AI Governance Framework for Financial Institutions

Führungsrahmen für AI Governance in Finanzinstituten

Dokumenttyp: Führungsrahmen für AI Governance in Finanzinstituten

Version: V2.0 · Status: Published Version · Datum: 12.07.2026

Owner: Martin Bischoff · 4T2 Solutions GmbH

Public · Frei verfügbares Rahmenwerk

Dokumentkontrolle

Feld	Wert
Dokumentname	AI Governance Framework for Financial Institutions
Version	V2.0
Status	Published Version
Vertraulichkeit	Public · Frei verfügbares Rahmenwerk
Owner	Martin Bischoff
Review-Zweck	Publizierte Version

Inhaltsverzeichnis

Executive Summary

Artificial Intelligence (AI) verändert zunehmend, wie das Institut mit Kundinnen und Kunden interagiert, Wissen nutzt, Software entwickelt und Entscheidungen vorbereitet. Damit ist AI nicht nur ein Technologiethema, sondern eine Führungsaufgabe. Das AI-Governance-Modell schafft den Rahmen, um den geschäftlichen Nutzen von AI gezielt zu realisieren und gleichzeitig Verantwortung, Sicherheit, Nachvollziehbarkeit und regulatorische Konformität sicherzustellen.

Das Modell verbindet unternehmerische Steuerung mit einer verhältnismässigen Kontrolle. Jede relevante AI-Nutzung wird einem klaren Geschäftszweck und einem verantwortlichen Owner zugeordnet, risikobasiert eingestuft und vor ihrem Einsatz nachvollziehbar freigegeben. Massgeblich ist dabei der tatsächliche Einfluss auf Kundinnen und Kunden, Mitarbeitende, Geschäftsprozesse und kritische Funktionen: Je grösser dieser Einfluss, desto höher die Anforderungen an Kontrolle, Dokumentation und laufende Überwachung.

Für die Geschäftsleitung entsteht dadurch eine konsolidierte Sicht auf das AI-Portfolio. Sie kann Investitionen und Prioritäten am erwarteten Nutzen ausrichten, Risiken frühzeitig erkennen und wirksam adressieren sowie AI-Nutzungen stoppen, deren Einsatz nicht verantwortbar oder nicht verhältnismässig ist. Das AI Governance Framework definiert die Führungs- und Entscheidungslogik. Die drei zugehörigen Dokumente konkretisieren diese Vorgaben für Kontrollen, technische Mindestanforderungen und den operativen Betrieb.

Executive Management View

Teil I: Das Fundament

Wirksame AI Governance beginnt mit einem gemeinsamen Verständnis. Bevor Verantwortlichkeiten, Entscheidungswege und Kontrollen festgelegt werden, muss klar sein, welche AI-Nutzungen vom Framework erfasst werden und wie die zugehörigen Governance-Dokumente zusammenwirken. Dieses Fundament schafft eine einheitliche Grundlage für Entscheidungen über den gesamten AI-Lebenszyklus.

1. Zweck, Zielbild und Einordnung

Das Framework richtet sich an Finanzinstitute im europäischen Regulierungsumfeld. Es berücksichtigt die Anforderungen der deutschen Bankenaufsicht und des europäischen Rechtsrahmens.

Der primäre Geltungsbereich sind Finanzinstitute mit Rechtseinheiten oder AI-Nutzung in Deutschland. Die Dokumentenfamilie ist auf eine vollständige und prüfbare Operationalisierung der anwendbaren Deployerpflichten ausgelegt. Ob dieses Ziel erreicht ist, wird erst durch das genehmigte Applicability Record und die tatsächliche Evidence des einzelnen AI Use Case belegt. Eine Vollständigkeitsfeststellung gilt nur für die dokumentierte Rechtseinheit, das Einsatzland, das AI-System und seine Version, die Operatorrolle, die Risikoklasse und die ausgelösten Pflichten. Sie ist keine pauschale Konformitätsaussage für ein Institut oder den gesamten AI Act.

AI wird im Institut eingesetzt, um Dienstleistungen, Prozesse und Entscheidungen zu unterstützen. Ihr Einsatz kann Vorteile schaffen, aber auch Kundinnen und Kunden, Mitarbeitende und den Geschäftsbetrieb unmittelbar beeinflussen. Deshalb muss das Institut jederzeit wissen, wo AI eingesetzt wird, welchem Zweck sie dient und wer dafür verantwortlich ist.

Das Framework macht diese Verantwortung verbindlich. Es sorgt dafür, dass geschäftlicher Nutzen, Risiken und Kontrollanforderungen gemeinsam beurteilt werden. Entscheidungen über AI werden damit nicht isoliert getroffen, sondern im Rahmen der bestehenden Unternehmenssteuerung.

Jede relevante AI-Nutzung benötigt einen klaren Zweck, einen verantwortlichen Owner und eine dokumentierte Freigabe. Die Intensität der Kontrolle richtet sich nach der Bedeutung und dem möglichen Einfluss der Nutzung. Kapitel 5 beschreibt, wie dieser Grundsatz angewendet wird.

Das Framework ersetzt keine bestehenden Zuständigkeiten. Es ergänzt die heutige Governance nur dort, wo AI besondere Anforderungen schafft, die mit den bestehenden Verfahren nicht ausreichend abgedeckt werden.

2. Geltungsbereich

Das AI Framework gilt für jede AI-Nutzung, für die das Institut Verantwortung trägt. Massgeblich ist nicht die Bezeichnung eines Produkts oder einer Plattform, sondern ihre tatsächliche Wirkung auf Daten, Prozesse, Entscheidungen oder den Geschäftsbetrieb.

Erfasst sind sowohl selbst entwickelte als auch eingekaufte oder extern betriebene AI-Systeme und

AI-Funktionen. Dazu gehören Funktionen in Fachanwendungen und Cloud-Diensten ebenso wie generative Assistenten, automatisierte Entscheidungsunterstützung und agentenbasierte Systeme. Der Geltungsbereich schliesst Pilotierungen, Proof of Concepts und interne Experimente ein, sobald sie Daten, Prozesse oder technische Umgebungen des Instituts berühren.

Auch Shadow AI fällt in den Geltungsbereich. Nicht gemeldete oder dezentral eingeführte AI-Nutzungen können erhebliche Risiken verursachen, weil sie der etablierten Steuerung entzogen sind. Das Framework soll solche Nutzungen deshalb früh sichtbar machen und in einen geregelten Prozess überführen. Bis dahin begrenzt das Institut erkannte Risiken mit angemessenen technischen und organisatorischen Massnahmen.

Der AI Governance Lifecycle beginnt, sobald eine AI-Idee oder ein möglicher Einsatz erkennbar wird. Die erste Orientierung schafft Transparenz und führt konkrete AI-Vorhaben in den formalen Intake. Der Lifecycle endet erst mit der kontrollierten Stilllegung. Auch ein Pilotprojekt oder eine ausschliesslich interne Nutzung bleibt damit governancepflichtig.

Eigenentwicklung, Entwicklung im Auftrag und Inbetriebnahme unter dem Namen oder der Marke des Instituts können unmittelbar eine Providerrolle begründen. Rebranding, eine wesentliche Änderung oder eine geänderte Zweckbestimmung können zusätzlich einen Rollenwechsel nach Art. 25 AI Act auslösen. Ein positiver Provider-Aktivierungsentscheid blockiert eine reine Deployer-Freigabe, bis ein Provider Compliance Plan genehmigt ist. Importeur-, Distributor-, Annex-I-Produkthersteller-, eigene GPAI-Modellprovider- und formelle Real-World-Testing-Signale führen zu Stop und spezialisiertem Legal-/Compliance- Handoff. Das Framework baut dafür bewusst kein allgemeines Herstellerhandbuch auf.

Nicht jede AI-nahe Funktion erfüllt später die Definition eines AI-Systems im Sinne des AI Acts.[1] Solange diese Einordnung offen ist, behandelt das Institut Funktionen mit möglicher Wirkung auf Daten, Prozesse oder Entscheidungen als prüfpflichtig. So wird verhindert, dass relevante AI-Nutzungen ohne angemessene Bewertung produktiv eingesetzt werden.

Das Framework verwendet vier Begriffe mit unterschiedlicher Bedeutung. **AI-Nutzung** bezeichnet den beabsichtigten oder tatsächlichen fachlichen Einsatz von AI für einen bestimmten Zweck. Der **AI Use Case** ist die konkret abgegrenzte AI-Nutzung, die das Institut ab dem formalen Intake als Governance-Gegenstand führt. Ihr Eintrag im AI Register verbindet Zweck, Owner, Einstufung, Freigaben und Evidenzen. **AI-System** bezeichnet das technische System und, wo anwendbar, das rechtlich einzuordnende System nach AI Act. Ein **AI-Vorhaben** ist die zeitlich begrenzte Initiative, mit der eine AI-Nutzung eingeführt oder wesentlich geändert wird. Die AI-Idee löst die Orientierung aus, bildet aber weder einen eigenen Governance-Gegenstand noch eine Lifecycle-Phase.

3. AI Framework und Dokumentenfamilie

Das Framework beschreibt die Führungslogik für AI. Es legt fest, welche Entscheidungen das Institut

1 Der Begriff folgt Art. 3 № 1 der Verordnung (EU) 2024/1689. Entscheidend sind insbesondere Maschinenbasiertheit, ein gewisser Grad an Autonomie und die Fähigkeit, aus Eingaben Ausgaben abzuleiten, die physische oder virtuelle Umgebungen beeinflussen können.

treffen muss, wer dafür verantwortlich ist und wie diese Entscheidungen über den AI Governance Lifecycle hinweg verbunden bleiben.

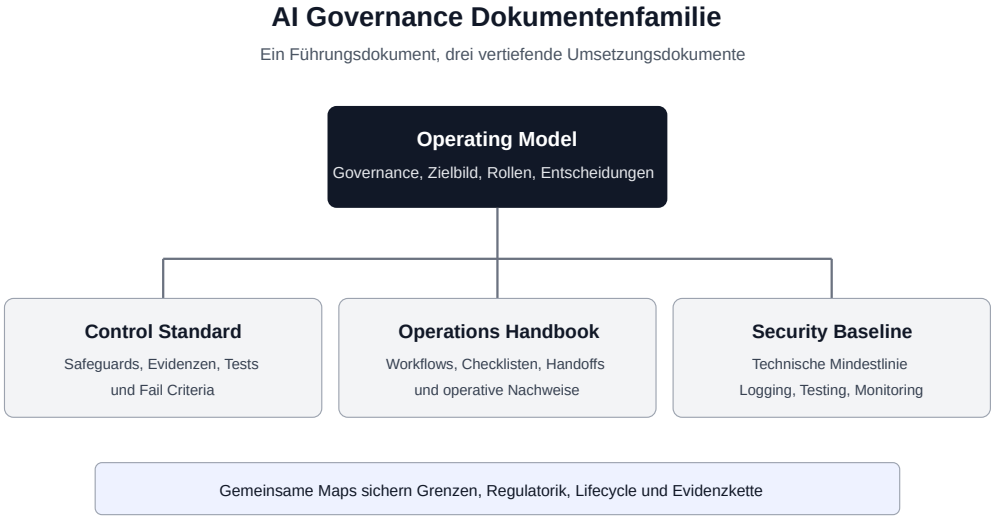

Abbildung 1: Die Dokumentenfamilie trennt Governance-Modell, prüfbare Safeguards, operative Workflows und technische Mindestanforderungen.

Die vier Dokumente werden als Familie geführt. Ändert sich die Führungslogik, prüft das Institut die Auswirkungen auf Abläufe, Kontrollen und technische Mindestanforderungen. Erkenntnisse aus Betrieb, Prüfungen oder regulatorischen Änderungen wirken umgekehrt auf das Governance-Modell zurück.

3.1 AI Framework und Governance-Modell

Das AI Framework und Governance-Modell legt die Führungs- und Entscheidungslogik für AI fest. Es verbindet Grundsätze, Verantwortlichkeiten, Entscheidungswege und den AI Governance Lifecycle. Im Mittelpunkt steht die Frage, welche Entscheidungen das Institut treffen muss, um AI verantwortbar und wirtschaftlich einzusetzen.

3.2 AI Control Standard

Der AI Control Standard übersetzt die Governance-Vorgaben in prüfbare Anforderungen. Er definiert die erforderlichen Safeguards, Evidenzen und Prüfkriterien für unterschiedliche Risikoprofile. First Line, Second Line, Interne Revision und externe Prüfstellen können damit beurteilen, ob eine Governance-Entscheidung angemessen umgesetzt wurde.

3.3 AI Operations Handbook

Das AI Operations Handbook führt die Beteiligten durch die operative Umsetzung. Es beschreibt vom ersten erkennbaren Einsatz bis zur Stilllegung, wer welchen Schritt ausführt, welche Entscheidung erforderlich ist und welcher Nachweis entsteht. Damit werden die Vorgaben im täglichen Betrieb einheitlich anwendbar.

3.4 AI Security Baseline

Die AI Security Baseline definiert den technischen Mindestzustand für AI-Plattformen, Modelle und Integrationen. Sie macht die Sicherheitsziele des Governance-Modells für Architektur, Identitäten, Datenflüsse, Überwachung und AI-spezifische Bedrohungen prüfbar. Damit zeigt sie, unter welchen technischen Voraussetzungen ein sicherer und resilienter Betrieb möglich ist.

Teil II: Das Governance-Modell

Teil II beschreibt, wie das Institut AI führt. Er verbindet Grundsätze, Verantwortlichkeiten und Entscheidungswege zu einem Modell, das über den gesamten Lebenszyklus trägt. Der Fokus bleibt auf den Entscheidungen des Managements; Abläufe, Kontrollen und technische Anforderungen werden in den zugehörigen Dokumenten konkretisiert.

4. Überblick über das Governance-Modell

Das Governance-Modell bettet AI in die bestehende Unternehmenssteuerung ein. Es verbindet Verantwortung, Risikoeinstufung, Freigaben und Überwachung so, dass eine AI-Nutzung nicht als isoliertes Technologievorhaben behandelt wird.

Sein verbindendes Element ist der AI Governance Lifecycle. Er führt das Institut von der ersten Orientierung bis zur kontrollierten Stilllegung durch eine Folge klarer Entscheidungen. Fachbereiche, IT und Kontrollfunktionen bringen ihre jeweilige Verantwortung in diesen gemeinsamen Prozess ein.

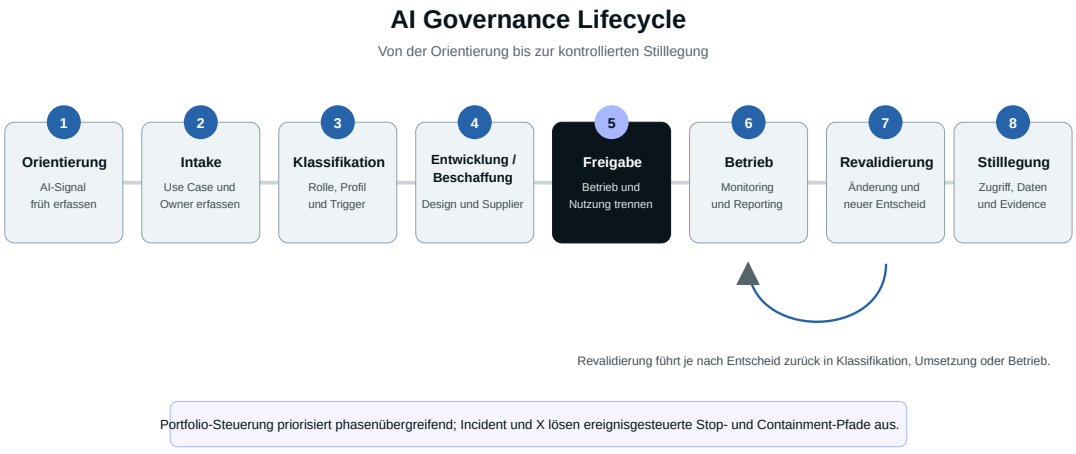

Abbildung 2: Das Governance-Modell führt den AI Governance Lifecycle und verweist Detailtiefe an die passenden Familiendokumente.

Die Tiefe der Governance richtet sich nach der tatsächlichen Wirkung, nicht nach dem technischen Etikett eines Systems. Eine AI-Nutzung für allgemeine interne Wissensarbeit benötigt weniger Steuerung als ein System, das Kundinnen und Kunden beeinflusst, regulatorische Pflichten auslöst oder kritische Geschäftsprozesse unterstützt.

Das Governance-Modell verfolgt dabei vier Grundziele:

- AI-Nutzung frühzeitig sichtbar machen.
- Verantwortlichkeiten eindeutig zuordnen.

- Entscheidungen nachvollziehbar dokumentieren.
- Risiken während des gesamten Lebenszyklus steuern.

Diese vier Ziele schaffen einen konsistenten Entscheidungsrahmen. Risikoappetit, Rollen, Freigaben und die weiteren Governance-Disziplinen richten sich an ihm aus. So kann das Institut Innovation ermöglichen, ohne Transparenz, Kontrollfähigkeit oder regulatorische Konformität preiszugeben.

5. Leitprinzipien, Risikoappetit und Portfolio-Steuerung

Ein gemeinsamer Risikoappetit übersetzt die strategische Haltung des Instituts in konkrete Entscheidungen. Er sorgt dafür, dass vergleichbare AI-Nutzungen nach denselben Grundsätzen beurteilt werden und Managementaufmerksamkeit dort eingesetzt wird, wo die Wirkung am grössten ist.

Verantwortung bleibt beim Institut

AI unterstützt Entscheidungen, übernimmt jedoch keine institutionelle Verantwortung. Für jede relevante AI-Nutzung muss deshalb eindeutig festgelegt sein, wer den fachlichen Zweck verantwortet, wer den technischen Betrieb sicherstellt und wer regulatorische sowie sicherheitsrelevante Fragestellungen beurteilt.

Governance ersetzt bestehende Verantwortlichkeiten nicht. Sie verbindet die bestehenden Funktionen zu einem gemeinsamen Entscheidungsmodell.

Die Kontrolltiefe richtet sich nach der Wirkung

Nicht jede AI-Nutzung benötigt denselben Umfang an Governance. Die erforderliche Kontrolltiefe ergibt sich aus den möglichen Auswirkungen auf Kunden, Mitarbeitende, Geschäftsprozesse, regulatorische Verpflichtungen und kritische Funktionen.

Eine interne Unterstützung für allgemeine Wissensarbeit stellt andere Anforderungen als ein AI-System, das Kreditentscheidungen vorbereitet, Kunden berät oder operative Prozesse automatisiert.

Freigaben sind keine einmaligen Entscheidungen

Governance endet nicht mit einer Produktivsetzung. Änderungen an Modellen, Daten, Prompts, Tool-Rechten, Integrationen oder regulatorischen Anforderungen können eine Neubewertung erforderlich machen. Jede Freigabe muss deshalb während des gesamten Lebenszyklus überprüfbar und bei Bedarf anpassbar bleiben.

AI Risk Appetite

Der AI Risk Appetite beschreibt den Rahmen, innerhalb dessen AI genutzt werden soll. Er legt fest, welche Nutzungen aktiv gefördert werden, welche zusätzliche Kontrollen benötigen und welche grundsätzlich ausgeschlossen bleiben.

Für die Portfolio-Steuerung genügt eine einfache Dreiteilung.

Förderzone

AI-Nutzungen mit geringer Wirkung und überschaubarem Risiko sollen möglichst einfach eingeführt werden können. Voraussetzung sind freigegebene Plattformen sowie definierte Basiskontrollen. Ziel

ist es, Innovation und Produktivität nicht unnötig zu behindern.

Kontrollierte Zone

AI Use Cases mit vertraulichen Daten, Kundenbezug, erheblicher Entscheidungswirkung, kritischen Prozessen oder weitreichenden Systemrechten bleiben grundsätzlich zulässig. Sie erfordern jedoch eine ihrem Risiko angemessene Governance, nachvollziehbare Freigaben und zusätzliche Managementaufmerksamkeit.

Stop-Zone

Nutzungen, die gegen gesetzliche Vorgaben, interne Grundsätze oder den festgelegten Risikoappetit verstossen, sind nicht freigabefähig. Dazu gehören insbesondere verbotene Praktiken nach AI Act sowie nicht autorisierte AI-Nutzung mit Kunden-, Beschäftigten- oder anderen besonders schützenswerten Daten.

Der Risikoappetit wird mindestens jährlich sowie bei wesentlichen regulatorischen, technologischen oder strategischen Änderungen überprüft.

Portfolio-Steuerung

Governance bewertet nicht nur Risiken. Sie unterstützt auch Investitions- und Priorisierungsentscheidungen.

Neben regulatorischen Anforderungen berücksichtigt die Portfolio-Steuerung insbesondere den erwarteten Geschäftsnutzen, den tatsächlichen Wertbeitrag, die Betriebs- und Kontrollkosten, technische Abhängigkeiten sowie das verbleibende Risiko.

Dadurch entsteht für die Geschäftsleitung eine belastbare Grundlage, um AI-Initiativen gezielt auszubauen, zu konsolidieren oder einzustellen.

Abweichungen vom definierten Risikoappetit sind als Risk Acceptance zu behandeln. Sie benötigen eine nachvollziehbare Begründung, einen verantwortlichen Owner, geeignete Kompensationsmassnahmen sowie eine zeitliche Befristung.

6. Rollen, Verantwortlichkeiten und Gremien

Wirksame AI Governance braucht sichtbare Verantwortung. Für jede wesentliche Entscheidung muss klar sein, wer sie vorbereitet, wer sie trifft und wer ihre Umsetzung überwacht. Die Rollen ergänzen bestehende Zuständigkeiten; sie schaffen keine parallele Organisation für AI.

Rolle	Governance-Verantwortung
Geschäftsleitung	Legt AI Risk Appetite fest, entscheidet über strategische Plattformen und kritische AI-Nutzung sowie über die erforderlichen Ressourcen.
AI Owner	Verantwortet Zweck, Wirkung, Risikoeinstufung, fachliche Freigabe und die Steuerung des gesamten AI-Lebenszyklus.
System Owner / IT Operations	Verantwortet Betrieb, Verfügbarkeit, Zugriffsschutz, Logging, Monitoring, Change Management und Stilllegung der technischen Umgebung.

Rolle	Governance-Verantwortung
Fachbereich	Verantwortet Nutzen, Prozesswirkung, fachliche Kontrollen und Human Oversight.
Chief Information Security Officer (CISO) / Security	Bewertet Sicherheitsarchitektur, Plattformen, technische Risiken, Monitoring und Incident-Fähigkeit.
Compliance / Legal	Bewertet regulatorische Anforderungen, Vertragsfragen sowie Transparenz- und Haftungsthemen.
Datenschutz	Bewertet Rechtsgrundlagen, Datenschutz-Folgenabschätzung, Betroffenenrechte und datenschutzrechtliche Risiken.
Risk Management	Verankert AI-Risiken im unternehmensweiten Risikomanagement und begleitet wesentliche Risikoentscheidungen.
Data Owner	Verantwortet Datenquellen, Datenqualität, Klassifikation und zulässige Datennutzung.
Enterprise Architecture	Stellt die Einbindung in Zielarchitektur und Plattformstrategie sicher.
Einkauf / Auslagerungsmanagement	Verantwortet Lieferantensteuerung, Vertragsanforderungen und Drittrisiken.
Interne Revision	Prüft Design und Wirksamkeit der AI Governance unabhängig.

Der AI Owner und der Data Owner übernehmen unterschiedliche Verantwortlichkeiten. Während der AI Owner den fachlichen Einsatz und dessen Wirkung verantwortet, stellt der Data Owner sicher, dass Datenquellen, Qualität und Nutzung den festgelegten Anforderungen entsprechen.

Das etablierte Three-Lines-Modell bleibt auch für AI die organisatorische Grundlage. Die First Line verantwortet Nutzung und Betrieb, die Second Line bewertet Risiken und überwacht die Einhaltung der Governance-Vorgaben, während die Third Line die Wirksamkeit der Governance unabhängig beurteilt.

Bestehende Gremien sollen nach Möglichkeit genutzt werden. Neue AI-Gremien sind nur dann erforderlich, wenn vorhandene Entscheidungsstrukturen AI-Risiken nicht angemessen bewerten oder koordinieren können.

Kritische Abweichungen, ungeklärte Risikoeinstufungen oder schwerwiegende Verstösse gegen Governance-Vorgaben dürfen nicht innerhalb einzelner Projekte verbleiben. Sie sind entsprechend der definierten Eskalationswege an die zuständigen Governance-Funktionen und gegebenenfalls an die Geschäftsleitung weiterzuleiten.

7. AI Governance Lifecycle

Der AI Governance Lifecycle macht aus einer erkennbaren AI-Idee einen gesteuerten Lebenszyklus. Er verbindet die bestehenden Projekt-, Beschaffungs- und Betriebsprozesse mit den Entscheidungen, die für einen verantwortbaren Einsatz von AI erforderlich sind.

Der Lifecycle ist deshalb kein zusätzlicher Projektprozess. Er stellt sicher, dass Verantwortung, Risiko und Freigabestatus an den entscheidenden Übergängen geklärt sind. Governance begleitet das AI-Vorhaben von Beginn an und wird nicht erst vor der Produktivsetzung als Dokumentationspflicht ergänzt.

Orientierung

Die Orientierung beginnt, sobald eine AI-Idee, ein möglicher Einsatz oder eine bereits bestehende Nutzung erkennbar wird. Das Institut klärt zunächst Zweck, erwarteten Nutzen und mögliche Auswirkungen. Diese frühe Sichtbarkeit verhindert, dass AI-Vorhaben ausserhalb der Governance heranreifen.

Auf dieser Grundlage entscheidet das Institut, ob ein AI-Vorhaben weiterverfolgt werden soll und welcher Governance-Pfad erforderlich ist.

Intake und Klassifikation

Der formale Intake beginnt, sobald ein konkretes AI-Vorhaben oder eine bestehende AI-Nutzung beurteilt werden kann. Das Institut ordnet den verantwortlichen Owner und die betroffene Rechtseinheit zu, eröffnet den Eintrag im AI Register und beginnt mit der formalen Nachweisführung. Die anschliessende Klassifikation bestimmt die regulatorische Rolle, den anwendbaren Pfad und die erforderliche Kontrolltiefe.

Entwicklung oder Beschaffung

Während der Entwicklung oder Beschaffung werden Architektur, Daten, Lieferanten, Sicherheitsanforderungen und regulatorische Rahmenbedingungen konkretisiert. Gleichzeitig entsteht die Grundlage für die spätere Freigabe.

Ziel dieser Phase ist es, ein AI-System bereitzustellen, dessen Risiken verstanden und angemessen beherrscht werden können.

Freigabe und produktiver Einsatz

Vor der Produktivsetzung wird beurteilt, ob die technischen, organisatorischen und fachlichen Voraussetzungen für einen sicheren Betrieb erfüllt sind. Dabei werden Betriebs- und Nutzungsfreigabe bewusst getrennt betrachtet.

Mit der Freigabe beginnt der produktive Betrieb. Die bereits mit der Orientierung begonnene Governance wird durch Monitoring, Reporting und laufende Überwachung fortgeführt, damit die ursprünglichen Annahmen auch im Betrieb weiterhin gelten.

Laufende Steuerung

AI-Systeme verändern sich im Laufe ihres Betriebs. Modelle werden aktualisiert, Datenquellen erweitert, Anwendungsbereiche angepasst oder regulatorische Anforderungen verändert.

Governance stellt deshalb sicher, dass wesentliche Änderungen erkannt und erneut bewertet werden. Entscheidungen werden nicht dauerhaft konserviert, sondern kontinuierlich anhand der tatsächlichen Nutzung überprüft.

Revalidierung

Eine Revalidierung wird erforderlich, wenn sich das Risikoprofil oder die Wirkung einer AI-Nutzung wesentlich verändert.

Typische Auslöser sind:

- Änderungen an Modellen oder Modellversionen,
- neue oder geänderte Datenquellen,
- Anpassungen von Prompts oder Tool-Rechten,
- Erweiterungen des fachlichen Einsatzbereichs,
- neue Nutzergruppen,
- regulatorische Änderungen,
- Sicherheitsvorfälle,
- Findings aus Audits oder Kontrollen,
- wesentliche Änderungen bei Lieferanten.

Die Revalidierung stellt sicher, dass frühere Freigaben auch unter veränderten Rahmenbedingungen weiterhin tragfähig bleiben.

Stilllegung

Zur kontrollierten Stilllegung gehören:

- Beendigung der Nutzung,
- Entzug von Zugriffsrechten,
- Archivierung oder Löschung relevanter Daten,
- Behandlung von Modellartefakten,
- Abschluss regulatorischer Nachweise,
- Anpassung von Lieferantenbeziehungen,
- Aktualisierung aller Register.

Dadurch bleibt der gesamte Lebenszyklus einer AI-Nutzung nachvollziehbar.

8. Betriebsfreigabe und Nutzungsfreigabe

Das Institut trennt die technische Betriebsfähigkeit vom erlaubten fachlichen Einsatz. Eine sichere Plattform ist nicht automatisch für jeden Zweck und jede Datenkategorie geeignet. Umgekehrt kann ein fachlich sinnvoller AI Use Case nicht freigegeben werden, wenn die technische Umgebung seine Anforderungen nicht erfüllt. Erst Betriebsfreigabe und Nutzungsfreigabe zusammen schaffen eine tragfähige Grundlage für den produktiven Einsatz.

Betriebsfreigabe

Die Betriebsfreigabe beantwortet die Frage:

Kann diese Plattform oder dieses AI-System unter definierten Rahmenbedingungen sicher betrieben werden?

Im Mittelpunkt stehen technische und organisatorische Voraussetzungen. Dazu gehören insbesondere Architektur, Identitäts- und Berechtigungsmanagement, Datenhaltung, Logging, Monitoring, Change Management, Incident-Fähigkeit, Lieferantensteuerung, technische Evidenz sowie Resilienz und Fallback-Konzepte.

Die Betriebsfreigabe bestätigt, dass eine Plattform grundsätzlich kontrolliert betrieben werden kann. Sie trifft jedoch keine Aussage darüber, für welchen fachlichen Zweck sie eingesetzt werden darf.

Nutzungsfreigabe

Die Nutzungsfreigabe beantwortet die Frage:

Darf dieser konkrete AI Use Case mit diesen Daten, diesem Nutzerkreis und dieser Wirkung betrieben werden?

Hier stehen die fachlichen Auswirkungen im Vordergrund. Bewertet werden insbesondere Zweck der Nutzung, Prozesswirkung, Datenkategorien, Zielgruppen, regulatorische Anforderungen, Human Oversight, Verantwortlichkeiten sowie Auswirkungen auf Kunden oder Mitarbeitende.

Die Nutzungsfreigabe bezieht sich immer auf einen konkreten Anwendungsfall und kann nicht pauschal auf andere Einsatzbereiche übertragen werden.

Beispiel	Betriebsfreigabe	Nutzungsfreigabe	Ergebnis
Interne AI-Plattform für Wissensarbeit	Plattform erfüllt technische Anforderungen.	Zusammenfassung interner Richtlinien zulässig.	Betriebs- und Nutzungsfreigabe möglich.
Nutzung derselben Plattform mit Kundendaten	Plattform unverändert.	Neuer Datenschutz- und Verwendungszweck.	Neue Nutzungsfreigabe erforderlich.
Externer Kundenchatbot	Plattform und Supplier technisch geeignet.	Kundenkommunikation und Transparenz müssen separat bewertet werden.	Getrennte Freigaben erforderlich.
AI-Agent mit Administrations- oder Zahlungsrechten	Plattform erfüllt Sicherheitsanforderungen.	Fachliche Wirkung und Missbrauchsrisiko erfordern zusätzliche Kontrollen.	Erweiterte Governance und Human Oversight erforderlich.

Änderungen der Nutzung

Eine neue Freigabe ist nicht nur bei der Einführung einer neuen Plattform erforderlich.

Auch Änderungen an Zweck, Daten, Nutzerkreis, Prozessintegration, Tool-Rechten, Modell, Lieferanten oder Zielsystemen können das Risikoprofil wesentlich verändern und eine erneute Bewertung erforderlich machen.

Human Oversight

Menschliche Aufsicht ist nur wirksam, wenn verantwortliche Personen Entscheidungen tatsächlich beeinflussen können.

Dazu gehört insbesondere die Fähigkeit,

- Ergebnisse kritisch zu hinterfragen,
- Eingriffe vorzunehmen,
- AI-Ausgaben zu verwerfen,
- Prozesse zu unterbrechen und
- den Betrieb bei Bedarf einzuschränken oder zu stoppen.

Ein Mensch im Prozess erfüllt die Anforderungen an Human Oversight nur dann, wenn er über die erforderlichen Kompetenzen, Befugnisse und Handlungsmöglichkeiten verfügt.

Fallback und Degraded Mode

Für AI-Nutzungen mit wesentlichen Auswirkungen auf Geschäftsprozesse muss jederzeit nachvollziehbar sein, wie der Betrieb ohne AI fortgeführt werden kann.

Der Fachbereich definiert deshalb bereits vor der Produktivsetzung den alternativen Prozess, Verantwortlichkeiten, Auslösekriterien, Entscheidungswege sowie Voraussetzungen für die Wiederaufnahme des regulären Betriebs.

Bei kritischen oder wichtigen Funktionen verweist die Freigabe auf die Business Impact Analysis, die genehmigten Wiederanlauf- und Datenverlustziele sowie einen getesteten Wiederanlaufplan. Bestehende DORA- und Business-Continuity-Evidence wird genutzt, statt eine parallele AI-Resilienzakte aufzubauen.

Damit bleibt das Institut auch bei Störungen, Sicherheitsvorfällen oder regulatorischen Einschränkungen jederzeit handlungsfähig.

Teil III: Governance-Disziplinen

AI-Risiken lassen sich nicht entlang organisatorischer Grenzen beurteilen. Datenschutz, Informationssicherheit, Regulatorik, Modellrisiken, Lieferantensteuerung und menschliche Aufsicht beeinflussen dieselbe Nutzung aus unterschiedlichen Perspektiven. Teil III führt diese Perspektiven zu einer gemeinsamen Entscheidungsgrundlage zusammen.

Dieser Teil beschreibt die fachlichen Leitplanken, innerhalb derer AI genutzt werden darf. Er ersetzt weder regulatorische Detailvorgaben noch technische Standards oder operative Arbeitsanweisungen. Sein Ziel ist es, Managemententscheidungen auf eine einheitliche Grundlage zu stellen und die zuständigen Fachfunktionen entlang des gesamten AI-Lebenszyklus miteinander zu verbinden.

9. Regulatorik und Risikoklassifikation

Die Klassifikation beantwortet eine Führungsfrage: Welcher Steuerungspfad ist für diese AI-Nutzung erforderlich? AI Act, Digital Operational Resilience Act (DORA) und Datenschutz-Grundverordnung (DSGVO) bilden dafür den Kern. Bankaufsichts-, Arbeits- und Sicherheitsrecht ergänzen ihn, wenn der konkrete Einsatz dies auslöst.[234]

Zuerst wird die Rolle bestimmt. Rechtseinheit, Einsatzland und Betriebsmodell entscheiden, ob das Institut als Deployer handelt oder einen anderen Rollenpfad aktivieren muss. Rebranding, eine wesentliche Änderung oder ein neuer Verwendungszweck können die Providerrolle auslösen. In diesem Fall blockiert der Provider-Aktivierungsentscheid die Freigabe im Deployer-Pfad und übergibt

2 Verordnung (EU) 2024/1689 (AI Act), insbesondere Art. 2-6, 9-27, 43, 47-50, 72-73 und 86 sowie Anhänge I, III, IV, VI-VIII.

3 Verordnung (EU) 2022/2554 (DORA), insbesondere Art. 5-14 zu ICT-Risikomanagement sowie Art. 28-30 zu ICT-Drittparteirisiken; die verbindlichen Templates des Informationsregisters stehen in der Durchführungsverordnung (EU) 2024/2956.

4 Verordnung (EU) 2016/679 (DSGVO), insbesondere Art. 5, 6, 22, 25, 28, 32 und 35.

an den Provider Compliance Plan.

Danach wird die Wirkung eingeordnet. Verbotene Praktiken erhalten den Status X und bleiben ausgeschlossen.[5] Die Hochrisikoprüfung beginnt bei Anhang I und führt anschliessend durch die einschlägigen Nummern von Anhang III. Für Finanzinstitute sind vor allem Beschäftigung, Kreditwürdigkeitsprüfung sowie Risiko- und Preisbewertung in der Lebens- und Krankenversicherung relevant. Ein Ausnahmeentscheid nach Art. 6 Abs. 3 braucht eine substanzielle, versionierte Begründung.

Schliesslich werden die Folgepflichten aktiviert. Profilklasse und gesetzliche Trigger bestimmen Freigabe, Kontrolltiefe und Evidence. FRIA, EU-Datenbankregistrierung, Transparenz, Erläuterungsrecht und DORA-Eintrag bleiben eigenständige Entscheidungen. Keine dieser Prüfungen ersetzt eine andere.[6]

Klassifikationsergebnis	Governance-Folge
Deployerrolle und zulässige Nutzung	risikobasierter Freigabe- und Kontrollpfad
Hochrisiko- oder vergleichbar kritische Wirkung	erhöhte Kontrolle, Managementsicht und vollständige Triggerprüfung
Providerrolle oder anderer Spezialfall	Freigabestopp und verbindlicher Specialist Handoff
Verbotene oder nicht vertretbare Nutzung	Ausschlussstatus X, technische Blockierung und keine Risk Acceptance

Der Regulatory Watch hält diese Entscheidungslogik aktuell. Er dokumentiert Quelle, Rechtsstatus, betroffene Controls, Owner und Umsetzungsfrist. Erwartete Änderungen wie der Digital Omnibus on AI werden erst nach Prüfung von Veröffentlichung, Inkrafttreten und Anwendungszeitpunkt verbindlich in die Klassifikation übernommen.[7]

10. Model Risk Governance

Sobald AI ein bankfachliches Modell oder die Risikosteuerung beeinflusst, müssen AI- und Modellrisiko gemeinsam beurteilt werden. Massgeblich ist die fachliche Wirkung, nicht die verwendete Technologie. Die etablierte Modellsteuerung greift deshalb nicht bei jedem AI-Assistenten, wohl aber dort, wo ein System Scores, Empfehlungen oder Entscheidungen mit wesentlicher Wirkung erzeugt.

---

5 Art. 5 AI Act untersagt die dort genannten Praktiken. Eine interne Risk Acceptance kann ein gesetzliches Verbot nicht aufheben.

6 Art. 49 AI Act enthält keine pauschale Registrierungspflicht für private Finanzinstitute als Deployer. Providerpflichten und die besondere Deployerpflicht für öffentliche Stellen oder in deren Auftrag handelnde Personen werden rollen- und einsatzbezogen geprüft. Eine FRIA-Pflicht nach Art. 27 wird davon unabhängig beurteilt.

7 Der Rat der EU hat den Digital Omnibus on AI am 29. Juni 2026 endgültig angenommen. Die Europäische Kommission nennt nach der politischen Einigung den 2. Dezember 2027 für eigenständige Hochrisiko-Systeme und den 2. August 2028 für produktintegrierte Hochrisiko-Systeme. Der Regulatory Watch prüft Veröffentlichung, Inkrafttreten und spätere Leitlinien vor jeder verbindlichen Anwendung.

Model Risk Governance beantwortet im AI-Kontext vor allem fünf Fragen:

- Ist das AI-System ein bankfachlich relevantes Modell oder Bestandteil eines solchen Modells?
- Welche Wirkung haben Output, Score, Empfehlung oder Priorisierung auf Entscheidungen?
- Wie werden Performance, Drift, Bias, Robustheit und Overrides überwacht?
- Welche Modellversion, Datenbasis und Validierung galten zu einem bestimmten Zeitpunkt?
- Wann muss ein Modell eingeschränkt, revalidiert oder stillgelegt werden?

Das AI Register und das Modellinventar dürfen nicht widersprüchlich geführt werden. Das AI Register dokumentiert AI-spezifische Rollen, Nutzung, Betrieb, Freigaben und Kontrollen. Das Modellinventar dokumentiert bankfachlich oder aufsichtlich relevante Modelle, ihre Validierung, Performance, Drift und Stilllegung. Nicht jeder AI Use Case gehört automatisch in das Modellinventar. Die Verknüpfung wird erforderlich, wenn ein System Score- oder Entscheidungswirkung entfaltet, ein regulatorisches internes Modell berührt oder die institutsinternen Model-Risk-Kriterien erfüllt.

Kritische Modelle benötigen angemessene Validierung, Monitoring, Challenger-Ansätze oder Benchmarking, Drift-Auswertung, Bias- und Fairness-Prüfung, Override-Analyse und dokumentierte Stilllegung. Aus diesen Ergebnissen folgt die Managemententscheidung, ob Nutzung, Einschränkungen und Restrisiko weiterhin tragfähig sind.

11. Datenschutz, Vertraulichkeit und Beschäftigtenbezug

Datenschutz beginnt mit dem Zweck einer AI-Nutzung und nicht mit einer nachgelagerten Freigabeprüfung. Wie das System Daten erhält, kombiniert, speichert und in Ausgaben sichtbar macht, bestimmt einen wesentlichen Teil seines Risikos.

Bei personenbezogenen Daten sind Rechtsgrundlage, Zweckbindung und Datenminimierung zu prüfen. Transparenz, Betroffenenrechte und Art. 22 DSGVO gehören ebenfalls in die Prüfung. Das gilt auch für Auftragsverarbeitung, Drittlandtransfer, technische und organisatorische Massnahmen sowie die Datenschutz-Folgenabschätzung. Eine belastbare Beurteilung ist erst möglich, wenn Zweck, Daten, Anbieter und Wirkung des Use Case konkret genug beschrieben sind. Ein Mensch im Prozess beseitigt Art.-22-Risiken nicht automatisch. Entscheidend ist, ob der Mensch die AI-Ausgabe tatsächlich kritisch prüfen, abändern oder verwerfen kann.

Vertraulichkeit reicht über personenbezogene Daten hinaus. Das Institut schützt auch Geschäfts-, Bank- und Berufsgeheimnisse sowie sensible Prüfungs-, Risiko-, Vertrags- und Sicherheitsinformationen. Verantwortliche müssen deshalb erkennen können, welche Daten auf welcher Plattform verwendet werden dürfen und wie Ausnahmen nachgewiesen werden.

Bei Beschäftigtenbezug sind Arbeitsrecht und Mitbestimmung früh einzubinden. Relevant ist nicht nur eine finale Personalentscheidung, sondern bereits eine technische Einrichtung oder Auswertung, die Beschäftigtenverhalten oder Leistung beeinflussen kann. Auch Auswahl, Bewertung und Steuerung von Beschäftigten fallen in diese Risikosicht.

Vor der Freigabe muss deshalb feststehen, ob Daten und Verwendungszweck zulässig sind, welche Schutzinteressen berührt werden und welche Fachfunktionen der Entscheidung zustimmen müssen.

12. Supplier und Supply Chain Governance

AI erweitert das klassische Drittparteienrisiko. Modelle, Plattformen, Datenquellen und technische Komponenten können von unterschiedlichen Anbietern stammen und sich im Betrieb verändern. Das Institut muss deshalb nicht nur den direkten Vertragspartner, sondern die für den AI Use Case wesentlichen Abhängigkeiten verstehen.

Vor Beschaffung oder Nutzung muss klar sein, welche Rolle der Anbieter für Steuerung, Betrieb und Datenschutz spielt. Auch Nachweisführung, Modellleistung, Security, Incident Response und Exit müssen zugeordnet sein. Für Finanzinstitute ist zusätzlich zu prüfen, ob DORA, eine kritische oder wichtige Funktion, ein ICT-Drittparteienbezug oder ein bankaufsichtlicher Auslagerungsbezug einschlägig ist.

Die Lieferantenbeurteilung umfasst:

- Datenresidenz, Subprozessoren, Drittlandtransfer und Trainingsnutzung,
- Modellherkunft, Lizenz, Integrität, Updatepolitik und Dokumentation,
- Audit-, Informations-, Incident- und Meldepflichten,
- technische und organisatorische Security-Nachweise,
- Exit, Datenlöschung, Export, Portabilität und Wechselkosten,
- Konzentrationsrisiken und strategische Anbieterabhängigkeit.

Eine AI-Drittpartei ist nicht freigabefähig, wenn das Institut wesentliche Risiken nicht prüfen, begrenzen, überwachen oder beenden kann. Kritische Befunde führen zu Auflagen, Risk Acceptance, Ablehnung oder Exit-Anforderungen.

Open-Source-Modelle und öffentliche Modellplattformen sind nicht automatisch risikoarm. Herkunft, Integrität, Lizenz, Schwachstellen und fachliche Eignung müssen soweit beurteilbar sein, dass eine verantwortbare Beschaffungs-, Nutzungs- oder Exit-Entscheidung möglich bleibt.

13. Secure AI Governance

Sicherheit ist eine Voraussetzung für die Freigabe und kein nachgelagerter technischer Prüfschritt. Je stärker eine AI-Nutzung auf Daten, Entscheidungen oder Systeme einwirkt, desto wichtiger werden Begrenzung, Überwachung und kontrollierte Abschaltbarkeit.

Bei externer Nutzung von General-Purpose AI (GPAI) muss zusätzlich entschieden werden, über welchen kontrollierten Datenpfad Prompts, Dateianhänge, Quellcode und automatisch bereitgestellter Kontext das Institut verlassen dürfen. Eingaben mit Geheimnissen oder unzulässigen Datenklassen werden vor der Übertragung maskiert, blockiert oder einer genehmigten Ausnahme zugeführt.

Aus Managementsicht stehen vier Sicherheitswirkungen im Vordergrund: Unzulässige Daten dürfen das Institut nicht verlassen; Eingaben und Quellen dürfen das Verhalten des Systems nicht unkontrolliert manipulieren; technische Rechte dürfen keine unbeabsichtigten Handlungen ermöglichen; und das Institut muss Fehlverhalten erkennen, begrenzen und stoppen können. Die AI Security Baseline ordnet diesen Wirkungen die technischen Bedrohungen und Mindestanforderungen zu.

Kritische AI-Nutzungen benötigen deshalb eine freigegebene Architektur und einen kontrollierten Datenpfad. Berechtigungen und Handlungsmöglichkeiten werden auf das erforderliche Mass begrenzt. Überwachung und Sicherheitstests müssen relevante Fehler und Missbrauch erkennbar machen. Fallback und Abschaltmechanismen stellen sicher, dass das Institut auch bei einem Versagen des AI-Systems handlungsfähig bleibt.

Die Geschäftsleitung entscheidet auf Grundlage der verbleibenden technischen Risiken, ob der erwartete Nutzen, die Schutzmassnahmen und die Abschaltbarkeit für den vorgesehenen Einsatz ausreichen.

14. AI Literacy und Human Oversight

AI Literacy zeigt sich in handlungsfähigen Personen, nicht in absolvierten Schulungen. Wer AI nutzt, verantwortet, kontrolliert oder beaufsichtigt, muss die für seine Rolle relevanten Möglichkeiten, Grenzen und Risiken verstehen. Die erforderliche Tiefe richtet sich nach der Wirkung der Nutzung und den Entscheidungen, welche die Person treffen muss.

Die erforderliche Entscheidungsfähigkeit unterscheidet sich nach Rolle.

Rolle	Erforderliche Entscheidungsfähigkeit
Geschäftsleitung und Führungskräfte	Risikoappetit, Portfolio, Aufsichtsfähigkeit und wesentliche AI-Entscheidungen verstehen und steuern.
AI Owner	Lifecycle, Freigaben, Register, Revalidierung, Kontrollen und Evidenzen für die verantwortete AI-Nutzung beherrschen.
Fachbereiche	Zulässige Nutzung, Datenregeln, Prüfung von Ergebnissen und Meldewege im eigenen Prozess anwenden.
IT- und Plattformteams	Betriebsmodell, Security, Überwachung, Änderungen und Incident Response sicher ausführen.
Kontrollfunktionen	Einstufung, Rechtsgrundlagen, Transparenz sowie Datenschutz-, Supplier- und Modellrisiken wirksam hinterfragen.

Die in Kapitel 8 definierte menschliche Aufsicht ist nur tragfähig, wenn die zuständigen Personen Systemgrenzen, typische Fehlerbilder, Automation Bias, Eskalationswege und ihre Stop-Rechte verstehen.[8] Rollenbezogene Kommunikation und Kompetenznachweise machen diese Fähigkeit überprüfbar.

Wirksamkeit wird nicht nur über Schulungsquoten gemessen. Zusätzlich relevant sind Shadow-AI-Fälle, wiederholte Fehlverwendung, Qualität der Intake-Meldungen, Incident-Ursachenanalysen, offene Oversight-Findings und kurze rollenbezogene Kompetenznachweise für kritische Rollen.

Teil IV: Transparenz, Steuerung und kontinuierliche Verbesserung

Eine Freigabe schafft keinen dauerhaften Sollzustand. Die Geschäftsleitung benötigt auch im Betrieb eine verlässliche Sicht auf Portfolio, Risiken, Ausnahmen und wesentliche Veränderungen. Teil IV zeigt, wie Register, Reporting, Incident Governance und Assurance diese Sicht herstellen und

8 Human Oversight bezeichnet die wirksame menschliche Aufsicht über ein AI-System. Für Hochrisiko-AI konkretisiert Art. 14 AI Act Anforderungen an Verständnis, Eingriff und Stop-Möglichkeit.

Entscheidungen in die laufende Steuerung zurückführen.

15. Register, Evidenzen und Dokumentation

Die Geschäftsleitung benötigt eine verlässliche Sicht auf Bestand, Verantwortung, Einstufung, Freigabestatus und Handlungsbedarf. Konsistente Referenzen zwischen den bestehenden Registern schaffen diese Sicht ohne zusätzliche Monolith-Datenbank.

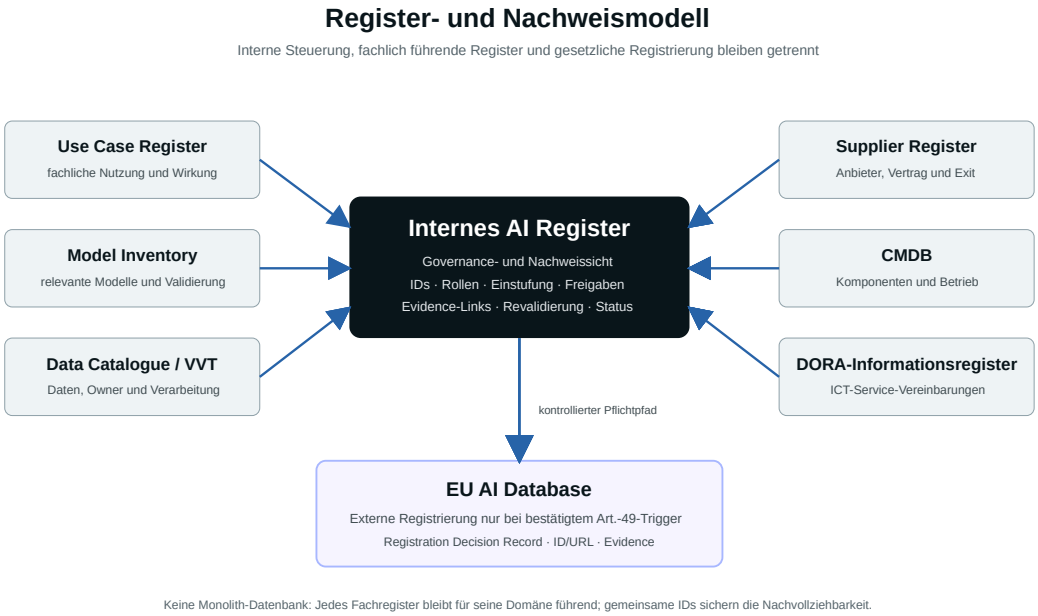

Abbildung 3: Das interne AI Register verbindet die fachlich führenden Register und führt erforderliche Angaben kontrolliert in die externe EU-Datenbankregistrierung über.

Das AI Register schafft die gemeinsame Steuerungssicht auf alle AI Use Cases. Es zeigt, wer verantwortlich ist, welche Einordnung gilt und ob die erforderlichen Entscheidungen und Nachweise vorliegen. Die fachlich führenden Inventare bleiben bestehen; das AI Register verbindet sie über eindeutige Referenzen, statt ihre Inhalte zu kopieren.

Die rechtliche Nachvollziehbarkeit entsteht im AI Act Requirements and Applicability Register. Für jeden Use Case hält ein genehmigtes Applicability Record fest, welche Pflichten gelten, welche nicht anwendbar sind und welcher Spezialpfad gegebenenfalls aktiviert wurde. Dadurch wird bei einer Änderung von Zweck, Systemversion oder Betriebsmodell sofort sichtbar, welche Entscheidungen erneut geprüft werden müssen.

Diese interne Registerarchitektur ist weder die gesetzliche EU-Datenbank noch eine Konformitätserklärung. Eine externe Registrierung erfolgt nur nach einem bestätigten gesetzlichen Trigger. Bei personenbezogenen Daten bleibt zugleich die konsistente Verknüpfung mit dem Verzeichnis von Verarbeitungstätigkeiten gewahrt.

Registerebene	Rechts- oder Governance-Grundlage	Scope	Führendes System
Internes AI Register	institutsweite Governance-Anforderung	AI-Systeme, AI Use Cases, Rollen, Einstufung, Freigaben und Revalidierung	AI Governance
EU AI Database	Art. 49 und 71 AI Act sowie Annex VIII	gesetzlich registrierungspflichtige Systeme und Nutzungen	EU-Datenbank; interner Registration Decision Record als Nachweis
DORA-Informationsregister	Art. 28 Abs. 3 DORA und Durchführungsverordnung (EU) 2024/2956	vertragliche Vereinbarungen über Informations- und Kommunikationstechnologie-Dienstleistungen (ICT-Dienstleistungen)	DORA Register of Information
Model Inventory	bankfachliche, aufsichtlich relevante oder intern als Modellrisiko eingestufte Modelle	Modellsteuerung, Validierung, Performance, Drift und Stilllegung	Model Risk Management
Fachregister	DSGVO, Supplier-, Daten-, Asset- und Betriebsgovernance	Verarbeitungstätigkeiten, Daten, Supplier, Komponenten und Services	Verzeichnis von Verarbeitungstätigkeiten (VVT), Data Catalogue, Supplier Register und CMDB

16. KPIs, KRIs und Management Reporting

Geschäftsleitung und Verwaltungsrat benötigen eine verdichtete Sicht auf Wertbeitrag, Risiko und Kontrollwirksamkeit. Key Performance Indicators (KPIs), Key Risk Indicators (KRIs) und Key Control Indicators (KCIs) liefern dafür unterschiedliche Signale; entscheidend ist ihre gemeinsame Aussage über das Portfolio und nicht die einzelne Kennzahl.

Typische Managementsignale sind:

- Anzahl und Entwicklung registrierter AI-Systeme und AI Use Cases,
- Anteil klassifizierter und freigegebener AI Use Cases,
- überfällige Revalidierungen,
- Hochrisiko-, GPAI-, Kundendaten-, Personal- oder Kreditbezüge,
- Shadow-AI-Funde,
- offene Red-Team-, Supplier-, Datenschutz-, Modell- oder Kontroll-Findings,
- AI-Incidents nach Schweregrad,
- Risk Acceptances, Temporary Approvals und ablaufende Ausnahmen,
- offene Provider-Aktivierungsentscheide und Spezial-Handoffs,
- aktive Hochrisiko-Deployerfälle und Art.-26-Suspendierungen,
- offene oder überfällige FRIA- und Art.-50-Entscheide,
- Art.-86-Anfragen und überfällige Antwort-SLAs,
- überfällige Rechts- oder Regulatory-Watch-Entscheide,
- erwarteter und realisierter Wertbeitrag kritischer AI Use Cases.

Jeder Geschäftsleitungsbericht beantwortet fünf Fragen: Was hat sich verändert? Was liegt ausserhalb des Risikoappetits? Welche Wirkung hat dies auf Nutzen und Risiko? Welche

Massnahmen sind vorgesehen? Welche Entscheidung wird jetzt benötigt? Operative Detaildaten verbleiben in Dashboards und Arbeitslisten.

Die Geschäftsleitung erhält den konsolidierten Bericht mindestens quartalsweise. Der Verwaltungsrat erhält mindestens jährlich eine Gesamtsicht. Ereignisse mit rotem Trigger werden unabhängig von dieser Cadence unverzüglich eskaliert. Dazu gehören eine positive Provider-Aktivierung ohne freigegebenen Compliance Plan, eine Art.-26-Suspendierung, eine überfällige Revalidierung eines kritischen oder Hochrisiko-Systems, eine offene freigabeblockierende FRIA- oder Transparenzpflicht und ein kritischer Supplierbefund ohne wirksame Begrenzung.

Eine klar definierte Ampellogik unterstützt die Priorisierung: Rot verlangt eine Entscheidung oder unmittelbare Risikobehandlung, Gelb eine befristete Massnahme und Grün bestätigt den Betrieb innerhalb des festgelegten Rahmens.

17. Incident Governance und Meldepflichten

AI-Vorfälle können technische Störungen, fehlerhafte fachliche Wirkungen, Datenabfluss, unzulässige Nutzung oder Versagen der menschlichen Aufsicht betreffen. Für die Führung steht zuerst die Begrenzung des Schadens im Vordergrund, unabhängig davon, wie der Vorfall später rechtlich eingeordnet wird.

Incident Governance verfolgt vier Ziele:

- Schaden begrenzen,
- Evidenz sichern,
- Meldepflichten parallel prüfen,
- Ursachen dauerhaft in Prozesse und Kontrollen zurückführen.

Neben der technischen und fachlichen Behandlung prüft das Institut gleichzeitig alle regulatorischen und vertraglichen Meldepflichten. Die juristische Endbewertung darf weder Schadensbegrenzung noch Beweissicherung oder die Wiederherstellung eines kontrollierten Betriebs verzögern.

Bei einem Hochrisiko-System folgt der Deployer der Meldekette nach Art. 26 AI Act. Eine Provider-Meldepflicht nach Art. 73 wird nicht pauschal auf den Deployer übertragen. Sie wird nur bei bestätigter Providerrolle über den Provider Compliance Plan aktiviert.

Wesentliche Vorfälle werden der Geschäftsleitung mit Auswirkung, Sofortmassnahmen, Meldebeurteilung, Restrisiko und erforderlicher Entscheidung vorgelegt. Lessons Learned gelten erst als abgeschlossen, wenn Ursachen in Verantwortlichkeiten, Freigaben oder Kontrollen wirksam behandelt sind.

18. Assurance, Reifegrad und Interne Revision

Assurance verbindet Managementaufsicht, laufende Kontrolltests und unabhängige Prüfung zu einem geschlossenen Steuerungskreislauf. Die First Line belegt, dass Nutzung und Betrieb innerhalb der Freigabe bleiben. Die Second Line beurteilt Einstufungen, Risikoentscheidungen und Kontrollwirksamkeit. Die Interne Revision prüft unabhängig, ob das Zusammenspiel verlässlich funktioniert und wesentliche Abweichungen wirksam eskaliert werden.

Die Beurteilung führt Register, Evidenzen, Freigaben, Kontrollen, Ausnahmen, Vorfälle und Managementmassnahmen zusammen. Dadurch erkennt die Geschäftsleitung nicht nur, ob Vorgaben dokumentiert sind, sondern ob das Verantwortungsmodell in der Praxis trägt.

Die Interne Revision prüft unabhängig, ob Design und Wirksamkeit der AI-Steuerung angemessen sind. Sie ist nicht operativ verantwortlich und trifft keine Freigabeentscheidungen. Ihr Fokus liegt risikoorientiert auf Führungsstruktur, Kontrollwirksamkeit, Nachweisführung, Managementsteuerung und Umgang mit Abweichungen.

Für die Reifegradmessung kann ein einfaches AI Governance Maturity Model verwendet werden:

Reifegrad	Beschreibung	Typische Evidenz
1 Ad hoc	Nutzung ist nicht vollständig erkannt oder gesteuert.	Shadow AI, unklare Owner.
2 Initial kontrolliert	Intake, Register und Grundregeln sind eingeführt.	Erste Freigaben und manuelle Prüfungen.
3 Definiert	Rollen, Lifecycle und Kontrollen sind etabliert.	Freigaben, Entscheide und Revalidierungen.
4 Wirksam gesteuert	Kontrollen werden getestet und Findings verfolgt.	Reporting, Kontrolltests und Validierungen.
5 Integriert und optimiert	Steuerung ist in die Führungsprozesse integriert.	Vernetzte Register und risikobasiertes Monitoring.

Das Maturity Model ist kein Selbstzweck. Es hilft der Geschäftsleitung zu entscheiden, wie weit das Institut in Steuerung, Security, Tooling und Assurance investieren muss. Der erforderliche Reifegrad hängt vom AI-Portfolio und dessen Wirkung auf kritische Geschäftsprozesse ab.

19. Glossar und Quellen

Ein gemeinsames Führungsmodell braucht eine gemeinsame Sprache. Dieses Glossar hält deshalb nur die Kernbegriffe und zentralen Quellen des Hauptdokuments fest. Detailbegriffe zu Workflows, Controls oder technischen Anforderungen können in den jeweiligen Companion-Dokumenten vertieft werden.

Begriff	Erklärung
AI-System	Maschinenbasiertes System, das aus Eingaben Ausgaben wie Inhalte, Empfehlungen, Vorhersagen, Scores oder Entscheidungen ableitet.
AI	Artificial Intelligence; in der Dokumentfamilie einheitlich verwendete Bezeichnung für künstliche Intelligenz.
AI-Nutzung	Beabsichtigter oder tatsächlicher fachlicher Einsatz von AI für einen bestimmten Zweck. Sie ist der zentrale Gegenstand der Governance.
AI Use Case	Konkret abgegrenzte AI-Nutzung, die das Institut ab dem formalen Intake als Governance-Gegenstand führt und im AI Register dokumentiert.
AI-Vorhaben	Zeitlich begrenzte Initiative zur Einführung oder wesentlichen Änderung einer AI-Nutzung.

Begriff	Erklärung
AI Owner	Fachlich verantwortliche Rolle für Zweck, Wirkung, Freigabe und lifecyclebezogene Steuerung eines AI Use Case.
Data Owner	Verantwortliche Rolle für Datenquelle, Datenqualität, Klassifikation, zulässige Nutzung und Datennachweise.
GPAI	General-Purpose AI; Modell, das für viele unterschiedliche Aufgaben eingesetzt werden kann.
FRIA	Fundamental Rights Impact Assessment; Grundrechte-Folgenabschätzung nach Art. 27 AI Act für die dort bestimmten Betreiber und Nutzungen.
DSGVO	Datenschutz-Grundverordnung; europäischer Rechtsrahmen für die Verarbeitung personenbezogener Daten.
KPI / KRI / KCI	Leistungs-, Risiko- und Kontrollindikatoren für die verdichtete Managementsteuerung.
ICT	Information and Communication Technology; Informations- und Kommunikationstechnologie im DORA-Kontext.
VVT	Verzeichnis von Verarbeitungstätigkeiten nach Datenschutzrecht.
CMDB	Configuration Management Database; führendes Betriebsinventar für technische Konfigurationselemente und Beziehungen.
Hochrisiko-AI	AI-System mit besonderer Wirkung auf Sicherheit, Grundrechte oder wesentliche Lebensbereiche nach AI Act.
Human Oversight	Qualifizierte menschliche Aufsicht mit Verständnis, Befugnis, Override- und Stop-Möglichkeit.
Risk Appetite	Vom Institut festgelegter Rahmen, welche AI-Risiken eingegangen, begrenzt oder ausgeschlossen werden.
Risk Acceptance	Befristete und begründete Annahme eines Restrisikos durch die zuständige Entscheidungsrolle.
RAG	Retrieval-Augmented Generation; Kombination von Wissensabruf und generativer Antwort.
Agent	AI-System, das mit Tools oder Schnittstellen Aktionen vorbereiten oder auslösen kann.
Prompt Injection	Angriff oder Eingabe, die ein AI-System zu unerwünschtem Verhalten bringt.
DORA	EU-Verordnung zur digitalen operationalen Resilienz im Finanzsektor.
AI Register	Internes Governance- und Nachweisinstrument für AI-Systeme, AI Use Cases, Operatorrollen, Einstufung, Owner, Freigaben und Revalidierung.
Provider-Aktivierungsentscheid	Verbindlicher Entscheid, ob das Institut Providerpflichten aktivieren und eine reine Deployer-Freigabe blockieren muss; englische Referenz: Provider Activation Gate.
AI Act Requirements and Applicability Register	Kontrollierte Traceability-Schicht zwischen atomarer Pflicht, Rolle, Trigger, Control, Workflow und Evidence.
EU AI Database Registration	Gesetzliche Registrierung nach Art. 49 und 71 AI Act; nicht mit dem internen AI Register gleichzusetzen.

Begriff	Erklärung
DORA Register of Information	Informationsregister über vertragliche Vereinbarungen mit ICT-Drittdienstleistern nach Art. 28 Abs. 3 DORA.
CISO	Chief Information Security Officer; Führungsrolle für Informationssicherheit und zugehörige Sicherheitsrisiken.
Model Inventory	Fachlich führendes Inventar für bankfachlich, aufsichtlich oder nach internen Kriterien relevante Modelle.
Digital.gov Plain Language	Offene Methode der U.S. General Services Administration für zielgruppenorientierte, auffindbare, verständliche und testbare Inhalte.
4T2 Management Prose	4T2 Schreibmethode für klare Managemententscheidungen, Verantwortung und Leserführung.

Zentrale Quellen sind der EU AI Act, DORA, DSGVO, das Kreditwesengesetz (KWG) und die Mindestanforderungen an das Risikomanagement (MaRisk). Hinzu kommen einschlägige Veröffentlichungen der Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin), der Bundesbank und des Bundesamts für Sicherheit in der Informationstechnik (BSI) zu Cloud, Cybersecurity und AI. Das Open Worldwide Application Security Project (OWASP), dessen Top 10 for Large Language Model Applications (LLM Top 10), das OWASP GenAI Security Project, GeschGehG, BetrVG sowie aktuelle Veröffentlichungen zur deutschen Durchführung der KI-Verordnung bleiben ebenfalls relevante Quellen.

Zum Rechtsstand vom 12. Juli 2026 gehört auch der vom Rat am 29. Juni 2026 endgültig angenommene Digital Omnibus on AI. Für seine Anwendung bleiben Veröffentlichung und Inkrafttreten der Änderungsverordnung sowie die finalen Leitlinien zu Hochrisiko-Systemen im Regulatory Watch.

Vor verbindlicher Veröffentlichung müssen Rechtsstand, Aufsichtspraxis, nationale Zuständigkeiten, Meldewege und interne Policies aktuell geprüft werden.

Methodische Quellen sind die Digital.gov Plain Language Guide Series der U.S. General Services Administration sowie der 4T2 Governance Writing Standard. Digital.gov ist unter https://digital.gov/guides/plain-language öffentlich verfügbar. Der 4T2 Standard ist die dokumentierte interne Ausführungsregel für managementorientierte Governance-Prosa.

Versionshistorie

Version	Status	Inhalt
V1.0	Review Draft	Erste vollständige Writer-Fassung der Dokumentenfamilie.
V1.2	Review Draft	Managementorientierte Schlussredaktion und sprachliche Verdichtung.
V1.3	Review Draft	Neuer Titel und Executive Management View.
V1.4	Release Candidate	Zielmarktkontext, Shadow-AI-Begrenzung, VVT-Referenz, GPAI-Ingress-Kontrolle und Release-Layout.
V1.5	Release Candidate	Konsistente Familienversion mit redaktionell harmonisierten Companion Documents.

Version	Status	Inhalt
V1.6	Release Candidate	Vereinheitlichter Lifecycle, getrennte Register- und X-Logik sowie überarbeitetes Dokumentdesign.
V1.7	Editorial Review	Managementorientierte Sprache, kontrollierte Begriffswelt und offene Zielgruppen-Gates.
V1.8	Family Review	Familienweit harmonisierte Sprache, Methoden, Abkürzungen und Querverweise.
V1.9	Family Review	Deutschland-Scope, atomare Deployerpflichten und schlankes Provider-Sicherungsnetz.
V2.0	Published Version	Veröffentlichungsreife Redaktion, konsistentes Layout und präzisierte Rollen- und Incident-Handoffs.

About 4T2 Solutions

4T2 Solutions GmbH unterstützt Organisationen an der Schnittstelle von Informationssicherheit, AI Engineering, Governance und Umsetzung. Dieses Whitepaper wurde als praxisorientierter Führungsrahmen für Finanzinstitute entwickelt.

4T2 Solutions GmbH · Wildensteinstrasse 10 · 9404 Rorschacherberg · Schweiz

20. Methodische Grundlage

Dieses Framework verwendet die offene Plain-Language-Methode von Digital.gov für Zielgruppenorientierung, Informationsstruktur und Verständlichkeitsprüfung. Die 4T2 Management Prose überträgt diese Grundsätze auf Governance-Entscheidungen, Verantwortlichkeiten und Leserführung. Die Methode verbessert die Darstellung, ersetzt aber keine fachliche, regulatorische oder rechtliche Prüfung.

21. Haftungs- und Anwendungshinweis

Dieses Framework wurde mit fachlicher Sorgfalt und gegen den angegebenen Rechtsstand erstellt. Eine Gewähr für Vollständigkeit, Fehlerfreiheit oder dauerhafte Aktualität wird nicht übernommen. Das Dokument ist weder Rechtsberatung noch regulatorische Freigabe, Zertifizierung oder Konformitätsbescheinigung.

Die Anwendbarkeit ist für jede Rechtseinheit, jedes Einsatzland, jede Operatorrolle, Systemversion und konkrete AI-Nutzung gesondert zu prüfen. Verantwortung für Einführung, Freigabe, Umsetzung und laufende Wirksamkeitskontrolle bleibt beim anwendenden Institut.

Copyright und offene Lizenz

Transparenz zur Erstellung

Martin Bischoff hat dieses Dokument mit Unterstützung generativer AI erstellt. AI wurde für Formulierungsentwürfe, redaktionelle Überarbeitung, grafische Gestaltung sowie die strukturierte Prüfung auf inhaltliche Vollständigkeit und Konsistenz eingesetzt. Auswahl, fachliche Bewertung und Freigabe erfolgten durch Martin Bischoff. Die AI-Unterstützung ersetzt keine unabhängige Rechts-, Wirksamkeits- oder Konformitätsprüfung.