

Antwort der Bundesregierung

**auf die Kleine Anfrage der Abgeordneten Ulla Jelpke, Dr. André Hahn,
Gökay Akbulut, weiterer Abgeordneter und der Fraktion DIE LINKE.
– Drucksache 19/5744 –**

Ergebnisse des Pilotprojekts zur biometrischen Gesichtserkennung am Bahnhof Berlin Südkreuz

Vorbemerkung der Fragesteller

Von August 2017 bis Juli 2018 hat die Bundespolizei am Berliner Bahnhof Südkreuz ein Pilotprojekt zur biometrischen Gesichtserkennung durchgeführt. Am Abschlussbericht, der im Oktober 2018 veröffentlicht wurde (www.bundespolizei.de/Web/DE/04Aktuelles/01Meldungen/2018/10/181011_abschlussbericht_gesichtserkennung_down.pdf?__blob=publicationFile&v=1), haben Computer- und Datenschutzexperten zum Teil massive Kritik geäußert. So kommentierte der Chaos Computer Club (CCC), die Ergebnisse seien „nicht überzeugend und absichtlich geschönt worden“, die Ergebnisse seien „manipuliert worden, um sie nicht ganz so desaströs aussehen zu lassen“ (www.ccc.de/en/updates/2018/debakel-am-suedkreuz).

Auch nach Ansicht der Fragestellerinnen und Fragesteller gibt der Abschlussbericht in vielfacher Hinsicht Anlass zur kritischen Nachfrage, sowohl hinsichtlich der Versuchsanordnung, der Ergebnisstatistiken und nicht zuletzt der grundsätzlichen politischen Fragestellungen, die mit dem Einsatz einer sogenannten intelligenten Videotechnik verbunden sind.

In der ersten Testphase wurden Fotos von 312 Testpersonen in eine Datenbank eingespeist, um die Fähigkeit der Systeme zu erproben, diese Personen zu detektieren.

Im Abschlussbericht heißt es, dass zwei der von drei Herstellern getesteten Systeme jetzt „für den praktischen polizeilichen Einsatz geeignet seien“. Dabei wird positiv auf die Trefferrate (76,7 Prozent) bzw. Falschtrefferrate (FAR; 0,67 Prozent) des sogenannten Gesamtsystems Bezug genommen. Das Gesamtsystem besteht aus einer Kombination der Ergebnisse aller drei Systeme. Betrachtet man nur die Einzelsysteme, liegt die durchschnittliche Trefferrate bei lediglich 68,5 Prozent (bezogen auf die erste Testphase). Die Bundespolizei lässt im Abschlussbericht eine Tendenz zur Nutzung des Gesamtsystems erkennen: „Die zu erwartende höhere Trefferrate könnte – unter Inkaufnahme einer anzunehmenden höheren Falschakzeptanzrate – die Wahrscheinlichkeit eines Fahndungstreffers signifikant erhöhen.“

Die Antwort wurde namens der Bundesregierung mit Schreiben des Bundesministeriums des Innern, für Bau und Heimat vom 26. November 2018 übermittelt.

Die Drucksache enthält zusätzlich – in kleinerer Schrifttype – den Fragetext.

Eine bewusst in Kauf genommene Zunahme „falscher Treffer“ bei polizeilichen Maßnahmen halten die Fragestellerinnen und Fragesteller allerdings für den Ausdruck eines problematischen Grundrechtsverständnisses. Bei einer angenommenen Nutzerzahl des Bahnhofs Südkreuz von 90 000 Personen pro Tag würden, bei einer FAR von 0,67 Prozent, rund 600 Personen täglich zu Unrecht als „Gesuchte“ identifiziert. Die Bundespolizei selbst führt (hinsichtlich der zweiten Testphase) aus, dass im Realbetrieb eines im „ODER“-Modus interkonnektierenden Systems beim Einsatz von 15 Kameras und angenommenen 15 000 Personen pro Tag täglich 1 020 Falschtreffer erfolgen würden. Die FAR würde steigen, wenn die Zahl der gespeicherten Fotos stiege (im Abschlussbericht wird als Beispiel auf 600 auf freiem Fuß befindliche „Gefährder“ Bezug genommen).

Der Bericht hebt zwar hervor, dass in Testphase 2 eine höhere Trefferrate erreicht wurde, jedoch erscheint den Fragestellerinnen und Fragestellern die dabei erzielte Fehlerrate von 0,34 Prozent (im ODER-Modus des Gesamtsystems; bei über 80 Prozent Trefferrate) immer noch bedenklich hoch. Lediglich im UND-Modus wurde eine niedrige Fehlerrate erzielt (0,00018 Prozent), wobei sich allerdings die Trefferrate wiederum auf rund zwei Drittel absenkte.

Allerdings gibt es an der Verlässlichkeit der zweiten Testphase Zweifel: Denn dabei ging es ausschließlich um die Detektionsfähigkeit solcher Bilder, die von den Kameras im Bahnhof Südkreuz angefertigt worden waren, und die sie anschließend auch selbst wieder erkennen sollten. Aussagekräftiger wäre nach Einschätzung der Fragestellerinnen und Fragesteller die Frage gewesen, ob die getesteten Systeme auch in der Lage wären, Personen anhand solcher Bilder zu identifizieren, die von anderen Videokameras (nicht vom Bahnhof Südkreuz) stammen.

Unabhängig von Fragen der Technik, der Versuchsanordnung und der Statistik sind die Fragestellerinnen und Fragesteller allerdings darüber besorgt, dass die Auswirkung der biometrischen Gesichtserkennung auf die Grundrechte bzw. deren mögliche Einschränkung im Abschlussbericht kaum thematisiert wird. Denn der Bericht streift die Frage, was mit irrtümlich identifizierten Personen im Wirkbetrieb passieren würde, nur kurz: „Im günstigsten Fall“ führe ein visueller Abgleich durch einen Polizisten dazu, dass keine weiteren polizeilichen Maßnahmen ergriffen würden. Weniger günstige Fälle werden hingegen nicht thematisiert. Die Fragestellerinnen und Fragesteller sehen das Risiko, dass im Wirkmodus eines solchen Systems täglich Hunderte von Personen – an nur einem Bahnhof – fälschlich erkannt würden und der Großteil von ihnen sehr wohl zumindest mit einer Identitätsfeststellung durch die Polizei rechnen müsste. Bei flächendeckendem Einsatz an allen Bahnhöfen bundesweit dürfte diese Zahl an die Zehntausende reichen. Das bedeutet erhebliche Grundrechtseingriffe für Passanten und erheblichen Mehraufwand auf Seiten der Polizei. Dies muss berücksichtigt werden, wenn es um die Frage geht, ob ein solcher flächendeckender Biometrie-Einsatz zu mehr Sicherheit führen kann.

Vorbemerkung der Bundesregierung:

Im Rahmen des gemeinsamen Projektes „Sicherheitsbahnhof Berlin Südkreuz“ von Bundesministerium des Innern, für Bau und Heimat, Bundespolizei, Bundeskriminalamt und Deutsche Bahn AG ist vereinbart worden, den Nutzen von intelligenter Videoanalysetechnik für polizeiliche und unternehmerische Zwecke zu erproben.

Das Projekt „Intelligente Videoanalyse“ gliedert sich in zwei Teilprojekte:

In einem ersten Teilprojekt wurde ohne Beteiligung der Deutsche Bahn AG der Nutzen von biometrischer Gesichtserkennungstechnik in Live-Videoströmen der Überwachungskameras der Deutschen Bahn AG für polizeiliche Zwecke getestet.

Dieses Teilprojekt hatte am 1. August 2017 begonnen und endete nach einem Jahr am 31. Juli 2018. In diesem einjährigen Erprobungszeitraum konnten sämtliche Jahreszeiten, Witterungsbedingungen und Lichtverhältnisse dargestellt werden. In den ersten sechs Monaten wurde zum Abgleich eine Referenzdatenbank mit qualitativ hochwertigen Lichtbildern von Freiwilligen verwendet. In den folgenden sechs Monaten sollten die Gesichtserkennungssysteme den Beweis erbringen, dass sie auch mit Fahndungsfotos ungünstigerer Qualität (Auflösung, Schärfe, Position des Gesichts, Lichtverhältnisse zum Zeitpunkt der Aufzeichnung) eine hohe Trefferquote zu generieren imstande sind. Dafür wurden mehrere Bilder der Probanden aus den am Bahnhof Berlin Südkreuz vorhandenen Videoüberwachungssystemen in die Referenzdatenbanken der Gesichtserkennungssysteme eingestellt. Bereits jetzt werden Bilder aus vorhandenen Videoüberwachungsanlagen auf den Bahnanlagen des Bundes genutzt, z. B. für Öffentlichkeitsfahndungen.

Bei der biometrischen Gesichtserkennung wurde nicht von jeder Person die Identität festgestellt, sondern es erfolgte ein Abgleich des bzw. der Bilder mit einer Testdatenbank. Der Test wurde ausschließlich mit freiwilligen Teilnehmerinnen und Teilnehmern auf Basis einer Einwilligung durchgeführt. Als Testperson konnte sich jede volljährige Person melden, die den Bahnhof Berlin Südkreuz im Testzeitraum regelmäßig nutzt, so dass ein "Spiegelbild" der Reisenden im Rahmen des Projektes abgebildet werden konnte. Es gab keine weiteren Vorgaben.

Der Abschlussbericht des Bundespolizeipräsidiums über den Test der Gesichtserkennungssysteme am Bahnhof Berlin Südkreuz wurde am 11. Oktober 2018 durch das Bundesministerium des Innern, für Bau und Heimat veröffentlicht. Als Bilanz wurde durch die Bundespolizei festgestellt, dass die Gesichtserkennungssysteme nach dem Stand der Technik in der Zukunft einen wesentlichen Mehrwert für die polizeiliche Arbeit, insbesondere der Bundespolizei, darstellen können. Nach polizeifachlicher Bewertung soll die Technik die Arbeit der Polizeivollzugsbeamtinnen und Polizeivollzugsbeamten unterstützen und nicht ersetzen, da bereits nach 20 Minuten Monitorbeobachtung die menschliche Aufmerksamkeit für Videodetails sinkt und den Mehrwert von Videoüberwachung in signifikanter Weise beeinträchtigt.

Im Anschluss an den Test der Gesichtserkennungssysteme sollen in einem zweiten Projekt voraussichtlich ab Januar 2019 intelligente Videoanalysesysteme für die Behandlung und Auswertung verschiedener Gefahrenszenarien erprobt werden. Dabei sollen u.a. Gefahrensituationen wie das Erkennen hilfloser Personen oder stehengelassener Gegenstände automatisiert erkannt und gemeldet werden. Die Federführung für dieses Teilprojekt liegt bei der Deutsche Bahn AG und befindet sich in der Vorbereitung.

1. Macht sich die Bundesregierung die wesentlichen Schlussfolgerungen aus dem Abschlussbericht der Bundespolizei zu eigen, und inwiefern bestehen aus ihrer Sicht noch Vorbehalte oder Prüfbedarf (bitte begründen)?

Die Meinungsbildung in der Bundesregierung hierzu ist noch nicht abgeschlossen.

2. Inwiefern teilt die Bundesregierung die Auffassung des Chefs der Bundespolizei, Dr. Dieter Romann, der anführte, die Technik erleichtere es, „Straftäter ohne zusätzliche Polizeikontrollen zu erkennen und festzunehmen“ (Pressemitteilung des Bundesministeriums des Innern, für Bau und Heimat – BMI – vom 11. Oktober 2018)?

Inwiefern hält sie es für grundrechtskonform, eine Person festzunehmen nur weil sie von einem automatisierten System identifiziert wurde, ohne wenigstens noch zuvor eine Identitätsfeststellung („Polizeikontrolle“) vorzunehmen?

Maßnahmen der Bundespolizei müssen stets im Rahmen der gesetzlichen Vorgaben erfolgen.

Die Bundesregierung weist darauf hin, dass in der Pressemitteilung des Bundesministeriums des Innern, für Bau und Heimat direkt im Anschluss an die Aussage von Herrn Dr. Romann die Pressemitteilung wie folgt fortführt: „Die getesteten Systeme sollen dabei jedoch nur alarmieren, nicht entscheiden. Einer Alarmierung folgt in jedem Fall eine Bewertung durch Polizistinnen oder Polizisten, ob der automatisch generierte Treffer ein polizeiliches Einschreiten erfordert und welche Maßnahmen ergriffen werden müssen.“ Die Aussage von Herrn Dr. Romann ist in diesem Zusammenhang zu verstehen.

3. Wie müsste sich die Polizei nach Einschätzung der Bundesregierung verhalten, wenn eine zur Fahndung ausgeschriebene Person von einem Gesichtserkennungssystem „erkannt“ wird?
 - a) Inwiefern darf diese Person ohne weitere Identifizierungskontrolle in Gewahrsam genommen werden?
 - b) Inwiefern wäre es legitim, sie zum Zweck der Identitätsfeststellung in Gewahrsam zu nehmen?

Die Fragen 3 bis 3b werden gemeinsam beantwortet.

Auf die Antwort zu Frage 2 wird verwiesen.

4. Teilt die Bundesregierung die Einschätzung im Abschlussbericht, Falschtreffer könnten „im Rahmen der unmittelbaren Validierung“ mittels visuellem Abgleich durch Angehörige der Bundespolizei „unverzüglich als solche erkannt werden“, und wenn ja, auf welcher Grundlage bzw. welchen Erfahrungswerten?

Etwaige Fehler eines möglichen künftigen Gesichtserkennungssystems können durch den anschließenden visuellen Abgleich durch Polizeivollzugsbeamtinnen und Polizeivollzugsbeamte erkannt und als solche festgestellt werden. Für den Betroffenen ergeben sich dann keine weiteren Folgen. Irrtümliche Verwechslungen sind im Übrigen im polizeilichen Alltag auch ohne Gesichtserkennungstechnik nicht gänzlich auszuschließen. Im Übrigen wird auf die Antwort zu Frage 2 verwiesen.

- a) Inwiefern war die Validierung von Falschtreffern auf den Monitoren durch Bundespolizistinnen und Bundespolizisten Gegenstand des Projektes, welche Vorgaben waren den im Rahmen des Projektes eingesetzten Beamten gemacht worden, und welche Ergebnisse kann die Bundesregierung hierzu mitteilen?

Dies war nicht Bestandteil des Tests.

- b) Teilt die Bundesregierung die Annahme der Fragestellerinnen und Fragesteller, dass die Fähigkeit bzw. schieere Möglichkeit zur unmittelbaren Validierung durch visuellen Abgleich in dem Maße abnimmt, in dem die Gesamtzahl gespeicherter (und abzugleichender) Bilder zunimmt, und dadurch entsprechend auch das Risiko steigt, dass Personen, bei denen das System einen Falschtreffer auslöst, einer polizeilichen Maßnahme unterzogen werden (falls nicht, bitte begründen)?

Die Fähigkeit eines Menschen, zwei Gesichtsbilder visuell abzugleichen und deren Übereinstimmung zu überprüfen, ist nicht von der Anzahl der in der Referenzdatenbank gespeicherten Bilder abhängig. Bei der unmittelbaren Validierung durch visuellen Abgleich hat der automatisierte Datenabgleich schon stattgefunden und es wäre nur noch der mögliche Treffer zu bestätigen.

5. Inwiefern ist die Bundesregierung der Ansicht, eine ODER-Verknüpfung von Gesichtserkennungssystemen unter Inkaufnahme einer erhöhten Falschtrefferrate und der dadurch bedingten irrtümlichen Identifizierung von mehreren Hundert Personen an einem Bahnhof pro Tag sei geeignet, in besonderen polizeilichen Lagen die Sicherheit zu erhöhen?

Wie ist nach ihrer Einschätzung hierbei der polizeiliche Mehraufwand zu beziffern und zu berücksichtigen, der durch die Falschtreffer entsteht?

In einer polizeilichen Sonderlage, wie z. B. bei einer Fahndung nach einem Tatverdächtigen eines kürzlich ausgeübten oder unmittelbar bevorstehenden terroristischen Anschlags, kann die Sicherheit durch den Einsatz eines möglichst effektiven Instruments zur Fahndungsunterstützung erhöht werden. Ein „ODER-verknüpftes“ System zur biometrischen Gesichtserkennung wäre rein technisch gesehen als ein solches Werkzeug in Betracht zu ziehen. Der polizeiliche Mehraufwand kann dabei nicht abstrakt beziffert werden.

6. Welche konkreten gemessenen Auswirkungen auf die Trefferrate bzw. Falschtrefferrate hatte die witterungsbedingt verwendete Kleidung der Testpersonen (insbesondere Schals, Mützen usw.)?

Inwiefern wurde in Testphase 1 erfasst, wie viele Testpersonen an wie vielen Tagen Schals, Mützen oder andere Verdeckungen trugen und inwiefern dadurch eine biometrische Gesichtserkennung erschwert bzw. unmöglich war, und inwiefern flossen diese Werte in die Ergebnisbewertung ein?

Im Erprobungszeitraum konnten sämtliche Jahreszeiten und Witterungsbedingungen dargestellt werden, so dass auch witterungsbedingt entsprechende Kleidung durch die Freiwilligen verwendet wurde. Bei dem Test handelte es sich um eine praxisnahe Erprobung der Technik, eine umfassende Auswertung aller möglichen Einflussfaktoren war nicht geplant. Insbesondere wurden keine Messreihen zur Ermittlung der in Frage stehenden Korrelation der getragenen Mützen und Schals mit der Güte der Treffer durchgeführt. Gleichwohl wurde bei einzelnen Proben festgestellt, dass das Tragen von Mützen und Schals die Gesichtserkennung nicht grundsätzlich beeinträchtigt.

7. Welche konkreten gemessenen Auswirkungen auf die Trefferrate bzw. Falschtrefferrate hatte es, wenn Personen, insbesondere bei Erfassung der Kamera an der Rolltreppe des Bahnhofs, nach unten bzw. zur Seite sahen, statt frontal geradeaus?

Alle getesteten Systeme waren grundsätzlich in der Lage, Gesichter auch bei ungünstigen Blickwinkeln zu detektieren und zu identifizieren. Im Übrigen wird auf die Antwort zu Frage 6 verwiesen.

8. Inwiefern wurde während des Versuchs bzw. seiner Auswertung die Möglichkeit berücksichtigt, bzw. inwiefern geht die Bundesregierung davon aus, dass das Wissen der Testpersonen um den Test Auswirkungen auf ihr Verhalten und damit das Testergebnis hatte?
 - a) Inwiefern wurde während des Projektes erfasst oder zumindest auf einer theoretischen Ebene berücksichtigt, dass die Testpersonen über eine überdurchschnittlich hohe Motivation verfügen könnten, ihre Identifikation zu ermöglichen, und deswegen möglicherweise häufiger als andere Reisende direkt in Kameras zu sehen, bzw. weniger häufig beispielsweise auf ihre Smartphones (nach unten) oder zur Seite, was ihre Identifikation erschwert hätte?
 - b) Wurden die Testpersonen vor Versuchsbeginn oder danach zu ihrer Motivation und ihrem diesbezüglichen Verhalten befragt (bitte ggf. Ergebnisse angeben)?
 - c) Inwiefern geht die Bundesregierung davon aus, dass ein Wirkbetrieb (ohne zuvor ausgewählte und über den Versuch informierte Probanden) zu gleichen Ergebnissen hinsichtlich Treffer- und Falschtrefferrate führen würde (bitte begründen)?

Die Fragen 8 bis 8c werden gemeinsam beantwortet.

Es gab keine Vorgaben durch die Bundespolizei in Bezug auf das Verhalten der Freiwilligen bei Durchschreiten der Testbereiche. Auf Nachfrage wurde Freiwilligen mitgeteilt, dass diese sich entsprechend natürlich verhalten sollen. Es gibt keinen Grund zur Annahme, dass ein etwaiger Wirkbetrieb zu anderen Ergebnissen hinsichtlich Treffer und Falschtrefferrate führen würde.

9. Warum wurde die im Abschlussbericht getroffene Feststellung, dass sich besondere Ansprüche an Gesichtsdetektoren ergeben durch Alter, Hautfarbe und Geschlecht, unterschiedliche Emotionen bzw. Gesichtsausdrücke, das mögliche Vorhandensein von Verdeckungen z. B. durch Brillen und Bärte, sowie eine Vielzahl kosmetischer Aspekte wie beispielsweise Frisuren, bei der Auswahl der Testpersonen nicht berücksichtigt, sondern nur darauf abgestellt, mindestens 84 Personen (als Minimum für eine repräsentative Auswahl) zu gewinnen bzw. letztlich sämtliche Bewerberinnen und Bewerber zu akzeptieren?
 - a) Wie gliedern sich die Testpersonen in Hinsicht auf Alter, Hautfarbe, Geschlecht, Brillen- und Barträger auf, und inwiefern ist eine Repräsentativität hinsichtlich der Gesamtbevölkerung bzw. der durchschnittlichen Nutzerinnen und Nutzer des Bahnhofs Südkreuz gegeben (bitte entsprechendes Bestimmungsverfahren erläutern)?

- b) Inwiefern wurde auf mögliche diskriminierende Algorithmen bzw. Wirkungen einer biometrischen Gesichtserkennung geachtet, insbesondere hinsichtlich der Treffer- bzw. Falschtreffermeldungen bei Menschen mit dunklerer Hautfarbe?

Falls solche Aspekte nicht berücksichtigt wurden, warum nicht, und welche Auswirkungen hatte dies nach Auffassung der Bundesregierung auf die Aussagekraft der Testergebnisse?

Die Fragen 9 bis 9b werden gemeinsam beantwortet.

Bei der Auswahl der Testpersonen wurde besonderer Wert daraufgelegt, ein „Spiegelbild“ der Reisenden, die den Bahnhof Berlin Südkreuz nutzen, darzustellen (so realitätsnah wie möglich). Durch diese Zielsetzung wurden viele der personengebundenen Einflussfaktoren, wie z. B. Geschlecht, Alter sowie das Tragen einer Brille oder Kosmetik abgedeckt. Zum Abdecken anderer Faktoren, wie abweichende Mimik oder Emotionen, ist keine erhöhte Anzahl der Testpersonen notwendig. Insofern war zur Berücksichtigung der genannten Faktoren keine abweichende Mindestanzahl der Testpersonen notwendig.

Im Übrigen wird auf die Antwort zu Frage 21 verwiesen.

10. Wie viele Falschpositivtreffer hat es insgesamt während des Versuchs gegeben und wie viele Falschnegativtreffer (für Testphase 1 bitte zusätzlich jeweils pro Woche aufliedern)?

Da keine vollständige Validierung des gesamten Zeitraums durchgeführt wurde, kann diese Frage nur in Bezug auf die validierten Testwochen beantwortet werden.

Im ersten Halbjahr wurden von den drei Gesichtserkennungssystemen insgesamt 4397 falsch positive sowie 767 falsch negative Treffer festgestellt, geprüft wurden insgesamt ca. 930 000 Personen. Im zweiten Halbjahr wurden von den drei Gesichtserkennungssystemen insgesamt 6760 falsch positive Treffer sowie 567 falsch negative festgestellt, geprüft wurden insgesamt ca. 2 010 000 Personen.

Eine Aufschlüsselung nach Kalenderwochen ist nicht möglich, da auch die zugrundeliegende Validierung nicht nach Kalenderwochen durchgeführt wurde.

11. Nach welchen Kriterien wurden von den 41 000 vorliegenden Datensätzen der Transponderlogdateien in Testphase 1 lediglich 6 000 mit einem Treffer sowie 4 500 mit einem Falschtreffer überprüft (5.1.1. im Abschlussbericht), und wie viele Logdateien der Testphase 2 fielen an, und wie viele davon wurden überprüft (bitte auch hier Kriterien angeben)?

In beiden Testphasen wurden pro Monat in der Regel mindestens 7 Tage validiert. Aus der Auswahl dieser Testtage resultiert auch die Wahl der an diesen Tagen angefallenen Transponderdatensätze. Während der zweiten Testphase sind ca. 18 000 Transponderdatensätze erfasst worden, 6 835 wurden davon validiert.

12. Welche Ergebnisse erbrachten die Einzelprüfungen solcher vom System gemeldeten Treffer, die ohne Anwesenheit eines Transponderträgers gemeldet wurden, und inwiefern flossen diese Prüfungen in die im Abschlussbericht genannte Falschtrefferrate ein (falls sie nicht einfließen, bitte nachholen und den Wert angeben)?

Bei wie vielen richtig positiven Treffern hatten die identifizierten Personen einen Transponder dabei, und bei wie vielen nicht?

Die im Bericht genannte Falschtrefferrate beinhaltet alle validierten und als falsch beurteilten Treffer unabhängig von der Transponderanwesenheit. Die zur Beantwortung der zweiten Teilfrage notwendige Auswertung wurde während des Testzeitraums nicht vorgenommen und sie kann nachträglich nicht erfolgen, da die hierzu notwendigen personenbezogenen Daten gelöscht wurden.

13. Wie viele Personen passieren nach Kenntnis der Bundesregierung durchschnittlich pro Tag den Bahnhof Südkreuz (ohne durchreisende Personen), und wie viele Personen wurden während des Versuchszeitraums durchschnittlich pro Tag von den installierten Kameras erfasst?

Nach Kenntnis der Bundesregierung passieren durchschnittlich 106 000 Personen pro Tag den Bahnhof Berlin Südkreuz. Im Durchschnitt wurden während des Versuchszeitraums ca. 26 000 Personen pro Tag von den für den Test der Systeme zur Gesichtserkennung verwendeten Kameras erfasst.

14. Warum enthält die Auswertung der Trefferrate (bezogen auf das „Gesamtsystem“) für die Testphase 1 neun Datenpunkte für den November 2017, die Auswertung der Falschtrefferrate aber lediglich einen Datenpunkt für den gleichen Monat?

Welche Werte für Trefferrate und Falschtrefferrate wurden insgesamt erreicht (bitte pro Testwoche, wie in den Diagrammen im Anhang 3 des Abschlussberichts, darstellen)?

In die Berechnung der Falschtrefferrate wird die Gesamtanzahl der Personen einbezogen. An den Tagen im November 2017, an denen im Diagramm keine Datenpunkte angezeigt werden, ist es zu technischen Fehlern bei der Übertragung der Logdateien mit Personenzahlen gekommen, daher konnten auch aus den vorhandenen absoluten Werten der Falschtrefferanzahl keine Prozentzahlen berechnet werden. Die absoluten Werte dieser Tage weichen jedoch nicht signifikant von den vergleichbaren Wochentagen des Monats ab, insofern hat dieser Fehler keine Auswirkung auf das Gesamtergebnis.

15. Welche Auswirkungen auf die Übertragbarkeit der Testergebnisse auf andere Jahreszeiten hat nach Einschätzung der Bundesregierung die zeitliche Eingrenzung der Auswertung in Testphase 1 auf die Monate Oktober bis Januar (Fußnote 22 auf S. 31 des Abschlussberichtes) vor dem Hintergrund der Aussage im Abschlussbericht, dass sich der Effekt negativer Auswirkungen auf die Bildqualität durch Gegenlicht „insbesondere in den Sommermonaten bemerkbar“ mache?

In diesem Erprobungszeitraum konnten auch sämtliche Jahreszeiten, Witterungsbedingungen und Lichtverhältnisse dargestellt werden. Im Übrigen wird auf die Vorbemerkung der Bundesregierung verwiesen.

16. Für wie realistisch hält die Bundesregierung die Ausgangslage in Testphase 1 (Vorliegen qualitativ hochwertiger Gesichtsbilder) in Hinsicht auf einen Wirkbetrieb?
- a) Von wie vielen als Gefährder eingestuften Personen liegen den Sicherheitsbehörden nach Kenntnis der Bundesregierung qualitativ hochwertige Gesichtsbilder vor?
- b) Von wie vielen Personen, die derzeit wegen schwerer Gewalttaten zur Fahndung ausgeschrieben sind, liegen den Sicherheitsbehörden nach Kenntnis der Bundesregierung qualitativ hochwertige Gesichtsbilder vor?

Die Fragen 16 bis 16b werden gemeinsam beantwortet.

Es findet keine statistische Erhebung im Hinblick auf die Anzahl vorliegender qualitativ hochwertiger Bilder statt. Unabhängig davon sind in einzelnen polizeilichen Datenbanken auch entsprechend qualitativ hochwertige Bilder vorhanden.

17. Warum wurde in Testphase 2 darauf verzichtet, in die Referenzbilddateien solche Bilder der Testpersonen einzupflegen, die aus anderen, nicht im Bahnhof Südkreuz eingesetzten, Videokameras aufgenommen wurden?
- Welche Auswirkung hatte nach Einschätzung der Bundesregierung die Einschränkung der Versuchsanordnung auf die Identifikation der „eigenen“ Bilder auf das Testergebnis?

Die in der zweiten Testphase verwendeten Bilder entstammen den am Bahnhof Berlin Südkreuz vorhandenen Videoüberwachungssystemen. Diese Anlage entspricht modernen Videoüberwachungsanlagen ohne Systeme zur Gesichtserkennung, wie sie bereits auf mehreren Bahnhöfen der Deutsche Bahn AG verwendet werden. Entsprechend hätten die Videobilder auch aus diesen Videoüberwachungsanlagen anderer Bahnhöfe entnommen werden können. Insoweit wird keine Auswirkung auf den Test gesehen.

18. Gibt es in der Bundesregierung tatsächlich Überlegungen, sämtliche auf freiem Fuß befindlichen islamistischen „Gefährder“ in einer Bilddatenbank zu speichern, um sie im möglichen künftigen Wirkbetrieb biometrischer Gesichtserkennung auf Bahnhöfen identifizieren zu können, und wenn ja, auf welcher Rechtsgrundlage, warum, und warum werden im Abschlussbericht nur islamistische „Gefährder“ genannt und keine Neonazis?

Die Meinungsbildung in der Bundesregierung über einen etwaigen Wirkbetrieb ist insgesamt noch nicht abgeschlossen.

19. Hat es im Vorfeld des Versuches Kostenschätzungen gegeben bzgl. der Entwicklung einer eigenen Software (durch Sicherheitsbehörden des Bundes; bitte ggf. angeben)?

Die Durchführung des Projektes war auf die Erprobung vorhandener technischer Möglichkeiten ausgelegt. In diesem Sinne war eine Eigenentwicklung von Softwarelösungen durch die Bundespolizei nicht Gegenstand des Projektes und demzufolge wurden keine Kostenbetrachtungen im Vorfeld angestellt.

20. Liegen der Bundesregierung bereits Kostenprognosen hinsichtlich einer Nutzung der biometrischen Gesichtserkennung im Wirkbetrieb vor (bitte erläutern, welche Anbieter und wie die Prognose zusammengesetzt ist), und welche Auswirkungen auf die Kosten hätte ein kombinierter Einsatz dreier verschiedener Systeme (als „Gesamtsystem“) gegenüber dem Einsatz nur eines Systems?

Der Bundesregierung liegen keine Kostenprognosen im Sinne der Fragestellung vor.

21. Hatte die Bundespolizei (oder andere Behörden des Bundes) die Möglichkeit, die für die Systeme der verschiedenen Anbieter verwendeten Algorithmen einzusehen?

Über welche Fähigkeiten verfügt die Bundespolizei (oder hilfsweise hinzugezogene andere Bundesbehörden), mögliche Manipulationen der Systeme durch die Anbieter oder durch (auch unbefugte) Dritte zu erkennen, insbesondere solche, die es ermöglichen, bestimmte Personen(typen) mit höherer oder geringerer Wahrscheinlichkeit richtig positiv zu erkennen als andere?

Die Bundespolizei oder andere Behörden des Bundes hatte keine entsprechende Möglichkeit, da es sich bei den Algorithmen grundsätzlich um Firmengeheimnisse handelt. Der Bundesregierung liegen keine Erkenntnisse darüber vor, dass Hersteller der Gesichtserkennungsalgorithmen absichtlich die Algorithmen manipulieren, um die Erkennungsleistung zu beeinflussen. Zu den im Test verwendeten Gesichtserkennungssystemen hatten „Dritte“ keinen Zugang.

Der Test fand auf einem abgeschotteten System der Deutsche Bahn AG statt ohne Zugriff auf das Internet. Lediglich die Bundespolizei hatte Zugriff auf dieses Netz. Eine Möglichkeit zur Manipulation war somit nicht gegeben.

22. Wie und auf welcher Rechtsgrundlage wurden Daten (Fotos) jener Personen gespeichert und bearbeitet, bei denen das System einen Falschtreffer ausgelöst hatte?

Wer hatte für welche Zeiträume Zugriff auf diese Daten?

Sind alle diese Daten mittlerweile gelöscht?

23. Welche Daten (incl. Logdateien und transponderbezogene Daten) aus dem Versuch sind derzeit noch nach Kenntnis der Bundesregierung gespeichert, wo sind sie gespeichert, wer hat Zugriff auf diese Daten, und inwiefern erlauben diese Daten eine Zuordnung zu konkreten Personen?

Haben die Anbieter aller drei Systeme (Auftragnehmer) mittlerweile sämtliche in ihren Besitz gelangte Unterlagen, Verarbeitungs- und Nutzungsergebnisse sowie Datenbestände (sowie hiervon gefertigte Kopien oder Reproduktionen) zurückgegeben bzw. vernichtet, und liegt hinsichtlich der Vernichtung ein Lösungsprotokoll vor?

Die Fragen 22 und 23 werden gemeinsam beantwortet:

Die Rechtsgrundlage für die Bundespolizei für die Erprobung von Systemen zur intelligenten Videoanalyse am Bahnhof Berlin Südkreuz ist nach Auffassung des Bundesministeriums des Innern, für Bau und Heimat § 27 Satz 1 Nummer 2 BPolG. Die Testpersonen haben am abgeschlossenen Test der Gesichtserkennungssysteme auf Basis einer Einwilligung teilgenommen. An den Abgleich mit einer Testdatenbank wurden keinerlei Rechtsfolgen geknüpft und die erhobenen

Daten wurden spätestens innerhalb der gesetzlichen Speicherfristen wieder gelöscht. Ein grundrechtsrelevanter Eingriff, der über den mit der konventionellen Videoüberwachung einhergehenden Eingriff wesentlich hinausgeht, wird daher nicht gesehen.

Es hatte nur die Bundespolizei Zugriff auf die entsprechenden Daten. Alle personenbezogenen Daten sowie sämtliche Daten, die einen Rückschluss auf eine Person ermöglichen würden, wurden innerhalb der gesetzlichen Speicherfristen bereits gelöscht.

Lediglich die nicht-personenbezogenen Daten der Logdateien sowie der transponderbezogenen Daten (ohne Möglichkeit der Zuordnung zu einer Person) liegen der Bundespolizei weiterhin vor.

24. Inwiefern will die Bundesregierung der Handlungsempfehlung der Bundespolizei folgen, an weiteren Bahnhöfen eine biometrische Gesichtserkennung einzuführen und hierzu entweder neue technische Anlagen einzurichten oder vorhandene nachzurüsten?

Es wird auf die Antwort zu Frage 18 verwiesen.

25. Ist die Bundesregierung der Ansicht, dass auf den Einsatz biometrischer Gesichtserkennungssysteme auch im Wirkbetrieb explizit hinzuweisen wäre, oder ist sie der Ansicht, es genüge ein allgemeiner Hinweis auf Videoüberwachung (bitte begründen)?

Es wird auf die Antwort zu Frage 18 verwiesen.

26. Wer soll in die vom BMI angekündigten Gespräche über Bedingungen und Umfang eines Wirkbetriebs biometrischer Gesichtserkennung (Presseerklärung des BMI vom 11. Oktober 2018) eingebunden werden, und inwiefern gibt es hierfür bereits einen Zeitplan?

In die Gespräche sollen alle innerhalb der Bundesregierung betroffenen Ressorts sowie die betroffenen Interessensvertreter und Datenschutzbeauftragte einbezogen werden. Die Planungen hierzu sind noch nicht abgeschlossen.

