@aiimi.com

Aiimi
100 Avebury Boulevard
Milton Keynes, United Kingdom
MK9 1FH

aiimi

WE ARE AIIMI.

POWERING AI
EVERYWHERE
AT WORK.

AIIMI.COM