CATALYST GROWTH COACHING · THE NUMBERS, IN PLAIN ENGLISH · ARTIFICIAL INTELLIGENCE · DIRECTOR'S EDITION · PART ONE OF THREE

The Language

For directors in large organisations — the vocabulary of enterprise AI, without the theatre

A quiet truth first: the AI agenda item is where a lot of capable directors quietly stop following. The slides carry vocabulary invented eighteen months ago, the claims are enormous, the governance language is unfamiliar, and questioning any of it risks looking like the person who doesn't get it.

Two reassurances. First, the room is faking more than it admits — this field is young enough that genuine expertise is rare and confident opinion is abundant. Second, unlike finance or law, there is no settled body of knowledge you are missing. There is a vocabulary, a set of trade-offs, and a small number of questions that reliably separate real initiatives from theatre. That is learnable in an afternoon, which is what this guide is.

The stake is unusually direct. AI is now the fastest route by which a director either becomes visibly valuable to the executive layer, or becomes the person whose function was reorganised around a capability they didn't understand. Every function is being asked what it is doing about this. The quality of your answer is being noticed.

The respectable sentence, usable in any room: "translate that into what changes for the business, and what could go wrong."

John Obidipe · catalystgrowthcoach.co.uk

01

The engine — what these systems are and aren't

Foundation models, LLMs, and model providers

In plain English: A foundation model is a large general-purpose system trained on vast data, adaptable to many tasks. LLM (large language model) is the text-based kind behind most business AI. The major providers are a small group (OpenAI, Anthropic, Google, Meta and a handful of others), and almost every AI product your organisation buys is built on top of one of them.

Why it matters at your level: The strategic implication most decks omit: your vendor's capability, cost, and roadmap are downstream of a model provider they don't control. "Whose model is underneath this, and what happens to us if that changes?" is a supplier-risk question wearing technical clothing, and it is squarely yours to ask.

Generative vs predictive AI

In plain English: Predictive AI (long established in most large organisations) forecasts and classifies: churn scores, demand forecasts, fraud flags. Generative AI produces new content and reasoning: drafts, summaries, analysis, code.

Why it matters at your level: Your organisation almost certainly had the first for years, quietly, in the data team. The second is what changed in 2023 and what everyone means now. Knowing the difference prevents a common executive confusion — treating a mature, well-governed predictive capability and a novel, weakly-governed generative one as the same risk category.

Hallucination and model risk

In plain English: Hallucination: confident, plausible, wrong output — an inherent property of probabilistic systems, not a defect awaiting a patch. Model risk is the governance term for the consequences at scale.

Why it matters at your level: At enterprise volume the arithmetic changes character: a 2% error rate is a curiosity in a pilot and a serious incident across fifty thousand transactions. The director-grade instinct is to ask not 'is it accurate?' but 'what is the error rate, on our data, and what does it cost when it happens at volume?'

Reasoning models and extended thinking

In plain English: Models that work through a problem in explicit steps before answering, rather than responding immediately. Slower and more expensive per use, materially better on genuinely hard analytical tasks. Most providers now sell them as a distinct tier.

Why it matters at your level: It is a cost and fit question rather than a technical one. Vendors will often price and propose the most capable tier by default, when most enterprise volume, summarising, drafting, classifying, does not need it. "Which tier is this running on, and does the task need it?" is the kind of question that saves real money at scale, and it is entirely askable without technical knowledge.

Non-deterministic — the property underneath the risk

In plain English: The same input can produce a different output on different runs. Traditional software is deterministic: given identical inputs it returns identical results, every time. Generative AI does not, by design.

Why it matters at your level: This is the precise word for what people are circling when they say AI 'needs proper testing', and it is more useful than 'hallucination' in a governance conversation because it names a property rather than an error. It explains why you cannot test an AI system once and declare it correct, why identical customer queries can receive different answers, and why controls have to be designed around distributions of behaviour rather than fixed outputs. Say it in a risk conversation and the room will place you accurately.

Tokens, context windows, and inference cost

In plain English: Text is processed in tokens; the context window is how much a model can hold at once; inference is the cost of each run. Unlike a licence, this scales with usage.

Why it matters at your level: The FinOps lesson from the technology series, repeating: AI cost is variable and grows silently with adoption. A successful pilot that scales tenfold scales its bill too. "What does this cost at full adoption, not at pilot volume?" is a question finance will thank you for asking first.

02

How enterprises actually deploy it

RAG — grounding AI in your own knowledge

In plain English: Retrieval-augmented generation: connecting a model to your organisation's documents, policies and data so it answers from your material rather than general knowledge. The architecture behind virtually every internal 'ask our knowledge base' initiative.

Why it matters at your level: The most common enterprise pattern, and the one whose success depends least on the model and most on something unglamorous: whether your documents are current, findable, and correct. Which is why the honest first question about any RAG project is a data-quality question, not an AI question.

Fine-tuning, prompting, and when each is warranted

In plain English: Prompting: instructing a general model well. RAG: giving it your knowledge. Fine-tuning: further training it on your examples so it adopts a task or style. Cost and permanence rise across the three.

Why it matters at your level: You need this to evaluate proposals, not to choose. The pattern to watch: fine-tuning proposed early, expensively, and often unnecessarily — because prompting plus retrieval solves most enterprise use cases at a fraction of the cost and stays current as models improve. "What does fine-tuning achieve that context and retrieval cannot?" is a fair, non-hostile question.

Context engineering — the term that replaced prompt engineering

In plain English: The discipline of managing everything a model can see when it answers: which documents are retrieved, how current they are, what the system knows about the user and the task. Prompt engineering, the crafting of instructions, is now understood as one component inside it.

Why it matters at your level: Worth knowing for two reasons. First, if you learned 'prompt engineering' as the skill to have, that framing has moved on, and using the older term marks you as a year or two behind. Second, and more usefully, it relocates the problem: when an AI initiative underperforms it is now far more often a context failure than a model failure, which means the fix is usually data quality, access and currency rather than a better model or a cleverer prompt. That is a budget conversation, and it is one you can lead.

Copilots and per-seat AI economics

In plain English: AI embedded in tools your people already use — productivity suites, CRM, service desks — licensed per user per month on top of existing costs.

Why it matters at your level: This is where most enterprise AI spend quietly lands, and where the shelfware risk is highest: seats bought in enthusiasm, unused by month four. Two director disciplines: audit usage against licences quarterly, and check whether the AI features being upsold are already included in a tier you hold. Both are exactly the licensing hygiene from the technology series, applied to a faster-growing line.

Agents and orchestration

In plain English: Agents take multi-step actions toward a goal rather than producing output for a human. Orchestration platforms coordinate them across systems. The defining enterprise shift of 2026 — from AI that answers to AI that does.

Why it matters at your level: The governance stakes rise steeply here, and the vocabulary to hold onto is scope of action: what can it do without a human, in which systems, with what limits, and how would we know if it did something wrong quietly? An agent initiative without crisp answers to those four is a pilot pretending to be a programme.

Build, buy, and embedded

In plain English: Build: your own solution on a provider's model. Buy: a vendor product. Embedded: AI arriving inside software you already own, whether or not you asked for it.

Why it matters at your level: The third is the one that catches organisations out — capability (and data flows, and risk) arriving through existing suppliers without passing any AI governance gate. Worth knowing which of your function's current tools have quietly switched features on.

MCP — the emerging integration standard

In plain English: Model Context Protocol: an open standard, originated by Anthropic and now supported across the major providers and a growing share of enterprise software vendors, that lets AI tools connect to systems and data without bespoke integration work for each pairing. It moved onto executive agendas during 2026.

Why it matters at your level: Integration complexity, not model capability, has been the practical brake on enterprise AI, and this is the industry's answer to it. Two things follow for you. Ask whether the vendors in your function support it, because it changes the cost and timeline of anything you are told is a custom build. And know that the security community's attention has moved here too: an over-permissioned connection is a real exposure, so "who reviewed what this connection can reach?" is a fair question to put alongside the enthusiasm.

03

Governance, risk, and the regulatory floor

AI governance — what the word actually covers

In plain English: The framework of decisions: which tools are approved, what data may enter them, which use cases require review, who signs off, who is accountable for outcomes, and how it is all recorded.

Why it matters at your level: Every function is now running AI whether sanctioned or not, which makes this your vocabulary regardless of your role. The director who asks "what's our approved toolset, our data boundary, and our checking requirement — on one page?" is contributing governance rather than resisting progress, and is usually asking a question nobody has yet answered simply.

Human in the loop, and meaningful human oversight

In plain English: A person reviews or approves before an AI output takes effect. 'Meaningful' is the operative word — a reviewer approving four hundred items an hour is providing decorative oversight, not real control.

Why it matters at your level: This is the phrase that separates genuine control from compliance theatre, and it is exactly where regulators and auditors probe. When a proposal claims human oversight, the question is: how many items, in how much time, with what information in front of them?

Sycophancy — the failure mode of a compliant assistant

In plain English: The documented tendency of these models to agree with the user: to accept the framing they are given, soften disagreement, and confirm what the questioner appears to want confirmed.

Why it matters at your level: This is the quiet risk inside 'human in the loop' and it deserves naming. A reviewer checking AI output is a control; a reviewer asking an AI whether their own analysis is sound is not, because the answer will tend to agree. It matters wherever AI is used in analysis, business cases or option appraisal: the tool will support the case it is shown. The mitigation is procedural rather than technical, ask for the strongest counter-argument, the assumptions that would have to hold, or the case against, and treat unprompted agreement as no evidence at all.

The EU AI Act and the risk-based approach

In plain English: The EU's regime classifies AI uses by risk, with the heaviest obligations on 'high-risk' applications — notably including employment decisions, credit, and access to essential services — and outright prohibitions on a small set of practices. It reaches organisations outside the EU whose systems affect people in it, and obligations have been phasing in across 2025–2026. The UK has taken a lighter, regulator-led route so far, but sectoral regulators are active.

Why it matters at your level: You are not expected to know the detail. You are expected to know that AI touching hiring, promotion, credit, or customer eligibility sits in a different regulatory category from AI drafting a marketing email — and to say so before an enthusiastic pilot heads in that direction. Route to legal and compliance early; that instinct is the whole requirement at your level.

Data protection and AI

In plain English: Personal data in AI systems still sits under UK GDPR: lawful basis, purpose limitation, transparency, and rights of access and erasure. Plus a newer question — whether your data is used to train a provider's models, which differs sharply between consumer and enterprise agreements.

Why it matters at your level: The practical exposure at your level is straightforward and common: staff pasting personal or client data into consumer tools with no agreement in place. That is not an AI problem; it is a data protection incident with an AI-shaped origin.

Shadow AI — and why it's a symptom

In plain English: Unsanctioned AI use across the organisation: personal accounts, browser extensions, free tiers, quietly enabled features.

Why it matters at your level: It is happening in your function now; assume rather than hope. Read it as the technology series taught: people route around slow official channels, so the tools they chose tell you where the friction is. The response is a fast, credible sanctioned path plus a clear boundary, never a policy nobody can enforce.

AI-washing

In plain English: Existing capability — rules engines, statistical models, chatbots — rebadged as AI, sometimes repriced accordingly. Now sufficiently common that financial regulators have warned about AI claims in investor and customer communications.

Why it matters at your level: Two uses: evaluating vendors ("what can it do that it couldn't two years ago?"), and evaluating your own organisation's statements. A director who notices their function's work being overstated in an AI narrative, and says so early, is doing genuine risk work.

Agent washing

In plain English: This year's variant: existing chatbots, scripted automations and rules engines rebadged as 'agents' because agentic AI is what attracts funding, attention and premium pricing in 2026.

Why it matters at your level: The distinction is not pedantry, it is a governance one. A genuine agent chooses its next action, which is why it needs scope limits, reversibility rules and an audit trail; a scripted automation does not. Approving the second under the governance regime of the first wastes control effort; approving the first as though it were the second is how organisations end up with unsupervised action in production. "What decides what it does next, the model or a rule somebody wrote?" separates them in one question.

04

The five things to know about AI in your own function

- **Your sanctioned and shadow usage** — what your team is officially licensed for, and what they're actually using. Ask openly; the answer is the fastest audit available.
- **Your AI-touched data** — which of your function's data flows into AI tools, and whether any of it is personal or client-confidential.
- **Your embedded AI** — which existing vendors have switched AI features on inside tools you already run.
- **Your copilot seats versus actual usage** — the number, the cost, and the honest utilisation after three months.
- **Your one high-value use case** — the narrow, measurable job in your function where AI would save the most hours at the least risk. Have it ready before you're asked.

Three questions for your next conversation with IT, data, or your CIO's team:

- "What's our approved toolset and data boundary — and can I have it on one page for my team?"
- "Which use cases need review before we pilot them, and who decides?"
- "What's the cost model at full adoption rather than pilot volume?"

Where Part Two takes you

You now have the vocabulary — the engine, the deployment patterns, the governance and regulatory language — which means the AI agenda item no longer requires live translation.

Part Two, "Reading the Story," goes deeper: reading AI business cases and vendor claims the way executives read forecasts, the five patterns of enterprise AI failure — the proof-of-concept graveyard, the productivity claim nobody can find in the P&L, the oversight that isn't — and how to read what AI is doing to your function's position in the organisation.

One observation from inside large organisations: the AI conversation is currently the least evidence-based conversation happening at senior level. Which is precisely why the director who brings evidence, error rates and cost-at-scale into it becomes disproportionately influential, disproportionately fast.

John Obidipe is a business transformation specialist and Gallup-Certified CliftonStrengths coach. He works with founders and leadership teams on the structural side of growth — the part the numbers can only hint at. ·
coach@catalystgrowthcoach.co.uk · catalystgrowthcoach.co.uk