A PRACTITIONER'S BLUEPRINT FOR STANDING UP A CENTRAL AI FUNCTION THAT SHIPS

Building an Internal AI Center of Excellence (CoE)

Executive briefing — C-suite & senior technology leaders | ~20 min

Why a CoE Matters

- Without central ownership, AI investment fragments into dozens of disconnected pilots
- A CoE turns one-off wins into repeatable delivery capacity across the enterprise
- Concentrates scarce skills (MLOps, prompt/agent engineering) instead of diluting them
- Gives the board a single accountable owner for AI risk, spend, and outcomes
- Shortens time-to-value by reusing infrastructure, data access, and governance patterns

What Goes Wrong Without One

- Shadow AI: business units buy overlapping tools with no security or data review
- Pilot purgatory: dozens of POCs launch, few reach production, none get decommissioned
- Duplicated spend on similar models, licenses, and vendor contracts across teams
- No shared evaluation standard, so "it works" is never independently verified
- Compliance exposure surfaces only after a regulator or auditor asks first

CoE Mandate & Scope

- Owns AI strategy, standards, and the approved technology stack — not every project
- Sets the bar: model/vendor evaluation, security review, and responsible-AI sign-off
- Builds and maintains shared platform capital: infrastructure, tooling, reusable components
- Explicitly out of scope: routine BI/analytics work already served by existing data teams
- Mandate reviewed and renewed annually against delivered business outcomes

Organizational Structure & Roles

- Hub-and-spoke model: central hub sets standards, embedded "spoke" leads sit in business units
- AI Product Owner — translates business problems into scoped, fundable initiatives
- Prompt/Agent Engineer — designs, tests, and hardens LLM-based workflows
- MLOps Engineer — owns deployment, monitoring, cost, and model lifecycle
- Responsible AI Lead — coordinates risk, privacy, and compliance review with legal

Use Case Intake & Prioritization

- Single intake form: every idea, from any business unit, enters one shared pipeline
- Impact/effort rubric scores value (revenue, cost, risk) against build complexity (1-5 each)
- Data readiness gate: no scoring past intake until data access is confirmed feasible
- Monthly triage board (CoE + business sponsors) approves the next quarter's build slate
- Kill criteria set at approval time — pilots without a path to production aren't approved

Reusable Components: Build Once, Deploy Many

- Shared component library: auth, logging, evaluation harnesses, RAG retrieval templates
- Agent "skills" catalog — packaged, tested capabilities teams compose rather than rebuild
- Standard project scaffold cuts new-project setup from weeks to days
- Central prompt/eval registry with version control and regression testing
- Every new build must justify not reusing an existing component before custom work starts

Governance & Compliance, Built In

- Risk-tiering at intake: low, medium, high — review depth scales with tier
- Every production model gets a documented owner, purpose, and decommission date
- Legal, privacy, and security are co-reviewers at design time, not gatekeepers at launch
- Audit trail: model version, training/reference data lineage, and evaluation results logged
- Aligns to emerging AI regulation (e.g., EU AI Act risk categories) without waiting for mandates

Training & Capability Building

- Tiered curriculum: awareness for all staff, hands-on for builders, deep-dive for engineers
- Internal certification path signals who is cleared to build production AI systems
- Office hours and an internal Slack/Teams channel for real-time troubleshooting support
- Rotational program embeds business analysts in the CoE for 3-6 months to build capability
- Capability, not headcount, is the real bottleneck — training budget treated as core, not optional

Vendor & Tool Management

- CoE owns the approved vendor list; business units select from it, don't source independently
- Structured evaluation: accuracy, cost per unit of work, data handling, and exit terms
- Multi-vendor strategy for core model access avoids single-provider lock-in and outages
- Contract terms mandate no training on enterprise data without explicit opt-in
- Quarterly cost review — usage-based AI spend is reallocated faster than typical software

Success Metrics & KPIs

- Adoption: active users and workflows touched, not just pilots launched
- Delivery: cycle time from intake to production, and percentage of pilots that ship
- Value: hours saved, cost avoided, or revenue influenced, validated by business sponsors
- Quality: model accuracy/error rates and human override rates in production
- Efficiency: reuse rate of shared components across new projects

Illustrative Example: CoE-Driven Pilot Rollout

- Representative scenario, not a specific client — a global industrial conglomerate
- Problem: field service teams spent hours searching manuals for equipment troubleshooting
- CoE reused its RAG template and eval harness to cut build time to roughly six weeks
- Piloted with one region, measured technician resolution time before wider rollout
- Result pattern: material reduction in average resolution time, then scaled to additional regions

Common Pitfalls to Avoid

- Treating the CoE as a bottleneck approval office instead of an enabling platform team
- Chasing flagship demos while ignoring unglamorous production reliability work
- Understaffing MLOps — pilots stall in production without operational ownership
- No sunset process, so retired experiments quietly keep consuming budget and risk
- Measuring activity (models built) instead of outcomes (value delivered)

Maturity Roadmap

- Stage 1 (0-6 mo): Foundation — governance, intake process, first 2-3 production pilots
- Stage 2 (6-12 mo): Scale — reusable component library, embedded spoke model live
- Stage 3 (12-18 mo): Optimize — cost governance, formal KPIs, vendor consolidation
- Stage 4 (18mo+): Federate — business units self-serve within CoE guardrails
- Each stage gated by evidence of the prior stage's outcomes, not the calendar

The Ask & Next Steps

- Approve initial CoE charter, budget, and a founding team of 4-6 roles for year one
- Sponsor executive to co-chair the monthly use-case triage board
- Commit first two business units as spoke partners for the initial pilot wave
- Review progress against the maturity roadmap at 90 days and 6 months
- Decision requested by [date] to hit the next planning cycle's funding window