

Nr.	Beschreibung Anforderung / Kriterium	Art der Anforderung: Ausschluss (A) Bewertung (B) Ohne Bewertung (O)	Maximal Erreichbare Punkte (B-Kriterien)	Anbieter 1		Erreichte Punkte (B-Kriterien)
				Anforderung erfüllt (ja/nein)	Vom Bieter auszufüllen: Erläuterung zum Kriterium	

Kapitel 1.1: Lieferung und Leistungen - Übersicht						
1	Bereitstellung einer Speech-to-Text- Engine über eine API (STT-Engine) für die deutsche und englische Sprache als „Software as a Service“ in einem DSGVO (GDPR) konformen Rechenzentrum mit Standort innerhalb der Europäischen Union	A	nicht ausfüllen			nicht ausfüllen
2	Bereitstellung der Dokumentation der STT-Engine API	A	nicht ausfüllen			nicht ausfüllen
3	Durchführung eines minutengenauen Reportings und Abrechnung der Nutzung des STT-Service im aktuellen Rechnungsmonat per „Dashboard“	A	nicht ausfüllen			nicht ausfüllen
4	Optional: Versand einer Email durch den AN an den AG bei Erreichen bestimmter Schwellenwerte im aktuellen Rechnungsmonat	B	1			Wird vom ZDF aufgefüllt
5	Mengengerüste: Gleichzeitige Entgegennahme und gleichzeitiges Processing von mindestens 5 A/V Dateien durch die STT-Engine API	A	nicht ausfüllen			nicht ausfüllen
6	Processing-Zeit (bei 5 gleichzeitigen Tasks) für die Erstellung von Transkripten: ½ Echtzeit (d.h. das Transkript eines Videos mit 10 Minuten Länge soll in 5 Minuten fertiggestellt sein)	A	nicht ausfüllen			nicht ausfüllen
7	Response-Zeiten (Anforderungen): - Beginn der Erstellung des Transkripts: max. 60 Sekunden nach vollständiger Übergabe des A/V Materials an die STT-API des AN	A	nicht ausfüllen			nicht ausfüllen
8	Response-Zeiten (Anforderungen): - Übergabe des erstellten Transkripts per JSON oder XML Dokument an den AG: max. 60 Sekunden nach Fertigstellung (Processing) des Transkripts	A	nicht ausfüllen			nicht ausfüllen
9	Der AN verbessert das deutsche und englische Sprachmodell kontinuierlich weiter und stellt es für den AG bereit ohne Ausfallzeiten bereit. Sollten Ausfallzeiten erforderlich sein, sind diese mit einem Vorlauf von 7 Tagen dem AG mitzuteilen	A	nicht ausfüllen			nicht ausfüllen
10	Der AN stellt während der Vertragslaufzeit und insbesondere in der Implementierungsphase einen Projektkoordinator bzw. Ansprechpartner bereit.	A	nicht ausfüllen			nicht ausfüllen
11	Optional: Übergabe von Status Informationen über die STT-API zur Errechnung des Abarbeitungsgrads bzw. zur grafischen Anzeige des Fortschritts (z.B. „Übertragung gestartet“, „Processing gestartet“, „Aufgabe abgeschlossen“)	B	3			Wird vom ZDF aufgefüllt

Kapitel 1.2: Technische Anforderungen						
12	Die STT-Engine muss über eine REST-basierte API mit folgenden Eigenschaften verfügen und mit Angebotsabgabe dem AG bereitgestellt werden: - Nutzung von https für sämtliche Kommunikation mit der STT-API - Nutzung von URLs / URI - Verwendung von Standard http Methoden (GET/POST/PUT/PATCH/DELETE)	A	nicht ausfüllen			nicht ausfüllen
13	Die erstellten Transkripte müssen Timecodegenau sein (Basis: 25fps).	A	nicht ausfüllen			nicht ausfüllen
14	Übergabe/Bereitstellung der Timecode-genauen Transkripte als JSON oder XML Dokumente über die STT-API an die ZDF MAM Schnittstelle	A	nicht ausfüllen			nicht ausfüllen
15	Optional: Möglichkeit der Auswahl der zu transkribierenden Tonspuren via API (sofern mehrere Tonspuren im Video vorhanden sind)	B	2			Wird vom ZDF aufgefüllt
16	Unterstützung der folgenden Audio/Video Formate - MP4 Video-Files mit bis zu 8 Audiospuren (für die Transkription relevante Audiospuren i.d.R. Spuren 1 + 2) - WAV - MP3	A	nicht ausfüllen			nicht ausfüllen
17	Optional: Unterstützung von Videofiles mit h.264 Codec im MXF-Container mit bis zu 16 Tonspuren (für die Transkription relevante Audiospuren i.d.R. Spuren 1 + 2)	B	1			Wird vom ZDF aufgefüllt
18	Optional: Möglichkeit zur Anbindung eines Cloud Object Stores	B	1			Wird vom ZDF aufgefüllt

Kapitel 1.3: Funktionale Anforderungen						
19	Erstellung einer möglichst korrekten Interpunktion der Transkripte für die deutsche und englische Sprache (Kommata werden nicht erwartet).	A	nicht ausfüllen			nicht ausfüllen
20	Timecodegenaue Transkription der gesprochenen Inhalte im A/V Material, d.h. je Satz wird ein Timecode-Bereich bereitgestellt, der Beginn und Ende des Satzes umfasst - „TC Start des Satzes“ - „TC Ende des Satzes“	A	nicht ausfüllen			nicht ausfüllen
21	Optional: Automatische Erkennung der Sprache des zu transkribierenden Videos (Deutsch oder Englisch)	B	2			Wird vom ZDF aufgefüllt

Kapitel 1.4: Qualitätskriterien zur Bewertung der erstellten Transkripte						
22	Mindestens 3 der deutsch- und englischsprachigen Videos (von den jeweils 5 durch das ZDF bereitgestellten Videos) müssen eine quantitative Erkennungsquote von mindestens 80% bis höchstens 120 % aufweisen. Ist dies nicht der Fall, wird der AN aus dem Verfahren ausgeschlossen.	A	nicht ausfüllen	Wird vom ZDF aufgefüllt	nicht ausfüllen	nicht ausfüllen
23	Vorausgesetzt, dass mindestens 3 der deutsch- und englischsprachigen Videos die Schwellenwerte für die quantitative Erkennungsquote nicht unter- bzw. überschritten haben, werden Punkte nach folgendem Schema vergeben: - Erkennungsquote 80% = 0 Punkte - Erkennungsquote 81% = 1 Punkt - Erkennungsquote 82% = 2 Punkte - Erkennungsquote 83% = 3 Punkte -usw.... - Erkennungsquote 100% = 20 Punkte Ab 101% Erkennungsquote erfolgt je Prozentpunkt ein Punkt Abzug von 20 Punkten, d.h. - Erkennungsquote 101% = 19 Punkte - Erkennungsquote 102% = 18 Punkte - Erkennungsquote 103% = 17 Punkte - ... usw bis maximal Abzug von -20 Punkten = 0 Punkte bei einer Erkennungsquote von 120% Hinweise: - Auch wenn alle 5 bereitgestellten Videos die oben angegebenen Schwellenwerte erreichen, werden zur Sicherung der Vergleichbarkeit je Bieter nur die besten 3 Videos zur Bewertung herangezogen. - Eine Erkennungsquote über 100% liegt dann vor, wenn die STT-Engine mehr Wörter erkennt und transkribiert, als ursprünglich im Video vorhanden (gesprochen).	B	20	Wird vom ZDF aufgefüllt	nicht ausfüllen	Wird vom ZDF aufgefüllt
24	Mindestens 3 der deutsch- und englischsprachigen Videos (von den jeweils 5 durch das ZDF bereitgestellten Videos) müssen eine qualitative Erkennungsquote von mindestens 80% aufweisen. Ist dies nicht der Fall, wird der AN aus dem Verfahren ausgeschlossen.	A	nicht ausfüllen	Wird vom ZDF aufgefüllt	nicht ausfüllen	nicht ausfüllen

Nr.	Beschreibung Anforderung / Kriterium	Art der Anforderung: Ausschluss (A) Bewertung (B) Ohne Bewertung (O)	Maximal Erreichbare Punkte (B-Kriterien)	Anbieter 1		Erreichte Punkte (B-Kriterien)
				Anforderung erfüllt (ja/nein)	Vom Bieter auszufüllen: Erläuterung zum Kriterium	
25	Vorausgesetzt, dass mindestens 3 der deutsch- und englischsprachigen Videos die Schwellwerte für die qualitative Erkennungsquote nicht unterschritten haben, werden Punkte nach folgendem Schema vergeben: - Erkennungsquote 80% = 0 Punkte - Erkennungsquote 81% = 2 Punkte - Erkennungsquote 82% = 4 Punkte - Erkennungsquote 83% = 6 Punkte -usw.... - Erkennungsquote 100% = 40 Punkte Hinweis: Auch wenn alle 5 bereitgestellten Videos die oben angegebenen Schwellwerte erreichen, werden zur Sicherung der Vergleichbarkeit je Bieter nur die besten 3 Videos zur Bewertung herangezogen.	B	40	Wird vom ZDF aufgefüllt	nicht ausfüllen	Wird vom ZDF aufgefüllt

Summe Punkte B-Kriterien		70			0
Bewertung A-Kriterien (Alle erfüllt ja/nein) (Nein führt zum Ausschluss vom Verfahren)					

Kapitel 1.5: Optionales Webfrontend und optionale mobile App						
26	Bitte beschreiben Sie ob und wenn ja in welcher Form browserbasierte Webfrontends und/oder mobile Apps für iOS und Android für den Upload und den Erhalt von Transkripten bei Ihnen vorhanden sind und dem ZDF perspektivisch zur Verfügung gestellt werden können. Bitte fügen Sie ihre Beschreibung als Anhang zu dieser Ausschreibung bei. Hinweis: Die Bereitstellung von Webfrontends und/oder mobilen Apps für die Erstellung von Transkripten hat lediglich informativen Charakter und geht nicht in die Bewertung ein.	O	nicht ausfüllen			nicht ausfüllen

Kapitel 1.6: Preise und Abrechnung						
Die Bezahlung der Transkriptionsleistung soll nach tatsächlicher Nutzung abgerechnet werden. Der AG garantiert keine Mindestnutzung des STT-Services. Der AG bittet, Preise für die folgenden erwarteten monatlichen Nutzungskontingente (Mengengerüste) unten anzugeben:		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		nicht ausfüllen
0-100 Stunden zu transkribierendes Audio- und Videomaterial pro Monat		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
101-200 Stunden zu transkribierendes Audio- und Videomaterial pro Monat		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
201-300 Stunden zu transkribierendes Audio- und Videomaterial pro Monat		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
301-400 Stunden zu transkribierendes Audio- und Videomaterial pro Monat		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
401-500 Stunden zu transkribierendes Audio- und Videomaterial pro Monat		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
501-600 Stunden zu transkribierendes Audio- und Videomaterial pro Monat		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
601-700 Stunden zu transkribierendes Audio- und Videomaterial pro Monat		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
701-800 Stunden zu transkribierendes Audio- und Videomaterial pro Monat		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
801-900 Stunden zu transkribierendes Audio- und Videomaterial pro Monat		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
901-1000 Stunden zu transkribierendes Audio- und Videomaterial pro Monat		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		nicht ausfüllen
Durchschnittspreis		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		0,00 €
Durchschnittspreis günstigster Anbieter		Wird vom ZDF aufgefüllt	nicht ausfüllen	nicht ausfüllen		nicht ausfüllen
		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		nicht ausfüllen
Preisunterschied zum günstigsten Anbieter (in Prozent)		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		#WERT!
		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		nicht ausfüllen
Punkte Preis Berechnung (sofern kleiner 0 wird die Punktzahl in der nachfolgenden Zeile auf 0 gesetzt)		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		#WERT!
Punkte Preis		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		#WERT!
Hinweis: der günstigste Durchschnittspreis erhält die volle Punktzahl (30 Punkte), alle weiteren Bieter erhalten einen Punktabzug in Höhe des prozentualen Unterschieds, ab 100% Preisunterschied erhält der teurere Bieter 0 Punkte (es werden keine "negativen" Punkte vergeben).		nicht ausfüllen	nicht ausfüllen	nicht ausfüllen		nicht ausfüllen

Gesamtpunktzahl					#WERT!
------------------------	--	--	--	--	---------------

Hinweis:

Eintragungen durch den Bieter sind ausschließlich in den gelb markierten Feldern zu machen. Das Dokument wird elektronisch weiterverarbeitet und ist dem Angebot daher im Excel-Format (.xlsx) beizufügen.

Für die anteilige preisliche Wertung des Angebotes ist der Durchschnittspreis 0 bis 1000 Stunden / Monat (Zelle G64) maßgeblich.

Erläuterungen von A-Kriterien durch den Bieter sind wünschenswert. Für B-Kriterien sind Erläuterungen durch den Bieter obligatorisch. Die Nichterfüllung eines A-Kriteriums führt zum Ausschluss des Angebotes.