Zwei Jahrzehnte nach 9/11 – Höchste Zeit für ein empirisch basiertes Monitoring staatlicher Überwachungsmaßnahmen

Ralf Poscher 2022-04-29T13:29:49

Eine der nachhaltigsten Veränderungen, die die Anschläge des 11. September 2001 in der westlichen Welt nach sich gezogen haben, ist die spürbare Beschleunigung der seit den späten 1980er Jahren zu beobachtenden Akzentverschiebung von einer reaktiven hin zu einer präventiv orientierten Sicherheitspolitik. Sinnbild dieser Entwicklung ist die kontinuierliche Ausweitung der Kompetenzen der Sicherheitsbehörden zur Überwachung verschiedenster Lebensbereiche der Bürgerinnen und Bürger. Diese langfristige „*Versicherheitsrechtlichung*“ wird durch den rasanten technischen Fortschritt in der Digitalisierung wesentlich erleichtert – wenn nicht sogar befördert. Sowohl die Verfügbarkeit potenziell sicherheits- und damit zugleich auch überwachungsrelevanter Daten als auch die Möglichkeiten für deren technisch unkomplizierten Transfer und ihre systematische/tiefe Auswertung durch staatliche Behörden (Sicherheitsbehörden ebenso wie Nachrichtendienste) haben sich signifikant verändert. Noch in den 1980er Jahren lag der Schwerpunkt staatlicher Überwachung zu einem wesentlichen Teil im Bereich der „klassischen“ Telefonüberwachung; digital erfasste (Massen-) Daten etwa zur Mobilität, zu den Kommunikationsverläufen oder zum Surfverhalten im Internet, aus denen sich vielfältige Informationen mit potenzieller Sicherheitsrelevanz generieren lassen, waren entweder gar nicht verfügbar oder mussten einzelfallbezogen und personalaufwändig erhoben werden, etwa durch längerfristig angelegte Observationsmaßnahmen.

Drei Beispiele intensivierter Überwachung seit 9/11

Vor allem in drei Bereichen wurde die Überwachung in Reaktion auf die islamistischen Terroranschläge systematisch ausgeweitet; dabei war jeweils die Europäische Union Impuls- beziehungsweise Taktgeberin. Der erste Anwendungsbereich betrifft den erweiterten Zugriff auf die Telekommunikations-Verkehrsdaten. Die bis heute gerade in dem konkreten Kontext der Verkehrsdatenüberwachung besonders kontrovers diskutierte Pflicht zur anlasslosen Vorratsdatenspeicherung wurde in Deutschland und einigen anderen Ländern überhaupt nur auf Druck der EU eingeführt. Die EU-Kommission wollte das Instrument damals bekanntlich um fast jeden Preis eingeführt wissen und hatte die Speicherpflicht mangels eigener Kompetenz zum Erlass sicherheitsrechtlicher Rechtsakte in der Prä-Lissabon-Ära ungeachtet der Kritik aus einigen Mitgliedsstaaten in eine wettbewerbsrechtliche Richtlinie gegossen (RL 2006/24/EG). Die beiden Begriffe Verkehrsdatenüberwachung und Vorratsdatenspeicherung

werden in der öffentlichen Debatte seither mitunter fast wie Synonyme gebraucht. Auch die Rechtsprechung des BVerfG zur Vorratsdatenspeicherung und die daraus entwickelten ersten Ideen einer Überwachungsgesamtrechnung (s.u.) beziehen sich im Wesentlichen auf diesen konkreten sachlichen Kontext.

Das zweite, ebenfalls europarechtlich determinierte Feld massenhafter Überwachung betrifft die präventive Geldwäschekontrolle. Dieses ursprünglich eng begrenzte, auf die (organisierte) Drogenkriminalität zugeschnittene Instrument wurde in Reaktion auf 9/11 auf die Bekämpfung der Terrorismusfinanzierung ausgedehnt. Die Idee der Verhinderung terroristischer Aktivitäten durch Überwachung der weltweiten Finanzströme wird seitdem zur Legitimation der kontinuierlichen und in immer kürzeren Zeitintervallen erfolgenden Erweiterungen (zuletzt 2021; eine weitere Geldwäsche-Richtlinie – es wird dann bereits die sechste sein – ist auf EU-Ebene bereits in Vorbereitung) angeführt. Sukzessive wurde daher auch in Deutschland ein umfangreiches Regularium zur flächendeckenden anlasslosen Speicherung von Finanztransaktionsdaten implementiert, die jeden und jede von uns nahezu unausweichlich trifft. Zusätzlich zu den Kundenstammdaten müssen jeder unbare Bezahlvorgang und jede Kontobewegung zusammen mit den dazugehörenden Begleitdaten (quasi eine Mischung aus ‚Verkehrs-' und Inhaltsdaten) fünf Jahre lang gespeichert werden, nach einem Wechsel der Bankverbindung noch weitere fünf. In wenigen Lebensbereichen dürfte die Metapher des gläsernen Menschen der Wirklichkeit mittlerweile so nahe kommen wie hier.

Ein dritter, hier ebenfalls nur exemplarisch aufgezeigter Lebensbereich mit einer weitreichenden Überwachung auf Vorrat betrifft die Erfassung von Fluggastdaten auf der Basis der sog. PNR-Richtlinie der EU (RL (EU) 2016/681). Das zu ihrer Umsetzung eingeführte Fluggastdatengesetz (2017) verpflichtet die Airlines zur Übermittlung von bis zu 60 individuellen (20 größeren Merkmalsgruppen zuzuordnenden; vgl. § 2 FlugDaG) Informationen über sämtliche an deutschen Flughäfen abfliegenden und ankommenden Passagiere vor. Das betrifft neben den allgemeinen Flugdaten (Flugnummer, Flugziel, Abflug-/Ankunftszeit etc.) eine Vielzahl personenbezogener Informationen wie zum Beispiel Familienname, Geburtsname, Vornamen, Doktorgrad (*sic!*), Wohnadresse und Ausweisdaten des Fluggastes, sowie Vielfliegerstatus, Sitzplatznummer, Informationen zur Buchung und gegebenenfalls zu dem Buchungsportal beziehungsweise Reisebüro, Flugpreis und Kreditkartendaten, mitgeführtes Gepäck, mitreisende Personen und viele weitere. Anders als bei den TK-Verkehrs- und Finanztransaktionsdaten erfolgt die Speicherung nicht bei den privaten Dienstleistern, sondern unmittelbar beim Bundeskriminalamt (in seiner zusätzlichen Funktion als Fluggastdatenzentralstelle: Passenger Information Unit – PIU). Eine Reihe weiterer, originär nationalstaatlicher Kompetenzen wie die sogenannte Online-Durchsuchung ergänzen das aktuelle Anti-Terror-Instrumentarium.

Verfassungsrechtliche Diskussion zur Überwachungsgesamtrechnung

Die kritische fachliche Auseinandersetzung mit der hier nur bruchstückhaft skizzierten Entwicklung konzentriert sich im Wesentlichen auf die relevanten (verfassungs-)rechtlichen Aspekte. Dies gilt für Beiträge aus der systemisch-dogmatischen beziehungsweise legislativen Perspektive ebenso wie für Beiträge, die den Fokus eher auf die Anwendungsebene und die individuelle Betroffenenperspektive richten. Vergleichsweise wenig Aufmerksamkeit wird hingegen der Frage nach dem *tatsächlichen Umfang der Überwachung* zuteil. Auch in der Rechtsprechung des BVerfG spielt die Häufigkeit einer Maßnahme bislang allenfalls eine indirekte Rolle, wenn es etwa darum geht, durch die Aufstellung hoher rechtlicher Hürden die Anwendung besonders eingriffsintensiver Maßnahmen faktisch zu begrenzen. Unmittelbarer Bezugspunkt seiner Rechtsprechung ist bislang aber stets die Verhältnismäßigkeit der konkreten Maßnahme.

Spätestens mit dem Urteil zur Vorratsdatenspeicherung vom März 2010 (1 BvR 256/08) hat das Gericht auch eine andere Perspektive ins Spiel gebracht, indem es, über die konkrete Überwachungsmaßnahme hinaus, die Notwendigkeit einer Gesamtbetrachtung aller wesentlichen überwachungsrelevanten Kompetenzen der Sicherheitsbehörden ins Spiel gebracht hat. Im Hinblick auf die freiheitliche Verfassungsidentität der Bundesrepublik, zu deren Kernbestandteilen das Gericht ausdrücklich das Verbot einer Totalerfassung und Registrierung der Freiheitswahrnehmung der Bürgerinnen und Bürger zählt, dürften die verschiedenen Kompetenzen in ihrer Summe nicht zu einer umfassenden Überwachung führen. Eine unzulässige (Total-)Überwachung sähe das Gericht bereits im Falle einer bloß theoretischen Rekonstruierbarkeit als erfüllt an. Daher müsse der Gesetzgeber bereits bei der Planung (das Gericht spricht wörtlich von „Erwägung“) neuer Speicherpflichten und Befugnisse zum behördlichen Zugriff auf bereits (irgendwo) gespeicherte personenbezogene Daten die Gesamtheit der verschiedenen bereits existierenden Datensammlungen und ihrer Nutzungsvoraussetzungen zu berücksichtigen. Dieses Prinzip ist generell zu beachten, auch jenseits des Bereichs der Vorratsdatenspeicherung.

Diese erweiterte Perspektive wurde in der Wissenschaft aufgegriffen und in ein Konzept zur ganzheitlichen Betrachtung des Überwachungsgeschehens übertragen, für das sich der auf Roßnagel zurückgehende Topos der „*Überwachungsgesamtrechnung*“ (ÜGR) durchgesetzt hat. Mit dem etwas sperrigen Begriff wird auf die Notwendigkeit einer Gesamtbetrachtung des (jeweils aktuellen) Standes staatlicher Überwachung verwiesen, die alle einschlägigen präventiven und repressiven Überwachungsmaßnahmen quasi aufaddiert. Der Koalitionsvertrag der neuen Bundesregierung greift den Ansatz der ÜGR wieder auf und betrachtet diese als ein wichtiges Element vorausschauender, evidenzbasierter und grundrechtsorientierter Sicherheits- und Kriminalpolitik.

Anders als der Begriff der Gesamtrechnung eigentlich impliziert, wurde die ÜGR in der Vergangenheit vor allem auf einer qualitativ dogmatischen Ebene

diskutiert und nur in rudimentären Ansätzen operationalisiert. Beiträge aus der verfassungsrechtlichen Literatur halten sich meist im Vagen und begnügen sich weitgehend mit Vorschlägen, wie man die Gesamtheit der rechtlichen Befugnisse zur Überwachung abstrakt fassen und bewerten könnte. Völlig ausgeblendet wurde bislang die Frage, ob und gegebenenfalls wie häufig eine bestimmte Überwachungsmaßnahme und der damit verbundene Grundrechtseingriff zum Einsatz kommt; hier stochern wir bildlich gesprochen im Nebel. Wir können bislang nicht annähernd quantifizieren, in welchem Umfang sich die „Überwachungslast" in Deutschland seit 9/11 tatsächlich verändert hat, noch lässt sich deren Gesamtumfang bestimmen. Erst mit der Ausübung der verfügbaren rechtlichen Kompetenzen materialisiert sich der damit verbundene Grundrechtseingriff. Daher ist die *Kernfrage* nach dem – verfassungsrechtlich vertretbaren – Maß staatlicher Überwachung eben *auch eine quantitative*. Denn mit der Häufigkeit solcher Maßnahmen steigt auch die statistische Wahrscheinlichkeit und damit das Risiko der eigenen Betroffenheit. Der Blick auf zwei der eingangs genannten Beispiele macht es deutlich: während das individuelle Risiko, tatsächlich in den Fokus einer Online-Durchsuchung zu geraten, aufgrund der wenigen Einsatzfälle faktisch nahe Null liegt, betrifft die Fluggastüberwachung jeden und jede, die von einem deutschen Flughafen abfliegen und dort wieder ankommen (bzw. umgekehrt). Gleichwohl nimmt die Online-Durchsuchung in den wissenschaftlichen und (rechts-) politischen Diskussionsforen deutlich breiteren Raum ein als die Fluggastüberwachung.

»Überwachungsbarometer« – ein neues, empirisch unterlegtes Instrument zur Erfassung der realen Überwachungslast

Es erscheint mithin zwingend, das Überwachungsgeschehen nicht nur durch die dogmatische Brille zu betrachten, sondern parallel auch die empirische Realität in die Bewertung mit einzubeziehen. Dass das bislang nicht geschehen ist, kann jedenfalls partiell auch damit erklärt werden, dass belastbare statistische Informationen zur Häufigkeit der durchgeführten Überwachungsmaßnahmen insbesondere im präventiven Anwendungskontext lange Zeit gar nicht oder nur sporadisch verfügbar waren. Das ist eine entscheidende Lücke, die sukzessive geschlossen werden muss. Das Freiburger Max-Planck-Institut zur Erforschung von Kriminalität, Sicherheit und Recht arbeitet aktuell an dem Konzept für ein periodisches Überwachungsbarometer, das die aufgezeigten Defizite aufgreift und damit ein empirisch unterlegtes Instrument zur Weiterentwicklung der ÜGR entwickelt. Ziel dieses neuartigen Modells ist es, die verschiedenen rechtlichen Überwachungskompetenzen und ihre normative Ausgestaltung (verfassungsrechtliche Perspektive) mit der realen Anwendungspraxis (empirische Perspektive) zu kombinieren. Auf dieser Basis kann das Überwachungsgeschehen gemessen und damit die Überwachungslast, der die Bürgerinnen und Bürger in Deutschland in einem bestimmten Referenzzeitraum (z.B. Kalenderjahr) ausgesetzt sind, sichtbar gemacht werden. Unterstützt wird das Vorhaben durch die absehbar zunehmende Verfügbarkeit aggregierter statistischer Daten.

Zur Erstellung eines realistischen Abbildes der Überwachungssituation und ihrer verfassungsrechtlichen Einordnung reicht es jedoch nicht aus, Zugriffsnormen und Anwendungszahlen rein quantitativ zu erfassen. Staatliche Überwachungsmaßnahmen und Zugriffe auf datenförmig hinterlegte Informationen müssen jeweils spezifiziert und im Hinblick auf ihre Zielsetzung und ihre Eingriffswirkung gewichtet werden. Beispielsweise dürfte ein nach abstrakter Bewertung eingriffsintensiver präventiver Echtzeit-Zugriff auf mobile Standortdaten einer in einem weitläufigen Waldgebiet vermissten Person oder ihrer Begleitperson zur Abwendung einer konkreten Gefahr für Leib oder Leben anders zu bewerten sein als die repressive Abfrage von Kontodaten zur Aufklärung einer mutmaßlichen Geldwäsche-, Steuer- oder Vermögensstraftat. Beide könnten ihrerseits schwerer wiegen als etwa die massenhafte, potenziell Hunderttausende betreffende Verkehrsüberwachung mittels kennzeichenbasierter Abschnittskontrolle. Als entscheidende Parameter müssen daher sowohl die verfassungsrechtliche als auch die empirische Eingriffsintensität berücksichtigt und zueinander ins Verhältnis gesetzt werden.

Um die Eingriffsintensität der verschiedenen Überwachungsmaßnahmen bestimmen zu können, müssen diese nach einheitlichen verfassungsrechtlichen Kriterien typisiert und gewichtet werden. Hierfür wurde ein komplexes Kategoriensystem entwickelt, das alle abstrakt bestimmbaren Kriterien der Eingriffsintensität berücksichtigt und nach deren verfassungsrechtlicher Bedeutung jeweils unterschiedlich quantifiziert. Auch das BVerfG rekurriert in seiner Rechtsprechung sehr häufig auf die Schwere der Eingriffe und bewertet diese beispielsweise als nur „gering“ oder „geringfügig“ am einen, sowie „tiefgreifend“ oder „besonders stark“ am anderen Ende einer angedeuteten Skala. Eher in der Mitte einzuordnen sind wohl Maßnahmen von „erheblichem“ beziehungsweise „nicht unerheblichen Gewicht.“ Die genannten Beispiele sind weit entfernt von einer systematischen Kasuistik. Löffelmann spricht diesbezüglich in seiner Besprechung zum Bestandsdatenauskunft-II-Beschluss nicht zu Unrecht von „Begriffssynkretismus“ (GSZ 2020, 182, 185). In der Tat wirken die zitierten Beschreibungen mitunter fast ein wenig hilflos. Was unterscheidet etwa einen *nicht un*erheblichen von einem erheblichen Eingriff? Auch die begrifflichen Unschärfen lassen erkennen, dass eine ausschließlich normative Bewertungsmethode zur Berechnung der Überwachungslast nicht geeignet ist. Bei dem Überwachungsbarometer geht es – anders als in der traditionellen Verhältnismäßigkeitsdogmatik – nicht um die abstrakte (verfassungs-)rechtliche Zulässigkeit oder Unzulässigkeit einer Maßnahme, sondern um ihre konkrete Eingriffswirkung bei den Adressaten. Aus dieser Perspektive ist *jede* Maßnahme ein Eingriff – auch die *verhältnismäßige.*

Um die Eingriffsintensität der verschiedenen Überwachungsmaßnahmen quantifizieren zu können, müssen die jeweiligen Umstände und potenziellen Folgewirkungen ihrer Durchführung in die Bewertung einfließen. Dies betrifft etwa die Voraussetzungen und Zielsetzung einer Maßnahme, die Durchführungsmodalitäten, Dauer und Streubreite, die Art der erhobenen Daten, ihre Verwendung einschließlich einer möglichen Weitergabe, ihre spätere Löschung, und viele weitere Umstände. Jedem dieser Parameter wird auf der

Basis eines einheitlichen Kategoriensystems ein individueller Intensitätswert zugeordnet. Im Ergebnis kann die gleiche Maßnahme, zum Beispiel die Telekommunikationsüberwachung, in Bundesland A eine andere Eingriffsintensität haben als in Bundesland B oder C. Dasselbe gilt, wenn solche Maßnahmen einerseits auf der Grundlage der StPO und andererseits im präventiven Kontext auf der Grundlage beispielsweise des BKAG oder eines Landespolizeigesetzes zur Anwendung kommen, wir auch das Bundesverfassungsgericht in seinem jüngsten Urteil zu den Überwachungsbefugnissen des Verfassungsschutzes (1 BvR 1619/17) noch einmal betont hat. Mit unserer aktuell getesteten Formel errechnet sich zum Beispiel für die präventive Abfrage von TK-Verkehrsdaten aufgrund der unterschiedlichen Ausgestaltung der aktuellen landesgesetzlichen Rechtsgrundlagen für fast jedes Bundesland ein anderer Intensitätswert; die Werte variieren auf der vorläufigen zehnstufigen Intensitätsskala um bis zu einen ganzen Punkt (das entspricht einer Varianz von 10 Prozent). Selbst im fiktiven Fall identischer Häufigkeit trügen die landesrechtlichen Maßnahmen in unterschiedlichem Maße zur Gesamtüberwachungslast in Deutschland bei.

Ausblick

Das neue Überwachungsbarometer versteht sich als ein rechts- und gesellschaftspolitisches Transparenzprojekt, das der interessierten Öffentlichkeit ebenso wie den verantwortlichen Akteuren in Wissenschaft, Politik und Justiz erstmals aussagekräftige, verständliche und leicht zugängliche Informationen zu der realen Überwachungslast der Bürgerinnen und Bürger im täglichen Leben zur Verfügung stellt. Die methodischen Kernelemente wurden an anderer Stelle bereits ausführlich dargestellt. Im Zentrum der Projektarbeit steht aktuell die Feinjustierung der Formeln für die quantitativen und die qualitativen Parameter. Hierfür wurden zunächst zwei exemplarische Indexformeln entwickelt, mit denen die beiden Parameter rechnerisch unterschiedlich gewichtet werden; eine Formel ist nach oben offen konstruiert und orientiert sich stärker an der Häufigkeit, die andere ist stärker indexiert und fokussiert eher die Intensität der Zugriffe. Die verschiedenen Berechnungsmöglichkeiten sind noch im experimentellen Stadium und sollen in den kommenden Monaten anhand erster Realdaten überprüft und auf ihre jeweiligen Effekte hin überprüft werden, wenn es etwa darum geht, wie quantitativ sehr wenige Maßnahmen mit sehr hoher Grundrechtsrelevanz wie die Online-Durchsuchung einerseits mit den massenhaften Zugriffen wie bei den Fluggastdaten andererseits, deren Eingriffsintensität nach unseren bisherigen Rechenmodellen relativ niedriger als die der Online-Durchsuchung anzusetzen ist, ins Verhältnis gesetzt werden können. Parallel hierzu werden auch verschiedene Darstellungsarten getestet, die das Überwachungsgeschehen erkennbar und die daraus resultierende(n) Überwachungslast(en) greifbar machen sollen. Neben der traditionellen Darstellung in absoluten Zahlen kann die Zahl der jeweiligen Maßnahmen etwa in Form der durchschnittlichen Anzahl von Datenzugriffen pro Tag oder als Inzidenzwert bezogen auf 100.000 Einwohner angezeigt werden. So hat sich beispielweise die Gesamtzahl der behördlichen Kontoabfragen bei Kreditinstituten zwischen 2005 und 2018 von durchschnittlich 290 auf 3.758 Abfragen pro Tag beziehungsweise von 107,0 auf 1.353,3 Abfragen pro 100.000 Einwohner vervielfacht (siehe Projektbericht).

In der ersten Zeit wird das Barometer die Vielzahl an relevanten Überwachungstatbeständen zunächst nur selektiv abdecken können. Mit zunehmender Datendichte wird es immer besser in der Lage sein, Entwicklungen bereichs- und maßnahmenspezifisch zu identifizieren und Trends frühzeitig zu erkennen. Dabei können Änderungen des normativen Rahmens ebenso eine Rolle spielen wie Veränderungen der rechtstatsächlichen Rahmenbedingungen, etwa durch praxisrelevante gerichtliche Intervention und nicht zuletzt auch technologische Entwicklungen. Mit dem geplanten Modell dürfte es dann auch möglich sein, den faktischen Beitrag der verschiedenen Antiterrorismus-Befugnisse an der Überwachungsgesamtlast zielgenauer zu bestimmen.
