



Kompetenzzentrum  
Öffentliche IT

FRAUNHOFER-INSTITUT FÜR OFFENE KOMMUNIKATIONSSYSTEME FOKUS

# DIGITALE SOUVERÄNITÄT IM BEREICH DER KÜNSTLICHEN INTELLIGENZ (KI)

## ERHEBUNGSBOGEN ÜBER DIE NUTZUNG GROßER SPRACHMODELLE (LLMS) IN DER ÖFFENTLICHEN VERWALTUNG

Version für den Vorhabenstatus „In Verkehr/ in Betrieb“

Gabriele Goldacker, Jan Dennis Gumz, Maximilian Kupi, Nicole Opiela, Fabian Manzke, Dorian Wachsmann, Mike Weber & Jan Ziesing

## Verortung Ihres Vorhabens

Zunächst benötigen wir einige Informationen von Ihnen, um Ihr Vorhaben besser verorten zu können. So möchten wir Ihre Angaben mit denen des Marktplatzes der KI-Möglichkeiten verknüpfen und damit sicherstellen, dass wir nur die notwendigen Fragen stellen, um den Aufwand für Sie so gering wie möglich zu halten.

### V01

Damit wir Ihre Antworten um die Angaben des KI-Marktplatzes anreichern können, bitten wir Sie, den dort vermerkten Titel oder die ID anzugeben. Dadurch können wir die dort hinterlegten Informationen nutzen, ohne sie erneut bei Ihnen abfragen zu müssen.

---

### V02

Bitte wählen Sie aus, in welcher Phase der Entwicklung sich das KI-System befindet. Wir möchten den aktuellen Umsetzungsstand, der sich gegenüber der Abfrage zum Marktplatz möglicherweise verändert hat, erfahren. Zudem benötigen wir Ihre Angabe, um den weiteren Fragebogen auf Ihr Projekt abzustimmen.

- Bedarf festgestellt / Idee 1
- In Planung / in Vorstudie (PoC) 2
- In Entwicklung 3
- **In Verkehr / in Betrieb** 4
- Nicht mehr in Verkehr / in Betrieb 5
- Zurückgezogen 6

### V03

Kommen bei Ihrem KI-System große Sprachmodelle (Large Language Models, LLMs) zum Einsatz?

- Ja
- Nein
- Keine Antwort

### V3a

[Filter V03 == „N“] Aus Ihrer Projektbeschreibung im Marktplatz der KI-Möglichkeiten haben wir abgeleitet, dass bei Ihrem Vorhaben möglicherweise ein LLM zum Einsatz kommt respektive kommen soll. Kommt möglicherweise ein kleineres Sprachmodell zum Einsatz? Auch in diesem Fall bitten wir Sie, die vorhergehende Frage zu bejahen. Falls kein Sprachmodell zum Einsatz kommt, beschreiben Sie bitte kurz, wie Ihr Vorhaben technisch umgesetzt wird.

---

## Allgemeine Projektaspekte

Zu Beginn würden wir uns freuen, wenn Sie einige allgemeine Fragen zu der Größe Ihres Projektes, den anvisierten Nutzenden und zu weiteren projektübergreifenden Aspekten beantworten. Unter Projekt verstehen wir die Beauftragung oder Eigenentwicklung einer Anwendung, die in irgendeiner Form ein oder mehrere große Sprachmodelle (LLMs) nutzt.

### Q01

1.1) Welches finanzielle Volumen hat Ihr Projekt in etwa?

- Unter 10.000 € **1**
- 10.000 € bis unter 100.000 € **2**
- 100.000 € bis unter 1 Mio. € **3**
- 1 Mio. € und mehr **4**
- Keine Antwort

### Q02

1.2) Wen adressieren Sie mit ihrem Projekt?

[Hinweis] Bitte wählen Sie den größten der adressierten Nutzendenkreise aus.

- Spezifische Referate **1**
- Gesamtes Haus **2**
- Ressortübergreifend **3**
- Gesamte Bundesverwaltung **4**
- Gesamte öffentliche Verwaltung **5**
- Verwaltung und andere Sektoren **6**
- Sonstige Nutzengruppen, und zwar: \_\_\_\_\_
- Keine Antwort

### Q03

1.3) In welcher Form sind Sie in den Erfahrungsaustausch mit anderen KI-Projekten im öffentlichen Sektor eingetreten?

(Mehrfachantworten möglich)

- Institutionell (z. B. im Rahmen interministerieller Arbeitsgruppen) **C1**
- Informell (z. B. über NExT-Arbeitsgruppen) **C2**
- Ad-hoc (z. B. Gespräche auf Fachkonferenzen) **C3**
- Kollaborativ (z. B. über Codesharing-Plattformen) **C4**
- Sonstiges, und zwar: \_\_\_\_\_
- Nichts davon [Exklusiv] **C5**
- Keine Antwort [Exklusiv] **C6**

**Q04**

1.4) Wie würden Sie die in Ihrem Projekt vorliegenden Kompetenzen hinsichtlich rechtlicher, technischer, fachlicher und spezifisch KI-bezogener Expertise einschätzen? Bitte verorten Sie die Kompetenzen im Projekt auf einer Skala von 1 (fast keine Kompetenzen) bis 10 (exzellente Kompetenzen).

- \_ Rechtliche Kompetenzen **C1**
- \_ Technische Kompetenzen **C2**
- \_ Fachliche Kompetenzen **C3**
- \_ KI-spezifische Kompetenzen **C4**

**Q05**

1.5) Mussten Sie an einem Zeitpunkt aufgrund fehlender Ressourcen (Geld, Personal, Kompetenzen) das Projekt respektive die Projektplanung anpassen?

- Ja, substanzial: Es wurden Abstriche am Funktionsumfang gemacht werden. **4**
- Ja, deutlich: Es wurden Abstriche an der Verfügbarkeit der Anwendung (z. B. des Nutzendenkreises) gemacht. **3**
- Ja, minimal: Es wurden kleinere Abstriche an Funktionsumfang oder Verfügbarkeit gemacht. **2**
- Nein: Es wurden keine Abstriche gemacht. **1**
- Sonstiges, und zwar: \_\_\_\_\_
- Keine Antwort

## 2. Anwendung: Software und Infrastruktur

Im diesem Themenblock interessieren uns die zum Einsatz kommende Software und Hardware für die Anwendung. Die Anwendung ist das Bindeglied zwischen Mensch (Nutzerin/Nutzer) und KI. Mit Hilfe einer Anwendung kann die Nutzerin / der Nutzer die KI (einfach) bedienen.

Beispielsweise ist ChatGPT eine Webanwendung für die Nutzung der Sprachmodelle von OpenAI, also etwa von GPT-4. Bei Anwendungen kann es sich um Bürossoftware wie Textverarbeitungsprogramme oder E-Mail-Dienste, Weboberflächen oder eigens entwickelte Anwendungen handeln.

Explizit **ausgenommen** ist das zum Einsatz kommende Sprachmodell selbst, das im nächsten Themenblock betrachtet wird.

### Q06

2.1) Wo ist die Anwendung (also die Software, über die mit der KI interagiert wird) installiert? *Die Frage bezieht sich auf das Backend der Anwendung.*

- Lokal bzw. eigener Betrieb (im Haus) C1
- Öffentlicher IT-DL (z. B. ITZ Bund, Deutsche Verwaltungscloud) C2
- Anwendungsanbieter C3
- Public-Cloud-Anbieter C4
- Andere, und zwar \_\_\_\_\_
- Keine Antwort C5

### Q07

[Filter: Q06 C4 == „Y“]

Bei welchem Public-Cloud-Anbieter läuft die Anwendung?

- AWS 1
- Google 2
- Microsoft 3
- Deutsche Telekom 4
- IONOS 5
- StackIT 6
- OVH 7
- Anderer: \_\_\_\_\_
- Keine Antwort

### Q08

2.2) Wie einfach sind Teile und Funktionen Ihrer Anwendung austauschbar (Modularität der Anwendung)

- Sehr einfach: sehr hoher Modularisierungsgrad und Standardkonformität aller Module 4
- Einfach: hoher Modularisierungsgrad und Schnittstellenkonformität aller Module, funktionale Standardkonformität 3

- Weniger einfach: geringer Modularisierungsgrad, geringe Schnittstellenkonformität wesentlicher Module, grundlegende funktionale Standardkonformität **2**
- Kaum möglich: monolithische Lösung (Anwendungsteile und -funktionen sind faktisch untrennbar integriert und mit vertretbarem Aufwand nicht modularisierbar) oder geringe Schnittstellenkonformität oder nicht vorhandene funktionale Standardkonformität **1**
- Keine Antwort

### **Q09**

2.3) Steht der Programmcode Ihrer Anwendung als Open Source zur Verfügung?

- Ja (z. B. auf OpenCode.de) **3**
- Nein, ist aber geplant **2**
- Nein und ist nicht geplant **1**
- Keine Antwort

### **Q10**

2.4) Ist die Anwendung für andere nachnutzbar?

(Mehrfachantworten möglich.)

- Ja, kopierbar (z. B. weil Open Source) **C1**
- Ja, skalierbar (z. B. weil cloudbasiert und zentral bereitgestellt) **C2**
- Sonstiges, und zwar:\_\_\_\_\_
- Nein [Exklusiv] **C3**
- Keine Antwort [Exklusiv] **C4**

### **Q11**

2.5) Stellen Sie sich vor, eine neue Mitarbeiterin oder ein neuer Mitarbeiter soll zukünftig Ihre Anwendung nutzen. Wie ausführlich ist die Dokumentation zur Nutzung der Anwendung, die ihr oder ihm zur Verfügung steht?

- Sehr ausführlich **4**
- Ausführlich **3**
- Wenig ausführlich **2**
- Es gibt keine Dokumentation zur Nutzung **1**
- Keine Antwort

### **Q12**

2.6) Liegen Zertifikate oder Garantien (z. B. vertraglich festgehalten) des Anbieters hinsichtlich der digitalen

Souveränität der Software vor – etwa hinsichtlich des Betriebsortes oder der Nutzungsrechte an Ein- und Ausgabedaten?

- Ja, und zwar: \_\_\_\_\_ **1**
- Geplant, und zwar: \_\_\_\_\_ **2**
- Nein **3**
- Keine Antwort

**Q13**

2.7) Ist die Anwendung Teil eines bereits bestehenden Ökosystems von Programmen?

Ein Beispiel für ein solches Ökosystem ist etwa Microsoft Office 365 mit seinen Einzelanwendungen Word, Excel, PowerPoint etc.

- Ja, es handelt sich um eine in ein Ökosystem eingebettete Lösung. **1**
- Nein, die Anwendung läuft unabhängig. **2**
- Keine Antwort

**Q14**

[Filter: Q13 == 1]

Um welches Ökosystem handelt es sich?

---

**Q15**

2.8) Lässt sich in der Anwendung zwischen verschiedenen LLMs flexibel nach Bedarf wechseln (z. B. durch ein LLM-Gateway oder Proxies)?

- Ja, auch zwischen verschiedenen Anbietern **3**
- Ja, aber nur zwischen Modellen eines Anbieters **2**
- Nein **1**
- Andere Situation, und zwar: \_\_\_\_\_
- Keine Antwort

**Q16**

[Filter: Q15 == 2 OR Q15 == 3]

Bitte erläutern Sie nach Möglichkeit kurz, wie die technische Umsetzung der freien oder begrenzten Modellwahl realisiert ist:

---

---

### 3. Großes Sprachmodell (LLM): Software und Infrastruktur

In diesem Themenblock möchten wir Ihnen einige Fragen zu dem Sprachmodell stellen, das in Ihrer Anwendung zum Einsatz kommt. Ein großes Sprachmodell (LLM) ist ein (vor-)trainiertes statistisches Modell zur Generierung von Text. Im Rahmen des Fragebogens interessieren das oder die in dem Projekt zum Einsatz kommenden LLMs, beispielsweise GPT4, Llama, Claude oder Teuken.

Möglicherweise läuft das von Ihnen verwendete LLM auf derselben Infrastruktur wie der Rest Ihrer Anwendung. In diesem Fall beantworten Sie bitte dennoch die folgenden Fragen zur Infrastruktur, auch wenn sich die Antworten dann doppeln.

#### Q16

3.1) Wo ist das LLM respektive wo sind die LLMs installiert, also auf welcher Infrastruktur läuft es / laufen sie?

[Mehrfachantworten möglich]

- Lokal bzw. in eigenem Betrieb (im Haus) **C1**
- Öffentlicher IT-DL (z. B. ITZ Bund, Deutsche Verwaltungscloud) **C2**
- LLM-Anbieter (z. B. Open-AI) **C3**
- Public-Cloud-Anbieter **C4**
- Sonstiges, und zwar: \_\_\_\_\_
- Keine Antwort **C5**

#### Q17

[Filter Q16 C4 = „Y“]

Bei welchem Public-Cloud-Anbieter?

- AWS **1**
- Google **2**
- Microsoft **3**
- Deutsche Telekom **4**
- IONOS **5**
- StackIT **6**
- OVH **7**
- Anderer: \_\_\_\_\_
- Keine Antwort

#### Q18

[Filter Q16 C3 == „Y“ OR Q16 C4 = „Y“]

Stehen die Server für das LLM-Modell in der EU?

- Ja **1**
- Nein **2**

- Kann ich (noch) nicht beantworten **3**
- Keine Antwort

**Q19**

3.2) LLMs von welchen Anbietern werden genutzt?  
(Mehrfachantwort möglich)

- GPT von OpenAI **C1**
- Claude von Anthropic **C2**
- Gemini von Google **C3**
- LLaMA von Meta **C4**
- Bert von Google **C5**
- Le Chat von Mistral **C6**
- Teuken-7B von OpenGPT-X **C7**
- DeepSeek von DeepSeek **C8**
- Andere, und zwar: \_\_\_\_\_
- Keine Antwort [Exklusiv] **C9**

**Q19c1 bis Q19c8:** [Filter Q19 C1 bis 8 == „Y“] Welche Version(en)? \_\_\_\_\_

**Q20**

3.3) Welche LLMs von welchen Anbietern wurden (darüber hinaus) auf Ihre Eignung geprüft?  
(Mehrfachauswahl möglich)

- GPT von OpenAI **C1**
- Claude von Anthropic **C2**
- Gemini von Google **C3**
- LLaMA von Meta **C4**
- Bert von Google **C5**
- Le Chat von Mistral **C6**
- Teuken-7B von OpenGPT-X **C7**
- DeepSeek von DeepSeek **C8**
- Andere, und zwar: \_\_\_\_\_
- Keine [Exklusiv] **C9**
- Keine Antwort [Exklusiv] **C10**

**Q20c1 bis Q20c8** [Filter Q20 C1 bis C8 == „Y“] Prüfung welcher Version(en)? \_\_\_\_\_

### Q21

3.4) Mit welchem Gewicht sind die nachfolgenden Kriterien bei der LLM-Auswahl eingegangen?  
Bitte bewerten Sie analog zur Gewichtung von Zuschlagskriterien bei einer Vergabe auf 100 Punkte. Nach dieser Frage wird Ihnen Ihre aktuell vergebene Gesamtpunktzahl angezeigt.

- |                                   |            |
|-----------------------------------|------------|
| _ Geopolitische Aspekte           | <b>C01</b> |
| _ Kosten                          | <b>C02</b> |
| _ Erfahrungen                     | <b>C03</b> |
| _ Leistungsfähigkeit              | <b>C04</b> |
| _ Lizenzmodell                    | <b>C05</b> |
| _ Wechselfähigkeit                | <b>C06</b> |
| _ Gestaltungsfähigkeit            | <b>C07</b> |
| _ Vorgabe des Anwendungsanbieters | <b>C08</b> |
| _ Einfluss auf Anbieter           | <b>C09</b> |
| _ Hausinterne Vorgaben            | <b>C10</b> |
| _ Sonstiges                       | <b>C11</b> |

**Q21Kommentar** Sie haben bisher **[Summe C01 bis C11] Punkte** vergeben.

Wenn Sie die Frage nicht beantworten möchten, dann belassen (respektive stellen) Sie alle Regler auf null.

### Q21Sonstiges

[Filter Q21 C11 > 0]

Sonstige Kriterien für die LLM-Auswahl: \_\_\_\_\_

### Q22

[Filter Q19 C9 != „Y“]

3.5) Für wie ausführlich halten Sie die Informationsbereitstellung Ihres präferierten LLM-Anbieters, um Mitarbeitenden zu ermöglichen, die Funktionsweise für die Nutzung nachzuvollziehen?

- |   |          |
|---|----------|
| <input type="radio"/> Sehr ausführlich            | <b>4</b> |
| <input type="radio"/> Ausführlich                 | <b>3</b> |
| <input type="radio"/> Wenig ausführlich           | <b>2</b> |
| <input type="radio"/> Es gibt keine Informationen | <b>1</b> |
| <input type="radio"/> Kann ich nicht beurteilen   | <b>0</b> |
| <input type="radio"/> Keine Antwort               |          |

### Q23

3.6) Wenn Sie ein kostenpflichtiges LLM nutzen: Nehmen Sie ein bestehendes Preismodell des Anbieters in Anspruch oder handeln Sie einen dedizierten Vertrag für die Nutzung aus?

- Bestehendes Preismodell 1
- Dedizierter Vertrag 2
- Weder noch, sondern: \_\_\_\_\_
- Kein kostenpflichtiges Modell 3
- Keine Antwort

#### 4. Daten

Im letzten Themenblock interessieren uns die Daten, die für die LLM-Anwendung benötigt bzw. genutzt werden. Das können sowohl Eingabedaten wie auch Datenbanken für Fine-Tuning, Retrieval-augmented Generation (RAGs) oder Ähnliches sein. Nicht betrachtet werden Daten, die für das Pre-Training des LLMs Verwendung fanden, es sei denn, es wurde eigenes Pre-Training im Rahmen Ihres Projektes betrieben.

##### Q24

4.1) [Welche Datenquellen werden in Ihrem Projekt genutzt?

(Mehrfachantworten möglich)

- Eigene Daten wie interne Dokumente etc. **C1**
- Öffentlich zugängliche Daten wie die der amtlichen Statistik etc. **C2**
- Kommerzielle Daten von entsprechenden Datenanbietern wie etwa Adressdaten **C3**
- Sonstige Daten, und zwar: \_\_\_\_\_
- Es werden keine Daten im Sinne dieses Fragebogens verwendet. [Exklusiv] **C5**
- Keine Antwort. [Exklusiv] **C6**

##### Q25

[Filter Q24 C5 != „Y“ AND Q24 C6 != „Y“]

Welchen personellen Aufwand bedeutet es, Daten auch nach Einführung der Anwendung zu aktualisieren oder zu ergänzen?

- Nahezu keinen Aufwand **5**
- Geringer Aufwand **4**
- Mittleren Aufwand **3**
- Eher hoher Aufwand **2**
- Sehr hoher Aufwand **1**
- Anders, und zwar: \_\_\_\_\_
- Keine Antwort

##### Q26

[Filter Q24 C5 != „Y“ AND Q24 C6 != „Y“]

Liegen die Daten in einem standardkonformen Format vor?

- Ja, vollständig **3**
- Ja, teilweise **2**
- Nein (weil z. B. proprietär) **1**
- Keine Antwort

## Abschluss

Vielen Dank für Ihren Beitrag zur Steigerung der digitalen Souveränität der öffentlichen Verwaltung im Bereich der KI-Anwendungen!

### **Q27**

Möchten Sie uns noch etwas mitgeben, was wir bei der Bearbeitung der Fragestellung berücksichtigen sollten?

---

### **Q28**

Wir planen, zur Vertiefung einzelner Aspekte qualitative Expert:inneninterviews durchzuführen. Wären Sie grundsätzlich bereit, uns als Expert:in für ein solches vertiefendes Interview zur Verfügung zu stehen? Wenn ja, dann geben Sie hier bitte Ihre Kontaktdaten (Name und E-Mailadresse) an.

---

## Letzte Seite

Herzlichen Dank, dass Sie sich die Zeit genommen haben, unser gemeinsames Vorhaben zu unterstützen.

Die Veröffentlichung der dann frei zugänglichen Studie ist für das späte Frühjahr geplant. Wenn bis dahin noch Fragen aufgekommen sind, wenden Sie sich gerne an [uns](#).