Universitätsklinikum Frankfurt

Anlage 4 – Leistungsverzeichnis

Begleitende Beratungs-, Architektur-, Governance-, Security-, Audit- und Providersteuerungsleistungen für die KI-Compute-, Entwicklungs- und Integrationsplattform

Verfahrensgegenstand	Begleitende KI-Beratungsleistungen
Laufzeit	36 Monate ab Leistungsbeginn
Stand	13.11.2026

1. Ausgangslage und Zielsetzung

Das Universitätsklinikum Frankfurt (UKF) beschafft parallel eine flexible KI-Compute-, Entwicklungs- und Integrationsplattform. Die Plattform soll über 36 Monate für heterogene KI-, Analyse-, Entwicklungs- und Integrationsszenarien genutzt werden. Für die fachlich-technische Steuerung, sichere Einbindung in die bestehende UKF-Infrastruktur, wirtschaftliche Nutzung und kontrollierte Weiterentwicklung wird ein vom Compute-Anbieter unabhängiger Beratungspartner beauftragt.

Gesucht wird kein reiner KI-Strategieberater und kein Personalpool, sondern ein persönlich arbeitendes Spezialistenteam, das Krankenhaus-IT, Enterprise-/Campus-/Rechenzentrumsinfrastrukturen, KI-Compute, Cloud/Hybrid, Container/Kubernetes, Informationssicherheit, Auditbegleitung und Providersteuerung fachlich miteinander verbindet.

Ziel ist ein wiederholbares Vorgehen von der Use-Case-Idee über technische Architektur, Ressourcen- und Sicherheitsbewertung sowie Pilotierung bis zur Entscheidung über Skalierung, Regelbetrieb oder Beendigung. Der Auftragnehmer bildet die unabhängige technische Beratungs- und Steuerungsschnittstelle zwischen UKF, Fachbereichen, IT, ISB/DSB, Compute-Anbieter und weiteren Technologie- bzw. Implementierungspartnern.

2. Rollenmodell und Interessentrennung

2.1 Trennung der Vergabeverfahren und Vermeidung wirtschaftlicher Interessenkonflikte

Der Auftraggeber führt die Beschaffung von KI-Compute-Plattform und „Begleitende Beratungs-, Architektur-, Governance-, Security-, Audit- und Providersteuerungsleistungen für die KI-Compute-, Entwicklungs- und Integrationsplattform“ in zwei eigenständigen Vergabeverfahren durch.

Zwischen beiden Leistungsbereichen besteht eine unmittelbare fachliche und wirtschaftliche Abhängigkeit. Insbesondere können die im Rahmen der KI-Beratungsleistungen erarbeiteten Empfehlungen, technischen Konzepte und Bedarfsbewertungen maßgeblichen Einfluss auf Art, Umfang und Abruf der KI-Compute-Leistungen nehmen.

Zur Sicherstellung einer unabhängigen, objektiven und wirtschaftlich neutralen Beratung sowie zur Vermeidung potenzieller Interessenkonflikte sieht der Auftraggeber eine personelle und wirtschaftliche Trennung der Auftragnehmer beider Leistungsbereiche vor.

Eine gleichzeitige Beteiligung desselben Wirtschaftsteilnehmers an beiden Vergabeverfahren ist daher nicht zugelassen. Dies gilt entsprechend für Bietergemeinschaften sowie für wirtschaftlich verbundene Unternehmen, soweit aufgrund ihrer konkreten Verflechtungen eine unabhängige Leistungserbringung nicht gewährleistet werden kann. Dies gilt darüber hinaus auch für Nachunternehmer.

Mit der Angebotsabgabe erklärt der Bieter, dass er die vorgenannten Anforderungen einhält und keine Beteiligung am jeweils anderen Vergabeverfahren besteht, die zu einem unzulässigen Interessenkonflikt führen kann.

Der Auftraggeber behält sich vor, mögliche Interessenkonflikte unter Beachtung der vergaberechtlichen Grundsätze der Transparenz, Gleichbehandlung und Verhältnismäßigkeit zu prüfen und erforderliche Maßnahmen zu ergreifen.

Rolle	Verantwortung
UKF – Auftraggeber	Priorisierung von Use Cases; fachliche Entscheidungen und Freigaben; medizinische, organisatorische und datenschutzrechtliche Verantwortung; Benennung interner Service- und Fachverantwortlicher.
KI-Compute-Dienstleister	Bereitstellung und Betrieb der beschafften KI-Compute-, Entwicklungs- und Integrationsplattform einschließlich Compute, Runtime, Storage, Netzwerk und der dort vertraglich geschuldeten Serviceleistungen.
Beratungspartner	Bedarfsspezifikation, Architektur- und Infrastrukturberatung, Compute-Dimensionierung, POC-/Evaluierungsbegleitung, Security/Audit, Providersteuerung, Governance, FinOps und Wissenstransfer. Keine Übernahme des Plattformbetriebs, der DSB-/ISB-Funktion, medizinischer Freigaben oder Rechtsberatung.

Leitprinzip: Infrastruktur liefern und ihre Nutzung, Integration, Sicherheit und Wirtschaftlichkeit auf Auftraggeberseite steuern werden bewusst getrennt. Der Beratungspartner muss seine Empfehlungen ergebnisoffen und hersteller-/produktunabhängig erbringen.

3. Standardprozess für KI-Vorhaben

Schritt	Inhalt
1	Idee aufnehmen – Problem, Nutzen, Owner
2	Use-Case-Passport – Daten, Risiken, Abhängigkeiten, KPIs
3	Architektur & Ressourcenbudget – Compute, Runtime, Storage, Netzwerk, Security
4	Pilot/POC – begrenzte technische Umsetzung und Validierung
5	Evaluation – Qualität, Ressourcenverbrauch, Kosten-/Nutzenlogik, Risiken
6	Go / Scale / Stop – Zielbetrieb, Skalierung, Nachsteuerung oder Exit

4. Leistungsbausteine A–H

A – AI Portfolio & Use-Case Factory

Intake, Strukturierung und Priorisierung von Use Cases; Use-Case-Passports; technische Machbarkeit; Daten- und Integrationsabhängigkeiten; Ressourcenbedarf; KPI-/Nutzenlogik; Roadmap; Vorbereitung dokumentierter Go/Scale/Stop-Entscheidungen.

B – Architektur & Platform Stewardship

UKF-seitige Referenz- und Lösungsarchitekturen; Integration der KI-Plattform in Enterprise-, Campus- und Rechenzentrumsstrukturen; Netzwerk-, Server-, Storage-, Virtualisierungs-, Cloud- und Hybrid-Schnittstellen; Daten- und Modellflüsse; Container/Kubernetes; APIs; IAM; Portabilität; technische Standards und Architekturentscheidungen.

C – Compute Demand & FinOps

Technische Bewertung und Dimensionierung von GPU/Accelerator, CPU, RAM, Storage und Netzwerk; Einordnung von Compute-Klassen und Workload-Anforderungen; CU-/Ressourcenplanung; Forecast; Verbrauchs- und Benchmarkplausibilisierung; Kapazitätsplanung; Effizienzanalysen und Management-Reporting.

D – POC & Technical Validation

Planung und Begleitung von POCs und Integrationsprototypen; technische Erfolgskriterien; Ressourcen- und Integrationsplanung; Container-/Kubernetes- und Schnittstellenvalidierung; Bewertung von Machbarkeit, Skalierbarkeit, Security und Betriebsfähigkeit; dokumentierte Go/Scale/Stop-Empfehlung.

E – AI Governance & Compliance

Unterstützung bei KI-Register, Rollenmodell, Freigabepfaden, Nachweisdokumentation und Entscheidungsprozessen; Vorbereitung fachlicher Unterlagen für DSB, ISB, Recht und Fachverantwortliche. Keine Rechtsberatung und keine Übernahme gesetzlich oder organisatorisch zugewiesener Kontrollfunktionen.

F – Security & Audit

Security-by-Design; Schutzbedarfs- und Risikoanalysen; technische Security Reviews; IAM/PAM, Logging/SIEM, Secrets, Netzwerk- und Schnittstellensicherheit; Cloud-/Provider-Security; Bewertung von Sicherheitsnachweisen; Audit-Readiness, Auditvorbereitung, fachliche Auditbegleitung und Maßnahmenverfolgung; Unterstützung an Schnittstellen zu ISO 27001, BSI, C5, KRITIS und NIS2, soweit für den jeweiligen Use Case relevant.

G – Provider & Service Management

Technische Schnittstelle zum Compute-Anbieter; SLA-/Leistungsreview; Bereitstellungszeiten; Kapazitäten; Ressourcen- und Verbrauchsplausibilisierung; Benchmark-/Leistungsbewertung; Eskalations- und Änderungsmanagement; Roadmap; Prüfung technischer Providerangaben; Exit- und Transition-Unterstützung.

H – Kompetenzaufbau & Service Owner AI

Rollenbasierte Workshops, Coaching, Methoden, Vorlagen, Betriebshandbücher, Entscheidungsleitfäden und strukturierter Wissenstransfer. Ziel ist die dauerhafte UKF-seitige Architektur-, Bewertungs- und Steuerungsfähigkeit ohne dauerhafte Abhängigkeit vom Beratungspartner.

5. Mengengerüst und Abrufmodell

Der Gesamtpreis beträgt 100.000,00 EUR brutto für die gesamte Vertragslaufzeit. Der Bieter muss hierfür mindestens 56 Personentage (PT) verbindlich garantieren. Ein PT entspricht 8 Arbeitsstunden.

Über 56 PT hinaus angebotene PT sind ebenfalls verbindlicher Vertragsbestandteil und werden nach Anlage 2 gewertet.

Das angebotene Gesamtvolumen bildet einen gemeinsamen, über 36 Monate flexibel einsetzbaren PT-Pool. Jeder tatsächlich eingesetzte Personentag einer zugelassenen Person verbraucht genau 1 PT. Rollenbezogene Verbrauchsfaktoren oder unterschiedliche Abrechnungssätze werden nicht angewendet.

Rolle	Kalkulatorischer Referenzanteil
SP1 Lead KI-/IT-Consultant	35 %
SP2 KI-/IT-Consultant	40 %
SP3 Security-/Audit-Spezialist	25 %

Der Referenzmix dient ausschließlich der einheitlichen Kalkulation. Er ist weder Abnahmegarantie noch starre Rollenquote. Der tatsächliche Abruf richtet sich nach dem Projektbedarf. Das Angebot ist so zu kalkulieren, dass sämtliche zugelassenen Rollen innerhalb des Festpreises eingesetzt werden können.

6. Schlüsselpersonen, Vertretung und persönliche Leistungserbringung

Mit dem Angebot sind SP1 Lead KI-/IT-Consultant, SP2 KI-/IT-Consultant und SP3 Security-/Audit-Spezialist namentlich zu benennen. Diese Personen sind für die Vertragsausführung verbindlich. Die Mindest- und Wertungsanforderungen ergeben sich aus Anlage 2.

Mindestens 70 % der insgesamt abgerufenen PT sind persönlich durch SP1–SP3 zu erbringen. Maximal 30 % dürfen durch weitere, für die jeweilige Aufgabe fachlich geeignete Teammitglieder erbracht werden. Aufgaben, die vom UKF ausdrücklich einer Schlüsselperson zugewiesen werden, sind durch diese persönlich zu verantworten.

Für jede Schlüsselrolle ist eine qualifizierte Vertretungsregelung vorzuhalten. Ein dauerhafter Austausch einer Schlüsselperson ist nur aus sachlichem Grund und nach Zustimmung des UKF zulässig. Eine Ersatzperson muss sämtliche MUSS-Anforderungen erfüllen und hinsichtlich der für den Zuschlag berücksichtigten Wertungsmerkmale mindestens gleichwertig sein.

7. Parallelleistungsfähigkeit

Der Auftragnehmer muss personell und organisatorisch gewährleisten, dass mindestens zwei voneinander unabhängige Arbeitspakete parallel bearbeitet werden können und gleichzeitig die vereinbarte Reaktionsfähigkeit für einen kritischen Security-/Providerfall erhalten bleibt. Die Erfüllung darf nicht allein von der Verfügbarkeit einer einzelnen Person abhängen.

8. Hersteller- und Produktunabhängigkeit

Der Auftragnehmer muss nachweislich hersteller- und produktunabhängig beraten. Im für diesen Auftrag relevanten Leistungsbereich dürfen keine Exklusivverträge oder vergleichbaren Bindungen bestehen, durch die die Beratung auf einen oder wenige Hersteller, Plattformanbieter oder Technologieanbieter beschränkt wird. Herstellerpartnerschaften sind zulässig, sofern sie die ergebnisoffene Beratung nicht einschränken.

Mit dem Angebot sind eine Eigenerklärung sowie mindestens ein geeigneter Nachweis über ein Projekt vorzulegen, in dem unterschiedliche Hersteller-, Plattform- oder Technologiealternativen ergebnisoffen bewertet wurden. Provisionen, Reselling-Interessen oder sonstige wirtschaftliche Beziehungen, die den Auftragsgegenstand berühren können, sind offenzulegen.

9. Leistungssteuerung, Dokumentation und Reviews

- Abrufe durch autorisierte UKF-Ansprechpartner in Textform.
- Fortlaufendes PT-Konto mit Auftrag, Rolle, Datum, Aufwand, Ergebnis und Restkontingent.
- Monatlicher Leistungsnachweis und quartalsweiser Statusbericht zu Portfolio, Ressourcenverbrauch, Risiken, Provider-Themen, Security/Audit und offenen Entscheidungen.
- Arbeitsergebnisse in üblichen editierbaren Formaten; Entscheidungen und Annahmen nachvollziehbar dokumentieren.
- Nicht abgerufene PT verfallen innerhalb der 36-monatigen Vertragslaufzeit nicht.

10. Verfügbarkeit und Reaktionsfähigkeit

Regelabrufe sowie qualifizierte Reaktionen auf kritische Security-/Providerfälle werden in Anlage 3 verbindlich angeboten und nach Anlage 2 bewertet. Eine qualifizierte Reaktion umfasst fachliche Aufnahme, erste Bewertung und Abstimmung des weiteren Vorgehens; eine automatisierte Ticketbestätigung genügt nicht.

11. Initialisierungsphase und Vergütung

Die Vergütung erfolgt auf Abruf gemäß Anlage 5. Der Abruf erfolgt nach Bedarf des Auftraggebers. Der maximale Abrufwert beträgt 100.000,00 EUR brutto. Ein Preiswettbewerb findet nicht statt. Die Rechnungstellung und Zahlung erfolgen nach Abruf.

Mit Vertragsschluss (Zuschlagserteilung) entsteht der vertraglich verbindliche Anspruch des Auftraggebers auf das vollständig angebotene Beratungskontingent für die 36-monatige Vertragslaufzeit. Nicht verbrauchte PT verfallen während der Vertragslaufzeit nicht.

Die Initialisierungsphase umfasst mindestens: Governance-/Steuerungsmodell, initiale Architektur- und Infrastruktur-Baseline, Security-Baseline, Providersteuerungsprozess, Use-Case-Verfahren, FinOps-/Reportingverfahren, Audit-/Security-Vorgehensmodell sowie eine initiale 36-Monats-Roadmap. Einer formalen Abnahme der Initialisierungsphase bedarf es nicht. Die in der Initialisierungsphase erbrachten Leistungen werden aus dem angebotenen Gesamtkontingent abgerufen und abgerechnet; eine gesonderte Vergütung erfolgt nicht.

12. Laufzeit, Exit und Nutzbarkeit der Arbeitsergebnisse

Die Leistung wird für 36 Monate erbracht. Zum Vertragsende unterstützt der Auftragnehmer die strukturierte Übergabe aller offenen Themen, Methoden, Templates, Architekturentscheidungen, Provider- und Security-Maßnahmen. Sämtliche für die weitere Steuerung erforderlichen Arbeitsergebnisse sind UKF in editierbaren Formaten zu übergeben und durch UKF für eigene Zwecke dauerhaft nutzbar.