

# 514413-2024 - Vorankündigung – Direktvergabe

Deutschland – Datenbankdienste – Datenzugang zur Twitterdatenbank

OJ S 166/2024 27/08/2024

Freiwillige Ex-ante-Transparenzbekanntmachung

Dienstleistungen

## 1. Beschaffer

---

### 1.1. Beschaffer

Offizielle Bezeichnung: Universität der Bundeswehr München

E-Mail: [zv.i.3@unibw.de](mailto:zv.i.3@unibw.de)

Rechtsform des Erwerbers: Einrichtung des öffentlichen Rechts

Tätigkeit des öffentlichen Auftraggebers: Bildung

## 2. Verfahren

---

### 2.1. Verfahren

Titel: Datenzugang zur Twitterdatenbank

Beschreibung: Die UniBw M beabsichtigt, im Rahmen des Projekts SPARTA einen Zugang zu Twitter-Daten zu beschaffen. Die UniBw M plant hierfür den Einkauf von Daten sowie der Programmier-schnittstellen, mit denen sie auf Daten von Twitter, insbesondere Tweets, zugreifen kann. Die UniBw M benötigt einen direkten und unmittelbaren Zugang zu historischen und Echtzeit-Rohdaten der Tweets, die weder aggregiert, noch vorverarbeitet sind und alle relevanten Metadaten enthalten. Nicht aggregierte Daten sind insoweit wesentlich, da nur diese im vollen Maße für die Forschungszwecke ausgeschöpft werden können. Die Twitter-Daten müssen in großen Mengen automatisiert über eine fest definierte programmierbare Schnittstelle abgerufen werden können. Reine Anzeige von Tweets mit allen Metadaten über eine Benutzeroberfläche, bei der eine menschliche Interaktion erforderlich ist, ist nicht ausreichend. Der automatisierte Abruf soll über die Schnittstelle in großen Batches von Tweets – im Falle der historischen Daten – und in Streams von Tweets – im Falle von Echtzeitdaten – an-hand der dedizierten Filterkriterien erfolgen, so dass diese Daten in weiteren Schritten maschinell verarbeitet und abgespeichert werden können. Darüber hinaus ist erforderlich, dass die Abfrage (Lookup) von einzelnen Tweets und Benutzern sowie deren Followern anhand eindeutiger ID möglich ist. Die bereitgestellte Schnittstelle muss fähig sein, die Daten in JSON-Format auszuliefern. Die Schnittstelle muss zudem ermöglichen, mittels offener und nicht proprietärer Protokolle zu kommunizieren (z.B. HTTP). Die ausgelieferten Tweets bzw. Benutzer müssen Datenfelder gem. der Twitter-Definition des Tweet- bzw. User-Objektes enthalten. Dabei sind bei den Tweets insbesondere hervorzuheben: Textinhalte des Tweets (Kurznachricht), eingeschlossene Entities wie Hashtags, Medien, URLs etc.; Autordaten, Geolocation, Daten oder Referenz des originalen Tweets bei einem Retweet, Reply oder Quote etc. Zur Filterung der Daten im historischen und Echtzeitmodus muss das System eine dedizierte Sprache zur Bildung von Filterregeln anbieten. Diese Abfragesprache (Query Language) muss neben den Operatoren zum Filtern nach Hashtags und Textausdrücken auch diverse andere Operatoren unterstützen: Abfrage von binären Eigenschaften (Bsp: Ist Retweet, Hat Geolocation etc.), Filtern nach erkannter Sprache des Tweets und der Geolocation nach Namen (Bsp: Land, Region etc.) und Radius, Abfrage von erkannten Eigennamen für Personen und Orten (Named-entity), Zähleroperatoren (z.B. Anzahl der Tweets des Autors) u.w. Die Filter-Operatoren müssen mit booleschen Operatoren UND /

ODER / NICHT gruppiert wer-den können, so dass deren Verbindung eine Filterregel mit einer maximalen Länge von ca. 2000 Zeichen ergeben kann. Das System muss fähig sein, mehrere gleichzeitige 100 Filterregeln für jeden Batch im Rahmen der historischen Abfragen und mehrere 1000 Filterregeln für jede einzelne Streaming-Verbindung im Rahmen der Echtzeitabfragen verarbeiten zu können. Die Filterregeln müssen einen eindeutigen Bezeichner tragen können, so dass nach der Abfrage ersichtlich ist, welcher Tweet durch welche Regel eingeholt wurde. Im Hinblick auf den Datendurchsatz muss die Schnittstelle für die historischen Daten bis zu 50 Abfragen pro Tag erlauben sowie mindestens 2 gleichzeitig laufende Abfragen bedienen können. Die Schnittstelle für Echtzeitdaten sollte bis zu 50 gleichzeitige aktive Streaming-Verbindungen bedienen. Die UniBw M muss mittels der APIs spontan und flexibel auf gegenwärtige Entwicklungen reagieren und neue Keywords und Hashtags in die API-Abfrage minutiös einspeisen können. Die geladenen Daten müssen unmittelbar und direkt in die internen Analysepipelines eingeführt werden können. Eine historische Langzeitbetrachtung der Twitter-Daten, insbesondere von Tweets, seit der Gründung von Twitter 2006, sowie das Abrufen von Echtzeit-Tweets muss im großen Umfang (> 10 Mio. Tweets pro Monat) möglich sein. Daten-Löschkonzepte müssen umsetzbar sein, die den Löschanforderungen von Twitter entsprechen.

Kennung des Verfahrens: faf1d432-d018-4066-80a2-4b02701d341f

Interne Kennung: M/HSB1/OT746/SOWI2\_DS

Verfahrensart: Verhandlungsverfahren ohne Aufruf zum Wettbewerb

#### 2.1.1. Zweck

Art des Auftrags: Dienstleistungen

Haupteinstufung (cpv): 72320000 Datenbankdienste

#### 2.1.2. Erfüllungsort

Postanschrift: Werner-Heisenberg-Weg 39

Stadt: Neubiberg

Postleitzahl: 85579

Land, Gliederung (NUTS): München, Landkreis (DE21H)

Land: Deutschland

#### 2.1.4. Allgemeine Informationen

**Rechtsgrundlage:**

Richtlinie 2014/24/EU

## 5. Los

---

### 5.1. Los: LOT-0001

Titel: Datenzugang zur Twitterdatenbank

Beschreibung: Die UniBw M beabsichtigt, im Rahmen des Projekts SPARTA einen Zugang zu Twitter-Daten zu beschaffen. Die UniBw M plant hierfür den Einkauf von Daten sowie der Programmier-schnittstellen, mit denen sie auf Daten von Twitter, insbesondere Tweets, zugreifen kann. Die UniBw M benötigt einen direkten und unmittelbaren Zugang zu historischen und Echtzeit-Rohdaten der Tweets, die weder aggregiert, noch vorverarbeitet sind und alle relevanten Metadaten enthalten. Nicht aggregierte Daten sind insoweit wesentlich, da nur diese im vollen Maße für die Forschungszwecke ausgeschöpft werden können. Die Twitter-Daten müssen in großen Mengen automatisiert über eine fest definierte programmierbare Schnittstelle abgerufen werden können. Reine Anzeige von Tweets mit allen Metadaten über eine Benutzeroberfläche, bei der eine menschliche Interaktion erforderlich ist, ist nicht ausreichend. Der automatisierte Abruf soll über die Schnittstelle in großen Batches

von Tweets – im Falle der historischen Daten – und in Streams von Tweets – im Falle von Echtzeitdaten – an-hand der dedizierten Filterkriterien erfolgen, so dass diese Daten in weiteren Schritten maschinell verarbeitet und abgespeichert werden können. Darüber hinaus ist erforderlich, dass die Abfrage (Lookup) von einzelnen Tweets und Benutzern sowie deren Followern anhand eindeutiger ID möglich ist. Die bereitgestellte Schnittstelle muss fähig sein, die Daten in JSON-Format auszuliefern. Die Schnittstelle muss zudem ermöglichen, mittels offener und nicht proprietärer Protokolle zu kommunizieren (z.B. HTTP). Die ausgelieferten Tweets bzw. Benutzer müssen Datenfelder gem. der Twitter-Definition des Tweet- bzw. User-Objektes enthalten. Dabei sind bei den Tweets insbesondere hervorzuheben: Textinhalte des Tweets (Kurznachricht), eingeschlossene Entities wie Hashtags, Medien, URLs etc.; Autordaten, Geolocation, Daten oder Referenz des originalen Tweets bei einem Retweet, Reply oder Quote etc. Zur Filterung der Daten im historischen und Echtzeitmodus muss das System eine dedizierte Sprache zur Bildung von Filterregeln anbieten. Diese Abfragesprache (Query Language) muss neben den Operatoren zum Filtern nach Hashtags und Textausdrücken auch diverse andere Operatoren unterstützen: Abfrage von binären Eigenschaften (Bsp: Ist Retweet, Hat Geolocation etc.), Filtern nach erkannter Sprache des Tweets und der Geolocation nach Na-men (Bsp: Land, Region etc.) und Radius, Abfrage von erkannten Eigennamen für Personen und Orten (Named-entity), Zähleroperatoren (z.B. Anzahl der Tweets des Autores) u.w. Die Filter-Operatoren müssen mit booleschen Operatoren UND / ODER / NICHT gruppiert werden können, so dass deren Verbindung eine Filterregel mit einer maximalen Länge von ca. 2000 Zeichen ergeben kann. Das System muss fähig sein, mehrere gleichzeitige 100 Filterregeln für jeden Batch im Rahmen der historischen Abfragen und mehrere 1000 Filterregeln für jede einzelne Streaming-Verbindung im Rahmen der Echtzeitabfragen verarbeiten zu können. Die Filterregeln müssen einen eindeutigen Bezeichner tragen können, so dass nach der Ab-frage ersichtlich ist, welcher Tweet durch welche Regel eingeholt wurde. Im Hinblick auf den Datendurchsatz muss die Schnittstelle für die historischen Daten bis zu 50 Abfragen pro Tag erlauben sowie mindestens 2 gleichzeitig laufende Abfragen bedienen können. Die Schnittstelle für Echtzeitdaten sollte bis zu 50 gleichzeitige aktive Streaming-Verbindungen bedienen. Die UniBw M muss mittels der APIs spontan und flexibel auf gegenwärtige Entwicklungen reagieren und neue Keywords und Hashtags in die API-Abfrage minutiös einspeisen können. Die geladenen Daten müssen unmittelbar und direkt in die internen Analysepipelines eingeführt werden können. Eine historische Langzeitbetrachtung der Twitter-Daten, insbesondere von Tweets, seit der Gründung von Twitter 2006, sowie das Abrufen von Echtzeit-Tweets muss im großen Umfang (> 10 Mio. Tweets pro Monat) möglich sein. Daten-Löschkonzepte müssen umsetzbar sein, die den Löschanforderungen von Twitter entsprechen.

#### **5.1.1. Zweck**

Art des Auftrags: Dienstleistungen  
Haupteinstufung (cpv): 72320000 Datenbankdienste  
Menge: 1 Stück

#### **5.1.6. Allgemeine Informationen**

Die Beschaffung fällt unter das Übereinkommen über das öffentliche Beschaffungswesen: ja

#### **5.1.10. Zuschlagskriterien**

##### **Kriterium:**

Art: Preis

Bezeichnung: Preis

Beschreibung: Zuschlagskriterium ist der Preis

Kategorie des Gewicht-Zuschlagskriteriums: Gewichtung (Prozentanteil, genau)

#### 5.1.16. Weitere Informationen, Schlichtung und Nachprüfung

Überprüfungsstelle: Bundeskartellamt - Vergabekammer des Bundes

Informationen über die Überprüfungsfristen: Die Frist zur Einreichung eines Nachprüfungsantrages beträgt zehn Kalendertage ab dem Tag nach der Veröffentlichung dieser ex-ante Bekanntmachung. Auf § 135 Abs.3 GWB wird hiermit hingewiesen.

Organisation, die zusätzliche Informationen über das Vergabeverfahren bereitstellt: Universität der Bundeswehr München

Organisation, die einen Offline-Zugang zu den Vergabeunterlagen bereitstellt: Universität der Bundeswehr München

Organisation, die weitere Informationen für die Nachprüfungsverfahren bereitstellt: Bundeskartellamt - Vergabekammer des Bundes

## 6. Ergebnisse

---

Wert aller in dieser Bekanntmachung vergebenen Verträge: 640 000,00 EUR

### Direktvergabe

:

Begründung der Direktvergabe: Der Auftrag kann nur von einem bestimmten Wirtschaftsteilnehmer ausgeführt werden, da aus technischen Gründen kein Wettbewerb vorhanden ist

Sonstige Begründung: Zur Bearbeitung benötigt die UniBw M einen direkten und unmittelbaren Zugang zu historischen und Echtzeit-Rohdaten der Tweets, die weder aggregiert, noch vorverarbeitet sind und alle relevanten Metadaten enthalten. Nicht aggregierte Daten sind insoweit wesentlich, da nur diese im vollen Maße für die Forschungszwecke ausgeschöpft werden können. Jegliche Aggregation durch etwaige Vermittler schränkt die Datenanalyse stark ein, was nicht im Sinne des Forschungsprojektes ist. Daher ist die Bereitstellung der nicht aggregierten Tweets-Daten ein Alleinstellungsmerkmal der hier zu beschaffenden Twitter APIs, die ausschließlich durch Twitter direkt bereitgestellt werden können. Twitter verbietet explizit jede Weitergabe bzw. Publikation von nicht aggregierten Twitter-Daten in großen Maßen bei der Zulassung von Dritten zu seiner API, so dass damit automatisch ein Monopol auf diese Art von Daten besteht. Aus technischen Gründen steht der Auftragnehmer – Twitter – von vornherein fest. Denn nur Twitters Datenzugang kann die Anforderungen der UniBw M kumulativ erfüllen. Ein Wettbewerb scheidet von vornherein aus, da ein Monopol auf nicht aggregierte Twitter-Rohdaten besteht, die im Rahmen des Forschungsprojektes erforderlich sind.

### 6.1. Ergebnis, Los— Kennung: LOT-0001

#### 6.1.2. Informationen über die Gewinner

##### Wettbewerbsgewinner:

Offizielle Bezeichnung: Twitter Inc.

##### Angebot:

Kennung des Angebots: Angebot Twitter

Kennung des Loses oder der Gruppe von Losen: LOT-0001

Wert der Ausschreibung: 640 000,00 EUR

Das Angebot wurde in die Rangfolge eingeordnet: ja

Rang in der Liste der Gewinner: 1

##### Informationen zum Auftrag:

Kennung des Auftrags: M/HSB1/OT746/SOWI2\_DS

Datum der Auswahl des Gewinners: 26/08/2024

**Angaben zu Mitteln der Europäischen Union:**

Organisation, die den Auftrag unterzeichnet: Universität der Bundeswehr München

## 8. Organisationen

---

### 8.1. ORG-0001

Offizielle Bezeichnung: Universität der Bundeswehr München

Abteilung: ZV I.3

Postanschrift: Werner-Heisenberg-Weg 39

Stadt: Neubiberg

Postleitzahl: 85579

Land, Gliederung (NUTS): München, Landkreis (DE21H)

Land: Deutschland

E-Mail: [zv.i.3@unibw.de](mailto:zv.i.3@unibw.de)

Telefon: +49 8960044015

**Rollen dieser Organisation:**

Beschaffer

Zentrale Beschaffungsstelle, die öffentliche Aufträge oder Rahmenvereinbarungen im Zusammenhang mit für andere Beschaffer bestimmten Bauleistungen, Lieferungen oder Dienstleistungen vergibt/abschließt

Organisation, die zusätzliche Informationen über das Vergabeverfahren bereitstellt

Organisation, die einen Offline-Zugang zu den Vergabeunterlagen bereitstellt

Organisation, die den Auftrag unterzeichnet

### 8.1. ORG-0002

Offizielle Bezeichnung: Bundeskartellamt - Vergabekammer des Bundes

Stadt: Bonn

Postleitzahl: 53123

Land, Gliederung (NUTS): Bonn, Kreisfreie Stadt (DEA22)

Land: Deutschland

Telefon: +49 22894990

Fax: +49 2289499163

**Rollen dieser Organisation:**

Überprüfungsstelle

Organisation, die weitere Informationen für die Nachprüfungsverfahren bereitstellt

### 8.1. ORG-0003

Offizielle Bezeichnung: Twitter Inc.

Postanschrift: 1355 Market Street Suite 900

Stadt: San Francisco

Postleitzahl: CA 94103

Land: Vereinigte Staaten

Telefon: +49 22894990

Fax: +49 2289499163

**Rollen dieser Organisation:**

Bieter

**Gewinner dieser Lose: LOT-0001**

## Informationen zur Bekanntmachung

---

Kennung/Fassung der Bekanntmachung: 09d91e1b-6377-41e7-a364-eb4b40b4e022 - 01  
Formulartyp: Vorankündigung – Direktvergabe  
Art der Bekanntmachung: Freiwillige Ex-ante-Transparenzbekanntmachung  
Unterart der Bekanntmachung: 25  
Datum der Übermittlung der Bekanntmachung: 26/08/2024 07:25:46 (UTC+00:00)  
Westeuropäische Zeit, GMT  
Sprachen, in denen diese Bekanntmachung offiziell verfügbar ist: Deutsch  
Veröffentlichungsnummer der Bekanntmachung: 514413-2024  
ABl. S – Nummer der Ausgabe: 166/2024  
Datum der Veröffentlichung: 27/08/2024