

## **Kleine Anfrage**

**der Abgeordneten Ulla Jelpke, Andrej Hunko, Christine Buchholz, Jens Petermann, Dr. Petra Sitte, Halina Wawzyniak und der Fraktion DIE LINKE.**

### **Automatisierte Strafverfolgung, Data Mining und sogenannte erweiterte Nutzung von Daten in polizeilichen Informationssystemen**

Polizeibehörden beklagen sich über eine immense Zunahme digitaler Datenbestände. Von 2007 bis 2009 hätten sich laut Moritz Aly vom Bundeskriminalamt (BKA) bestehende Datenmengen jährlich verdoppelt, „ab 2010 dann stieg das Aufkommen explosionsartig“ (Vortrag auf der Veranstaltung „Europäischer Polizeikongress“, 14. Februar 2012). Behörden wollen die Informationen zunehmend automatisiert durchforsten und aufbereiten. Auf dem Europäischen Polizeikongress widmete sich ein eigenes Panel einer effektiveren digitalen Informationsverarbeitung. Das BKA stellte dort dessen „Inhaltliche Datenträgerauswertung“ (IDA) vor, die das Amt in Eigenregie entwickelt hat. Die Software erkennt Sprachen, sucht doppelte Einträge, filtert relevante Daten heraus und erstellt die für andere Programme notwendigen Statistiken. Dies wurde kürzlich von der Bundesregierung in der Antwort auf die Schriftliche Frage 19 des Abgeordneten Jan Korte auf Bundestagsdrucksache 17/10925 bestätigt. Die „Inhaltliche Datenträgerauswertung“ visualisiert Zusammenhänge und schlägt Hypothesen vor.

Eine derart automatisierte Strafverfolgung wurde vom BKA auch auf den von Bürgerrechtsgruppen heftig kritisierten Schulungen in Weißrussland vorgeführt (Bundestagsdrucksache 17/10713). Gezeigt wurde „Analyst’s Notebook“ der von IBM aufgekauften Firma i2. Die Firma i2 stellte auf dem letzten Europäischen Polizeikongress „Werkzeuge zur Detailanalyse und Visualisierung von Beziehungen zwischen physischen Merkmalen und allen dazu verfügbaren Informationen“ vor. Die Bundesregierung bestätigt, dass „Analyst’s Notebook“ „Hypothesen“ ausgibt und Zusammenhänge visualisiert. Analyst’s Notebook wurde bislang vom deutschen Polizeizulieferer rola Security Solutions vermarktet. Weitere Analysesoftware wurde mit „InfoZoom“ durch das BKA von der Firma humanIT Software GmbH beschafft.

Auf Ebene der Europäischen Union werden Standards für polizeilich genutzte Informationstechnik erarbeitet. Die Europäische Kommission hat das Forschungsprogramm „Composite“ gestartet, das technologische Veränderungsprozesse von Polizeibehörden in zehn europäischen Ländern erforscht. Auch die Bundesrepublik Deutschland ist beteiligt. Im August 2012 hatte die EU im dänischen Odense die zweite „European Intelligence and Security Informatics Conference“ (EISIC) ausgerichtet, die sich mit Verfahren der „Social Network Analysis“ und des „Data Mining“ befasste. Für sein „Pre-frontier Intelligence Picture“, das auch auf der intelligenten Auswertung von statistischem Material beruht, forscht auch die Europäische Agentur für die operative Zusammenarbeit an den Außengrenzen (FRONTEX) an Verfahren zum „Text Mining“. FRONTEX

nimmt am Forschungsprojekt „OPensource Text Information Mining and Analysis“ (OPTIMA) teil. Das Europäische Polizeiamt (Europol) betreibt eine sogenannte Social Network Analysis, wie sie von IBM verkauft wird, schon länger. Europol nutzt laut der EU-Innenkommissarin Cecilia Malmström ebenso Software von i2 „für Analysen und Datamining“ sowie der Firma Themis „für Textmining“ (Kommissionsdokument E-000171/2012). Eine „akademische Software“ analysiert demnach „Subnetze und Schnittpunkte“.

Problematisch ist, wenn auf diese Weise bislang „unstrukturierte Daten“ aus mehreren Quellen verarbeitet werden. Durch dieses sogenannte Data Mining könnten Ermittler/Ermittlerinnen nach neuen Erkenntnissen suchen, indem Einträge auf „Kreuztreffer“ analysiert werden. Die Definition, ab welcher Eingriffstiefe es sich um ein polizeiliches „Data Mining“ oder eine Rasterfahndung handeln würde, ist allerdings umstritten. Auch über die Funktionalität einer „Vorhersage“ gibt es unterschiedliche Auffassungen. Laut dem Staatssekretär im Bundesministerium des Innern (BMI), Klaus-Dieter Fritsche, nutzen oder testen Bundesbehörden keine „prediktiven Analyseprogramme“ (Antwort der Bundesregierung auf die Schriftliche Frage 13 des Abgeordneten Andrej Hunko auf Bundestagsdrucksache 17/10503). Jedoch erklärte der BKA-Vizepräsident Prof. Dr. Jürgen Stock auf Fachkonferenzen, das BKA teste jede auf dem Markt befindliche „prediktive Software“ und berichte hierzu den Landeskriminalämtern.

Dass sich Behörden auf zunehmendes „Data Mining“ vorbereiten, zeigt die Verankerung einer „erweiterten Nutzung“ im „Gesetz zur Errichtung einer standardisierten zentralen Datei von Polizeibehörden und Nachrichtendiensten von Bund und Ländern zur Bekämpfung des gewaltbezogenen Rechtsextremismus“. Gemeint sind „das Herstellen von Zusammenhängen zwischen Personen, Personengruppierungen, Institutionen, Objekten und Sachen, der Ausschluss von unbedeutenden Informationen und Erkenntnissen, die Zuordnung eingehender Informationen zu erkannten Sachverhalten sowie die statistische Auswertung der gespeicherten Daten“. Im Gesetzestext heißt es, dass Behörden „Beziehungen zwischen Personen und Zusammenhänge zwischen Personen, Personengruppierungen, Institutionen, Objekten und Sachen darstellen“ dürfen. In der Antwort auf die Mündliche Frage 39 des Abgeordneten Andrej Hunko (Plenarprotokoll 17/194) erklärt der Parlamentarische Staatssekretär beim Bundesminister des Innern, Dr. Ole Schröder, die „erweiterte Nutzung“ der „Rechtsextremismusdatei“ sei ab einer nun folgenden Projektphase vorgesehen.

Wenn aber Computer Zusammenhänge zwischen Personen, Sachen und Orten analysieren, bedeutet dies eine grundsätzliche Veränderung der Polizeiarbeit. Der erhoffte „Mehrwert“ unstrukturierter Information stellt eine automatisierte Wissensproduktion durch Polizeien und Geheimdienste dar. Die ohnehin zu beobachtende Verlagerung der Strafverfolgung ins Vorfeld bekommt digitale Schützenhilfe. Indikatoren und Routinen, nach denen die Datenbestände durchsucht werden, folgen der polizeilichen Definition von „Risiken“. Wenn eine Software Personendaten auf etwaige Risiken untersucht, muss deren Funktionsweise, mithin ihr Quellcode, bekannt sein. Dies gilt nicht nur für staatliche Trojaner und Programme zum „Herstellen von Zusammenhängen zwischen Personen, Personengruppierungen, Institutionen, Objekten und Sachen“.

Wir fragen die Bundesregierung:

1. Teilt die Bundesregierung die Ansicht der EU-Innenkommissarin Cecilia Malmström, dass es sich bei den im Kommissionsdokument E-000171/2012 beschriebenen Werkzeugen bzw. Vorgängen um ein „Data Mining“ durch die EU-Polizeiagentur Europol handelt?
  - a) Sofern es sich aus Sicht der Bundesregierung nicht um „Data Mining“ handelt, wie würde sie die beschriebene Nutzung „akademische[r] Soft-

ware“ und eines „speziellen Analyserahmen[s]“ ansonsten klassifizieren, der demnach zum Einsatz kommt, um „Netzkomponenten, Subnetze und Schnittpunkte zu analysieren, Zentralität zu ermitteln und die wichtigsten Akteure herauszufiltern“?

- b) Inwiefern sind Bundesbehörden über den Einsatz von Software der Firmen i2 und Themis bzw. des „Freeware-Tools (Pajek)“ bei Europol informiert?
  - c) Inwiefern haben auch Bundesbehörden bereits von Software der Firmen i2 und Themis mittelbar oder unmittelbar bei Europol Gebrauch gemacht?
2. Welche datenschutzrechtlichen Vorgaben gelten für Europol bei der Analyse von Netzkomponenten, Subnetze[n] und Schnittpunkte[n]?
- Inwiefern ist nach Ansicht der Bundesregierung hiervon auch ein „Data Mining“ erfasst?
3. Inwieweit nutzt die FRONTEX Verfahren zur automatisierten Analyse von Personen- oder Sachdaten?
- a) Inwieweit trifft es zu, dass FRONTEX für das „Common Pre-frontier Intelligence Picture“ an Verfahren zum „Text Mining“ forscht?
  - b) Worin besteht die Aufgabe der FRONTEX innerhalb des Forschungsprojekts OPTIMA?
4. Teilt die Bundesregierung die Ansicht der Fragesteller, dass es sich beim „Herstellen von Zusammenhängen zwischen Personen, Personen-gruppierungen, Institutionen, Objekten und Sachen, d[em] Ausschluss von unbedeutenden Informationen und Erkenntnissen, d[er] Zuordnung eingehender Informationen zu erkannten Sachverhalten sowie d[er] statistische[n] Auswertung der gespeicherten Daten“ um „Data Mining“ handelt?
- Falls nein, worin läge eine andere Definition von „Data Mining“ bzw. der beschriebenen Vorgänge?
5. Welche datenschutzrechtlichen Vorgaben gelten im Zuständigkeitsbereich des BMI bei einem „Data Mining“ bzw. der Verknüpfung von im Internet oder anderen öffentlichen Quellen ermittelten Informationen mit anderen Datensätzen für Behörden des Bundes (Bundestagsdrucksache 17/6587)?
- a) Welche weitergehenden Arbeitsanweisungen existieren bei welchen Behörden?
  - b) Inwieweit sehen Errichtungsanordnungen für andere vom BKA geführte Datensammlungen bereits eine „erweiterte Nutzung“ im Sinne der Fragestellung vor?
  - c) Inwieweit stimmt die Bundesregierung der Ansicht der Fragesteller zu, dass für eine „Gewichtung“ von Suchkriterien oder einer Suche nach „unstrukturierten Informationen“ in anderen vom BKA geführten Datensammlungen eine neue Errichtungsanordnung erforderlich wäre?
6. Inwieweit nutzen Behörden des BMI „Werkzeuge zur Detailanalyse und Visualisierung von Beziehungen zwischen physischen Merkmalen und allen dazu verfügbaren Informationen“, wie sie die Firma i2 auf dem Europäischen Polizeikongress bewarb?
- a) Seit wann kommt das von der Firma vertriebene „Analyst’s Notebook“ bei Behörden des BMI zum Einsatz?
  - b) Welche Kosten für Beschaffung, Wartung und Lizenzen fielen seitdem an?

- c) Von wem wurde die Software konkret beschafft?
- d) Inwieweit werden auch die Anwendungen „Social Network Analysis“, „Quick-Navigator“, „Google Earth-Plug-in“ sowie weitere „Filter und Histogramme“ eingebunden?
7. Auf welche beim BKA geführte Datensammlungen darf „Analyst’s Notebook“ zugreifen?
- a) Welche Art von „Hypothesen“ gibt „Analyst’s Notebook“ aus, und wie werden diese im Regelfall weiter prozessiert?
- b) Hat die Nutzung von „Analyst’s Notebook“ nach Ansicht der Bundesregierung dazu beigetragen, die Strafverfolgung oder Gefahrenabwehr bei Bundesbehörden zu erleichtern?
8. Auf welche Art und Weise unterstützt die Inhaltliche Datenträgerauswertung (IDA) des BKA auch eine „Auswertung von Texten“ (Antwort auf die Schriftliche Frage 19 auf Bundestagsdrucksache 17/10925)?
- a) Auf welche Art und Weise und nach welchem mathematischen Verfahren werden Dateien „bewertet“?
- b) Was ist gemeint, wenn die Bundesregierung diesbezüglich von einer „Erkennung unterschiedlicher Sprachen“ schreibt?
- c) In welchem Umfang wurde die IDA bereits in der „Zusammenfassung mehrerer Ermittlungsverfahren“ eingesetzt?
- d) Da die Bundesregierung die Auffassung vertritt, die IDA sei keine rasternde Analysesoftware, auf welche Weise nehmen Bundesbehörden dann derartige Ermittlungen vor?
9. Inwiefern hat die Nutzung von Werkzeugen wie „Analyst’s Notebook“ oder „Infozoom“ der Firma humanIT Software GmbH bei Polizeien des Bundes in den letzten fünf Jahren zugenommen (Bundestagsdrucksache 17/8089)?
- In welchem Umfang werden „Analyst’s Notebook“ und „Infozoom“ auch zur Gefahrenabwehr eingesetzt?
10. Mit welchen Verfahren sollen die in der „Rechtsextremismusdatei“ gesammelten Daten einer „erweiterten Nutzung“ unterzogen werden?
- a) Welche „Projektphasen“ existieren für den Aufbau der Datensammlung, wie es die Bundesregierung in der Fragestunde des Deutschen Bundestages am 26. September 2012 berichtet?
- b) Welche Schritte enthält die entsprechende „Projektphase“ zur „erweiterten Nutzung“?
- c) Wer ist an der Definition, Entwicklung und Beschaffung der notwendigen technischen sowie sonstigen Mittel beteiligt?
- d) Über welche Funktionalitäten verfügt die „Anwendungssoftware“, mit der das BKA bereits jetzt Daten in der „Rechtsextremismusdatei“ verarbeitet?
- e) Inwiefern dient „Inpol-Fall“ wie im Falle der „Antiterrordatei“ auch bei der „Rechtsextremismusdatei“ als „Quellsystem“ (Bundestagsdrucksache 17/8544 – neu)?
11. Inwieweit ist eine „erweiterte Nutzung“ auch für andere polizeiliche oder geheimdienstliche Datensammlungen beabsichtigt?

12. Mit welchen technischen Mitteln sind Behörden der Bundesregierung bereits in der Lage, wie für die „Rechtsextremismusdatei“ geplant, „räumliche und sonstige Beziehungen zwischen Personen und Zusammenhänge zwischen Personen, Personengruppierungen, Institutionen, Objekten und Sachen“ darzustellen sowie „die Suchkriterien [zu] gewichten“?
  - a) Wie gestalten Behörden die Suche nach phonetischen oder unvollständigen Daten, bzw. welche Kapazitäten werden hierfür entwickelt?
  - b) Wie gestalten Behörden die Suche über eine Mehrzahl von Datenfeldern, bzw. welche Kapazitäten werden hierfür entwickelt?
  - c) Wie gestalten Behörden die Verknüpfung von Personen, Institutionen, Organisationen, Sachen, bzw. welche Kapazitäten werden hierfür entwickelt?
  - d) Wie gestalten Behörden die zeitliche Eingrenzung der Suchkriterien, bzw. welche Kapazitäten werden hierfür entwickelt?
  - e) Sofern die Bundesregierung zur in den Fragen 12a bis 12d beschriebenen Suche noch nicht in der Lage ist, auf welche Art und Weise wird die entsprechende technische Umsetzung betrieben, bzw. welche Kapazitäten werden hierfür entwickelt?
13. Welche Software mit welchen Zusatzfunktionen wird bei Bundesbehörden normalerweise für eine Rasterfahndung eingesetzt?
  - a) Welche Definition existiert bei Bundesbehörden für den Begriff „Rasterfahndung“, und wie wird diese zu „Data Mining“ abgegrenzt?
  - b) Inwieweit wurden die Massendaten im Zusammenhang mit der (damals noch nicht bekannten) Täterschaft des NSU auch mit „Rasterfahndungen“ prozessiert?
  - c) Wie viele Anordnungen zur Rasterfahndung sind nach Kenntnis der Bundesregierung von Gerichten ergangen, und wie viele Ersuchen wurden zurückgewiesen?
  - d) In welchen weiteren Datenbeständen durfte nach Erlass der Beschlüsse jeweils gesucht werden?
14. Mit welcher Software mit welchen Funktionalitäten und welchen Zusatzfunktionen werden bei Bundesbehörden „Massendaten“ verarbeitet?
  - a) Welche Löschfristen existieren für die verschiedenen „Massendaten“ jeweils im Regelfall und im hier behandelten Fall?
  - b) Wie viele sogenannte Massendaten wurden bzw. werden in den Ermittlungen wegen der ungeklärten Mordfälle bis zur Entdeckung der Täterschaft des „Nationalsozialistischen Untergrunds“ von den damit befassten Bundes- und Landesbehörden nach Kenntnis der Bundesregierung erhoben, gespeichert oder verarbeitet?
  - c) Welche Software kam hier nach Kenntnis der Bundesregierung mit welchen Funktionalitäten zum Einsatz?
15. Welche „verbundrelevante[n] Daten oder im Zusammenhang mit Großschadenslagen gewonnene[n] Daten“ werden derzeit automatisiert an „Inpol-Fall“ übertragen?
  - a) Wann und von wem wurden die Bund-Länder-Datei-Schnittstelle (BLDS) und die Bund-Länder-Online-Schnittstelle (BLOS) im BKA entwickelt?

- b) Inwiefern trifft es zu, dass die BDLS bzw. die BLOS im Auftrag der damals erfolglos ermittelnden Besondere Aufbauorganisation Bosphorus eingerichtet wurde?
  - c) Welche Kosten sind hierfür entstanden, und wer hat diese getragen?
  - d) Inwiefern kann die Bundesregierung die Lieferung der BDLS bzw. BLOS durch die Firma rola security solutions GmbH als erfolgreich bewerten?
  - e) Sofern die Bundesregierung Nachbesserungsbedarf sieht, wo läge dieser?
  - f) Worin besteht die „Weiterentwicklung“ von „Inpol-Fall“, die „auf Basis der Landesfallsoftware ‚Crime‘“ aus Hamburg und Hessen „durch das BKA selbst“ vorgenommen wurde?
16. Auf welche Art und Weise und mit welchen Funktionalitäten wurde die Analysesoftware des BKA „B-case“ der Firma rola Security Solutions GmbH „BKA-spezifisch angepasst“ (Bundestagsdrucksache 17/8544 – neu)?  
Welche „grundsätzlich alle beschafften Module“ sind in der vom BKA beschafften Analysesoftware „B-case“ eingebunden?
17. In wie vielen Fällen wurden in den letzten fünf Jahren „auf Antrag der ermittlungsführenden Abteilung“ über eine Schnittstelle Daten aus der Telekommunikationsüberwachung (TKÜ-Daten) an „rsCase“ bzw. „B-case“ von Bundespolizei und BKA übertragen (bitte als Tabelle aufschlüsseln)?
18. Wie ist es gemeint, wenn die Bundesregierung erklärt, eine Suche von unstrukturierten Datenbeständen der Bundespolizei sowie in „B-case“ und „Inpol-Fall“ könne „nur mittels Volltextrecherche“ erfolgen?
19. Inwieweit haben Behörden des BMI im Jahr 2012 „prediktive Software“ oder Software zum „Data Mining“ getestet oder Testberichte anderer deutscher Stellen hierzu erhalten?
- a) Welche Kosten fielen für eine etwaige Beschaffung an?
  - b) Inwieweit nahm das BMI 2012 eine „Marktbeobachtung“ für „prediktive Software“ oder Software zum „Data Mining“ vor?
  - c) Welche Erkenntnisse wurden hierfür von welcher Abteilung zusammengetragen?
  - d) Was wurde den Landeskriminalämtern hierzu berichtet?
  - e) Inwieweit trifft es – wie von BKA-Vizepräsident Prof. Dr. Jürgen Stock auf Fachkonferenzen vorgetragen wurde – zu, dass das BKA jede auf dem Markt befindliche „prediktive Software“ testen würde?
20. Wann ist die Inbetriebnahme der geheimdienstlichen Datensammlung „NADIS WN“ geplant?
- a) Welche Errichtungsanordnungen werden hierfür überarbeitet, und welchen Stand kann die Bundesregierung hierfür mitteilen?
  - b) Welche Firmen haben hierfür Aufträge zur Errichtung des Datenbanksystems sowie für die dort genutzte Software erhalten?
  - c) Inwieweit kommt in „NADIS WN“ Analysesoftware zum Einsatz?
21. Auf welche Art und Weise werden bei Bundesbehörden Informationen in Fall- oder Vorgangsbearbeitungssystemen verarbeitet, die im Zuge von verdeckten Maßnahmen gewonnen wurden?
- a) Auf welche Weise und in welchem Format werden derart erlangte Informationen aus den Systemen an andere Stellen weitergegeben?

- b) Welche Bestimmungen existieren hierfür hinsichtlich des Datenschutzes?
22. Welche Regelungen existieren für Fall- oder Vorgangsbearbeitungssysteme hinsichtlich der Anonymisierung von Volltextinformationen?
- a) Trifft es zu, dass diese keinen Bezug zu einer bestimmten Person erlauben dürfen?
- b) Inwiefern kann die Bundesregierung garantieren, dass dies bei ihren Polizeibehörden vollumfänglich umgesetzt ist?
23. Wie ist die Aussage des BKA-Angehörigen Moritz Aly im Vortrag auf der Veranstaltung Europäischer Polizeikongress am 14. Februar 2012 gemeint, wonach sich die beim BKA bestehenden Datenmengen in den letzten Jahren jährlich verdoppelt hätten?
- a) Welche Datenbestände sind hiermit gemeint?
- b) Auf welche Untersuchungen oder Prognosen stützt sich Moritz Aly mit der Aussage?
24. Inwieweit nutzen Bundesbehörden in Ermittlungsverfahren Bilder aus der Satellitenaufklärung?
- a) Welche zivilen oder militärischen Behörden oder Firmen werden hierfür angefragt?
- b) Wie beurteilt die Bundesregierung die gegenwärtige Qualität und Aussagekraft derartiger Satellitenbilder?
- c) Wie beurteilt die Bundesregierung den Nutzen der genutzten Satellitenbilder für die jeweiligen Ermittlungsverfahren?
- d) Wie beurteilt die Bundesregierung einen etwaigen Nutzen der Satellitenbilder auch zur Gefahrenabwehr?
- e) Für welche Zwecke hält die Bundesregierung den Einsatz von Satellitenaufklärung für zweckmäßig und sinnvoll?
25. Auf welche Art und Weise ist die Bundesregierung am EU-Forschungsprogramm „Composite“ beteiligt, bzw. welche Beiträge mit welchem Inhalt wurden hierfür erbracht (Antworten auf Fragebögen bitte als Anlage beifügen)?
26. Inwieweit will Europol die Suche nach allgemeinen und spezifischen „(Bedrohungs-)Indikatoren“ mit der „EU-SOCTA-Methode“ automatisieren?
- a) Auf welche technische Art und Weise soll die „Trendanalyse“ bewerkstelligt werden?
- b) Welche Behörden, Unternehmen oder sonstigen Personen sind an der Entwicklung beteiligt?
- c) Welche Datenquellen würde die „EU-SOCTA-Methode“ auswerten?
- d) Aus welchen deutschen Datensammlungen könnten hierfür (statistische) Daten angeliefert werden?
- e) Was ist mit der betriebenen Identifizierung von „Crime Relevant Factors“ gemeint?
- f) Inwieweit soll dadurch eine Priorisierung von Kriminalitätsbereichen vorgenommen werden?
27. Welche „Anwendungstools“ werden den Teilnehmern/Teilnehmerinnen des Lehrgangs „Grundlagen und Methoden der polizeilichen Informationsverarbeitung“ (Operative Analyse) – wie im Falle von Polizeiangehörigen aus

Weißrussland – „mit Hilfe der Software Analyst’s Notebook“ vorgeführt (Bundestagsdrucksache 17/10713)?

- a) Auf welche Art und Weise werden Ermittlungen derart visualisiert, und welchen Vorteil haben Behörden hiervon?
- b) Was ist mit der „Erstellung von Hypothesen etc.“ gemeint?
- c) In welchem Umfang wurden auch Behörden aus der Türkei, Russland, der Ukraine, China, Estland, Georgien, Kroatien, Serbien, Aserbaidschan und Montenegro „Analyst’s Notebook“ bzw. ähnliche Analyseprogramme vorgeführt?

Berlin, den 19. Oktober 2012

**Dr. Gregor Gysi und Fraktion**