1. Overall Strategy

Objective
Position Airia as the "go-to secure platform for building AI agents" by educating founders, builders, and investors about AI risk – and then showing how Airia enables development without the risk.

Strategy: Education - thought leadership.
Key Message:
“We’re not saying ‘slow down and stop building’. We’re saying: build just as fast – but build differently, with guardrails.”

The conversation should feel like:

- An education piece, not a product demo.
- A gentle wake-up call about risks that most teams are currently ignoring.
- A story-driven dialogue where real incidents lead naturally to Airia’s approach as the sensible path forward.

Format & Logistics

- Recording
  - 75 DEAN STREET, SOHO, LONDON W1D 3SQ
  - Friday 12th December 14:00 - 15:00
- Duration
  - Target 40–45 minutes (the “sweet spot” for keeping energy and attention).
- Tone & structure
  - Not scripted; use a light framework plus short vignettes (1–2 minute stories).
  - Lead with stories and human drama, then draw out the lesson, then briefly connect to how Airia addresses that class of risk.

Audience

- Primary: Entrepreneurs, product & engineering leaders, and AI agent builders.
- Secondary: Investors/VCs and senior stakeholders (CISOs, CTOs) concerned with risk, governance, and defensibility.
- Target regions for promotion: Focus on US and UK for paid distribution and targeting.

---

2. Core Narrative

Main theme:

> “Is what you’re building actually safe – for your customers, your company, and your investors?”

Most people are building AI agents and workflows as fast as possible, often on open or public platforms, with:

- Little thought for security and governance
- Minimal understanding of regulatory risk
- No real strategy to reassure investors that the business is defensible

Airia is positioned as:

- The secure, governed platform for AI development and orchestration.
- A potential “Intel Inside”-style mark of trust for investors and boards: evidence that AI is being built responsibly.

---

3. Five Key Talking Points

1) Security & Data Leakage

Problem framing

- Teams enthusiastically build with tools like Replit, Copilot, ChatGPT, Lovable, etc., without guardrails.
- Sensitive code and data are routinely shared with public models.
- Agents are given broad powers with almost no constraints – the AI equivalent of “houses with no locks.”

Example story angles

- Replit “rogue agent” case
  - An AI agent in a coding environment drops production databases and takes actions a cautious human engineer likely wouldn’t.
  - It reportedly even misrepresented what it had done, which is more alarming than the mistake itself.
  - There were *no hard constraints* around what this agent could or could not do.

- Microsoft Copilot payroll leak
  - Over-permissive permissions allowed payroll data to be exposed to the wrong internal users.
  - Illustrates the danger of giving AI assistants broad visibility into internal systems without proper scoping.

- Samsung source-code leak
  - Engineers pasted confidential source code into a public AI chat.
  - That data was then incorporated into training, making elements of Samsung IP effectively queryable by outsiders.

- Healthcare PHI misuse
  - Doctors and healthcare staff using general-purpose LLMs to summarise patient notes, including PHI.
  - Risks serious HIPAA and other PHI violations if data is not controlled, redacted, or kept within compliant boundaries.

Key message
Security risks are not theoretical; they’re already happening at brand-name companies. Most builders are shipping agents with no meaningful locks on the doors.

---

2) Compliance, Regulation & Stacked Exposure

Problem framing

- AI introduces a new attack and compliance surface that didn’t exist a year or two ago.
- Incidents can trigger multiple overlapping regimes:
  - EU AI Act (dedicated AI regulation and fines)
  - GDPR and broader data-protection laws
  - Sector-specific frameworks (e.g. FCA in financial services, HIPAA for health data)

A single AI-driven misstep can lead to several fines at once: AI, data protection, and sector regulation.

Key message

Most organisations are still treating AI as if it’s “just another app,” not a regulated risk surface in its own right.

---

3) Investor Risk & Business Defensibility

Problem framing

- Investors don’t just ask, “Is this a good market?”; they ask, “How much risk and liability am I buying?”
- If a startup:
  - Builds on open, uncontrolled platforms
  - Casualy leaks IP into public LLMs
  - Has no clear approach to compliance or testing

  then:
  - Its product may be easily copied or commoditised.
  - It may face lawsuits and regulatory penalties.
  - At worst, it can become uninvestable or be forced to shut down after a major incident.

Analogy

- Like founders who never trademark their brand and only realise the risk when someone else files the mark first.
- These situations are more common than people admit, but are often kept quiet because they’re embarrassing.

Investor-oriented message

- VCs and boards should be asking:
  - “What are you doing to secure your AI stack?”
  - “How do you prevent data leakage and model misuse?”
  - “Do you have a way to prove your agents are tested and governed?”

A startup that can answer those questions credibly is more investable and more defensible.

---

4) Bad Actors & Red-Teaming AI Agents

Problem framing

- It’s not just about honest mistakes – there are deliberate attacks:
  - Prompt-injection
  - Data-exfiltration attempts
  - Attempts to subvert agent workflows
- Most teams:
  - Don’t realise their agents have a distinct, novel attack surface.
  - Don’t have the expertise or budget to run proper red-team exercises.

Talking points

- Airia offers the ability to “red-team” agents, whether they’re built on Airia or on other orchestration tools like n8n:
  - A swarm of agents tries to attack your agent.
  - Tests for DLP leakage, prompt injection, and other AI-specific security failures.
  - Can be scheduled and repeated – important because AI systems are non-deterministic and behaviour can change over time.

Key message

AI agents need the same seriousness we give to application security – plus new tests tailored to AI-specific failure modes.

---

5) Airia: Development Without the Risk

This is the resolution to the above risks: how to keep building ambitious AI products without gambling on security, compliance, or investability.

Framing

> “We’re not saying ‘slow down and stop building’. We’re saying: build just as fast – but build differently, with guardrails.”

Three simple promises

1. Build fast, inside guardrails
   - Airia lets teams orchestrate AI agents and workflows at speed, but within a platform deliberately designed for security and governance.

2. Know what your agents can and cannot do
   - Define permissions, constraints and policies:
     - e.g. “This agent can’t drop databases,”
     - “This workflow must not send PHI outside this region,”
     - “This assistant cannot access payroll systems.”
   - Turns free-form, unpredictable agents into governed components with clearly defined limits.

3. Prove you’re taking AI risk seriously
   - Logs, policies, and test results provide evidence for:
     - Regulators and auditors
     - Boards and CISOs
     - Investors and acquirers
   - This becomes part of the company’s risk and trust story.

Linking back to the stories

- Replit-style incidents → Airia can block or tightly control high-risk actions (e.g. destructive DB commands).
- Copilot-style internal leaks → Airia enables fine-grained access control and least-privilege agent design.
- Samsung-style IP exposure → Airia can detect and block sensitive code, keys, and secrets before they hit public models.
- Healthcare PHI → Airia can redact or route sensitive data so it’s handled within compliant boundaries.
- Bad actors → Airia’s red-teaming can continuously test agents for newly emerging attack patterns.

Aspirational positioning

- Over time, “Built on Airia” can function like “Intel Inside” once did:
  - A simple signal to investors, customers, and partners that this AI was not built recklessly.

---

4. Storytelling & Delivery Style

Story-first structure

For each theme:

1. Tell a short human story (Samsung engineers, Replit incident, Copilot leak, healthcare PHI, etc.).
2. Explain the underlying risk (security, compliance, investor, reputational).
3. Show what could have been done differently, and how that maps to Airia's approach and capabilities.

Vignettes

- Prepare multiple 1–2 minute vignettes under each talking point:
  - Security / data leaks
  - Compliance & regulation
  - Investor risk & business defensibility
  - Bad actors & red-teaming
  - Airia as the path to "development without the risk"
- These can be:
  - Used natively in the podcast, and
  - Clipped into short social and YouTube pieces.

---

5. Next Steps
- Dave (Airia)
  - Assemble 8–10 strong case stories:
    - Replit, Copilot, Samsung, PHI/healthcare examples, and
    - Any cases where AI risk seriously damaged or shut down a business or caused major regulatory pain.
  - Shape each into a clear 1–2 minute vignette (what happened, why it mattered, what could have been done differently).

- Jenna (Airia)
  - Confirm US + UK as the primary target geographies for the distribution plan.
  - Align PR and media messaging with the same storyline:
    - AI opportunity
    - Real-world risk
    - Airia as the way to develop without unnecessary risk.

FINALLY...
A Call to Action...What can you offer the listeners in order for them to take action that is greater than what they can get without listening to the podcast?
E.g. 6 months free trial? Time with a product expert?