

# **Datenschutzrechtliche Anmerkungen zum Artikel „Vom Datenangreifer zum zertifizierten Wissenschaftler“ von Ulrich Rendtel**

**Michael Smolle**

Eingegangen: 13. April 2015 / Angenommen: 15. April 2015 / Online publiziert: 1. Mai 2015  
© Springer-Verlag Berlin Heidelberg 2015

## **1 Einleitung**

Prof. Rendtel sieht sich in der Rolle des potenziellen Datenangreifers, weil er ein gewisses Misstrauen der amtlichen Statistik gegenüber der Wissenschaft, dass sie die Datenschutzverpflichtungen beachtet, als Ursache für die Zurückhaltung ausgemacht hat, Daten zur Forschung bereitzustellen. Allein mit der Forderung nach faktischer Anonymität werde dem Wissenschaftler die Absicht unterstellt, die Daten deanonymisieren zu wollen. Er plädiert für einen Wechsel auf das Paradigma des Vertrauens, womit schnellere und effizientere Möglichkeiten des Datenzugangs über das Fernrechnen eröffnet werden könnten.

Aus meiner Sicht als nicht auf ein bestimmtes Gebiet des öffentlichen Rechts spezialisierten Verwaltungsjurist einer Datenschutzaufsichtsbehörde liegt eine Ursache für eine solche Zurückhaltung eher in der Schwierigkeit, aufgrund der enormen technischen Veränderungen mit zunehmender Digitalisierung des Alltags die Risiken für eine Deanonymisierung Betroffener über Verschneidungen mit anderen Datensätzen kaum noch einschätzen bzw. überblicken zu können. Das in dieser Hinsicht vorhandene Gefahrenpotenzial wird in Zeiten von „Big Data“ weiter steigen.

Für mich stellte sich die Bewertung eines entsprechenden Anliegens aus dem Bereich der Bildungsforschung anhand der aktuellen Rechtslage jedenfalls als schwierig dar. Obwohl erstmals die Gelegenheit bestand, sich intensiver mit einem solchen Sachverhalt und der Rechtsmaterie zu beschäftigen, bleiben verschiedene Fragen aus meiner Sicht offen.

Außerdem dürfte in diesem Kontext auch zu berücksichtigen sein, dass von Seiten der amtlichen Statistik vorrangig das Statistikgeheimnis zu wahren ist, das zur Ver-

---

Der Beitrag gibt die persönliche Meinung des Autors wieder

M. Smolle (✉)  
Mainz, Deutschland  
E-Mail: m.smolle@datenschutz.rlp.de

trauensbildung beim Auskunftspflichtigen und somit der Gewinnung zuverlässiger Angaben dient.

## 2 Bundesstatistikgesetz

Einschlägig ist die Wissenschafts- oder Forschungsklausel gemäß § 16 Abs. 6 BStatG und somit geht es um faktisch anonymisierte Daten, d. h. um Daten, die dem Betroffenen nur zugeordnet werden können, wenn der Datenempfänger einen unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft erbringen müsste.<sup>1</sup> Mit dieser Vorschrift werden Erleichterungen für Wissenschaftler im Vergleich zu anderen Datennutzern festgeschrieben. Allerdings ist eine Forschungsklausel als eine Ausnahmeregelung möglichst restriktiv zu interpretieren.<sup>2</sup>

Weiterhin ist in diesem Zusammenhang § 3a Abs. 2 BStatG zu erwähnen, wonach zu den statistischen Arbeiten nach § 3a Abs. 1 BStatG auch die Bereitstellung von Daten für die Wissenschaft gehört. Nach der Gesetzesbegründung<sup>3</sup> umfasst eine solche Bereitstellung neben der Veröffentlichung von aggregierten Daten als klassischer Form der Verbreitung von statistischen Ergebnissen auch die Nutzbarmachung statistischer Daten, z. B. in Forschungsdatenzentren, in Form von anonymisierten Mikrodaten (Public und Scientific Use Files) oder auf andere geeignete Weise.

Weitere Informationen dazu, wie eine Bereitstellung im Rahmen der Forschungsklausel aussehen kann, finden sich in der Begründung zum Gesetz zur Förderung der elektronischen Verwaltung sowie zur Änderung weiterer Vorschriften, mit dem u. a. § 3 Abs. 1 Nr. 1 d) in das BStatG eingefügt wurde. Neben der Übermittlung absolut anonymisierter Einzelangaben nach § 16 Abs. 1 S. 2 Nr. 4 BStatG und der Übermittlung faktisch anonymisierter Einzelangaben nach § 16 Absatz 6 BStatG sieht der Gesetzgeber eine weitere Form der Bereitstellung darin, formal anonymisierte Einzelangaben, d. h. Einzelangaben ohne Hilfsmerkmale, durch kontrollierte Datenfernverarbeitung für Auswertungen nutzbar zu machen. Bei diesem Verfahren werden dem Auftraggeber von den statistischen Ämtern jedoch keine Einzelangaben übermittelt und der Auftraggeber erlangt bei diesem Verfahren auch keine Kenntnis von den Einzelangaben. Bei der kontrollierten Datenfernverarbeitung schreibt der Auftraggeber ein Auswertungsprogramm, das auf dem Server des statistischen Amts mit den „originalen“ Einzelangaben der amtlichen Statistik, die lediglich formal anonymisiert sind, die Berechnungen durchführt. Die so errechneten Ergebnisse werden vom statistischen Amt auf Geheimhaltung überprüft und dem Auftraggeber übermittelt.<sup>4</sup>

In dieser Weise wird ein Zugangsweg auch im Beitrag von Schimpl-Neimanns/Weiss<sup>5</sup> beschrieben.

<sup>1</sup> Bundestags-Drucksache (BT Drs.) 10/5345, S. 21.

<sup>2</sup> Simitis in Simitis (Hrsg.), Bundesdatenschutzgesetz (BDSG), § 40 Rn 35.

<sup>3</sup> BT Drs. 15/4696.

<sup>4</sup> BT Drs. 17/11473S. 54.

<sup>5</sup> Schimpl-Neimanns, Bernhard und Weiss, Felix (2014): Zur Bereitstellung amtlicher Mikrodaten für die Wissenschaft aus sozialwissenschaftlicher Perspektive. AStA: Wirtschafts- und Sozialstatistisches Archiv, 8 (4), 205-219. Hier: S. 208.

### 3 Auslegung des datenschutzrechtlichen Personenbezugs

Das Spannungsfeld zwischen dem Recht auf informationelle Selbstbestimmung des Einzelnen und dem Recht auf Freiheit von Wissenschaft und Forschung wurde in den bereits veröffentlichten Artikeln mehrfach angesprochen, weshalb an dieser Stelle auf die in Literatur und Rechtsprechung seit langem umstrittene Frage einzugehen ist, die für die Beurteilung, ob von einer ausreichenden Anonymisierung ausgegangen werden kann, von maßgeblicher Bedeutung ist:

Sind bei der Bestimmbarkeit einer Person nur die bei der verantwortlichen Stelle selbst verfügbaren (subjektiver Ansatz) oder auch bei Dritten gespeicherten Daten (objektiver Ansatz) zu berücksichtigen?

Oder anders gefragt: Wann ist eine Einzelangabe über persönliche oder sachliche Verhältnisse einer bestimmten oder zumindest bestimmbaren natürlichen Person so zuzuordnen, dass sie dem Schutzbereich des Datenschutzrechts unterfällt? Kommt es allein auf die Zuordnungsmöglichkeiten der verantwortlichen Stelle an (subjektiver Ansatz) oder sind auch die Zuordnungs- und Verknüpfungsfähigkeiten Dritter (objektiver Ansatz) zu berücksichtigen.<sup>6</sup>

### 4 Aktueller Vorgang

Nach diesen allgemeinen Ausführungen komme ich wieder auf den mir zur datenschutzrechtlichen Bewertung vorgelegten Fall aus dem Bereich der Bildungsforschung zurück. Die Bereitstellung bestimmter Daten wurde u. a. deswegen abgelehnt, weil eine faktische Anonymisierung aufgrund der Merkmalskombination nicht möglich sei und die Daten derzeit auch nicht für eine kontrollierte Datenfernverarbeitung angeboten werden könnten.

Die Definition der faktischen Anonymität gem. § 3 Abs. 8 LDSG RP entspricht § 16 Abs. 6 BStatG. Der „unverhältnismäßig hohe Aufwand“ als unbestimmter Rechtsbegriff bemisst sich an dem Verhältnis von Aufwand und Wert der zu erlangenden Information<sup>7</sup> und wird dabei folgendermaßen gehandhabt:

Im Rahmen einer Risikoabwägung wird der Grad der Sensibilität der erhobenen Daten berücksichtigt. Dem sind gegenüber zu stellen die rechtlich zulässigen Identifizierungsmöglichkeiten sowie das Identifizierungsinteresse der Stelle, der die Daten übermittelt werden. Risikofaktoren sind dabei das vorhandene oder erwerb- bare Zusatzwissen des Datenverwenders oder die technischen Möglichkeiten der elektronischen Datenverarbeitung. Führt die Risikoabwägung beispielsweise zu dem Ergebnis, dass zwar der Identifizierungsaufwand, aber auch die Empfindlichkeit der Daten gering sind, kann im Ergebnis gleichwohl von einer faktischen Anonymisierung ausgegangen werden.

Die Zugangswege der Wissenschaft zum Informationspotential der amtlichen Statistik werden in „Die Datennutzung bei den Forschungsdatenzentren“<sup>8</sup> als ver-

<sup>6</sup> Brink/Eckhardt, ZD 1/2015, S. 1, die Argumente für den jeweiligen Ansatz nennen.

<sup>7</sup> Bundesverfassungsgericht (BVerfG) Beschluss v. 28.9.1987 – 1 BvR 1063/87, S. 5.

<sup>8</sup> vgl. 20. Tätigkeitsbericht des LfDI RP, 2003–2005, Anlage 18.

schiedene Kombinationen von Datenanonymisierung und Zugriffsregulierung beschrieben. U. a. entscheidend ist, welches Zusatzwissen vorliegt und wo die Datennutzung stattfindet. Abhängig davon, ob die Mikrodaten extern oder in den statistischen Ämtern genutzt werden, kann die faktische Anonymität mit mehr oder minder starken Informations-einbußen erreicht werden.

Wendet man für einen Personenbezug den relativen Ansatz an, wäre bei dem zu beurteilenden Vorgang m. E. aber von einem faktischen anonymen Datensatz auszugehen, da einem Wissenschaftler in einem FDZ bei On-Site-Nutzung keine Hilfsmittel, wie z. B. maschinelles Zuspätspielen von Zusatzwissen zur Verfügung stehen, um einen Personenbezug herstellen zu können.

Auch mit dem absoluten Ansatz käme man hier nicht zur Annahme eines Personenbezugs, da illegal beim StaLa beschaffbare Daten in Auslegung des „vernünftigen Aufwandes“ eines Wissenschaftlers in einem FDZ nicht zu berücksichtigen sind.<sup>9</sup>

Aber wie könnte ein StaLa bei Annahme des objektiven Ansatzes vor der Bereitstellung eines Scientific Use File mit faktisch anonymisierten Daten zur externen Nutzung überhaupt noch einschätzen, welche Daten sich ein Wissenschaftler legal dazu über „Big Data“ o.a. Quellen beschaffen kann? Könnten solche Verknüpfungen vertraglich ausgeschlossen und sanktioniert werden?

In diesem Kontext stellen sich mir nun noch verschiedene Fragen:

1. Können lediglich in der oben beschriebenen Form formal anonymisierte Datensätze nach Risikoabwägung auch als faktisch anonyme Daten angesehen werden?
2. Kann „formal anonymisiert“ somit entweder noch „pseudonym“ oder bereits „faktisch anonym“ bedeuten?
3. Abhängig davon, ob die Mikrodaten extern oder in den statistischen Ämtern genutzt werden, kann die faktische Anonymität mit mehr oder minder starken Informationseinbußen erreicht werden?

Als Zugangsweg zu Mikrodaten der amtlichen Statistik wird auch die kontrollierte Datenfernverarbeitung (KDFV) genannt. Dabei wird mitunter zwischen Remote Batch- und Remote Access-Betrieb unterschieden bzw. von Seiten der Wissenschaft die Zulässigkeit eines echten Remote Access-Betriebes in Deutschland gefordert. Beim echten Remote Access-Betrieb hätte ein Wissenschaftler vom institutseigenen Remote Access-Terminal aus Zugriff auf lediglich formal anonymisierte Daten, an denen er Auswertungen vornehmen kann.

Hier stellen sich mir folgende Fragen:

- a) Könnte ein FDZ im Falle von Remote Batch-Betrieb für die Auswertung mit dem vom Wissenschaftler erstellten Programm die Originaldaten, die noch nicht einmal formal anonymisiert sind, verwenden? Der Wissenschaftler erhält ja keinen unmittelbaren Zugriff auf die Mikrodaten und der Output wird durch das FDZ auf Geheimhaltung überprüft.
- b) Muss die Output-Kontrolle anhand absoluter Anonymität (§ 16 Abs. 1 S. 2 Nr. 4 BStatG) als Maßstab vorgenommen werden, obwohl § 16 Abs. 6 BStatG nur faktische Anonymität vorsieht? Dies ginge dann in Richtung des objektiven An-

<sup>9</sup>Bundesgerichtshof (BGH), Beschluss v. 28.10.14– VI ZR 135/13.

satzes, bei dessen Anwendung es theoretisch nur noch personenbezogene Daten geben würde.<sup>10</sup>

- c) Könnte die im Rahmen der Empfehlung Nr. 21 des Statistischen Beirats<sup>11</sup> angesprochene Zertifizierung als Aspekt des Identifizierungsinteresses bei der Risikoabwägung berücksichtigt werden?

## 5 Informationsfreiheitsrecht

Erleichterungen beim Datenzugang kann auch das Landesinformationsfreiheitsgesetz (LIFG) RP nicht bringen. Denn zum Einen ist ein Antrag auf Informationszugang abzulehnen, wenn die amtliche Information einem besonderen Amtsgeheimnis unterliegt (§ 9 Abs. 1 Nr. 4 LIFG). Das Statistikgeheimnis bezieht sich aber gerade auch auf Daten, mit denen eine Reidentifizierung nicht ausgeschlossen ist.

Zum Anderen bezieht sich das Informationsrecht auf die bei einer Behörde vorhandenen amtlichen Informationen (§ 4 Abs. 1 S. 1 LIFG). Das LIFG kennt keine Pflicht zur Beschaffung oder Herstellung von Informationen.

Für das VG Wiesbaden handelt es sich aber erst dann um Einzelangaben, die einer natürlichen Person nicht mehr zugeordnet werden können, wenn ein neuer Datenbestand geschaffen wird, der keine personenbeziehbaren Daten mehr enthält. Zur Einhaltung des Statistikgeheimnisses gem. § 16 Abs. 1 BStatG bedarf es vorliegend auch mehr als dem einfachen Weglassen von personenbeziehbaren Datenteilen. Vielmehr müssten die Daten komplett neu berechnet und verändert werden, was die Herstellung neuer Datensätze bedeute. Dies sei vom Anspruch auf Informationsfreiheit nicht gedeckt.<sup>12</sup>

Bei faktisch anonymen Daten dürfte es sich also um vorhandene Daten im Sinne des LIFG handeln.

## 6 Fazit

Bevor der Gesetzgeber dazu aufgefordert wird, neue Zugangswege für die Forschung zu schaffen, sollte die Nutzung der vorhandenen Möglichkeiten verbessert werden. Zumal den begleitenden Beiträgen zum Artikel von Rendtel zu entnehmen ist, dass – auch im Vergleich mit der Situation in anderen Ländern – der wissenschaftlichen Forschung in Deutschland der Zugang zu einem breiten Feld von Mikrodaten geboten wird.

Die Notwendigkeit für Verbesserungen wird jedenfalls auch von der amtlichen Statistik gesehen. An der Ausarbeitung von Empfehlungen des Statistischen Beirats – Unterarbeitsgruppe „Zugang der Wissenschaft zu Einzelangaben der Bundessta-

<sup>10</sup> Brink/Eckhardt, a. a. O.

<sup>11</sup> Papier „Fortentwicklung der amtlichen Statistik“, November 2012. URL: [https://www.destatis.de/DE/UeberUns/LeitungOrganisation/StatistischerBeirat/ServiceVeroeffentlichungen/Fortentwicklung-Nov2012Teil1.pdf;jsessionid=49AA24C797711279C84810E23DC4E0A2.cae2?\\_\\_blob=publicationFile](https://www.destatis.de/DE/UeberUns/LeitungOrganisation/StatistischerBeirat/ServiceVeroeffentlichungen/Fortentwicklung-Nov2012Teil1.pdf;jsessionid=49AA24C797711279C84810E23DC4E0A2.cae2?__blob=publicationFile).

<sup>12</sup> Geschäftszeichen 6 K 1374/11.WI.

tistik“ – waren auch Mitarbeiter des StatBA beteiligt. Diesen Umstand findet auch Rendtel bemerkenswert.

Bei der On-Site-Nutzung könnten durch eine bessere Vernetzung, wie z. B. mit dem Projekt „FDZ im FDZ“, Anfahrtswege verkürzt und Kosten verringert werden. Diesem Fazit von S. Bender<sup>13</sup> kann ich mich uneingeschränkt anschließen. Nicht umsonst hat der Rat für Sozial- und Wirtschaftsdaten (RatSWD) 29 Institute akkreditiert und fordert als ein Kriterium für eine Akkreditierung als FDZ/DSZ, dass der Nutzungszugang den bereichsspezifischen Bestimmungen des Datenschutzes und der Datensicherheit unterliegt. Gegebenenfalls wäre eine Datenverarbeitung im Auftrag über § 16 Abs. 2 S. 2 BStatG und § 6 Abs. 1 LStatG RP denkbar.

Eine Idee könnte es auch sein, eine unabhängige „Clearing-Stelle“ einzurichten, der abgelehnte oder streitige Anfragen auf Zugang zu amtlichen Daten zur nochmaligen Bewertung vorgelegt werden könnten.

Vielleicht bleibt aber auch bei der Amtlichen Statistik ein (fortschreitender) Kulturwandel abzuwarten, der sich laut dem Memorandum des Fachkollegiums „Erziehungswissenschaft“ der DFG aus 03/2014 sowohl auf die Bereitstellung als auch auf die Nutzung von Datensätzen beziehen müsste.

Entsprechende Möglichkeiten müssten im Rahmen von §§ 3 Abs. 1 Nr. 1 d, 3a Abs. 2 BStatG vorhanden sein.

---

<sup>13</sup> Bender, Stefan (2014): Datenzugang in Deutschland: Der Paradigmenwechsel hat bereits stattgefunden, AStA: Wirtschafts- und Sozialstatistisches Archiv, 8 (4), 237–248. Hier: S. 247.