Deutscher Bundestag **Drucksache 21/9341**

21. Wahlperiode 20.11.2026

Vorabfassung - wird durch die lektorierte Version ersetzt.

Antwort

der Bundesregierung

auf die Kleine Anfrage der Abgeordneten Clara Bünger, Anne-Mieke Bremer, Agnes Conrad, weiterer Abgeordneter und der Fraktion Die Linke
– Drucksache 21/9007 –

Einsatz der Dialekterkennungssoftware DIAS im Bundesamt für Migration und Flüchtlinge

Vorbemerkung der Fragesteller

Seit 2017 setzt das Bundesamt für Migration und Flüchtlinge (BAMF) verschiedene IT-Assistenzsysteme ein, die zur Klärung der Herkunft und Identität von Asylsuchenden beitragen sollen. Dabei werden zum Beispiel mobile Datenträger von Geflüchteten ausgewertet oder Sprachproben von Asylsuchenden mittels der Dialekterkennungssoftware DIAS analysiert (Antwort der Bundesregierung auf die Kleine Anfrage auf Bundestagsdrucksache 20/3238).

Beim Einsatz der Sprachbiometrie müssen Asylsuchende circa zwei Minuten in ihrer Muttersprache ein Bild beschreiben und dabei in ein Telefon sprechen. Was sie sagen, wird aufgezeichnet. Die Sprechproben werden anschließend automatisch mit einem zugrundeliegenden Sprachmodell abgeglichen. So werden Wahrscheinlichkeiten für den gesprochenen Dialekt berechnet und in einem Ergebnisbericht festgehalten (www.heise.de/news/Sherlock-Bamf-Asyltaeuschern-auf-der-Spur-4257898.html?seite=all). Die Dialekterkennungssoftware wurde anfangs in erster Linie bei arabischsprachigen Asylsuchenden angewendet. Im Juli 2022 wurde sie um die Sprachen Dari, Persisch und Paschto erweitert (Antwort zu Frage 3 auf Bundestagsdrucksache 20/3238). Anfang 2025 teilte die Bundesregierung auf eine Anfrage nach dem Informationsfreiheitsgesetz mit, dass das Paschto-Sprachmodell zwar weiter aktiv sei, allerdings werde die Sprach- und Dialekterkennung bei Asylsuchenden, die als erste Muttersprache Paschto angeben, seit April 2023 nicht mehr eingesetzt. Der Grund liege in der Verfahrensökonomie. Es sei entschieden worden, bei der großen Zahl afghanischer Antragsteller auf den Einsatz zu verzichten, um die Verfahren zu beschleunigen (https://fragdenstaat.de/anfrage/use-of-the-dialektidentifizierungsassistent-dias-in-the-asylum-procedure/#nachricht-970273, Antwort zu Frage 2 auf Bundestagsdrucksache 20/3238). Das wundert die Fragestellenden, weil die Bundesregierung an anderer Stelle betonte, dass das Dialekterkennungstool ein „Unterstützungsinstrument“ sei, das den Entscheidern die Arbeit erleichtern solle (ebd., Antwort zu Frage 8).

Die Bundesregierung hebt hervor, dass die Ergebnisse der DIAS-Analyse lediglich ein Indiz neben anderen für die Herkunft der Antragstellenden seien. Sie dürften somit nicht als einziges Kriterium bei der Entscheidung für oder gegen den Schutzstatus eines Geflüchteten zugrunde gelegt werden (ebd.). Er-

---

Die Antwort wurde namens der Bundesregierung mit Schreiben des Bundesministeriums des Innern vom 20. November 2026 übermittelt.
Die Drucksache enthält zusätzlich – in kleinerer Schrifttype – den Fragetext.

fahrungen aus der Praxis und Einschätzungen von Expertinnen deuten jedoch darauf hin, dass das Ergebnis der Dialektanalyse häufig einen großen Einfluss auf die Asylentscheidung hat. Die Entscheider neigten dazu, solche Tools als unfehlbar anzusehen und die Ergebnisse nicht ausreichend kritisch zu hinterfragen (https://verfassungsblog.de/dialect-recognition-software-dias-law/).

Vorbemerkung der Bundesregierung

Das Bundesamt für Migration und Flüchtlinge (BAMF) setzt den Dialektidentifizierungsassistenten (DIAS) im Rahmen seiner Zuständigkeit für das Asylverfahren ein. Der DIAS liefert auf Grundlage einen Hinweis auf die gesprochene Sprache bzw. den Dialekt. Die Asylsachbearbeitenden des BAMF können anhand dieser Hinweise in der Anhörung gezielte Fragen zur möglichen Plausibilisierung des angegebenen Herkunftslands stellen. Die DIAS-Ergebnisse können die Angaben zur Herkunft der Antragstellenden weder bestätigen noch widerlegen und haben lediglich ergänzenden Hinweischarakter, an die Asylsachbearbeitenden des BAMF nicht gebunden sind. Die abschließende Feststellung im Einzelfall erfolgt auf Grundlage einer Gesamtwürdigung aller im Einzelfall vorliegenden Erkenntnisse.

1. In wie vielen Fällen wurde die Dialekterkennungssoftware des BAMF seit 2022 eingesetzt (bitte nach Jahren, BAMF-Standorten und den Herkunftsländern sowie dem Geschlecht der betroffenen Asylsuchenden aufschlüsseln, bitte zusätzlich für alle Jahre die Zahl der durchgeführten Asylverfahren als Vergleichsgröße nennen)?

Die Angaben sind den nachfolgenden Tabellen zu entnehmen.

Anzahl der Asylverfahren als Vergleichsgröße	
Jahr	Erstanträge
Jahr 2022	217.774
Jahr 2023	329.120
Jahr 2024	229.751
Jahr 2025	113.236
Jahr 2026 bis 31.10.2026	56.942

Anzahl der Einsätze des DIAS 2022 bis 31.10.2026	
Jahr	Anzahl
2022	28.595
2023	43.231
2024	5.376
2025	2.942
2026	2.983

Aufschlüsselung der DIAS-Einsätze nach BAMF-Außenstellen						
Außenstelle	2022	2023	2024	2025	2026	Gesamtergebnis
52C AS Heidelberg im AZ	7,84 %	8,29 %	30,78 %	24,10 %	16,55 %	10,59 %
52G AS Trier, LAS	5,24 %	7,30 %	0,12 %	0,23 %	2,61 %	5,39 %
51A AS Gießen im AZ, LAS	4,98 %	7,64 %	0,22 %	0,04 %	0,33 %	5,33 %
43C AS Eisenhüttenstadt, LAS	3,54 %	6,23 %	2,86 %	0,04 %	0,95 %	4,32 %
51G AS Halberstadt im AZ, LAS	3,33 %	1,93 %	12,98 %	19,23 %	7,21 %	4,22 %
43G AS Schwerin im AZ, LAS	2,83 %	1,41 %	18,35 %	13,03 %	7,02 %	3,93 %
42F AS Essen	3,52 %	5,08 %	0,02 %	0,45 %	1,56 %	3,69 %

Vorabfassung - wird durch die lektorierte Version ersetzt.

Aufschlüsselung der DIAS-Einsätze nach BAMF-Außenstellen						
Außenstelle	2022	2023	2024	2025	2026	Gesamt-ergebnis
42B AS Unna im AZ	3,67 %	4,55 %	0,20 %	1,02 %	0,19 %	3,55 %
42H AS Bonn im AZ	3,74 %	4,23 %	0,02 %	0,60 %	0,05 %	3,43 %
43D AS Chemnitz im AZ, LAS	3,81 %	3,69 %	0,06 %	0,00 %	3,79 %	3,34 %
42C AS Bielefeld im AZ	4,00 %	3,21 %	0,02 %	0,83 %	4,69 %	3,30 %
42E AS Mönchengladbach im AZ	3,53 %	4,01 %	0,00 %	0,15 %	0,47 %	3,24 %
53F AS Regensburg in AnkER	2,92 %	4,35 %	0,18 %	0,11 %	1,33 %	3,12 %
52I AS Lebach in AnkER, LAS	3,45 %	2,82 %	0,00 %	0,00 %	3,84 %	2,83 %
41F AS Braunschweig2 im AZ	1,80 %	0,47 %	13,38 %	12,39 %	5,55 %	2,65 %
51F AS Suhl im AZ	2,54 %	2,34 %	1,48 %	3,93 %	9,20 %	2,64 %
43F AS Dresden in AnkER	2,90 %	2,32 %	0,04 %	0,00 %	1,85 %	2,32 %
43B AS Berlin im AZ	2,34 %	2,73 %	0,08 %	0,00 %	2,42 %	2,24 %
53H AS Schweinfurt in AnkER	2,75 %	1,99 %	0,08 %	0,04 %	5,36 %	2,23 %
41C AS Neumünster, LAS	2,93 %	1,76 %	0,16 %	0,15 %	3,32 %	2,18 %
53D AS Augsburg in AnkER	2,43 %	2,02 %	0,34 %	0,53 %	3,65 %	2,08 %
43E AS Leipzig im AZ	1,01 %	1,13 %	10,37 %	7,82 %	3,27 %	2,07 %
52H AS Speyer	1,92 %	1,99 %	4,03 %	0,08 %	0,05 %	1,97 %
53A AS München	1,86 %	2,49 %	0,14 %	0,30 %	2,23 %	1,93 %
53G AS Deggendorf in AnkER	1,80 %	2,64 %	0,02 %	0,11 %	0,00 %	1,88 %
31D 31D Nürnberg	3,32 %	0,12 %	0,10 %	0,11 %	0,28 %	1,62 %
53B AS Manching in AnkER	1,49 %	1,67 %	0,22 %	0,00 %	0,24 %	1,37 %
41A AS Hamburg im AZ, LAS	1,43 %	1,51 %	0,04 %	0,23 %	0,66 %	1,29 %
41G AS Bramsche im AZ	1,15 %	1,90 %	0,04 %	0,04 %	0,09 %	1,28 %
42D AS Düsseldorf	1,03 %	1,92 %	0,02 %	0,23 %	0,47 %	1,26 %
53E AS Zirndorf in AnkER	1,55 %	0,82 %	0,04 %	0,45 %	5,93 %	1,25 %
51B AS Büdingen	2,14 %	0,46 %	0,08 %	0,00 %	0,00 %	1,19 %
53C AS Bamberg in AnkER	1,38 %	0,76 %	0,26 %	0,08 %	0,05 %	0,97 %
43A AS Berlin, LAS	0,29 %	1,64 %	0,22 %	0,08 %	0,09 %	0,80 %
41B AS Bremen im AZ, LAS	0,52 %	0,68 %	0,58 %	0,00 %	0,00 %	0,55 %
32F 32F Dublinzentrum Bayreuth	0,91 %	0,09 %	0,50 %	0,91 %	0,43 %	0,55 %
52D AS Ellwangen	0,38 %	0,29 %	0,66 %	4,00 %	0,24 %	0,50 %
41H AS Oldenburg	0,52 %	0,48 %	0,04 %	0,04 %	0,76 %	0,46 %
41D AS Friedland, LAS	0,17 %	0,53 %	0,40 %	0,79 %	0,90 %	0,38 %
42A AS Bochum, LAS	0,22 %	0,09 %	0,08 %	5,82 %	0,00 %	0,37 %
32E 32E Dublinzentrum Bochum	0,63 %	0,04 %	0,10 %	0,34 %	0,52 %	0,34 %
32D 32D Dublinzentrum Berlin	0,60 %	0,03 %	0,08 %	0,26 %	0,81 %	0,34 %
52E AS Freiburg	0,37 %	0,03 %	0,14 %	0,38 %	0,28 %	0,22 %
52F AS Sigmaringen	0,19 %	0,11 %	0,26 %	0,68 %	0,05 %	0,18 %
51E AS Jena/Hermsdorf, LAS	0,27 %	0,08 %	0,02 %	0,08 %	0,00 %	0,16 %
52B AS Karlsruhe, LAS	0,15 %	0,12 %	0,16 %	0,15 %	0,52 %	0,15 %
41E AS Braunschweig	0,30 %	0,00 %	0,02 %	0,04 %	0,09 %	0,14 %
51D AS Neustadt	0,17 %	0,00 %	0,00 %	0,00 %	0,00 %	0,08 %
61D/E 61D/E Nürnberg	0,04 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
Unbekannt	0,02 %	0,00 %	0,00 %	0,08 %	0,09 %	0,01 %
51C AS Frankfurt/Flughafen	0,02 %	0,00 %	0,00 %	0,00 %	0,00 %	0,01 %
31B 31B Nürnberg	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,01 %
Zustellzentrum Saarbrücken	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,01 %
32B 32B Nürnberg	0,01 %	0,00 %	0,00 %	0,04 %	0,00 %	0,01 %
32C 32C Dublinverfahren	0,00 %	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %
AS Kusel	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %

Vorabfassung - wird durch die lektorierte Version ersetzt.

Aufschlüsselung der DIAS-Einsätze nach BAMF-Außenstellen						
Außenstelle	2022	2023	2024	2025	2026	Gesamt-ergebnis
31A 31A Nürnberg	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
AS Ingelheim/Bingen	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
AS Hermeskeil	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Zustellzentrum Bonn	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
71F 71F Lagezentrum	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
AS Reutlingen/Eningen	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
32A 32A Nürnberg	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Gesamtergebnis	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

Aufschlüsselung der DIAS-Einsätze nach Herkunftsländern						
Herkunftsländer	2022	2023	2024	2025	202610	Gesamt-ergebnis
Afghanistan	23,19 %	23,22 %	19,31 %	25,61 %	34,19 %	23,35 %
Ägypten	1,21 %	1,04 %	1,38 %	0,94 %	1,42 %	1,15 %
Algerien	3,60 %	2,20 %	3,81 %	8,39 %	15,50 %	3,62 %
Armenien	0,00 %	0,00 %	0,04 %	0,00 %	0,00 %	0,01 %
Aserbaidschan	0,00 %	0,00 %	0,00 %	0,04 %	0,05 %	0,00 %
Äthiopien	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Bahrain	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Brasilien	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Bulgarien	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Burkina Faso	0,00 %	0,00 %	0,00 %	0,00 %	0,05 %	0,00 %
Dominica	0,00 %	0,00 %	0,00 %	0,04 %	0,09 %	0,00 %
Dschibuti	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Eritrea	0,07 %	0,04 %	0,08 %	0,11 %	0,05 %	0,06 %
Georgien	0,00 %	0,00 %	0,00 %	0,00 %	0,05 %	0,00 %
Guinea	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,01 %
Guinea-Bissau	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Indien	0,00 %	0,01 %	0,00 %	0,00 %	0,05 %	0,01 %
Irak	5,01 %	1,89 %	0,70 %	2,30 %	1,42 %	3,27 %
Iran, Islamische Republik	3,23 %	3,53 %	1,74 %	2,98 %	3,46 %	3,23 %
Israel	0,01 %	0,02 %	0,00 %	0,00 %	0,00 %	0,01 %
Italien	0,00 %	0,00 %	0,00 %	0,04 %	0,00 %	0,00 %
Jemen	1,57 %	1,89 %	0,04 %	0,11 %	0,85 %	1,50 %
Jordanien	0,33 %	0,33 %	0,28 %	0,23 %	2,99 %	0,40 %
Kamerun	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Kanada	0,00 %	0,00 %	0,00 %	0,00 %	0,05 %	0,00 %
Kasachstan	0,00 %	0,00 %	0,00 %	0,00 %	0,05 %	0,00 %
Kenia	0,00 %	0,00 %	0,00 %	0,04 %	0,00 %	0,00 %
Kolumbien	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Kongo	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Kuba	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Kuwait	0,03 %	0,02 %	0,00 %	0,00 %	0,00 %	0,02 %
Libanon	2,21 %	0,63 %	0,24 %	0,91 %	1,99 %	1,40 %
Libyen	0,60 %	0,18 %	0,32 %	1,47 %	2,32 %	0,50 %
Marokko	2,16 %	1,56 %	2,86 %	5,44 %	4,36 %	2,17 %
Mauretanien	0,02 %	0,02 %	0,00 %	0,04 %	0,00 %	0,02 %
Niger	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Nigeria	0,09 %	0,00 %	0,00 %	0,00 %	0,00 %	0,04 %

Vorabfassung - wird durch die lektorierte Version ersetzt.

Aufschlüsselung der DIAS-Einsätze nach Herkunftsländern						
Herkunftsländer	2022	2023	2024	2025	202610	Gesamt-ergebnis
Nordmazedonien	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Pakistan	0,05 %	0,01 %	0,10 %	0,00 %	0,14 %	0,04 %
Pers. aus palästinensischen Gebieten (nicht als Staat anerkannt)	0,71 %	0,26 %	0,58 %	1,74 %	0,52 %	0,56 %
Peru	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Russische Föderation	0,02 %	0,03 %	0,02 %	0,11 %	0,00 %	0,03 %
Saudi-Arabien	0,01 %	0,00 %	0,00 %	0,00 %	0,05 %	0,00 %
Schweden	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Serbien	0,00 %	0,00 %	0,00 %	0,08 %	0,00 %	0,00 %
Somalia	0,05 %	0,04 %	0,04 %	0,08 %	0,24 %	0,05 %
Spanien	0,00 %	0,00 %	0,00 %	0,00 %	0,05 %	0,00 %
Staatenlos	0,15 %	0,08 %	0,12 %	0,15 %	0,19 %	0,12 %
Sudan	0,38 %	0,57 %	2,44 %	5,59 %	4,36 %	0,93 %
Südsudan	0,03 %	0,03 %	0,12 %	0,00 %	0,05 %	0,04 %
Syrien, Arabische Republik	49,56 %	58,08 %	55,54 %	33,96 %	16,45 %	51,69 %
Tadschikistan	0,01 %	0,00 %	0,00 %	0,00 %	0,14 %	0,01 %
Tansania, Vereinigte Republik	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Tschad	0,04 %	0,02 %	0,00 %	0,00 %	0,14 %	0,03 %
Tunesien	2,73 %	2,23 %	4,01 %	4,19 %	2,42 %	2,67 %
Türkei	0,18 %	0,21 %	0,08 %	0,08 %	0,28 %	0,18 %
Ukraine	0,01 %	0,01 %	0,00 %	0,00 %	0,05 %	0,01 %
Ungeklärt	2,63 %	1,83 %	6,13 %	5,33 %	5,93 %	2,78 %
Venezuela	0,02 %	0,01 %	0,02 %	0,00 %	0,05 %	0,01 %
Vietnam	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Gesamtergebnis	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

Aufschlüsselung der DIAS-Einsätze nach Geschlecht						
Geschlecht	2022	2023	2024	2025	202610	Gesamt-ergebnis
divers	0,02 %	0,02 %	0,02 %	0,08 %	0,05 %	0,02 %
männlich	88,62 %	92,64 %	89,73 %	78,13 %	73,73 %	89,40 %
unbekannt	0,00 %	0,00 %	0,00 %	0,00 %	0,05 %	0,00 %
weiblich	11,35 %	7,33 %	10,25 %	21,80 %	26,17 %	10,57 %
Gesamtergebnis	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

2. In wie vielen Fällen wurden die Angaben der Asylsuchenden zu ihrer Herkunft bestätigt bzw. widerlegt (bitte nach Herkunftsländern, Sprache bzw. Dialekt, Geschlecht der Antragstellenden und Jahren aufschlüsseln)?

Die Angaben sind den nachfolgenden Tabellen zu entnehmen.

Es wird auf die Ausführungen in der Vormerkung der Bundesregierung verwiesen, wonach die DIAS-Ergebnisse die Angaben zur Herkunft der Antragstellenden weder bestätigen noch widerlegen können und lediglich ergänzenden Hinweischarakter haben.

Vorabfassung - wird durch die lektorierte Version ersetzt.

Im Übrigen:

Sprachindikation Auswertung Gesamt						
	2022	2023	2024	2025	2026	Gesamt-ergebnis
Report stützt Angaben der Antragstellenden	76,62 %	79,39 %	91,53 %	90,36 %	89,20 %	79,59 %
Report stützt Angaben der Antragstellenden nicht	23,38 %	20,61 %	8,47 %	9,64 %	10,80 %	20,41 %
Gesamtergebnis	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

Aufschlüsselung der Fälle, in denen der Report die Angaben der Antragstellenden stützt, nach Herkunftsländern						
Herkunftsländer	2022	2023	2024	2025	2026	Gesamt-ergebnis
Afghanistan	20,11 %	18,77 %	21,97 %	26,81 %	25,03 %	20,07 %
Ägypten	0,89 %	0,59 %	0,35 %	1,14 %	1,50 %	0,75 %
Algerien	2,74 %	2,51 %	2,77 %	9,66 %	21,09 %	3,42 %
Bahrain	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Belarus	0,02 %	0,05 %	0,00 %	0,00 %	0,00 %	0,03 %
Belgien	0,01 %	0,02 %	0,00 %	0,00 %	0,00 %	0,01 %
Bulgarien	0,05 %	0,02 %	0,05 %	0,00 %	0,00 %	0,03 %
Dänemark	0,03 %	0,02 %	0,00 %	0,00 %	0,00 %	0,02 %
Deutschland	0,00 %	0,03 %	0,05 %	0,00 %	0,14 %	0,02 %
Eritrea	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Frankreich	0,08 %	0,15 %	0,15 %	0,51 %	0,27 %	0,13 %
Ghana	0,00 %	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %
Griechenland	0,62 %	0,12 %	0,35 %	1,02 %	0,54 %	0,40 %
Großbritannien mit Nordirland	0,01 %	0,01 %	0,00 %	0,00 %	0,00 %	0,01 %
Indien	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Irak	4,59 %	1,88 %	1,36 %	1,78 %	1,63 %	3,00 %
Iran, Islamische Republik	4,55 %	4,28 %	2,16 %	3,56 %	4,22 %	4,20 %
Israel	0,01 %	0,01 %	0,00 %	0,00 %	0,00 %	0,01 %
Italien	0,03 %	0,17 %	0,10 %	0,38 %	0,14 %	0,11 %
Jemen	0,63 %	0,66 %	0,00 %	0,00 %	0,14 %	0,56 %
Jordanien	0,37 %	0,39 %	0,85 %	0,51 %	2,31 %	0,48 %
Kasachstan	0,00 %	0,00 %	0,05 %	0,00 %	0,00 %	0,00 %
Katar	0,01 %	0,00 %	0,00 %	0,25 %	0,14 %	0,02 %
Kuwait	0,06 %	0,03 %	0,15 %	0,13 %	0,14 %	0,06 %
Lettland	0,02 %	0,00 %	0,00 %	0,00 %	0,00 %	0,01 %
Libanon	3,85 %	2,12 %	1,26 %	1,02 %	6,12 %	2,88 %
Libyen	0,86 %	0,34 %	0,96 %	2,41 %	2,59 %	0,75 %
Litauen	0,04 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
Luxemburg	0,00 %	0,00 %	0,00 %	0,13 %	0,00 %	0,00 %
Malaysia	0,02 %	0,00 %	0,00 %	0,00 %	0,00 %	0,01 %
Malta	0,02 %	0,01 %	0,00 %	0,00 %	0,00 %	0,01 %
Marokko	2,13 %	1,34 %	2,92 %	6,86 %	6,26 %	2,14 %
Mauretanien	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Niederlande	0,01 %	0,01 %	0,00 %	0,00 %	0,14 %	0,01 %
Norwegen	0,02 %	0,00 %	0,00 %	0,00 %	0,00 %	0,01 %
ohne Angabe	0,55 %	0,01 %	0,00 %	0,00 %	0,00 %	0,24 %
Oman	0,00 %	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %
Österreich	0,06 %	0,02 %	0,00 %	0,00 %	0,00 %	0,03 %

Vorabfassung - wird durch die lektorierte Version ersetzt.

Aufschlüsselung der Fälle, in denen der Report die Angaben der Antragstellenden stützt, nach Herkunftsländern						
Herkunftsländer	2022	2023	2024	2025	2026	Gesamt-ergebnis
Pakistan	0,07 %	0,05 %	0,25 %	0,13 %	0,41 %	0,09 %
Pers. aus palästinensischen Gebieten (nicht als Staat anerkannt)	0,85 %	0,45 %	0,80 %	0,38 %	0,54 %	0,65 %
Polen	0,07 %	0,00 %	0,00 %	0,00 %	0,00 %	0,03 %
Portugal	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Russische Föderation	0,13 %	0,17 %	0,05 %	0,00 %	0,54 %	0,15 %
Saudi-Arabien	0,25 %	0,24 %	0,30 %	0,13 %	0,14 %	0,24 %
Schweden	0,12 %	0,07 %	0,25 %	0,51 %	0,00 %	0,12 %
Schweiz	0,02 %	0,01 %	0,05 %	0,13 %	0,00 %	0,02 %
Serbien	0,04 %	0,02 %	0,00 %	0,00 %	0,00 %	0,02 %
Somalia	0,03 %	0,02 %	0,00 %	0,00 %	0,14 %	0,02 %
Spanien	0,02 %	0,02 %	0,00 %	0,25 %	0,41 %	0,04 %
Staatenlos	0,04 %	0,00 %	0,00 %	0,00 %	0,14 %	0,02 %
Sudan	0,14 %	0,19 %	0,65 %	2,92 %	1,36 %	0,33 %
Südsudan	0,02 %	0,02 %	0,05 %	0,00 %	0,00 %	0,02 %
Syrien, Arabische Republik	45,64 %	52,59 %	48,82 %	24,02 %	13,20 %	47,15 %
Tschad	0,00 %	0,00 %	0,00 %	0,00 %	0,14 %	0,00 %
Tschechien	0,02 %	0,00 %	0,00 %	0,00 %	0,00 %	0,01 %
Tunesien	3,55 %	2,92 %	3,82 %	7,62 %	3,95 %	3,45 %
Türkei	4,19 %	8,67 %	6,59 %	4,07 %	4,49 %	6,28 %
Ukraine	0,15 %	0,02 %	0,00 %	0,00 %	0,00 %	0,07 %
Ungarn	0,00 %	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %
Ungeklärt	2,13 %	0,84 %	2,66 %	3,68 %	2,18 %	1,68 %
Venezuela	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Vereinigte Arabische Emirate	0,09 %	0,14 %	0,15 %	0,00 %	0,00 %	0,11 %
Zypern	0,02 %	0,01 %	0,05 %	0,00 %	0,00 %	0,02 %
Gesamtergebnis	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

Aufschlüsselung der Fälle, in denen der Report die Angaben der Antragstellenden stützt, nach Geschlecht						
Geschlecht	2022	2023	2024	2025	2026	Gesamt-ergebnis
divers	0,02 %	0,01 %	0,00 %	0,13 %	0,00 %	0,02 %
männlich	88,49 %	93,22 %	90,15 %	80,30 %	80,00 %	90,10 %
unbekannt	0,01 %	0,01 %	0,00 %	0,00 %	0,14 %	0,01 %
weiblich	11,48 %	6,76 %	9,85 %	19,57 %	19,86 %	9,87 %
Gesamtergebnis	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

Aufschlüsselung der Fälle, in denen der Report die Angaben der Antragstellenden stützt, nach Sprachen						
Sprache	2022	2023	2024	2025	2026	Gesamt-ergebnis
Algerisch-Arabisch	0,01 %	0,01 %	0,10 %	0,00 %	1,77 %	0,07 %
Arabisch	70,80 %	73,50 %	70,89 %	63,41 %	62,72 %	71,46 %
Armenisch	0,00 %	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %
Aserbaidschanisch	0,00 %	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %
Berbersprachen (Andere)	0,07 %	0,03 %	0,00 %	0,00 %	0,00 %	0,04 %
Dari	14,66 %	15,19 %	16,09 %	25,92 %	25,17 %	15,68 %
Deutsch	0,00 %	0,00 %	0,00 %	0,13 %	0,00 %	0,00 %

Vorabfassung - wird durch die lektorierte Version ersetzt.

Aufschlüsselung der Fälle, in denen der Report die Angaben der Antragstellenden stützt, nach Sprachen						
Sprache	2022	2023	2024	2025	2026	Gesamt-ergebnis
Englisch	0,03 %	0,03 %	0,05 %	0,13 %	0,00 %	0,03 %
Französisch	0,09 %	0,04 %	0,00 %	0,25 %	0,41 %	0,07 %
Italienisch	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Kabylisch	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Kurdisch-Badinani	0,21 %	0,11 %	0,15 %	0,13 %	0,00 %	0,15 %
Kurdisch-Kurmanci	2,49 %	2,43 %	2,92 %	2,67 %	2,99 %	2,52 %
Kurdisch-Sorani	0,33 %	0,31 %	0,10 %	0,00 %	0,41 %	0,30 %
Maghrebinisch	0,00 %	0,02 %	0,00 %	0,25 %	0,00 %	0,02 %
Marokkanisch-Arabisch	0,00 %	0,03 %	0,00 %	0,38 %	0,14 %	0,03 %
Panjabi	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Paschtu	6,27 %	3,55 %	7,44 %	2,67 %	2,72 %	4,99 %
Pashai-West	0,00 %	0,00 %	0,05 %	0,00 %	0,00 %	0,00 %
Persisch	4,22 %	4,32 %	1,71 %	3,18 %	3,27 %	3,99 %
Rifberberisch	0,02 %	0,00 %	0,00 %	0,00 %	0,00 %	0,01 %
Saho	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Somali	0,01 %	0,01 %	0,05 %	0,00 %	0,00 %	0,01 %
Spanisch	0,01 %	0,01 %	0,05 %	0,00 %	0,00 %	0,01 %
Tamil	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Tigre	0,00 %	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %
Türkisch	0,02 %	0,09 %	0,10 %	0,00 %	0,00 %	0,05 %
Turkmenisch	0,05 %	0,05 %	0,00 %	0,13 %	0,14 %	0,05 %
Ukrainisch	0,01 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
Unbekannt	0,54 %	0,01 %	0,00 %	0,25 %	0,14 %	0,25 %
Usbekisch	0,12 %	0,25 %	0,30 %	0,51 %	0,14 %	0,21 %
Gesamtergebnis	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

Aufschlüsselung der Fälle, in denen der Report die Angaben der Antragstellenden nicht stützt, nach Herkunftsländern						
Herkunftsländer	2022	2023	2024	2025	2026	Gesamt-ergebnis
Afghanistan	9,94 %	6,72 %	4,35 %	16,67 %	8,99 %	8,48 %
Ägypten	1,59 %	2,67 %	1,09 %	2,38 %	1,12 %	2,04 %
Algerien	0,50 %	0,11 %	1,63 %	1,19 %	14,61 %	0,58 %
Belarus	0,03 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
Belgien	0,03 %	0,04 %	0,00 %	0,00 %	0,00 %	0,03 %
Bulgarien	0,03 %	0,04 %	0,00 %	0,00 %	0,00 %	0,03 %
Dänemark	0,03 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
Eritrea	0,03 %	0,08 %	0,00 %	0,00 %	0,00 %	0,05 %
Frankreich	0,00 %	0,11 %	0,00 %	1,19 %	0,00 %	0,06 %
Georgien	0,00 %	0,00 %	0,00 %	0,00 %	1,12 %	0,02 %
Griechenland	0,28 %	0,08 %	0,00 %	5,95 %	3,37 %	0,31 %
Irak	2,74 %	1,13 %	0,54 %	0,00 %	4,49 %	1,97 %
Iran, Islamische Republik	1,03 %	1,05 %	1,63 %	5,95 %	5,62 %	1,19 %
Israel	0,06 %	0,00 %	0,00 %	0,00 %	0,00 %	0,03 %
Italien	0,00 %	0,04 %	0,54 %	0,00 %	1,12 %	0,05 %
Jemen	1,87 %	0,68 %	0,00 %	0,00 %	0,00 %	1,25 %
Jordanien	0,44 %	0,45 %	1,63 %	1,19 %	2,25 %	0,51 %
Kuwait	0,09 %	0,04 %	0,00 %	0,00 %	2,25 %	0,10 %
Libanon	2,37 %	1,69 %	3,26 %	1,19 %	2,25 %	2,09 %

Vorabfassung - wird durch die lektorierte Version ersetzt.

Aufschlüsselung der Fälle, in denen der Report die Angaben der Antragstellenden nicht stützt, nach Herkunftsländern						
Herkunftsländer	2022	2023	2024	2025	2026	Gesamt-ergebnis
Libyen	0,78 %	0,41 %	1,63 %	2,38 %	5,62 %	0,74 %
Litauen	0,03 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
Malaysia	0,06 %	0,00 %	0,00 %	0,00 %	0,00 %	0,03 %
Marokko	0,31 %	0,08 %	1,09 %	3,57 %	4,49 %	0,34 %
Norwegen	0,03 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
ohne Angabe	0,69 %	0,00 %	0,00 %	0,00 %	0,00 %	0,35 %
Pakistan	0,06 %	0,08 %	0,00 %	0,00 %	2,25 %	0,10 %
Pers. aus palästinensischen Gebieten (nicht als Staat anerkannt)	0,97 %	0,38 %	1,09 %	1,19 %	3,37 %	0,75 %
Russische Föderation	0,06 %	0,04 %	0,00 %	0,00 %	0,00 %	0,05 %
Saudi-Arabien	0,34 %	0,23 %	0,00 %	0,00 %	0,00 %	0,27 %
Schweden	0,09 %	0,00 %	0,00 %	0,00 %	0,00 %	0,05 %
Serbien	0,06 %	0,00 %	0,00 %	0,00 %	0,00 %	0,03 %
Somalia	0,03 %	0,08 %	0,00 %	0,00 %	0,00 %	0,05 %
Sudan	1,43 %	1,24 %	7,07 %	14,29 %	16,85 %	1,91 %
Südsudan	0,06 %	0,04 %	0,00 %	0,00 %	0,00 %	0,05 %
Syrien, Arabische Republik	64,74 %	69,44 %	58,15 %	29,76 %	7,87 %	65,28 %
Tansania, Vereinigte Republik	0,03 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
Tschad	0,06 %	0,04 %	0,00 %	0,00 %	0,00 %	0,05 %
Tunesien	1,22 %	0,34 %	1,63 %	3,57 %	0,00 %	0,87 %
Türkei	5,83 %	11,86 %	7,61 %	2,38 %	5,62 %	8,41 %
Ukraine	0,16 %	0,00 %	0,54 %	0,00 %	0,00 %	0,10 %
Ungeklärt	1,75 %	0,79 %	6,52 %	5,95 %	6,74 %	1,61 %
Vereinigte Arabische Emirate	0,12 %	0,04 %	0,00 %	1,19 %	0,00 %	0,10 %
Zypern	0,00 %	0,04 %	0,00 %	0,00 %	0,00 %	0,02 %
Gesamtergebnis	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

Aufschlüsselung der Fälle, in denen der Report die Angaben der Antragstellenden nicht stützt, nach Geschlecht						
Geschlecht	2022	2023	2024	2025	2026	Gesamt-ergebnis
divers	0,03 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
männlich	93,05 %	95,35 %	95,11 %	72,62 %	77,53 %	93,59 %
weiblich	6,92 %	4,65 %	4,89 %	27,38 %	22,47 %	6,39 %
Gesamtergebnis	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

Aufschlüsselung der Fälle, in denen der Report die Angaben der Antragstellenden nicht stützt, nach Sprachen						
Sprache	2022	2023	2024	2025	2026	Gesamt-ergebnis
Arabisch	84,01 %	89,30 %	89,13 %	69,05 %	73,03 %	86,07 %
Dari	4,52 %	4,95 %	2,17 %	17,86 %	11,24 %	4,91 %
Deutsch	0,03 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
Englisch	0,03 %	0,04 %	0,00 %	0,00 %	0,00 %	0,03 %
Französisch	0,03 %	0,00 %	0,00 %	0,00 %	1,12 %	0,03 %
Fur	0,00 %	0,00 %	0,00 %	0,00 %	1,12 %	0,02 %
Georgisch	0,00 %	0,00 %	0,00 %	0,00 %	1,12 %	0,02 %
Kanuri	0,03 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %

Vorabfassung - wird durch die lektorierte Version ersetzt.

Aufschlüsselung der Fälle, in denen der Report die Angaben der Antragstellenden nicht stützt, nach Sprachen						
Sprache	2022	2023	2024	2025	2026	Gesamt-ergebnis
Kurdisch-Badinani	0,56 %	0,15 %	0,00 %	0,00 %	0,00 %	0,35 %
Kurdisch-Kurmanci	3,12 %	2,36 %	4,35 %	3,57 %	1,12 %	2,81 %
Kurdisch-Sorani	0,65 %	0,15 %	0,00 %	0,00 %	1,12 %	0,42 %
Paschtu	5,74 %	1,54 %	3,26 %	5,95 %	4,49 %	3,85 %
Persisch	0,72 %	0,98 %	0,54 %	3,57 %	4,49 %	0,92 %
Russisch	0,06 %	0,00 %	0,00 %	0,00 %	0,00 %	0,03 %
Somali	0,00 %	0,04 %	0,00 %	0,00 %	0,00 %	0,02 %
Tigre	0,03 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
Tigrinya	0,09 %	0,00 %	0,00 %	0,00 %	0,00 %	0,05 %
Türkisch	0,03 %	0,04 %	0,54 %	0,00 %	0,00 %	0,05 %
Turkmenisch	0,00 %	0,19 %	0,00 %	0,00 %	0,00 %	0,08 %
Ukrainisch	0,03 %	0,00 %	0,00 %	0,00 %	0,00 %	0,02 %
Unbekannt	0,28 %	0,00 %	0,00 %	0,00 %	0,00 %	0,14 %
Urdu	0,00 %	0,08 %	0,00 %	0,00 %	0,00 %	0,03 %
Usbekisch	0,03 %	0,19 %	0,00 %	0,00 %	1,12 %	0,11 %
Gesamtergebnis	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

3. In wie vielen Fällen wurde das Ergebnis des Sprach- und Dialekterkennungstools seit 2022 in einem späteren Verfahrensstadium, etwa durch ein linguistisches Sachverständigengutachten, die Entscheidung des BAMF oder ein Gerichtsurteil, bestätigt bzw. widerlegt (bitte nach Sprache bzw. Dialekt, Geschlecht der Antragstellenden, Jahren sowie nach Art der späteren Bewertung aufschlüsseln)?

Der Bundesregierung liegen keine statistisch auswertbaren Daten im Sinne der Fragestellung vor.

4. In wie vielen Fällen stimmte das Ergebnis des Sprach- und Dialekterkennungstools seit 2022 mit der von den Antragstellenden angegebenen Sprache, dem angegebenen Dialekt bzw. der angegebenen Herkunftsregion überein bzw. wich es davon ab?

Es wird auf die Antwort der Bundesregierung auf Frage 2 verwiesen. Im Übrigen liegen der Bundesregierung keine statistisch auswertbaren Daten im Sinne der Fragestellung vor.

5. In wie vielen Fällen folgte die abschließende Feststellung des BAMF zu Sprache, Dialekt bzw. Herkunftsregion der Antragstellenden dem Ergebnis des Tools bzw. den Angaben der Antragstellenden (bitte jeweils nach Jahren, Sprache bzw. Dialekt und Geschlecht der Antragstellenden aufschlüsseln)?

6. In wie vielen der Fälle, in denen das Ergebnis des Tools von den Angaben der Antragstellenden abwich, wurde der Asylantrag abgelehnt bzw. Schutz zuerkannt, und wie verhalten sich diese Zahlen zu den Fällen, in denen das Ergebnis des Tools mit den Angaben der Antragstellenden übereinstimmte?

Vorabfassung - wird durch die lektorierte Version ersetzt.

7. In wie vielen der Ablehnungen wurde das Ergebnis des Tools oder wurden Zweifel an der angegebenen Herkunft in der Begründung des Bescheids angeführt (bitte jeweils nach Jahren, Herkunftsländern und Geschlecht der Antragstellenden aufschlüsseln)?

8. In wie vielen Fällen folgte seit 2022 auf den Einsatz des Sprach- und Dialekterkennungstools eine längere Sprachaufnahme oder eine linguistische Begutachtung durch Sachverständige, nach welchen Kriterien geschieht dies, und in wie vielen dieser Fälle hat das Gutachten das Ergebnis des Tools bestätigt bzw. widerlegt (bitte nach Jahren aufschlüsseln)?

Die Fragen 5 bis 8 werden gemeinsam beantwortet.

Der Bundesregierung liegen keine statistisch auswertbaren Daten im Sinne der Fragestellung vor.

9. Für welche Sprachen bzw. Dialekte kommt die Software aktuell zum Einsatz, und für welche weiteren Sprachen bzw. Dialekte ist eine Erweiterung geplant?

Derzeit sind in DIAS folgende Sprach- bzw. Dialektmodelle produktiv im Einsatz: Arabisch-Levantinisch, Arabisch-Maghrebinisch, Arabisch-Ägyptisch, Arabisch-Irakisch, Arabisch-Golf, Dari sowie auch Persisch (Farsi). Über die genannten Modelle hinaus sind derzeit keine weiteren Sprachmodelle produktiv im Einsatz oder geplant.

10. Wie kam es genau zu der Entscheidung, die Sprach- und Dialekterkennung bei Asylsuchenden, die als erste Muttersprache Paschto angeben, ab April 2023 aus Gründen der „Verfahrensökonomie“ nicht mehr einzusetzen?

Grund für das Aussetzen der Nutzung von DIAS für Personen, die als erste Muttersprache Paschto angegeben haben, war u. a. der anhaltend hohe Zugang von Antragstellenden aus Afghanistan, die als erste Muttersprache „Paschto“ angaben und es sich im Rahmen von Qualitätssicherungsmaßnahmen zum Produktiveinsatz des Dialekts Paschto gezeigt hat, dass eine Unterscheidung zwischen Antragsstellenden aus Afghanistan und aus Pakistan anhand des DIAS-Sprachmodells „Paschto“ kaum möglich ist.

11. Ist damit gemeint, dass Asylverfahren aufwendiger werden und länger dauern, wenn das Sprach- und Dialekterkennungstool verwendet wird (bitte erläutern)?

    a) Wenn ja, wie passt das damit zusammen, dass das Sprach- und Dialekterkennungstool an anderer Stelle als „Unterstützungsinstrument“ für die Entscheider im BAMF bezeichnet wird (bitte erläutern)?

    b) Wenn nein, was ist stattdessen gemeint?

Die Fragen 11 bis 11b wird gemeinsam beantwortet.

DIAS wird zur Unterstützung der Klärung von Herkunft und Identität von Antragstellenden ohne Identitätspapiere eingesetzt. Der Einsatz von DIAS nimmt eine gewissen Zeit im Rahmen der Aktenanlage in Anspruch, verlängert aber nicht die Dauer des Asylverfahrens. Die Bezeichnung von DIAS als Unterstützungsinstrument für die Entscheiderinnen und Entscheider bleibt davon unberührt. Im Übrigen wird auf die Antwort auf Frage 10 verwiesen.

Vorabfassung - wird durch die lektorierte Version ersetzt.

12. Welche Kosten sind mit aktuellem Stand seit der Pilotierung im Jahr 2017 für die Anschaffung und den Einsatz der biometrischen Sprach- bzw. Dialektanalysesoftware entstanden und im laufenden Haushalt vorgesehen (bitte nach Jahren aufschlüsseln)?

Für die Jahre 2017 bis 2021 wird auf die Bundestagsdrucksache 20/3238 (Antwort auf Frage 4) verwiesen.

Die seit 2022 für DIAS entstandenen Kosten sind der nachstehenden Tabelle zu entnehmen:

Jahr	Kosten
2022	763.747 Euro
2023	619.911 Euro
2024	410.393 Euro
2025	366.930 Euro
2026 (Stand 10.11.2026)	443.239 Euro

Im Haushaltsjahr 2026 werden voraussichtlich Ausgaben i. H. v. insgesamt 500.000 Euro geleistet.

13. Aus welchen Quellen stammen die für das Sprach- bzw. Dialekterkennungssystem verwendeten Sprach- und Hintergrundmodelle?

14. Für welche Sprachen liegen Modelle vor, und aus wie vielen Sprachbeispielen und welcher Gesamtdauer an Audiomaterial setzen sich die Trainingsdatensätze jeweils zusammen (bitte nach jeweiliger Sprache bzw. Dialekt aufschlüsseln)?

Die Fragen 13 und 14 werden gemeinsam beantwortet.

Es wird auf die Antwort der Bundesregierung zu Frage 6 auf die Kleine Anfrage der Fraktion Die Linke auf Bundestagsdrucksache 20/3238 verwiesen.

15. Wie haben sich die Erkennungsquoten der Sprach- bzw. Dialekterkennungssoftware seit 2017 entwickelt (bitte nach jeweiliger Sprache bzw. Dialekt aufschlüsseln), und wie werden die Erkennungsquoten berechnet?

Es wird auf die Antwort der Bundesregierung zu Frage 5b auf die Kleine Anfrage der Fraktion Die Linke auf Bundestagsdrucksache 20/9419 verwiesen. Die Erkennungsquote der arabischen Dialekte wird in Gesamttests gemessen und lag von 2017 bis 2020 bei rund 80 Prozent und erreichte nach den Trainings der Sprachmodelle im Jahr 2021 rund 85 Prozent. Durch weitere Trainings im Jahr 2023 wurde eine Erkennungsquote von rund 87 Prozent erreicht, die sich seitdem nicht signifikant verändert hat. Darüber hinaus liegen der Bundesregierung keine Erkenntnisse im Sinne der Fragestellung vor.

16. Wie viele Sprachproben umfasst der gelabelte Testdatensatz, mit dem die Sprachmodelle evaluiert werden, und wie viele davon wurden jeweils korrekt klassifiziert?

    a) Wie hoch sind Präzision (Precision) und Trefferquote (Recall) der einzelnen Modelle?

Die Fragen 16 und 16a werden gemeinsam beantwortet.

Vorabfassung - wird durch die lektorierte Version ersetzt.

Der Bundesregierung liegen hierzu keine statistisch auswertbaren Daten vor.

b) Welche durchschnittliche Wahrscheinlichkeit gibt das System für den jeweiligen Dialekt aus (bitte jeweils nach Sprache bzw. Dialekt aufschlüsseln)?

Die Erkennungsquote der arabischen Dialekte wird in Gesamttests gemessen und beträgt derzeit rund 87 Prozent. Dabei handelt es sich um die Erkennungsleistung der Sprachmodelle und nicht um die durchschnittliche vom System ausgegebene Wahrscheinlichkeit für einen einzelnen Dialekt.

Eine entsprechende differenzierte Erfassung der durchschnittlich ausgegebenen Wahrscheinlichkeiten ist daher nicht möglich. Darüber hinaus liegen der Bundesregierung keine Erkenntnisse im Sinne der Fragestellung vor.

17. Wie sieht für jede Sprach- bzw. Dialektgruppe die Konfusionsmatrix aus dem Test- bzw. Trainingsdatensatz aus, d. h. bei wie vielen Sprachproben hat das System für jede tatsächliche Sprache bzw. jeden tatsächlichen Dialekt der Testsprecherinnen und Testsprecher welcher Sprache bzw. welchem Dialekt die höchste Wahrscheinlichkeit zugewiesen, und falls ein Schwellenwert angewendet wird (vgl. Frage 18), wie viele Sprachproben je tatsächlicher Sprache bzw. tatsächlichem Dialekt lagen unterhalb dieses Schwellenwerts?

Der Bundesregierung liegen hierzu keine statistisch auswertbaren Daten vor. Hinsichtlich der Erkennungsquote wird auf die Antwort zu Frage 15 verwiesen.

18. In welcher Form wird den Entscheiderinnen und Entscheidern das Ergebnis des Tools angezeigt, etwa als die einzelne wahrscheinlichste Sprache bzw. der einzelne wahrscheinlichste Dialekt, als Wahrscheinlichkeit für jede Sprache bzw. jeden Dialekt oder als verbale Einstufung wie „hohe“ oder „geringe“ Wahrscheinlichkeit?

Den Entscheiderinnen und Entscheidern wird das Ergebnis in Form eines PDF-Reports angezeigt. Dieser weist die beiden wahrscheinlichsten Sprachen bzw. Dialekte jeweils mit ihrer prozentualen Ausprägung aus. Die übrigen erkannten Dialekte werden unter der Kategorie „Sonstige“ zusammengefasst.

a) Wird ein Wahrscheinlichkeitsschwellenwert angewendet, bevor ein Ergebnis ausgewiesen oder als belastbar behandelt wird, und wenn ja, wie hoch ist dieser?

Ein Wahrscheinlichkeitsschwellenwert wird nicht angewendet.

b) Wurde dieselbe Entscheidungsregel auch der Berechnung der in den vorangegangenen Fragen erfragten Werte zugrunde gelegt?

Da kein Wahrscheinlichkeitsschwellenwert angewendet wird, besteht insoweit keine entsprechende Schwellenwertregel im Sinnen der Fragestellung.

19. Wo, und wie lange werden Sprachaufzeichnungen der Antragstellenden gespeichert?

Die im Rahmen des Asylverfahrens gemäß § 16 Absatz 1 des Asylgesetzes (AsylG) erhobenen Sprachaufzeichnungen werden im Bundesamt für Migration

Vorabfassung - wird durch die lektorierte Version ersetzt.

und Flüchtlinge gespeichert. Die Löschung erfolgt gemäß § 16 Absatz 6 AsylG zehn Jahre nach unanfechtbarem Abschluss des Asylverfahrens.

20. Wurden die Sprachaufzeichnungen von Antragstellenden jemals an eine Institution außerhalb des BAMF und/oder an einen anderen Staat oder an sonstige Dritte weitergegeben, und wenn ja, an wen, bzw. welche Institutionen und/oder Staaten, und aus welchem Grund?

Es wird auf die Antwort der Bundesregierung zu Frage 9 auf die Kleine Anfrage der Fraktion Die Linke auf Bundestagsdrucksache 20/3238 verwiesen.

21. Welche Schulungen erhalten die Entscheiderinnen und Entscheider zur Anwendung des Sprach- bzw. Dialekterkennungsassistenten und zur Bewertung seiner Ergebnisse, und welche Inhalte umfassen diese Schulungen?

Die Entscheidenden werden gezielt dafür sensibilisiert, dass der DIAS-Bericht ausschließlich Hinweischarakter hat, keinen Beweis darstellt und seine Ergebnisse stets im Rahmen der Gesamtwürdigung aller vorliegenden Angaben und Indizien zu bewerten sind. Dies erfolgt im Rahmen der standardisierten Schulung zur Entscheiderin bzw. zum Entscheider.

22. Können Antragstellende gegen ein falsches Ergebnis der Spracherkennungssoftware Einspruch erheben, und wenn ja, welches Verfahren ist hierfür vorgesehen?

Es wird auf die Antwort der Bundesregierung zu Frage 10 auf die Kleine Anfrage der Fraktion Die Linke auf Bundestagsdrucksache 20/3238 verwiesen.

23. Falls ein solches Verfahren besteht, wie viele Antragstellende haben seit 2022 in jedem Jahr Einspruch gegen das Ergebnis eingelegt, und mit welchem Ergebnis wurden diese Einsprüche jeweils abgeschlossen (bitte jeweils nach Jahren, dem vom BAMF festgestellten Herkunftsland und dem Geschlecht der Antragstellenden aufschlüsseln)?

Der Bundesregierung liegen hierzu keine statistisch auswertbaren Daten vor.

24. Welches Unternehmen ist derzeit Eigentümer der Software, die für den Sprach- bzw. Dialekterkennungsassistenten verwendet wird?

Die eingesetzte Sprach- bzw. Dialekterkennungssoftware stammt vom Anbieter Nuance. Nuance Communications gehört seit März 2022 zum Microsoft-Konzern. Die Software wird vom BAMF über Lizenzen bezogen.

25. Welche aktuellen Angaben kann die Bundesregierung zur Zusammenarbeit des BAMF mit anderen EU-Staaten machen, um den Sprach- bzw. Dialekterkennungsassistenten gemeinsam zu entwickeln und zu nutzen (vgl. Antwort zu Frage 12 auf Bundestagsdrucksache 20/3238)?

Das BAMF initiierte im Rahmen des DEU-Vorsitzes der EU-Ratspräsidentschaft im zweiten Halbjahr 2020 das CELIA (Common European Language Indication and Analysis) Pilotprojekt. Das übergeordnete Ziel von CELIA war die Entwicklung und Implementierung eines gemeinsamen Verfahrens zur

Vorabfassung - wird durch die lektorierte Version ersetzt.

Sprachindikation und -analyse, das allen Mitgliedstaaten zugänglich sein soll. Die niederländische Migrationsbehörde IND hat in der Zwischenzeit die Federführung übernommen und führt das vom BAMF pilotierte CELIA-Projekt nun im Rahmen eines EU-Projektes, begleitete von der LADO (language assessment for determination of origin) Studie, weiter. Das BAMF ist Teil des Steuerungsgremiums.

a) Mit welchen EU-Staaten gibt es derzeit eine solche Zusammenarbeit?

Die Teilnahme an der LADO-Studie steht allen EU-Mitgliedstaaten offen.

b) Wird die Software des BAMF auch in anderen EU-Mitgliedstaaten verwendet, und wenn ja, in welchen?

Der Bundesregierung liegen keine Erkenntnisse im Sinne der Fragestellung vor.

c) In welchen Ländern haben Vertreterinnen und Vertreter des BAMF die Nutzung des Sprach- bzw. Dialekterkennungsassistenten demonstriert?

Nach Kenntnis der Bundesregierung wurde das Spracherkennungstool DIAS in den Ländern Niederlande, Norwegen und Schweiz demonstriert. Darüber hinaus liegen der Bundesregierung keine Erkenntnisse im Sinne der Fragestellung vor.

d) Mit welchen anderen Regierungen haben Vertreterinnen und Vertreter des BAMF über den Sprach- bzw. Dialekterkennungsassistenten gesprochen?

Es wird die Antwort der Bundesregierung zu Frage 12 auf die Kleine Anfrage der Fraktion Die Linke auf Bundestagsdrucksache 20/3238 verwiesen.

26. Ist es zutreffend, dass das BAMF eine unabhängige Evaluierung des Dialekterkennungsassistenten bei der Technischen Hochschule (TH) Nürnberg (TH-Ohm) in Auftrag gegeben hat, und wenn ja, wann wurde dieser Auftrag erteilt, und aus welchem Grund wurde er an diese TH vergeben?

Im Frühjahr 2023 wurde die Technische Hochschule Nürnberg (TH Ohm) mit der unabhängigen wissenschaftlichen Evaluierung des Dialekterkennungsassistenten beauftragt. Ziel der Beauftragung war es, das Tool durch einen wissenschaftlich unabhängigen und neutralen Dritten überprüfen und evaluieren zu lassen.

27. Ist die Evaluation bereits abgeschlossen, und wenn nein, wann ist mit einem Ergebnis zu rechnen, und wann soll der Abschlussbericht veröffentlicht werden?

Die Evaluation des Dialekterkennungsassistenten ist bereits abgeschlossen. Ein Zeitpunkt für die Veröffentlichung des Abschlussberichts kann derzeit nicht genannt werden.

Vorabfassung - wird durch die lektorierte Version ersetzt.

Vorabfassung - wird durch die lektorierte Version ersetzt.

Gesamtherstellung: H. Heenemann GmbH & Co. KG, Buch- und Offsetdruckerei, Bessemerstraße 83–91, 12103 Berlin, www.heenemann-druck.de
Vertrieb: Bundesanzeiger Verlag GmbH, Postfach 10 05 34, 50445 Köln, Telefon (02 21) 97 66 83 40, Fax (02 21) 97 66 83 44, www.bundesanzeiger-verlag.de
ISSN 0722-8333