

Deploying AI in Firms & Clinics: How to Avoid the Destructive Path

One-page checklist • v1.0 (Aug 2025) • Thorsten Meyer AI — thorstenmeyeraai.com

Purpose & Risk Boundary

- ☐ Define the problem and success metrics (quality, time, cost, equity); list harms to avoid.
- ☐ Classify risk: advice-only, draft generation, triage, decision support, autonomous; set forbidden uses.
- ☐ Publish clear stop-rules and human-second-opinion triggers.

Accountability & Governance

- ☐ Name a single accountable owner and map RACI across legal/clinical/data/security.
- ☐ Stand up a Safety/QA review with documented sign-offs for go-live and changes.
- ☐ Allocate a Training Dividend (\geq X% of savings) for apprenticeship, supervision, and audits.

Data & Privacy

- ☐ Map data flows; enforce minimization, consent/notice updates, and purpose limits.
- ☐ Protect PHI/PII: masking/de-identification with re-identification testing; encryption in transit/at rest.
- ☐ Define retention & deletion; review vendor boundaries and cross-border transfers.

Model & Prompt Controls

- ☐ Select model by risk; prefer retrieval-grounded outputs with citations and source controls.
- ☐ Build a test harness: accuracy, hallucination rate, robustness/OOD, bias/fairness, adversarial prompts.
- ☐ Register prompts/tools; lock golden prompts; guard against prompt injection.

Change Control & Monitoring

- ☐ Version model/prompt/dataset; run shadow/A-B before rollouts; document change tickets.
- ☐ Define rollback and on-call ownership; maintain a change logbook.
- ☐ Live monitors: drift, latency, alert precision/recall, override rate, near-miss rate.

Human Oversight & Workflow

- ☐ Two-person verification for high-risk outputs; mandatory human sign-off/attestation.
- ☐ Sample audits \geq X% weekly; escalate low-confidence/ambiguous cases via an exception queue.
- ☐ Make non-delegable decisions explicit (what AI must not decide).

Documentation & Auditability

- ☐ Maintain process logs: prompts, sources, model/version, checks, and responsible reviewer.
- ☐ Store per-decision evidence packs for N years with reproducible outputs.
- ☐ Client/patient disclosures and consent language reviewed by counsel/ethics.

Training & Apprenticeship

- ☐ Simulation cases, graded autonomy ladders, and skill checklists for juniors.
- ☐ Pair juniors with seniors on AI-assisted matters; protect training time in schedules.
- ☐ Track training hours/FTE and learning outcomes; report quarterly.

Procurement & Contracts

- ☐ Data ownership/return rights; no model training on your data without explicit opt-in.
- ☐ Security & audit rights; breach notification SLA; third-party pen-tests.
- ☐ Quality SLAs tied to metrics; indemnities/liability aligned to risk; exit & fallback plan.

Incident Response

- ☐ Kill switch and manual fallback rehearsed quarterly (fire-drills).
- ☐ Severity matrix with contact tree; regulator and public comms templates ready.
- ☐ Blameless 7-day postmortems with corrective actions and ownership.

Equity & Ethics

- ☐ Measure subgroup performance; mitigate disparities and document fixes.
- ☐ Accessibility/usability checks for patients/clients and staff.
- ☐ Guarantee rights: human second opinion and meaningful explanation.

KPI Dashboard (review cadence)

- ☐ Quality: error rate, near-misses, rework %.
- ☐ Safety: override rate, stop-rule triggers, audit findings.
- ☐ Operations: time saved, queue length, throughput without quality loss.
- ☐ Training: hours per FTE, audited cases, competency progression.
- ☐ Equity: disparity metrics across key subgroups.