

AI Operating Model. Von der Lösung zum Betrieb.

Portfolio, Product Ownership, Evaluation, Agent-Sicherheit, Governance und Operations für produktive KI.

Für Geschäftsführung, CIO und AI Product Owner

06

KI skaliert, wenn Verantwortung genauso systematisch wächst wie Technologie.

-
- 01 Jedes produktive KI-System braucht Business Owner, Product Owner und klare technische sowie fachliche Verantwortung.
 - 02 Evaluation wird Teil des Release-Prozesses: Qualität, Sicherheit, Kosten und Prozesswirkung werden vor und nach Änderungen gemessen.
 - 03 Agents benötigen kleine, eindeutig berechnete Tools, definierte Freigaben und eine nachvollziehbare Aktionsspur.
 - 04 Ein Portfolio Board priorisiert Investitionen; ein Plattform Team schafft wiederverwendbare sichere Standardwege.
-

Idee, Experiment, Pilot und Produkt trennen

Eine gute AI-Pipeline trennt Ideen, Experimente und produktive Systeme. Ideen werden anhand von Häufigkeit, Aufwand, Datenlage, Fehlerrisiko und wirtschaftlichem Nutzen bewertet. Experimente erhalten eine Zeitbox und eine klare Abbruchentscheidung.

Produktstatus erhält nur, was Owner, Nutzergruppe, Qualitätsmaß, Datenzugriff, Betriebskosten und Lifecycle-Budget besitzt. Exploration bleibt schnell, aber die Organisation verwechselt eine Demo nicht mit einer Lösung.

- Idee: Problem und erwartete Wirkung
 - Experiment: Hypothese, Daten und Zeitbox
 - Pilot: echte Nutzer und Evaluation
 - Produkt: Owner, SLA, Monitoring und Budget
-

Evaluation ist ein Release Gate

Modelle, Prompts, Quellen und Werkzeuge können sich unabhängig ändern. Jede Änderung kann Qualität, Verhalten, Latenz oder Kosten beeinflussen. Deshalb braucht jedes Produkt einen versionierten Evaluationssatz aus realen Aufgaben.

Automatisierte Tests prüfen Format, Quellen, Tool-Aufrufe und definierte Risiken. Fachliche Reviewer bewerten Nutzen, Vollständigkeit und kritische Grenzfälle. Beide Perspektiven ergeben die Release-Entscheidung.

- Ground Truth und repräsentative Aufgaben
 - Qualität, Sicherheit, Latenz und Kosten
 - Regressionstest vor Änderungen
 - Stufenweise Freigabe und Rollback
-

Mehr Agentenfähigkeit braucht engere Grenzen

Ein Agent verbindet Modell, Anweisungen und Tools. Er kann Daten verändern, Nachrichten vorbereiten oder Prozesse anstoßen. Damit wird er zu einer digitalen Identität mit Rechten und Verantwortung.

Tool-Zugriffe folgen Least Privilege. Kritische Aktionen benötigen menschliche Freigabe, Transaktionen idempotente Schnittstellen und alle Schritte eine nachvollziehbare Spur. Autonomie ist eine abgestufte Designentscheidung.

- Read, propose, approve und execute trennen
 - Kleinstmögliche Berechtigungen
 - Idempotenz vor Schreibzugriffen
 - Tracing, Kosten und Tool-Ergebnisse überwachen
-

Ein föderiertes Modell für den Mittelstand

Ein kleines AI Platform Team stellt Modellzugang, Retrieval, Evaluation, Logging, Security und wiederverwendbare Komponenten bereit. Fachnahe Product Teams verantworten Problem, Daten, Nutzer, Prozess und Wirkung.

Ein schlankes AI Portfolio Board entscheidet über Prioritäten, Risiken und Skalierung. Es genehmigt nicht jeden Prompt, sondern definiert klare Klassen und Standardwege für Fachteam, Security, Datenschutz und Architektur.

- Business Owner: Wirkung und Risiko
- AI Product Owner: Roadmap und Qualität
- Platform Team: sichere Standardwege
- Domain Expert: fachliche Evaluation
- Operations: Monitoring, Incident und Support

Sieben Fragen für die nächste interne Runde

01	Hat jedes System einen Owner?	[]
02	Sind Idee, Pilot und Produkt getrennt?	[]
03	Gibt es versionierte Evaluationen?	[]
04	Sind Agent-Tools minimal berechtigt?	[]
05	Sind Freigabe, Rollback und Incidents definiert?	[]
06	Werden Qualität, Nutzung und Kosten gemeinsam betrachtet?	[]
07	Gibt es ein Lifecycle-Budget?	[]

Quellen und Einordnung

- Microsoft Learn: Foundry Agent Service
- Microsoft Learn: What is Microsoft Foundry?
- Microsoft Learn: Agents hub

NÄCHSTER SCHRITT

Working Session mit NXTGN

Wir übersetzen das Thema in Rollen, Entscheidungswege und einen realistischen Implementierungsstart. info@nxtgn.de · +49 9181 5118073