

## **Leistungsbeschreibung für die Vergabe einer KI-Plattform**

Die Hamburgische Investitions- und Förderbank (im Folgenden: „IFB Hamburg“) ist eine rechtsfähige Anstalt des öffentlichen Rechts. Trägerin der IFB ist die Freie und Hansestadt Hamburg, deren zentrales Förderinstitut die IFB ist. Die IFB unterstützt den Senat bei der Erfüllung öffentlicher Aufgaben. Dabei führt sie im staatlichen Auftrag, auf dem Gebiet der Freien und Hansestadt Hamburg, Fördermaßnahmen, insbesondere Finanzierungen im Einklang mit den Beihilfavorschriften der Europäischen Union und unter Beachtung des gemeinschaftsrechtlichen Diskriminierungsverbotes im eigenen Namen in verschiedenen Bereichen durch (u. a. Wohnraumförderung, insbes. sozialer Wohnraum, Innovationsförderung etc.).

### **1. Ausgangslage**

Die IFB Hamburg möchte mit der Einführung einer KI-Plattform auf Basis von Large Language Models (LLM) die Potenziale dieser innovativen Technologie nutzen und Geschäftsprozesse damit unterstützend effizienter gestalten. In der Vergangenheit wurden bereits einige Use Cases in den vielfältigen Tätigkeitsbereichen identifiziert, womit sowohl ein Interesse als auch ein großer Nutzen festgestellt wurde. Das Ziel ist, allen knapp 390 Mitarbeitenden der IFB Hamburg den Zugriff auf die Plattform zu gewähren.

Für das Hosting darf als Alternative zum Selbst-Hosting durch den Auftragnehmer auch ein Unterauftrag geschlossen werden. Dieser muss alle sich aus der Auftragsbekanntmachung und den Vertragsbedingungen ergebenden Anforderungen erfüllen. Der Auftragnehmer ist verpflichtet, die Rechte und Anforderungen der IFB Hamburg im Unterauftragsverhältnis durchzusetzen. Der Unterauftragnehmer hat der IFB Hamburg alle relevanten Informationen und geltenden Anforderungsdokumente zur Verfügung zu stellen.

### **2.1 Anforderungen an die Dienstleistung: Betrieb, Weiterentwicklung, Workshops**

#### **2.1.1 Regulatorische Anforderungen**

Höchste Priorität muss dabei dem sicheren Einsatz unter Einbezug der regulatorischen Anforderungen, insbesondere der *EU-KI-VO*, des *Datenschutzes (DS-GVO, HmbDSG)* und der *Informationssicherheit (MaRisk/ DORA)* gelten. Näheres regeln die Vertragsbedingungen.

#### **2.1.2 Informationstechnologische Anforderungen & Kriterien**

Das Tool muss nicht in andere Software (MS Office-Produkte) integrierbar sein, sollte jedoch insbesondere bei Formatierungs- und Berechnungsproblemen, wie etwa bei der Bearbeitung von Tabellen im Office Umfeld, unterstützen. Angedacht ist ein browserbasiertes LLM, was via Login über den Browser für die Mitarbeitenden erreichbar ist. Die Authentifizierung sollte per Microsoft AD bzw. Azure AD Account in Kombination mit einem Multifaktor erfolgen und FIDO2 sowie gängige Authenticator-Apps unterstützen. Die Benutzerkontensteuerung muss in Account-, als auch Gruppenform über ein RBAC-Schema möglich sein.

#### **2.1.3 Gewünschte Eigenschaften**

Eine unternehmensinterne Zusammenarbeit zwischen verschiedenen Personen ist erwünscht, ebenso wie das Erstellen von Agenten. Agenten werden in diesem Fall als Chat mit vordefinierter Anweisung mit definiertem Ausgabeformat definiert. Der Upload/die

Verarbeitung von verschiedenen Dateitypen (Excel, PDF, ggf. Powerpoint, ggf. Bilder, auch in höherer MB-Anzahl > 30 MB) ist notwendig. Es muss sichergestellt sein, dass die Daten in einer abgesicherten Umgebung hochgeladen und verarbeitet werden. Eine abgesicherte Umgebung ist als eine durch technische und organisatorische Maßnahmen abgesicherte Infrastruktur definiert, in der Systeme, Daten und Prozesse gemäß den Vorgaben der Anlage Informationssicherheit gegen unbefugten Zugriff, Manipulation, Verlust oder Ausfall geschützt sind. Sofern ein Training des LLMs geplant ist, so sind sämtliche Daten und Eingaben der IFB ausschließlich zum Trainieren der IFB-Instanz zu verwenden.

Ebenso ist die Auswahl zwischen verschiedenen Large Language Modellen gewünscht. Um sowohl interne als auch externe Recherchen zu ermöglichen, ist eine Option denkbar, ein Modell für interne Zwecke in geschützter Umgebung zu hosten, während weitere Modelle für die externe Recherche genutzt werden können.

#### **2.1.4 Weiterentwicklung**

Im Rahmen der Weiterentwicklung soll eine Anbindung an interne Systeme diskutiert werden. Denkbar ist etwa eine Wissensdatenbank, in der zu internen Dokumenten recherchiert werden kann oder weitere interne Schnittstellen (RAG). Zur Weiterentwicklung zählt daher ebenso die Beratung und Unterstützung in der Umsetzung weiterer Anwendungsfälle und/oder Schnittstellen.

#### **2.1.5 Workshops und Einführungsunterstützung**

Der Auftragnehmer soll unterstützen, indem die Einführung der Plattform mit mehreren Workshops begleitet wird. Dazu sollen mindestens zwei allgemeine Workshop mit mehreren Terminoptionen zur Einführung ins Tool, Präsentation der Anwendungsmöglichkeiten sowie Prompting zählen. Zusätzlich soll der Auftragnehmer ein Handbuch über die grundlegenden Funktionen bereitstellen.

### **2.2 Anforderungen an die Dienstleistung: Hosting**

#### **2.2.1 Regulatorische Anforderungen**

Höchste Priorität muss dabei dem sicheren Einsatz unter Einbezug der regulatorischen Anforderungen, insbesondere der *EU-KI-VO*, des *Datenschutzes (DS-GVO, HmbDSG)* und der *Informationssicherheit (MaRisk/ DORA)* gelten. Näheres regeln die Vertragsbedingungen.

#### **2.2.2 Informationstechnologische Anforderungen & Kriterien**

Das angebotene System muss sich in die IT-Landschaft der IFB Hamburg integrieren. In Bezug auf Hosting (ggf. durch Unterauftragnehmer) sind on-premise oder Cloud-Lösungen mit Hosting-Standort in Deutschland oder in der EU/EWG mit entsprechender direkter Anbindung denkbar.

Das Tool muss nicht in andere Software (MS Office-Produkte) integrierbar sein, sollte jedoch insbesondere bei Formatierungs- und Berechnungsproblemen, wie etwa bei der Bearbeitung von Tabellen im Office Umfeld, unterstützen. Angedacht ist ein browserbasiertes LLM, was via Login über den Browser für die Mitarbeitenden erreichbar ist. Die Authentifizierung sollte per Microsoft AD bzw. Azure AD Account in Kombination mit einem Multifaktor erfolgen und FIDO2 sowie gängige Authenticator-Apps unterstützen. Die Benutzerkontensteuerung muss in Account-, als auch Gruppenform über ein RBAC-Schema möglich sein.

Der Netzwerk-Kommunikationsport sollte frei konfigurierbar und die Verbindung Proxy kompatibel sein. Ebenfalls muss die Möglichkeit bestehen ein eigenes Webserver-Zertifikat für die Verbindungen zu hinterlegen (bei Selbst-Hosting Option).

### **2.2.3 Performanz des Hostings**

Der Hostinganbieter muss eine Lösung vorhalten, die eine genügende Leistung aufweist, bei der ein Großteil der knapp 400 Mitarbeitenden der IFB Hamburg die Plattform täglich mehrfach in einer zufriedenstellenden Geschwindigkeit nutzen können. Der Anbieter garantiert eine durchschnittliche Generierungsgeschwindigkeit von mindestens 30 Tokens/s pro aktiver Nutzersitzung (Output-Tokens), gemessen am Server-Endpunkt unter einer Maximallast von bis zu 20 parallel verarbeiteten Anfragen (Concurrent Requests).

### **2.2.4 Nachhaltigkeit**

Der Bieter hat für den Betrieb der angebotenen KI-Lösung folgende Informationen bereitzustellen:

1. Geschätzter jährlicher Stromverbrauch der für die Leistungserbringung eingesetzten Rechenzentrumsressourcen (kWh/Jahr).
2. Angaben zur Energieeffizienz der eingesetzten Rechenzentren (PUE-Faktor)
3. Offenlegung der bezogenen Stromart einschließlich der jeweiligen Anteile der Energieträger sowie Angabe des verwendeten Treibhausgas-Emissionsfaktors (kg CO<sub>2</sub>e/kWh) für den bezogenen Strom

## **3 Vertragslaufzeit**

Die Laufzeit des zu schließenden Vertrags beträgt zunächst zwei Jahre ab Datum der Zuschlagserteilung. Die Auftraggeberin hat die Option, diese Laufzeit zweimalig um ein Jahr zu verlängern, sodass die maximale Vertragslaufzeit vier Jahre ab dem Tag der Zuschlagserteilung beträgt. Sollte die Auftraggeberin Gebrauch von der Option machen wollen, wird sie dem Auftragnehmer dies spätestens drei Monate vor Ablauf der Vertragslaufzeit in Textform mitteilen. Seitens des Auftragnehmers besteht kein Anspruch darauf, dass diese Option gezogen wird.