

Session 2: AI for Leadership & Managerial Effectiveness

Using AI to support sense-making, decisions, communication and execution

AI as a leadership copilot: sense, interpret, decide, communicate, execute

Leadership with AI – Rajendra M Sonar, IIT Bombay 

What we are going to learn?

Understand how AI is different than other techs so far. What it needs behind the scenes.

What are the advantages and issues of AI. What it means for leadership.

Think on possible tasks involved in AI leadership right from picking up right use-case or problem to POC, deployment to monitoring.

Understand AI tools landscape well using generic Gen AI tools like ChatGPT, Gemini, Claude etc.

Practice how AI and other tech can be leveraged to lead the job efficiently, productively and intelligently in autonomous mode

1. Need of Strategic & Digital Foundations

Pre-requisites & Spectrum of Intelligence

- Digital foundations (APIs, cloud, integration) and how AI readiness strictly depends on digital readiness.
- Use meta-data as much as possible about everything: software, APIs, data sources, ML models, skillsets etc.
- The spectrum of intelligence: from expert rules to machine learning, RAG, and fully autonomous Agentic AI.
- Operational failure modes: identifying hallucinations, data privacy leaks, bias, and the cost-of-error trade-offs.

2. Bounded 10-Step Template for reference

Reusable Enterprise Framework

- Understanding role of strategy, governance, performance, regulation, Risk, Culture & Trust
- Structuring business problems & evaluating standard software alternatives before selecting AI.
- Establishing governance boundaries: policy, acceptable use, data residency, and human accountability.
- Financial modeling for AI: total cost of ownership (TCO) and maximizing business value (Valuemaxxing).

3. AI as co-pilot

Using AI tools for leadership tasks

- Using reference template for specific use-case or problem to be solved.
- Walk through AI tools like ChatGPT to create a blueprint for problem or use-case based on the reference template.
- Try to automate all the steps using Agentic AI to great extent possible.
- Make engineering team to follow blueprint and implement AI solution to the problem or use-case and monitor on on-going basis.

Recap: Digital Readiness Defines AI Success - robust core infrastructure must precede model deployment

The Integration Challenge

Systemic Isolation: AI cannot work effectively in isolation. To drive real business value, it must plug directly into existing transactional databases, document stores, and system workflows.

The Dependencies: Data platforms, cloud hosting, secure API endpoints, and clean integration patterns represent the invisible foundation of any successful AI initiative.

Security Posture: Robust access controls, authorization, and data encryption must be designed into the infrastructure level from Day One.

Key Takeaway: AI readiness is digital readiness.

Layer 4: AI & Agentic Capabilities

Cognitive Models, Semantic Engines, Custom RAG, Action Tools

Layer 3: Workflows & Orchestration

Secure API Integrations, Role-Based Access, API Gateways

Layer 2: Cloud Data Platforms

Vector Databases, Elastic Compute, Clean ETL Pipelines

Layer 1: Enterprise Core Systems

ERP, CRM, Databases, Active Directory, Shared Drives

Recap: Selecting the Right Intelligence - choose the simplest model that delivers business value

Rules / Semantic

Rules Decide.

Deterministic decision trees, explicit business logic, and standard FAQs.

Classical ML

ML Predicts.

Statistical algorithms for forecasting, churn prediction, and numeric modeling.

Generative AI

GenAI Interprets.

LLMs for reading, writing, parsing, and interpreting unstructured text data.

Agentic AI

Agents Plan & Act.

Autonomous systems executing multi-step reasoning, tool usage, and self-correction.

THE LEADERSHIP PARADIGM

Rules decide → ML predicts → GenAI generates & interprets → Agents plan and act

Rule of thumb: Not every problem needs Generative AI or an Autonomous Agent.

Managing AI Downside Risk - accuracy alone is not enough to protect enterprise operations

Technical & Data Failure Modes

Hallucinations & Fabrication: Models generating confident but entirely incorrect statements.

Context Underflow: Poor or incomplete document parsing, leading to missed details.

Model Drift & Degradation: Output quality decaying as real-world data and usage shift.

Data Leakage: Confidential customer or corporate data escaping boundaries.

Operational & Human Risks

Bias & Toxic Outputs: Propagating historical prejudices or inappropriate recommendations.

Over-Automation: Removing human oversight from safety-critical decision paths.

Unexplained Logic: Black-box decisions that compliance/legal cannot audit.

Agent Misuse: Autonomous loops consuming thousands of dollars in unexpected API fees.

CRITICAL LEADERSHIP LESSON

Model accuracy alone is never enough. Leaders must always evaluate the relative economic and human cost of False Positives versus False Negatives within their operating domain.

Three Pillars of AI Governance - human accountability, token economics, and value alignment

1. Governance & Humans

Human-in-the-loop: Establish clear manual override gates for high-risk, customer-facing actions.

Policy Alignment: Assess and document data classifications and data residency before coding begins.

2. Token Economics

Infrastructure Budgets: Forecast monthly API consumption, hosting costs, and token routing patterns.

Valuemaxxing: Optimize strictly for net business value delivered, rather than minimizing or maximizing raw tech usage.

Classical AI: Use classical AI wherever possible.

3. Continuous Quality

Active Monitoring: Track accuracy, latency, and feedback loops in real-time post-deployment.

Freshness Audits: Establish automated routines to keep model knowledge bases synchronized with actual realities.

CORE STRATEGIC PRINCIPLE

AI success is not model success; it is business value delivered reliably, responsibly, and economically.

The AI Leadership Template*

Problem → Fit → Value → Governance → Capability → Economics → Risk → Pilot → Decision → Scale & Monitor

Step	Generic Leadership Question	Expected Output / Deliverable
1. Understand Problem	What specific business problem or opportunity are we addressing?	Business Problem Statement
2. Challenge AI Fit	Can process redesign or standard software solve it without AI?	AI Fit Assessment
3. Define Success Metrics	What baseline are we comparing against & what KPIs should improve?	Value Hypothesis + KPI Target Matrix
4. Check Policies & Gov.	Can this data leave the organization? Can cloud LLMs process it?	Governance & Policy Boundary Sign-off
5. Assess Capabilities	What deployment matches our existing skills and infrastructure?	Deployment Choice & AWS/Cloud Architecture
6. Evaluate Economics	Are we maximizing net business value rather than raw AI usage?	TCO Model & ROI Business Case
7. Identify Risks	Pre-mortem: If this AI initiative failed badly, what was the reason?	Stakeholder Matrix + Active Risk Register
8. Design the Pilot	How do we test assumptions safely with a narrow golden dataset?	Pilot Plan & Benchmark evaluation test
9. Management Decision	Proceed, Modify, Extend Pilot, Defer, or Stop based on evidence?	Management Decision / Go-No-Go Record
10. Scale & Monitor	How do we maintain quality, value, and safety after deployment?	Live Performance Dashboard & Sync Schedule

* This template (a kind of framework) is created to walk through how AI and other tools can help AI Leadership as co-pilots. Demos on are using ChatGPT. Participants are free to create their own template based on typical tasks followed for addressing a problem or solving use-case. They can use other AI other tools instead of ChatGPT. *Participants may follow this template for capstone project.*

The Strategic Bridge

“We now know what digital foundations are required, what forms of AI are available, and what their limitations are. The next question is:

How do we actually put these capabilities to work?”

Introducing a rigorous, 10-step template designed to systematically evaluate, pilot, and scale enterprise AI use cases.

RUNNING CASE BRIEF

Scenario: AI-Powered Enterprise Knowledge Assistant

Core Problem: Employees waste excessive time searching unstructured HR policies, travel rules, and SOPs scattered across legacy intranets.

The Vision: An AI RAG Assistant capable of searching documents, synthesizing precise answers with verified page-level citations.

The Method: To keep this practical, we will apply each step of our 10-step template directly to this case study side-by-side.

Step 1: Understand and Structure the Business Problem - define real needs before technology

Understand & Structure the Business Problem

The Mandate: Define the real business problem before discussing algorithms or models. Focus strictly on operational constraints and organizational pain.

Five Key Diagnostic Questions:

- What specific business problem or revenue/efficiency opportunity are we addressing?
- Who is affected (customers, partners, HR, operations, sales)?
- What is the current manual or system-based process?
- What are the major structural pain points and process delays?
- What specific decisions or human tasks are involved in the loop?

Core Output: A rigorous Business Problem Statement.

RUNNING CASE CO-PILOT

Enterprise Knowledge Assistant Context

The Pain Point: Employees waste hours looking for travel, expense, HR, and IT security policies scattered across legacy portals.

The Baseline: Average search takes 25 minutes with a 15% failure rate, driving 300+ repetitive support tickets to helpdesks weekly.

The Target: Enable rapid query answering in <1 min with direct source citations to minimize search time and ticket volumes.

Deliverable: Structured Business Problem Statement focused on employee productivity and helpdesk load.

Step 2: Challenge Whether AI Is Actually Required - avoid starting with AI as an assumed solution

Challenge Whether AI Is Actually Required

The Mandate: Avoid starting with AI as an assumed solution. Actively audit cheaper, more robust traditional alternatives.

Alternatives to Evaluate first:

- Process Redesign: Can simplifying the workflow solve the root cause?
- Conventional Software: Can structured databases, standard search, or indexing solve it?
- Rules/Expert Systems: Are standard hardcoded FAQ routing patterns sufficient?

Selecting the AI Tier: If traditional approaches fail, map the problem to the simplest suitable AI approach (Classical ML, RAG, GenAI, or Agentic).

Core Output: AI Fit Assessment (No AI / Classical / GenAI / Hybrid).

RUNNING CASE CO-PILOT

Enterprise Knowledge Assistant Context

Traditional Audit: Could better folder structures, keyword search, or an FAQ site solve the search issue?

Alternative Failures: Folders fail because policies are unstructured and long; standard keyword search fails due to synonyms and disjointed terms.

AI Fit Result: Retrieval-Augmented Generation (RAG) is selected. It reads unstructured PDFs, retrieves semantic paragraphs, and synthesizes answers safely without expensive custom model training.

Deliverable: RAG-based AI Fit Assessment approved by IT and operations.

Step 3: Define Expected Business Value and Success Criteria - establish why the project is worth pursuing

Define Expected Business Value and Success Criteria

The Mandate: Establish a hard value hypothesis and specific comparative KPIs before project authorization.

Business Value Metrics:

- Productivity: Time saved and human throughput increased.
- Operational Cost: Cost savings in manual review and support overhead.
- Quality & Turnaround: Turnaround time and process error rate reduction.
- CSAT/ESAT: Direct customer and employee satisfaction improvements.

Core Output: Value Hypothesis + Success Metrics Matrix.

RUNNING CASE CO-PILOT

Enterprise Knowledge Assistant Context

The Value Hypothesis: Giving employees instant access to policy answers will recover wasted search hours and lower helpdesk volumes.

Target KPI Metrics Matrix:

- Processing Time: Baseline 30m → Target 10m
- Search Accuracy: Baseline 85% → Target 95%
- Cost Per Case: Baseline ₹500 → Target ₹300
- User CSAT: Baseline 75% → Target 85%

Deliverable: Hard baseline-to-target KPI tracking matrix.

Step 4: Check Policies and Governance Constraints - define organizational limits before development

Check Policies and Governance Constraints

The Mandate: Define organizational, legal, and compliance boundaries before selecting or deploying technology.

Assess Constraints:

- **Data Classification:** Confidentiality, data privacy, and intellectual property limits.
- **Data Residency:** Legal limitations on cross-border data transfer.
- **Public Cloud Policies:** Rules governing cloud-hosted LLM usage.
- **Audit & Accountability:** Explanatory and citation requirements.

Core Output: Governance & Policy Boundary documentation.

RUNNING CASE CO-PILOT

Enterprise Knowledge Assistant Context

Data Classification: Policy documents contain proprietary business SOPs and personal HR information.

Governance Constraints: No corporate document can be sent to public, unvetted LLM models for training.

Policy Boundaries: Use enterprise cloud APIs with zero-data-retention agreements; sync role-based access to prevent unauthorized file reads; force page citations for auditing.

Deliverable: Data Privacy and Governance sign-off record.

Step 5: Assess Capabilities and Compare Solution Options - select an architecture that matches skills & stack

Assess Capabilities & Compare Deployment Options

The Mandate: Map the technical solution to existing capabilities, infrastructure, and organizational skills.

Three Pillars of Capability Evaluation:

- Existing Stack: Databases, cloud architecture, secure networks, identity systems.
- Team Skills: In-house capabilities in Python, cloud, APIs, data engineering, MLOps.
- Deployment Choice: Compare 'Buy' (SaaS), 'Configure' (managed cloud API), vs. 'Build' (custom development).

Core Output: Recommended Deployment & Architecture choice.

RUNNING CASE CO-PILOT

Enterprise Knowledge Assistant Context

Existing Stack: Corporate intranet relies on SharePoint; cloud resources are hosted on AWS.

Existing Skills: Strong internal Python development and AWS security skills, but zero deep model-training experience.

Architecture Decision: Configure and Integrate. Build a custom RAG pipeline on AWS using managed serverless database and LLM APIs. This fits existing skills perfectly and minimizes overhead.

Deliverable: AWS-hosted RAG System Architecture Diagram.

Step 6: Evaluate Benefits, Costs, and Business Value - ensure AI economics deliver net business ROI

Evaluate Benefits, Costs, and Business Value

The Mandate: Build a comprehensive Total Cost of Ownership (TCO) model compared against strategic business returns.

Three TCO Core Dimensions:

- CapEx/OpEx: Development, infrastructure, API token consumption, vector search hosting.
- Maintenance: Continuous data curation, drift monitoring, human review pipelines.
- Valuemaxxing: Optimizing purely for net business value delivered, rather than minimizing or maximizing raw tech usage.

Core Output: Business Case / Value Model.

RUNNING CASE CO-PILOT

Enterprise Knowledge Assistant Context

CapEx (Setup): ₹15,000 one-time development and AWS RAG infrastructure setup costs.

OpEx (Monthly): ₹2,000 monthly cloud hosting, vector database, and API token fees.

Valuemaxxing ROI: Reclaiming 2,400 search hours monthly across employees equates to ₹1.2 Lakhs in recovered productivity—delivering a massive 4.5x return on monthly cost.

Deliverable: TCO and ROI Business Case sheet.

Step 7: Identify Stakeholders, Risks, and Mitigations - conduct a pre-mortem to prevent failure

Identify Stakeholders, Risks, and Mitigations

The Mandate: Proactively identify execution friction points and run a rigorous pre-mortem exercise before development.

The Pre-Mortem Prompt:

- "Assume this AI initiative failed badly one year from now. What were the most likely reasons?"

Risks to Monitor: Hallucinations, outdated information, user adoption gaps, technical integration failures, or access policy violations.

Core Output: Stakeholder Matrix + Risk & Mitigation Register.

RUNNING CASE CO-PILOT

Enterprise Knowledge Assistant Context

Pre-Mortem Failure Causes:

1. Employees stopped using it because it cited outdated maternity policies.
2. HR team resisted deployment out of fear of job replacement.

Mitigation Strategies:

- Schedule automatic vector db syncs when policy docs update.
- Communicate the AI as an efficiency tool, not a human replacement.

Deliverable: Active Risk Mitigation Register.

Step 8: Design the Pilot / Proof of Concept - test critical assumptions in a safe, controlled sandbox

Design the Pilot / Proof of Concept

The Mandate: Test critical assumptions in a safe, low-consequence sandbox before committing to full organizational scale.

Core Elements of Pilot Design:

- Scope: Specific, narrow department, user group, and document subset.
- Evaluation: Construct a 'Golden Dataset' of 100 historical queries with human-verified correct answers.
- Go/No-Go: Set quantitative metrics for accuracy, latency, and costs.

Core Output: Pilot / Proof of Concept Plan.

RUNNING CASE CO-PILOT

Enterprise Knowledge Assistant Context

Pilot Scope: Only HR Benefits and Travel Policy documents, rolled out to 100 pilot users over 4 weeks.

The Golden Dataset: 100 historical queries with perfect, manually verified answers to benchmark model output.

Go/No-Go Gates: Pilot must achieve >90% answer accuracy and <5 seconds response latency before proceeding to production.

Deliverable: Pilot Plan and benchmarking test framework.

Step 9: Make the Management Decision - convert pilot evidence into formal corporate action

Make the Management Decision

The Mandate: Convert quantitative pilot evidence into a formal, recorded business option decision.

Five Formal Paths to Choose From:

- **Proceed:** Scale to full production immediately.
- **Modify:** Address specific gaps or user interface issues before expanding.
- **Extend Pilot:** Continue testing; evidence is currently insufficient.
- **Defer / Stop:** Organization, stack, or value is not ready.

Core Output: Management Decision / Go-No-Go Record.

RUNNING CASE CO-PILOT

Enterprise Knowledge Assistant Context

Pilot Performance: Achieved 92% retrieval accuracy, 3.8-second average latency, and high user CSAT of 4.2/5.0 during the trial.

Strategic Decision: Proceed with Modifications. Authorize enterprise-wide rollout to all 5 business lines.

Required Gaps to Fix: Build a visual query helper UI for complex lookups and mandate HR owners to pruneDeprecated document versions monthly.

Deliverable: Signed Go/No-Go Decision Document.

Step 10: Scale, Monitor, and Continuously Improve - maintain value, quality, and compliance post-launch

Communicate, Execute, Monitor & Improve

The Mandate: Deploy, communicate changes, and set up continuous monitoring dashboards to maintain quality and safety post-launch.

Four Lenses of AI Monitoring Dashboards:

- **Business:** Adoption rate, search hours saved, ticket deflection rate.
- **AI Quality:** Accuracy, hallucinations, human override rates.
- **Operations:** System availability, request latency, API token cost trends.
- **Governance:** Data leakage, policy violations, compliance audit logs.

Core Output: AI Dashboard + Continuous Improvement Plan.

RUNNING CASE CO-PILOT

Enterprise Knowledge Assistant Context

Execution & Training: Conduct town-halls framing the assistant as a helper. Establish HR ownership over policy vector stores.

The Dashboard: Track monthly API token budgets, user thumb ratings, latency spikes, and policy update triggers.

Continuous Improvement: Run automatic vector db index syncs every Friday to absorb newly approved or modified documents.

Deliverable: Live AI performance dashboard & maintenance schedule.

Automation of template steps (Agentic leadership management, a possibility!)

Phase 1: High-Value Targets for AI Automation: Several steps in your 10-step framework are highly analytical, making them perfect candidates for automation using specialized AI agents, LLMs, or custom diagnostic tools:

Step 1 (Problem Structuring) & Step 3 (Metrics Definition):

How to Automate: You can deploy a "Business Analyst Agent" to ingest raw, messy operational data (like search logs, customer support transcripts, or ticket histories) and automatically extract current baseline performance, pinpoint major bottlenecks, and suggest realistic target KPIs (like processing times, CSAT, or cost-per-case targets).

Step 2 (AI Fit Assessment):

How to Automate: A "Solution Architect Agent" can run a diagnostic routing pipeline. Given a structured problem statement, it can run a decision-tree analysis to evaluate if conventional software, rules-based engines, classical machine learning, or GenAI/Agentic architectures are the most cost-effective fit.

Step 4 (Governance & Policy Boundary Check):

How to Automate: AI-powered compliance tools can ingest your organization's data classification frameworks and acceptable-use policies. It can scan proposed data-flow diagrams to automatically flag potential data residency issues, regulatory infractions, or unauthorized external cloud LLM usage.

Step 6 (Economics & TCO Modelling):

How to Automate: You can run automated token-accounting scripts that simulate API payload sizes to generate precise estimates of token and infrastructure costs. These scripts can run "Valuemaxxing" scenarios to optimize model routing (e.g., routing 80% of simple queries to cheap, local models and only 20% to premium cloud LLMs).

Step 8 (Pilot Testing & LLM-as-a-Judge):

How to Automate: In your pilot, manually auditing every output is impossible. You can automate the evaluation of your **Golden Evaluation Dataset** by using a secondary LLM as a judge to grade response accuracy, hallucination rates, and retrieval quality, feeding a live dashboard.

Automation of template steps (Agentic leadership management!)

Phase 2: The Red Lines — Where Automation Fails

While AI can accelerate analysis, your sources explicitly warn against **over-automation** and outline several steps where human leadership, judgment, and political alignment are entirely irreplaceable:

Step 4 & Step 10 (Governance & Human Accountability):

Your sources raise a pivotal leadership question: **"Which actions must require human approval?"** and **"Who is accountable?"**. If an AI compliance tool approves a model that subsequently leaks proprietary data, the organization cannot blame the auditing agent. Humans must define the ethical and legal boundaries.

Step 7 (The Pre-Mortem Exercise):

While an LLM can generate a generic list of risks, it cannot replace the psychological safety and raw organizational intuition of a team-driven **Pre-Mortem exercise**. Experienced practitioners know that political friction, resistance from middle management, and change-management barriers are the primary killers of AI adoption—nuances that AI cannot feel or predict.

Step 9 (The Management Decision):

The transition from pilot evidence to a Go/No-Go decision (Proceed, Modify, Extend, Defer, or Stop) is a high-stakes, multi-variable judgment call. It involves balancing strategic fit, budget allocation across competing business units, and organizational readiness—which require human negotiation and executive ownership.

Step 10 (Scale & Continuous Improvement):

Your source emphasizes that **"AI success is not model success; it is business value delivered reliably, responsibly, and economically"**. While AI can track the metrics, only human leaders can judge if that delivered value is translating into actual market competitiveness, employee satisfaction, and strategic alignment.

Key Takeaways: Strategic AI Leadership Guidelines and AI as co-pilot

Understand digital and AI at core

AI sits on top of existing digital infrastructures and even integrated into them. AI leaders need to have fair understanding of technologies involved, their roles, benefits, costs, issues and risks to lead AI initiatives.

Need of a operational template

Create a reference template to for a complete leadership blue print. Right from problem definition, decisions to scale and monitor. All problems and use-cases can be put through this framework.

Understanding tools and using them

Leaders should know how to use AI tools to make their leadership job more effective, informed and productive. The good part is, AI can make leaders educated during the process telling in detail every tech piece, its roles, issues, risks etc.

Automate the template

With the help of AI/IT engineers leaders can possibly automate all steps in operational framework by integrating various tools and make the framework itself agentic.