Smart Data

Smart Data – Smart Privacy?

Impulse für eine interdisziplinär rechtlich-technische Evaluation

Technical Report des BMWi-Technologieprogramms „Smart Data – Innovationen aus Daten“

Impressum

Herausgeber
Smart-Data-Begleitforschung
FZI Forschungszentrum Informatik
Außenstelle Berlin
Friedrichstr. 60, 10117 Berlin
www.smart-data-programm.de

Konzeption und Gestaltung
LoeschHundLiepold Kommunikation GmbH, Berlin

Stand
November 2015

Druck
WIRmachenDRUCK GmbH

Bildnachweis
bluebay2014 – Fotolia.com (Titel)

Gefördert durch:

Bundesministerium
für Wirtschaft
und Energie

aufgrund eines Beschlusses
des Deutschen Bundestages

Inhalt

Einleitung: Smart Data als datenschutzrechtliches Phänomen 4

Editorial: Verfahrensrechtliche Anknüpfung und Gesetzessystematik des Schutzkonzeptes 5

Erste Impulse für eine interdisziplinäre rechtlich-technische Evaluation 7

1. Personenbezug als zentrales Tatbestandsmerkmal für regulatives Eingreifen 7
2. Datensparsamkeit 9
3. Datensicherheit 11
4. Transparenz und Information 13
5. Kontrolle 16
6. Zweckbindung 17
7. Rechte des Betroffenen 19

Fazit und Ausblick 20

Mitwirkung 22

Einleitung: Smart Data als datenschutzrechtliches Phänomen

Aus „Big Data“ wurde „Smart Data“. Ohne den Versuch einer exakten Definition zu wagen, könnte man dieses technische und sozioökonomische Phänomen verkürzt als die Generierung neuen Wissens aus mehr oder weniger strukturierten heterogenen Massen von Daten durch effektive Speicherung und fortgeschrittene Methoden des Datamining (z. B. Finden von unerwarteten aber potentiell interessanten Korrelationen) verstehen. Eine Besonderheit kann darin gesehen werden, dass sich auch neue, noch unbekannte Fragestellungen überhaupt erst aus der Analyse ergeben sollen. Im Ergebnis sollen auch auf Basis personenbezogener Daten individuell wirkende oder gesellschaftlich erwünschte Innovationen ermöglicht werden. Auf der anderen Seite sind die damit verbundenen Gefahren für den Betroffenen, dessen personenbezogene Informationen in die Datenverarbeitung einbezogen werden, zu berücksichtigen. Mit Blick auf das Recht auf informationelle Selbstbestimmung (Art. 2 Abs. 1, 1 Abs. 1 Grundgesetz), das Recht auf Achtung des Privat- und Familienlebens sowie das Recht auf Schutz personenbezogener Daten (Art. 7, 8 der EU-Grundrechtscharta) darf der bei Smart-Data-Anwendungen anvisierte Umgang mit personenbezogenen Daten nicht zu einer Abschwächung des Datenschutzniveaus führen.

Das vorliegende Dokument stellt thesenhaft die mit aktuellen und zukünftigen Smart-Data-Anwendungen verbundenen typischen Abläufe im Hinblick auf die bestehende gesetzliche Regelungssystematik und die anerkannten Schutzprinzipien des Datenschutzrechts dar. Es sollen Forschungsanstöße gegeben werden, wie technisch wirkende Mechanismen in die den Datenschutz absichernden Verfahren integriert werden können. Vor diesem Hintergrund werden auf Basis der einzelnen Schutzziele zur Absicherung der informationellen Selbstbestimmung von Menschen, die in der datenschutzrechtlichen Terminologie „Betroffene“ genannt werden, erste Forschungsansätze dargestellt. Diese sollen die Optionen technischer Datenschutzmechanismen nach dem aktuellen Stand der Technik reflektieren, aber auch, die Grenzen von technischen Schutzkonzepten aufzeigen.

Dieses Papier spiegelt sowohl unterschiedliche rechtliche Auffassungen als auch unterschiedliche Schwerpunktsetzungen zum regulatorischen Rahmen wider. Dies gilt insbesondere mit Blick auf die Betonung der technischen und organisatorischen Maßnahmen. Insoweit handelt es sich um ein Thesenpapier, das die unterschiedlichen Ansätze nicht negiert, sondern vielmehr die Ergebnisoffenheit der Begleitforschung betont.

Editorial: Verfahrensrechtliche Anknüpfung und Gesetzessystematik des Schutzkonzeptes

Dr. Oliver Raabe, Leiter Fachgruppe Rechtsrahmen, Smart Data Begleitforschung

Notwendigkeit der verfahrensrechtlichen Anknüpfung von Technikregulierung

Der Schutz der informationellen Selbstbestimmung im Kontext von Smart Data kann durch eine Orientierung an modernen technischen Regulierungsverfahren in kritischen Infrastrukturen gestärkt werden. Das Hauptproblem eines primär ordnungsrechtlich agierenden Schutzregimes besteht nach wie vor in der faktisch schwierigen Durchsetzung der ordnungsrechtlichen Obliegenheiten der Akteure mit den derzeit vorhandenen Mitteln. Diese Schwierigkeiten treten oftmals nicht deshalb auf, weil die Softwareentwickler nicht willens wären, einem Normapell zu folgen, sondern weil es an der Kenntnis der Obliegenheiten fehlt. Für die Fälle der Verwendung von Smart Data in den Fallgruppen „kritischer Infrastrukturen“ wird jedenfalls zukünftig, wie schon heute im Smart Grid, die verbindliche Festlegung von Protokollen und Standards unter dem Gesichtspunkt der notwendigen Interoperabilität eine bedeutendere Rolle einnehmen. Damit stehen in diesen Fällen – wie bspw. im Festlegungsverfahren der Bundesnetzagentur zur Marktkommunikation im Energiemarkt – wettbewerblich motivierte Verfahren zur Verfügung, die auch für eine bereichsspezifische Normierung technikwirksamer Datenschutzanforderungen genutzt werden können. Die Besonderheit dieser Verfahren ist, dass die Regulierungsbehörden lediglich Zielvorgaben hinsichtlich der einzustellenden funktionalen und nichtfunktionalen Anforderungen formulieren und sodann die betroffenen Branchen, die in der Regel den Stand der Technik besser einzuschätzen vermögen, die konkreten Protokolle und Formate entwickeln. Durch die branchenspezifische Verbindlichkeit (mit Widerrufsvorbehalt) wird damit in der Regel Rechtssicherheit für die Akteure generiert und eine Mitwirkung findet schon im Eigeninteresse der Marktakteure statt. Die Anknüpfung an den Begriff der „kritischen Infrastruktur“ könnte wegen der in diesen Bereichen erhöhten staatlichen Schutzpflichten zur umfassenden Ausgestaltung des technikwirksamen Rechtsrahmens in diesen Bereichen sinnvoll sein.

Erforderlichkeit bereichsspezifischer Differenzierung

Eine hohe staatliche Regulierungsdichte bei auch zukünftig bereichsspezifischen Regelungen dürfte für kritische Infrastrukturen und weitere sensible Bereiche (bspw. bei Gesundheitsdaten) angebracht sein. In gesetzessystematischer Hinsicht sind strengere Maßstäbe bei bislang schon besonders sensiblen Arten von personenbezogenen Daten anzulegen. Gleichzeitig sind gewichtige Aspekte wie bspw. eine durch den Einsatz von Smart Data angekündigte Effektivierung des Gesundheitsschutzes bei Erforderlichkeitsabwägungen und Risikoprognosen zu berücksichtigen. Dies zeigt, dass für spezifische Bereiche ein höheres Maß an „Verrechtlichung“ auch im Hinblick auf Technikgestaltungsvorgaben erforderlich ist, als bei Alltagstätigkeiten mit geringer Eingriffsintensität. Daher sollte es auch in einem System gut durchdachter technischer Schutzmechanismen bei einer bereichsspezifischen Detailregelung für bestimmte Sachbereiche bleiben, wie bspw. bei hochauflösenden Smart-Meter-Daten. Eine praxisferne Detailregulierung des technischen Datenschutzes in Sachbereichen, die weder den kritischen Infrastrukturen zuzurechnen noch mit besonderem grundrechtlichen Schutzauftrag versehen sind, erscheint meines Erachtens nicht sinnvoll. Die Regelung kann hier vermutlich mit einer geringeren Verbindlichkeit erfolgen, als es bei den vorgenannten Sachgestaltungen erforderlich ist. Gleichzeitig muss in einem differenzierten System zur Erhaltung eines grundsätzlich angemessenen Schutzniveaus, die Position des Betroffenen in informatorischer Hinsicht sowie seine Fähigkeit zur eigenen Risikoprognose erheblich gestärkt werden.

Die Vielgestaltigkeit von offenen Strukturen lässt die Befürchtung aufkommen, dass ein vergleichbares Maß an gesetzlicher Vorabwägung wie in geschlossenen

Infrastrukturen den gesetzlichen Prognosehorizont übersteigt. Zur Erreichung eines lückenlosen Datenschutzes gilt es hier insbesondere die Nutzer zusätzlich durch angemessene und auch technisch fundierte Informationen zu ertüchtigen.

Vom Datenschutz zur Wissenskontrolle

Die Datenquellen von Smart Data sind zukünftig so hochverteilt und heterogen, dass eine technikrechtliche Kontrolle des Einzelnen oder des Staates über die Datenquellen in offenen Strukturen an ihre Grenzen stößt. Insofern ist es angeraten, sich die Prinzipien des Volkszählungsurteils vor Augen zu führen, das in seiner Begründung vom Schutzgut der „Entscheidungsfreiheit über vorzunehmende oder zu unterlassende Handlungen einschließlich der Möglichkeit, sich auch entsprechend dieser Entscheidung tatsächlich zu verhalten“ ausgeht und in den Mittelpunkt der Schutzgutgefährdung mitnichten „Daten“, sondern vielmehr das mögliche „Wissen“ Dritter und die Rückwirkungen auf das Schutzgut stellt. Denn „wer nicht überschauen kann, welche ihn betreffenden Informationen in bestimmten Bereichen seiner sozialen Umwelt bekannt sind, und wer das Wissen möglicher Kommunikationspartner nicht einigermaßen abzuschätzen vermag, kann in seiner Freiheit wesentlich gehemmt werden, aus eigener Selbstbestimmung zu planen oder zu entscheiden. Mit dem Recht auf informationelle Selbstbestimmung wären eine Gesellschaftsordnung und eine diese ermöglichende Rechtsordnung nicht vereinbar, in der Bürger nicht mehr wissen können, wer was wann und bei welcher Gelegenheit über sie weiß.“ Auch wenn das eigentliche Schutzkonzept in der Folge eine datenzentrische Sicht einnimmt, ist es doch das „Wissen Dritter“, das hier im Mittelpunkt steht. Damit kann sich das Volkszählungsurteil auch insofern als zeitgemäß erweisen, als die eigentliche Bedrohung der informationellen Selbstbestimmung durch Smart Data insbesondere in der ungeplanten Generierung von neuem Wissen bestehen kann. Daher sollte untersucht werden, ob ein normativer und technologischer Ausgleich widerstreitender Positionen von verarbeitenden Stellen und Betroffenem in einem gesetzlichen Schutzprogramm nicht auch, neben der Kontrolle der Datenerfassung durch Elemente des Selbstdatenschutzes, ausdrücklicher an dem Prozess und den Methoden der Wissensgenerierung ansetzen könnte.

Die Zentrierung auf „Daten“ als Anknüpfungspunkt für rechtliche und technische Schutzmechanismen der informationellen Selbstbestimmung verkennt, dass es unter den Prämissen von Smart Data vermehrt ebenfalls um die Kontrolle über das geschöpfte Wissen, mithin auch die Vermeidung von unerwünschten Wissensasymmetrien, gehen muss. Die auch verfahrensrechtlich zu fixierenden rechtlichen und technischen Mechanismen müssen zukünftig zunehmend am Begriff des Wissens, den Akteuren der Wissensgenerierung und den dort greifenden Kontrollmechanismen ansetzen - ohne die Entstehung der Daten beim Betroffenen aus dem Blick zu verlieren.

Erste Impulse für eine interdisziplinäre rechtlich-technische Evaluation

Um künftigen, interdisziplinären Forschungsbedarf im Hinblick auf neue zu erwartende Phänomene von Smart Data aufzuzeigen, bedarf es eines umfassenden Verständnisses der Schutzoptionen, die schon gesetzlich adressiert oder als Innovationen der kommenden EU-Datenschutzgrundverordnung absehbar sind, sowie ggf. weitergehender technischer Schutzmechanismen und Analysemethoden der Informatik. Die interdisziplinäre Forschung ist an dem Ziel auszurichten, eine Abschwächung des Datenschutzniveaus bei Smart-Data-Anwendungen zu vermeiden. Vielmehr soll die Smart Data Begleitforschung einen Beitrag bei der Entwicklung von Lösungen leisten, innovative Formen der Datennutzung mit den aus Art. 2 Abs. 1, 1 Abs. 1 Grundgesetz und Art. 7, 8 der EU-Grundrechtscharta bestehenden Grundrechten in Einklang zu bringen. Die Vorstellung erster Impulse ist zwangsläufig generalisierend und berücksichtigt noch nicht die Vielzahl der bereichsspezifischen Schutzbedürfnisse und einfachgesetzlichen Ausprägungen, sondern soll vielmehr künftigen Forschungs- sowie Diskussionsbedarf aufdecken. Als Ordnungskriterium bieten sich die datenschutzrechtlichen Grundprinzipien an.

1. Personenbezug als zentrales Tatbestandsmerkmal für regulatives Eingreifen

Sowohl nach derzeit geltendem Recht als auch nach der voraussichtlich kommenden Datenschutzgrundverordnung knüpft die Anwendbarkeit des Datenschutzrechts und damit des Verbots mit Erlaubnisvorbehalt am Merkmal des Personenbezugs an. Dieser ist gegeben, wenn Informationen über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person vorliegen. Im Kontext von Smart Data stößt die Abgrenzung personenbezogener Daten und „datenschutzrechtlich unbedenklicher“ Sachdaten oder anonymer Daten zunehmend an Grenzen, wenn durch die Kombination von Daten aus unterschiedlichen Quellen die Wahrscheinlichkeit steigt, dass ein Datensatz zu einer bestimmbaren Person entsteht. Über Kontextwissen können Sachdaten und (anfänglich) anonyme Daten Personen zugeordnet werden und somit der Informationsgehalt der diese Person betreffenden Erkenntnisse erhöht werden. Unter diesem Gesichtspunkt können sämtliche Daten grundsätzlich geeignet sein, das Risiko der Persönlichkeitsausforschung zu verstärken.

These 1.1
Eine anfängliche Kategorisierung von Daten(quellen) als datenschutzrechtlich „relevant“ und „unbedenklich“ reflektiert die Generierung von zusätzlichem Kontext- und Zufallswissen nicht hinreichend.

Für eine verantwortliche Stelle stellt sich daher zunächst die Problematik, über die Einstufung der Daten als „personenbezogen“ die Anwendbarkeit des Datenschutzrechts zu prüfen. Die Verknüpfung von Daten – ob personenbezogen oder nicht-personenbezogen – kann grundsätzlich dazu führen, dass der neu entstehende Datensatz Personenbezug aufweist. Intensiviert wird diese nicht unumstrittene Abgrenzungsproblematik in offenen Datenverarbeitungsstrukturen (d. h. bei Datenweitergabe an Dritte), wenn die erhebende Stelle kaum noch abschätzen kann, mit welchem zusätzlichen Kontextwissen Daten kombiniert werden. Wenn Daten nach einer Anonymisierung weitergeleitet oder veröffentlicht werden, darf aus rechtlicher sowie technischer Sicht das Risiko der Re-Identifizierung nicht vernachlässigt werden.

These 1.2
Das Datenschutzrecht ist unter den Prämissen von Smart Data nur dann nicht anwendbar, wenn die verantwortliche Stelle auch eine Kombination der erhobenen Daten technisch-organisatorisch sicher ausschließen kann, sodass kein Wissen über eine bestimmbare, natürliche Person generiert wird.

Absolute und faktische Anonymität werden mehrheitlich als Gegensatz zum Personenbezug verstanden und

führen folglich zur Nichtanwendbarkeit des Datenschutzrechts. Im Gegensatz zur absoluten Anonymität ist bei faktischer Anonymität die Re-Identifizierbarkeit objektiv nicht vollständig ausgeschlossen, sondern ist mit einem unverhältnismäßigen bzw. unvernünftigen Aufwand verbunden. Gegenstand rechtlicher Diskussionen ist, ob sich die Bestimmung der Unverhältnismäßigkeit der De-Anonymisierung nur auf die Möglichkeiten und Ziele der verantwortlichen Stelle (relative Theorie) oder auch auf die von Dritten (absolute Theorie) bezieht. Zudem kann die Frage nach der Verhältnismäßigkeit nur kontextbezogen, anhand der verfolgten Ziele und zeitabhängig – auch anhand des Stands der Technik – aus objektiver Sicht beurteilt werden. Naturgemäß unterliegen diese Kriterien einem stetigen Wandel, sodass zuvor noch anonyme Daten durch technischen Fortschritt Personenbezug erhalten können. Dies gilt auch für eine vom Gesetzgeber in Form von widerleglichen Vermutungen und Regelfällen antizipiert vorgenommene Abwägung, die sich dem stetigen Wandel anzupassen vermögen.

These 1.3
Die faktische Anonymität eignet sich gerade für Rechtsunkundige kaum als rechtssicheres Abgrenzungskriterium.

Im Hinblick auf Vorstadien und Vorbereitungshandlungen zur Verwendung personenbezogener Daten könnte die Anknüpfung der datenschutzrechtlichen Regelungen am Begriff des Personenbezugs bei enger Auslegung dieses Begriffs die Gefahr von Schutzlücken hervorrufen. Wird der Personenbezug hingegen weit ausgelegt, kommt es u. U. zu einer Einschränkung der wirtschaftlichen Tätigkeit der datenverarbeitenden Stellen, selbst wenn im Einzelfall eine Gefährdung der Rechtsposition des Betroffenen nicht besteht. Bei einer Abwägung der widerstreitenden Rechtspositionen muss bedacht werden, ob der Schutz der Privatsphäre und das Recht auf freie wirtschaftliche Betätigung gleichberechtigt nebeneinander stehen. Aus der Ausstrahlung des Wertesystems des Grundgesetzes im Lichte der Rechtsprechung des Bundesverfassungsgerichts kann abgeleitet werden, dass dieses Wertesystem der Berufsausübungsfreiheit (Art. 12 GG) und der allgemeinen wirtschaftlichen Betätigungsfreiheit (Art. 2 Abs. 1 GG) im Verhältnis zum Schutz des Rechtes auf informationelle Selbstbestimmung als Ausprägung der Menschenwürde (Art. 2 Abs. 1 i. V. m. Art. 1 Abs. 1 GG) eine schwächere Stellung zuweise. Sind Smart-Data-Anwendungen aber auch dort unzulässig, wo das oder die verarbeitenden Unternehmen keine Kenntnis über die Identität des Betroffen haben, besteht in tatsächlicher Hinsicht kein Eingriff in das Recht auf informationelle Selbstbestimmung. Schwierigkeiten bereitet eine klare Grenzziehung, welche Kenntnisse der verantwortlichen Stelle zuzurechnen sind: als rechtliche Anknüpfungspunkte könnten berücksichtigt werden (A) welches Wissen die verantwortliche Stelle besitzt; (B) welches Wissen sie potentiell aus der weiteren Datenauswertung ermitteln kann; (C) zusätzlich welches Wissen Dritter ihr zugerechnet werden kann und (D) welches Wissen Dritte potentiell erlangen können. Grundsätzlich stellt sich in diesem Zusammenhang die Frage, wer das Risiko einer (nachträglichen) Re-Identifizierung tragen soll: bei (anfänglicher) Nichtanwendbarkeit des Datenschutzrechts ist zu befürchten, dass das Risiko beim Betroffenen verbleibt, obwohl Maßnahmen zur Risikoabwendung vielmehr von der verantwortlichen Stelle ergriffen werden sollten. Diesem Problem könnte eine differenzierte rechtliche Ausgestaltung entgegenwirken, welche explizit das Risiko bei Vorstadien der Verwendung personenbezogener Daten adressiert (z. B. das Re-Identifikationsrisiko und Datensicherheitsrisiken) – die Verwendung aber nicht per se verbietet, wenn eine Beeinträchtigung der informationellen Selbstbestimmung durch technisch-organisatorische Anforderungen auf ein akzeptables Maß verringert werden kann (z. B. durch Datentrennung, Verschlüsselung, weitergehende Anonymisierung/Pseudonymisierung, Weitergabeeinschränkungen, usw.)

2. Datensparsamkeit

Die in § 3a Bundesdatenschutzgesetz als Regelbeispiele vorgesehen Methoden zur Datensparsamkeit beziehen sich auf die (temporäre/personale) Aufhebung des Personenbezuges. Missverständnisse entstehen häufig durch die Tatsache, dass der technische Begriff der „Anonymisierung“ nicht mit der rechtlichen Definition des § 3 Abs. 6 Bundesdatenschutzgesetz übereinstimmt. Während „Anonymisierung“ im juristischen Sinn die Entfernung des Personenbezugs bedeutet, können technische Verfahren (Methoden aus dem Bereich Data Privacy), auch an der Aussage ansetzen und beispielsweise deren Informationsgehalt verringern. Diese Methoden werden oftmals ebenfalls als „Anonymisierungsverfahren“ bezeichnet. Greifen die aus den Daten gewinnbaren Erkenntnisse und Aussagen weniger in die Persönlichkeitssphäre der Betroffenen ein, kann damit, je nach konkretem Sachverhalt, ein verringertes Risiko für deren Rechtsposition enthalten sein. Dieses Potential der Senkung des Informationsgehalts wird rechtlich bisher kaum explizit adressiert.

Abbildung 1: Unterschiedliche Wirkungsmöglichkeiten von Anonymisierung im Rechtssinn und Methoden aus dem Bereich Data Privacy

These 2.1
Datensparsamkeit sollte nicht nur als Vermeidung der Datenerhebung als solche, sondern könnte auch als „Minimierung der Eingriffsintensität“ für den Betroffenen oder als „Wissenssparsamkeit“ verstanden werden, wenn Analysezwecke eine Erhebung legitimieren.

Bei Anonymisierungsmethoden, die den Personenbezug entfernen, werden die identifizierenden Attribute wie Namen und andere Identifikatoren gelöscht oder unterdrückt. Daneben bestehen Methoden, die an der Aussage (und damit an dem aus der Information generierbaren Wissen) ansetzen, d. h. die Aussage verallgemeinern oder unschärfer machen. Beispielsweise können numerische Attribute (wie Messwerte oder Angaben zur Kaufkraft von Personen) durch Hinzufügen von statistischem Rauschen (Verrauschen – wie es bspw. bei Verfahren, die differential privacy erfüllen angewandt wird) und kategorische Attribute (wie bspw. Krankheiten oder Religionszugehörigkeit) durch Ersetzung mit einer allgemeineren Kategorie (Generalisierung) in ihrer Aussage unschärfer gemacht werden. Es ist auch denkbar, solche Methoden auf den Personenbezug anzuwenden. Beispielsweise werden, um k-anonymity zu erreichen, Attribute, mit denen Personen indirekt identifiziert werden können, generalisiert. Aus rechtlicher Sicht wird hingegen kontrovers diskutiert, ob es sich um eine Verarbeitung handelt, welche einer Rechtsgrundlage bedürfte. Eine weitere Herausforderung besteht hierbei darin, diese recht einfachen abstrakten Methoden fallabhängig derart zu instanziieren und zu kombinieren, dass Datenschutzgarantien gegeben werden können, ohne die eigentliche Aussagepräzision so zu verändern, dass die Daten für die geplanten Analysen untauglich werden. Dies impliziert eine Abwägungsentscheidung zwischen

den Rechten der Betroffenen und den Interessen der verantwortlichen Stelle, unter Berücksichtigung der Wertentscheidung des Grundgesetzes. Diese Methoden können nicht als Substitut des derzeitigen rechtlichen Verständnisses des Anonymitätsbegriffs mit der Folge der Nichtanwendbarkeit des Datenschutzrechts gesehen werden, sondern lediglich als Absenkung der Eingriffsintensität. Für eine Datenverarbeitung bedarf es daher immer noch im zuvor beschriebenen Sinne einer Legitimationsgrundlage.

Ausgehend von der mit einem Wissenszugewinn verbundenen Eingriffsintensität für die Grundrechtsposition der betroffenen Bürger, ist zu untersuchen, ob das Absenken des Informationsgehaltes ein vergleichbares Schutzniveau erreichen kann, wie die Anonymisierung oder Pseudonymisierung im Rechtssinn. Datensparsamkeit im Wortsinne liegt nicht vor, da im Gegenteil mit dem bewussten Einsatz einer Vielzahl von Daten gearbeitet wird. Diese Betrachtung führt zur Frage, ob die Aufnahme neuer Begriffe (wie bspw. Verrauschen) in den gesetzlichen Rahmen sinnvoll ist. Eine Lösung der Problematik, dass durch die intelligente Kombination mehrerer entpersonalisierter („verrauschter“) Datenbestände der ursprüngliche Informationsgehalt wiederhergestellt werden kann, bleibt Gegenstand für die kommende Forschung.

These 2.2
Techniken mit dem Ziel der Wissensregulation können so eingesetzt werden, dass sie trotz großer Datenmengen wenig unerwünschtes Wissen über den Betroffenen generieren.

Ob der Einsatz von fortgeschrittenen Verfahren der Informatik zur Kontrolle und Absenkung der Eingriffsintensität bspw. durch Verringerung des Informationsgehalts, jenseits der Aufhebung des Personenbezuges von Daten, im gesetzlichen Rahmen eine Privilegierung erfahren sollte, könnte als Impuls für zukünftige Diskussionen weiter erforscht werden. In Fällen, in denen der Betroffene nicht mehr unmittelbar über die Verwendung der Informationen über ihn bestimmen kann, erhält die Feststellung des Bundesverfassungsgerichts in der „Mikrozensusentscheidung“, dass „dem Einzelnen um der freien und selbstverantwortlichen Entfaltung seiner Persönlichkeit willen ein ‚Innenraum‘ verbleiben muss, in dem er ‚sich selbst besitzt‘ und ‚in den er sich zurückziehen kann, zu dem die Umwelt keinen Zutritt hat, in dem man in Ruhe gelassen wird und ein Recht auf Einsamkeit genießt‘“, eine besondere Bedeutung.

These 2.3
Der Schutzauftrag aus dem Volkszählungsurteil wird im Recht verkürzt, wenn man ausschließlich auf den Personenbezug abstellt. Schutzzweck der informationellen Selbstbestimmung ist auch der Erhalt einer der Umwelt nicht zugänglichen Privatsphäre. Damit umfasst der Schutzauftrag auch den Schutz vor Wissensgenerierung über die Privatsphäre des Einzelnen.

Das für das aktuelle Datenschutzverständnis grundlegende Prinzip der Datensparsamkeit steht dann diametral mit dem Anspruch von Smart Data im Konflikt, wenn Erkenntnisse gerade aus großen Datenbeständen erlangt werden sollen. Mit Privacy-Preserving-Data-Mining-Technologien könnte eine Annäherung an die Vorgabe erreicht werden, möglichst wenig persönlichkeitsrelevante Daten zu verwenden. Bei diesen Technologien können Techniken zur Anonymisierung oder die aus der Kryptographie stammenden sicheren Mehrparteienberechnungen eingesetzt werden. Bei sicheren Mehrparteienberechnungen können mehrere Parteien, die jeweils eine geheime Eingabe besitzen, eine Funktion berechnen. Jede Partei lernt nur das Ergebnis der Berechnung, nicht jedoch direkt die Eingaben der anderen Parteien (auch wenn, je nach Funktion, aufgrund des Ergebnisses Rückschlüsse auf die Eingaben nicht vollständig auszuschließen sind). Solche Verfahren sind jedoch in Smart-Data-Anwen-

dungen, bei denen es in der Regel eine Vielzahl von Datenquellen gibt, kaum praktisch einsetzbar, da jede Partei an der Berechnung beteiligt sein muss. Erfolgversprechender wären Anonymisierung verwendende Privacy-Preserving-Data-Mining-Technologien. Hierbei können die erhobenen Daten, Zwischenergebnisse während der Verarbeitung oder auch nur das Analyseergebnis selbst anonymisiert werden. Aktuell ist jedoch ohne einen Erlaubnistatbestand die Verarbeitung personenbezogener Daten unzulässig, auch wenn das Ergebnis anonymisiert werden würde. Der Vorteil einer ausschließlichen Anonymisierung der Analyseergebnisse wäre jedoch ein exakter Ausgangsdatenbestand, der beispielsweise bei einer vertrauenswürdigen Instanz lagern könnte. Der Bestand würde somit für potentielle weitere Analysen zur Verfügung stehen, bei gleichzeitiger Minimierung der Eingriffsintensität für den Betroffenen durch die Analyse. Zu klären bleibt fallabhängig, ob die so entstandenen Ergebnisse noch genau genug für die Analysezwecke sind, sodass zwischen dem Grad der Anonymisierung (insbesondere dem Schutzbedürfnis der Betroffenen) und der Genauigkeit der Analyseergebnisse abzuwägen ist. Derartige Sachverhalte werden durch bestehende Gesetze nicht in Form der Anonymisierung im Rechtssinne, sondern im Rahmen von Abwägungsentscheidungen berücksichtigt. Die aktuelle Forschung sollte ausloten, ob und in welcher Form bei einer solchen, die Eingriffsintensität minimierenden Technikverwendung eine ausdrückliche (Teil-)Privilegierung im zukünftigen Rechtsrahmen erfolgen könnte. Für bestimmte sensible Daten könnte eine Genehmigungspflicht durch die Aufsichtsbehörden unter Beachtung des Grundsatzes der frühestmöglichen Anonymisierung, etwa in Form eines Verbots mit Genehmigungsvorbehalt durch zulassenden Verwaltungsakt, erwogen werden.

These 2.4
Privacy-Preserving-Data-Mining-Mechanismen und Methoden aus der Kryptographie können die datenschutzkonforme Gestaltung von Smart-Data-Anwendungen unterstützen. Hierbei besteht jedoch noch weiterer Bedarf interdisziplinärer Forschung insbesondere der Abgleich bestehender Verfahren mit dem Datenschutzrecht.

3. Datensicherheit

Für Smart-Data-Anwendungen könnten, wegen der großen Datenmengen und der benötigten Rechenleistung für Echtzeitanalysen vermehrt Cloud-Systeme zum Einsatz kommen. Bei der Nutzung von Cloud-Systemen, insbesondere von externen Dienstleistern, ist zu befürchten, dass aufgrund der dezentralen Speicher und der notwendigen Kommunikationsbeziehungen ein erhöhtes Angriffsrisiko besteht. Aus rein rechtlicher Perspektive sind als Maßnahme insbesondere die Erkenntnisse der Begleitforschung zu Trusted Cloud hervorzuheben, die durch eine modular gestaltete Auftragsdatenverarbeitung und ggf. Zertifizierung ein angemessenes Schutzniveau auch beim Cloud-Betreiber garantieren können. Im Einzelfall ist rechtlich zu prüfen, ob die Voraussetzungen der Auftragsdatenverarbeitung vorliegen. Die Zertifizierung von Auftragsdatenverarbeitungsvorgängen sind als integraler Bestandteil eines Schutzkonzeptes für cloud-basierte Smart-Data-Anwendungen zu berücksichtigen. Besondere Herausforderungen bestehen jedoch weiterhin bei Datenauslagerung in Nicht-EU-Staaten.

Verbindlich vereinbarte technische Schutzkonzepte können den Rahmen für eine Reihe von Datensicherheitsmechanismen bilden. Schutzkonzepte nach dem Stand der Technik sollten in den Anforderungskatalog überführt werden. In diesem Zusammenhang erhalten Instrumente der Zertifizierung, Gütesiegel und Auditierung, die bereits auf freiwilliger Basis genutzt wer-

den, eine besondere Bedeutung. Die zugrundeliegenden Anforderungskataloge müssen im Hinblick auf die besonderen Anforderungen von Smart-Data-Anwendungen und Anwendungsszenarien verfeinert und ihre Anwendung durch die verantwortlichen Stellen und Softwarehersteller als Mindeststandards verbindlich erklärt werden. Gleichzeitig gilt es, den Datenschutzbehörden für den Fall der Nichtbeachtung das notwendige Eingriffsinstrumentarium an die Hand zu geben, das ein frühzeitiges Tätigwerden gestattet.

Daneben bieten sich, wie in der Anlage zu § 9 Satz 1 Bundesdatenschutzgesetz explizit genannt, Verschlüsselungsverfahren an, um Daten während der elektronischen Übertragung und Speicherung vor unbefugtem Lesen zu schützen. Hierbei muss zwischen symmetrischen und asymmetrischen Verfahren unterschieden werden. Symmetrische Verfahren verwenden zur Ver- und Entschlüsselung den gleichen Schlüssel. Asymmetrische Verfahren verwenden zur Ver- und Entschlüsselung unterschiedliche Schlüssel und eröffnen dadurch der Verschlüsselung weitere Anwendungsszenarien.

Hierbei ist zu beachten, dass beide verwendeten Verfahren (symmetrische und asymmetrische), wie auch deren Kombination (Hybride Verfahren), wie beispielsweise in TLS(Transport Layer Security), nach dem Stand der Technik sicher sein müssen. TLS kommt zur sicheren Datenübertragung im Internet zum Einsatz. Aktuell nach dem Stand der Technik sichere symmetrische Verfahren sind bspw. Camellia und AES (Advanced Encryption Standard) ab einer Schlüssellänge von 192 Bit. Nach dem Stand der Technik unsichere Verfahren dürfen folglich laut § 9 Bundesdatenschutzgesetz nicht mehr verwendet werden, sind aber leider noch in vielen Anwendungen und Produkten zu finden. Veraltete Verfahren sind bspw. deterministische Verschlüsselungsverfahren, welche einen Klartext immer auf das gleiche Chiffrat abbilden. Sichere asymmetrische Verfahren sind aktuell ECC (Elliptic Curve Cryptography) 256 (NIST (National Institute of Standards and Technology) Kurve und Curve 25519), RSA 8192, ECC 384 (NIST) sowie ECC 512 (NIST). Diese spielen jedoch eine untergeordnete Rolle, da sie meist in hybriden Verfahren verwendet werden. Bei diesen gilt aktuell TLS 1.2 als Stand der Technik. Bei der Verwendung von kryptographischen Verfahren und Protokollen muss unbedingt beachtet werden, diese nicht selbst zu implementieren, da deren Implementierung extrem fehleranfällig ist. Vielmehr sollten verfügbare Implementierungen verwendet werden. An dieser Stelle sei auch auf das BSI verwiesen, welches eine regelmäßig aktualisierte technische Richtlinie zu kryptographischen Verfahren bereitstellt. Es sollte trotz der großen Anzahl an vorhandenen Verfahren beachtet werden, dass nicht jede aus dem Bundesdatenschutzgesetz abgeleitete Vorgabe mittels technischer Maßnahmen umgesetzt werden kann. Vielmehr müssen technische und organisatorische Maßnahmen aufeinander abgestimmt werden, so dass personenbezogene Daten in der gesamten Anwendung wirksam geschützt werden. Beispielsweise ist eine Verschlüsselung nicht sinnvoll, wenn die dabei verwendeten Schlüssel nicht entsprechend geschützt werden.

Bei Datensicherheits- und Datenschutzbetrachtungen von Smart-Data-Anwendungen nehmen die Systemgrenzen eine herausragende Rolle ein. Können dem Stand der Technik entsprechende Verschlüsselungstechnologien universal eingesetzt werden, so bieten Mechanismen wie bspw. durchsuchbare Verschlüsselung und Anonymisierungsverfahren meist anwendungsspezifische Sicherheitsgarantien. Diese gelten jedoch nur wenn von den der Garantie zugrundeliegenden Annahmen nicht abgewichen wird.

These 3.1
Sicherheitsgarantien ermöglichen eine nachvollziehbare Beurteilung von einzelnen Sicherheitsmaßnahmen. In vielen Fällen sind diese jedoch anwendungsspezifisch.

Können die Systemgrenzen nicht bestimmt werden, kann nicht ermittelt werden, ob eine anwendungsspezifische Sicherheitsgarantie noch gilt oder ob von den oben erwähnten Annahmen abgewichen wird. Dies ist bspw. bei einer Datenweitergabe der Fall ist, bei der die verantwortliche Stelle (technisch) nicht kontrollieren kann, wie die Daten verwendet werden und welche nachfolgenden Weiterleitungsvorgänge erfolgen („offene Systeme“). Ohne weitere Annahmen, bspw. die Annahme der Einhaltung vertraglicher Regelungen, können keine belastbaren Aussagen zu Sicherheit gegeben werden. Eine Überprüfung, ob vertragliche Regelungen auch eingehalten werden, kann sich unter Umständen technisch sowie organisatorisch als schwer beziehungsweise unmöglich erweisen.

These 3.2
Anwendungsspezifische Sicherheitsgarantien haben Grenzen, die in Datensicherheits- und Datenschutzbetrachtungen mit einbezogen werden müssen. Bei offenen Systemen kann unter Umständen nicht beurteilt werden, ob diese Grenzen erreicht werden.

Als technisch-organisatorische Maßnahme müssen datenverwendende Stellen zudem gewährleisten, dass zu unterschiedlichen Zwecken erhobene Daten getrennt verarbeitet werden. Zur Einhaltung dieser Vorgabe werden unterschiedliche Vorgehensweisen vorgeschlagen, beispielsweise softwareseitige Mandantentrennung, unterschiedliche Zugriffsberechtigungen, unterschiedliche kryptographische Schlüssel, Speicherung in getrennten Datenbanken, auf unterschiedlichen Servern, unterschiedliche Pseudonyme und Zuordnungsregeln, sowie Trennung des Test- und Routinebetriebs. Aus einer solchen Kombination unterschiedlichster Maßnahmen und Verfahren kann jedoch in vielen Fällen keine globale Garantie für die ganze Anwendung abgeleitet werden.

These 3.3
Um Datenschutzrisiken korrekt abschätzen zu können, ist zu untersuchen wie man von mehreren lokalen Sicherheitsgarantien (für einzelne Verfahren oder Maßnahmen) zu Aussagen bezüglich der Sicherheit der kompletten Anwendung gelangt.

4. Transparenz und Information

Für die Gewährleistung einer auf Informiertheit fußenden Selbstbestimmung der Betroffenen müssen diese über die mit der Datenfreigabe verbundene Risiken ausreichend informiert werden. Ohne Mithilfe der datenverarbeitenden Stellen sind diese Risiken schwer einschätzbar. Die Ausübung der Datenhoheit durch die Einwilligung bedarf eines deutlich höheren Informationsniveaus des Einzelnen, um die Chancen und Risiken bewerten und seine Rechte verwirklichen zu können. Eine Einholung von Einwilligungen zu Einzeldaten wird sich kaum als praktikabel erweisen. Bei einer vom Einzeldatum abstrahierenden Kategorisierung ist fraglich, ob ein noch hinreichendes Schutzniveau besteht. Andererseits ist bei hoher Granularität und damit verbundener Informationsflut eine Überforderung des Nutzers zu befürchten.

These 4.1
Die Risikoprognose des Bürgers als Grundlage seiner Entscheidung persönliche Daten preiszugeben, braucht ein rationales Fundament.

Die Forderung nach einer Stärkung der Datenhoheit des Bürgers verlangt gleichzeitig nach neuen Formen der Transparenzsicherung durch Information. In diesem Kontext könnte die Datenschutzfolgenabschätzung, welche ein Instrument der Risikobewertung darstellt, auch dafür genutzt werden die Transparenz für den Nutzer zu erhöhen und damit seine Selbstbestimmung zu stärken. Die Transparenz könnte durch die Verwendung eines einheitlichen, einfach verständ-

lichen Skalen- oder simplen Ampelsystems für die mit dem Datenumgang verbundenen Risiken unterstützt werden. Beispielsweise könnte man sich an der Darstellung der Energieeffizienzklassen bei der Kennzeichnung des Energieverbrauchs orientieren. Ergänzende Informationen bleiben zwar notwendig, um die Datenverarbeitung rechtlich nachvollziehbar zu machen. Eine simplifizierende Darstellung geht zwangsläufig mit Informationsverlusten einher. Ein Mehrwert bestünde aber gerade für Alltagsituationen, die eine schnelle Einschätzung erforderlich machen, darin das Bewusstsein über Risiken zu stärken. Die in den Entwurfsfassungen der Datenschutzgrundverordnung (Art. 33) bereits enthaltene Pflicht der Datenschutzfolgenabschätzung als interne Obliegenheit sollte erweitert werden und dem Betroffenen als Entscheidungsgrundlage über die Preisgabe von personenbezogenen Daten dienen. Darüber hinaus könnten einfach nutzbare Technologien, die automatisiert die Datenschutzpräferenzen der Betroffenen umsetzen, eine weitere sinnvolle Unterstützung sein. Derzeit befinden sich solche Technologien in der Erforschung und Erprobung.

These 4.2
Die Ergebnisse der Datenschutzfolgenabschätzung sollten zusätzlich in vereinfacht visualisierter Form dem Betroffenen als Entscheidungsgrundlage der Datenpreisgabe zugänglich gemacht werden.

Risikoprognosen im Bereich der IT-Sicherheit werden bereits in Form von Bedrohungsanalysen und im Rahmen des IT-Risikomanagements durchgeführt. Dabei werden Risiken für einen Prozess oder ein System identifiziert, beziffert und oft mit auf Erfahrungswerten basierenden Wahrscheinlichkeiten errechnet. Dies ermöglicht eine gewichtende Abwägung über die Ausgestaltung der technischen Systeme, insbesondere hinsichtlich des zusätzlichen Einsatzes von risikoverringernden Sicherheitsmechanismen, jeweils bezogen auf den zu erwartenden Schaden gegenüber dem Umsetzungsaufwand sowie den Kosten. Eine Datenschutzfolgenabschätzung geht über ein klassisches Risikomanagement hinaus. Im klassischen IT-Risikomanagement werden primär Risiken für das System oder mit dem System arbeitende Personen betrachtet. Bei einer Datenschutzfolgenabschätzung müssen über eine Werte-Betrachtung die persönlichen Rechte und Freiheiten der betroffenen Bürger in die Abschätzung miteinbezogen werden.

These 4.3
Die Verwendung von Risikoanalysen als Entscheidungsgrundlage bedarf gleichzeitig der Vermittlung der betrachteten Systemgrenzen und Wertekonstellationen. Es ist zu untersuchen, wie zukünftig laienverständliche Repräsentationen dieser komplexen Bewertungsprozesse umzusetzen sind.

Für personenbezogene Informationen in komplexen technischen Systemen lassen sich aktuell nur für sehr eingeschränkte Aspekte absolute und exakte Sicherheitsgarantien geben. Beispielsweise kann konzeptionell sichergestellt werden, dass bestimmte Daten, jenseits der prinzipiell nicht verhinderbaren analogen Lücke (z. B. das Abfotografieren oder Abschreiben von Bildschirminhalten), ein System nicht verlassen können. Jedoch besteht weiterhin die Gefahr von Fehlern in der Implementierung oder von menschlichem Fehlverhalten und den daraus resultierenden Sicherheitslücken. In kontrollierten Umgebungen müsste die verantwortliche Stelle das für den Betroffenen mit der Datenverwendung verbundene Risiko beurteilen können. Werden also Risikoanalysen grundsätzlich als Entscheidungsgrundlage des Betroffenen zu einer Einwilligung herangezogen, muss neben absoluten Werten, Mechanismen und Schutzgütern auch die Grenze dieser Risikoprognose in seine Entscheidung einfließen. Für Smart-Data-Anwendungen ist deshalb zu untersuchen, wie den Betroffenen die betrachteten Systemgrenzen und die verbleibenden Restrisiken vermittelt werden können.

These 4.4
Die Garantien von technischen oder organisatorischen Maßnahmen, insbesondere die von Anonymisierungsverfahren, sind anwendungsspezifisch und kontextabhängig. Aussagen zu offenen Systemen, bei denen nicht alle Kontexte bekannt sind, sind nur sehr eingeschränkt möglich. Dies sollte grundsätzlich zu einer Beurteilung als „hohes Risiko“ führen.

So besteht z. B. bei anonymisierten oder pseudonymisierten Daten eine kontextabhängige Gefahr der Re-Identifikation. Gegen Re-Identifikation robuste Anonymisierungstechniken und Anonymitätsgarantien sind nach wie vor Gegenstand der Forschung. Generell gilt, dass technische (und organisatorische Maßnahmen) das Risiko verringern können. Eine Abschätzung der Risiken selbst jedoch kann sich unter Umständen und abhängig vom zu beurteilenden System und der gewünschten Genauigkeit als schwierig bis unmöglich erweisen. Es ist oftmals zwar ersichtlich, welche Garantie eine einzelne Maßnahme bietet (Verschlüsselte Verbindungen verhindern bspw. das Abhören über ausschließlich ebendiese), aus einer Liste von Maßnahmen ist jedoch nicht ersichtlich, ob alle Datenschutzrisiken der Anwendung entsprechend abgedeckt sind.

These 4.5
Technische und organisatorische Maßnahmen dürfen nicht mit allgemeingültigen Garantien verwechselt werden. Eine Auflistung der technischen und organisatorischen Maßnahmen gibt den Betroffenen im Allgemeinen keine Garantien. Notwendig ist deshalb eine transparente, anwendungsumfassende Risikoabschätzung, die über eine Auflistung der verwendeten Maßnahmen hinaus geht.

Im Rahmen einer Risikofolgenabschätzung sind heute unter anderem Risikominimierungsmechanismen wie Anonymisierungs-, Pseudonymisierungs- und Verschlüsselungstechniken zu berücksichtigen. Auch bei sämtlichen Smart-Data-Anwendungen besteht grundsätzlich die gesetzliche Vorgabe die Auswahl und Gestaltung von Datenverarbeitungssystemen an dem Ziel der Datenvermeidung und Datensparsamkeit auszurichten. Es gibt jedoch keine effektive Sanktionierung bei Nichteinhaltung. Wenn dem Betroffenen die mit der Datenverwendung verbundenen Risiken offengelegt würden, könnte eine auf Risikominimierungsverfahren verzichtende Stelle den Nachteil erleiden, dass sich der Betroffene gegen eine Datenpreisgabe entscheidet. Das Schlagwort vom „Datenschutz als Wettbewerbsvorteil“ könnte neben den nun in der Datenschutzgrundverordnung angelegten Zertifizierungssystemen oder dem Selbstdatenschutz so zusätzlich belebt werden. Damit würde ein indirekt wirkender Anreiz geschaffen werden, Techniken zur Datenminimierung wie Anonymisierungs- und Pseudonymisierungsverfahren, aber ggf. auch zukünftig solche Techniken, die bei der Absenkung des Informationsgehaltes bzw. dem Wissen ansetzen, selbst bei gesteigerten Kosten oder zusätzlichem Aufwand einzusetzen. Dieser Aufwand könnte zusätzlich durch staatliche Unterstützung von Projekten, die den datenschutzkonformen Einsatz neuartiger Technologien erforschen oder neuartige Datenschutztechnologien entwickeln, für die Anwender von Smart-Data-Technologien verringert werden.

These 4.6
Anreizmechanismen zur Minimierung des mit dem Datenumgang verbundenen Risikos für die Rechtsposition des Betroffenen (bspw. Anonymisierung, Pseudonymisierung oder Verschlüsselung) sind erforderlich, um den Persönlichkeitsschutz zu verbessern. Die Verwendung der Ergebnisse von Datenschutzfolgenabschätzungen als Vorabinformation an den Betroffenen kann als Anreiz zur Datensparsamkeit wirken.

5. Kontrolle

Wie bereits ausgeführt, wurden klassische Risikoprognosen im Bereich der IT-Sicherheit bisher in Form von Bedrohungsanalysen und im Rahmen des IT-Risikomanagements durchgeführt. Die kommende Datenschutzgrundverordnung adressiert neben den klassischen Mechanismen des IT-Grundschutzes wie Verfügbarkeit, Integrität und Vertraulichkeit auch Aspekte wie den Schutz vor unbefugter Weitergabe, die primär die informationelle Selbstbestimmung zum Ziel haben. Weiter geht das von den Datenschutzbeauftragen des Bundes und der Länder verabschiedete Standarddatenschutzmodell (vgl. Abb. 2). Es ergänzt den Kanon der Schutzziele aus der IT-Sicherheit (Verfügbarkeit, Integrität, Vertraulichkeit) um Datenschutzziele wie Transparenz, Intervenierbarkeit (Nutzerrechte) und Nicht-Verkettbarkeit (Zweckbindung). Zudem müssen die Schutzbedarfsdefinitionen aus der Perspektive der Betroffenen formuliert werden und sämtliche Verfahren einbezogen werden (das Standarddatenschutzmodell umfasst Daten, Systeme und Prozesse). Diese Elemente werden also zukünftig auch integraler Bestandteil von Risikoabschätzungen.

Einheitliche, standardisierte Prüfmodelle erleichtern es, die Einhaltung datenschutzrechtlicher Vorgaben zu kontrollieren. Allgemeingültige Schemata und Methoden sollen den beteiligten Stellen, den Betroffenen sowie der Datenschutzaufsicht überprüfbare Kriterien und systematisierende Abwägungshilfen an die Hand geben, um für den jeweiligen Einzelsachverhalt den aufgrund der Sensibilität der Daten unterschiedlich zu beurteilenden Schutzbedarf für die Betroffenen anhand der verfolgten Schutzziele im Rahmen der einzelnen Datenverwendungsphasen zu bestimmen.

Standard-Datenschutzmodell

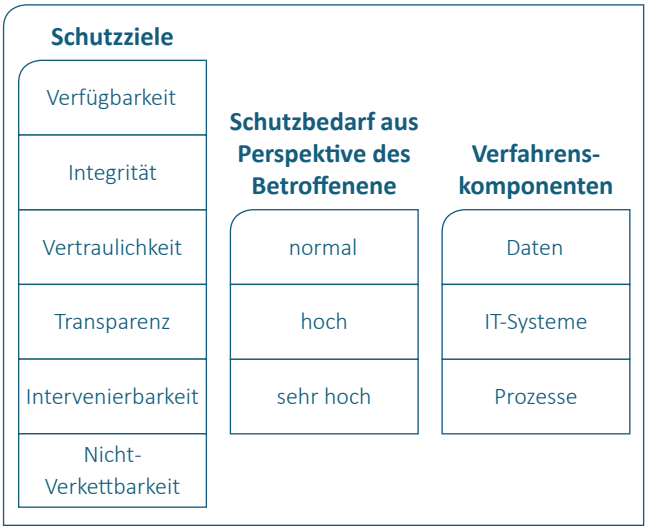

Abbildung 2 Vereinfachtes Standard-Datenschutzmodell (Quelle: Martin Rost, ULD / http://bit.ly/211knMH)

These 5
Das Standarddatenschutzmodell kann eine wichtige Konkretisierung und Entscheidungsbasis für die in der Datenschutzgrundverordnung enthaltenen Datensicherheitsanforderungen darstellen und sollte gleichzeitig im Rahmen der Datenschutzfolgenabschätzung nach Art. 33 DSGVO-E fortentwickelt werden.

6. Zweckbindung

Wenn explorative Analysen derart durchgeführt werden sollen, dass sich der Grund (d. h. das „Warum“) der Analyse erst im Nachhinein aus den Ergebnissen ergibt, oder sollen bereits bestehende Datenbestände „recycelt“ werden, sind Probleme in Bezug auf den Grundsatz der Zweckbestimmung und Zweckbindung zu erwarten. Um diese Form der Analysen zu unterstützen, ist eine Auflockerung der Zweckbindung im Disput. Festzuhalten ist, dass die Bestimmung der

Zwecke sowohl nach dem nationalen als auch nach dem europäischen Schutzkonzept eine notwendige Voraussetzung für die Bewertung der mit der Datenverarbeitung verbundenen Risiken für den Betroffenen und damit der darauf aufbauenden Schutzinstrumente ist. Sowohl eine im Rahmen der gesetzlichen Erlaubnistatbestände notwendige Interessenabwägung als auch die Reichweite der Einwilligung sind abhängig von der Bestimmung eines Zwecks. Der festgelegte Zweck bildet die Grundlage des „well informed consent". Nach der Rechtsprechung des EuGH verlangt der Schutz des Grundrechts auf Achtung des Privatlebens (Artt.7, 8 GrCharta), dass sich die Ausnahmen und Einschränkungen des Schutzes personenbezogener Daten auf das absolut notwendige Maß beschränken. Die Erforderlichkeit der Datenverwendung kann nur anhand des verfolgten Zwecks geprüft werden. Ein vollständiger Verzicht ist mit dem Gefüge des Datenschutzrechts demnach nicht möglich.

Die Zweckbestimmung und die Zweckbindung können als zwei unterschiedliche aufeinander aufbauende Regelungsinstrumente mit jeweils unterschiedlicher Regelungsfunktion betrachtet werden. Das deutsche Schutzkonzept sieht im Grundsatz die Identität der Zwecke für alle Phasen der Datenverwendung vor (strenge Zweckbindung), die durch Zweckänderungsmöglichkeiten für den öffentlichen Sektor nur bedingt und für den privaten Sektor großzügiger aufgelockert wird. Umstritten ist, ob das europäische Schutzkonzept Abweichungen zwischen den ursprünglichen Erhebungszwecken und den Zwecken der späteren Verwendung der Daten zulässt. Sind auf deutscher und europäischer Ebene die Schutzkonzepte der jeweils anwendbaren Grundrechte unterschiedlich, könnte folglich die Bedeutung und Funktion der Zweckbindung in Abhängigkeit dieser Schutzkonzepte variieren. Neben dieser Frage ist die zukünftige Ausgestaltung der Zweckbindung in der kommenden Datenschutzgrundverordnung Grundlage intensiver, disziplinübergreifender Diskussionen. Aus rechtlicher Sicht bestehen durchaus Zweifel, ob eine künftige Regelung, die die Zweckbindung der Datenerhebung sowie der Datenverarbeitung und / oder Weiterverarbeitung signifikant aufweicht, noch mit den Grundrechten aus Artt. 7 und 8 der Europäischen Grundrechtecharta vereinbar wäre.

Daneben wird auch im Hinblick auf Smart Data kritisch hinterfragt, welche Ansprüche an die Granularität der Zweckbestimmung zu stellen sind. Ein rechtlicher Maßstab für die Bestimmung der Verarbeitungszwecke, ihre Präzisierung und die darauf aufbauenden Schutzinstrumente könnte den ersten Schritt darstellen, derzeit bestehende Rechtsunsicherheit sowohl unter Datenverarbeitern als auch unter den Betroffenen zu verringern. Nichtsdestotrotz bleibt nach der bestehenden einfachgesetzlichen Ausgestaltung – auch nach den Entwürfen der Datenschutzgrundverordnung – die Zweckbestimmung der Einzelfallprüfung überlassen. Die Datenverarbeiter sind (in den Fällen zulässiger Zweckänderung) gezwungen, für jeden einzelnen Verarbeitungsschritt eine erneute Zweckbestimmung und damit Prüfung des Risikos für den Betroffenen und der erforderlichen Schutzinstrumente vorzunehmen. Die Menge der Einzelfallprüfungen, die aus der Grundlage der Daten für die Datenökonomie folgt, steht in diametralem Gegensatz zu dem Zwang, dem Datenverarbeiter unterliegen, unter Vermeidung von Transaktionskosten skalierbare Prozesse aufzusetzen. Auch die Betroffenen sind so gezwungen, für jeden Fall, in dem sie ein Datenverarbeiter über die Zwecke der Datenverarbeitung informiert, das damit für sie verbundene Risiko einzuschätzen. Die daraus folgende Informationsflut steht in einem vergleichbar diametralen Gegensatz zur Aufnahmemöglichkeit der Betroffenen in einer Welt, in der soziale Interaktion zunehmend auf der Verarbeitung von Daten beruht. Eine Lösungsmöglichkeit für dieses Problem könnte die Standardisierung und Zertifizierung von Verarbeitungszwecken darstellen.

These 6.1
Die Standardisierung von Zwecken könnte einen wichtigen Baustein zur Reduzierung von durch das Erfordernis von Einzelfallabwägungen hervorgerufene Rechtsunsicherheit darstellen. Als Voraussetzung für die Einstellung von Privacy-by-Design-Maßnahmen und ihre Einhaltung könnte die Zertifizierung Vertrauen stärken und einen Selbstregulierungsmechanismus für den internationalen Datenaustausch zur Verfügung stellen.

Technisch lassen sich, zumindest theoretisch, auf die Zweckbestimmung aufbauende Schutzinstrumente wie etwa die Zweckbindung mittels Usage Control-Methoden und Methoden aus dem Digital Rights Management (DRM) realisieren. Hierbei werden die Benutzungsbestimmungen zusammen mit den jeweiligen Daten gespeichert (Sticky Policies). Oft werden die eigentlichen Daten auch in verschlüsselten Containern gehalten. Diese können dann nur von zur Verarbeitung berechtigten Systemen (jene mit einer gültigen Lizenz) ausgelesen und verarbeitet werden. Bei einer Vielzahl berechtigter Systeme lässt sich dies effizient mittels Broadcast Encryption-Verfahren realisieren. Diese Verfahren erlauben es, Daten für eine Vielzahl möglicher Empfänger mit unterschiedlichen Schlüsseln zu verschlüsseln.

Die Sicherheit von DRM-Methoden ist umstritten. Sie werden allerdings bei digitalen Medien großflächig eingesetzt (Audio- und Video-Streaming, PayTV, Blu-ray, DVD), was zumindest in diesen Szenarien für sie spricht.

In Smart-Data-Szenarien gibt es jedoch zusätzliche praktische Herausforderungen:

- Wer setzt die Verarbeitungsrechte für die personenbezogenen Daten?
- Wer verschlüsselt die personenbezogenen Daten?
- Wie kann ein Schlüsselmanagement für DRM oder Usage Control im Smart-Data-Kontext umgesetzt werden?
- Wo und wie werden die zur Verarbeitung benötigten Lizenzen bzw. Zertifikate generiert und wie werden sie sicher in die Systeme eingebracht?

Denkbar ist hier die Hinzuziehung von vertrauenswürdigen Dritten (Trusted Third Party wie z. B. Notare oder Zertifizierungsstellen). Vor dem Einsatz solcher Techniken als technische Maßnahme zum Datenschutz gilt es jedoch abzuwägen, ob das damit erreichbare Schutzniveau den Aufwand rechtfertigt, oder ob sich auf die Zweckbestimmung aufbauende Schutzinstrumente wie die Zweckbindung mit organisatorischen Maßnahmen einfacher realisieren lassen.

These 6.2
Technische Maßnahmen, wie beispielsweise Methoden aus dem Bereich Usage Control, können Schutzinstrumente für die Risikokontrolle durch den Nutzer, die wie etwa die Zweckbindung auf der Zweckbestimmung aufbauen, für den Datenverarbeiter skalierbar sicherstellen und sollten auch aus diesem Grund erforscht werden.

7. Rechte des Betroffenen

Im Kontext von Smart Data wird der Analyse von Daten und der Generierung von neuem Wissen ein großes wirtschaftliches Potential zugemessen. Dadurch entstanden bereits spezialisierte Dienstleister, an die eine Datenverarbeitung ausgelagert werden kann. Haben die Betroffenen keine Kenntnis über die Identität dieses Dienstleisters oder die von ihm durchgeführten Verarbeitungsvorgänge, wird die Ausübung von Nutzer- und Kontrollrechten (z. B. §§ 33-35 BDSG) an praktischen Hürden scheitern. Dieses Problem unterliegt einer weiteren Eskalation, wenn die personenbezogenen Daten durch den Dritten an weitere Akteure weitergeleitet werden. Weitergehend besteht das

Problem, wenn anonymisierte, aggregierte Daten beim Dritt- oder Viertverwender durch Kombination mit anderen Daten plötzlich wieder Personenbezug erhalten. Sind die Weiterleitungsketten für Durchschnittsmenschen (insbesondere technische und/oder rechtliche Laien) kaum mehr nachvollziehbar, wird es technischer Mechanismen bedürfen, um die Gewährleistung von Transparenz und die Einhaltung der Nutzerrechte und Kontrollrechte in einer praktikablen Form zu ermöglichen, allen voran die Nachvollziehbarkeit der Datenströme zu gewährleisten. Auch hier besteht Forschungsbedarf.

Da die bisherigen gesetzlichen Rahmenbedingungen davon ausgehen, dass die die Daten ursprünglich erhebende (verantwortliche) Stelle als Kontaktperson des Betroffenen und Adressat der rechtlichen Verpflichtung zur Umsetzung von Nutzer- und Kontrollrechten zur Verfügung steht, stellt sich grundsätzlich die Frage, ob eine Neukonzeption erforderlich wird. Ist es auch für die ursprünglich verantwortliche Stelle nicht mehr nachvollziehbar oder kontrollierbar, wie die Daten weitergegeben werden und ob (insbesondere bei fehlender Legitimation) Daten wirksam anonymisiert werden, stößt das Schutzkonzept an seine Grenzen. Neben den ordnungsrechtlichen Mitteln der Aufsichtsbehörden gilt es die Rechte des Betroffenen zu stärken, indem sowohl die ursprünglich verantwortliche Stelle, als auch alle nachfolgenden Datenverarbeiter Aufzeichnungs- und Auskunftspflichten zu erfüllen haben. Hier bieten sich verschiedene Konzeptionen zur Erforschung an: Denkbar ist eine Art Meldesystem für die Weitergabe von Daten durch das, ähnlich einem Indossament auf einem Wechsel, unmittelbare Haftungsansprüche auf Auskunft, Unterlassung, Löschung, Schadensersatz, Gewinnabschöpfung usw. gegen jeden einzelnen der Indossanten begründet würden. Ein derartiges System ermöglichte auch die Überprüfung und ggf. Ahndung von Verstößen durch die Aufsichtsbehörden. Zum anderen wäre ein dem Produkthaftungsrecht nachempfundenes System denkbar. Diese Ansätze sollten in der kommenden Forschung untersucht werden.

These 7
Im Hinblick auf eine Erneuerung des Rechtsrahmens auf europäischer Ebene sollten Regelungen verstärkt auch bei datenverarbeitenden „Dritten“ ansetzen.

Fazit und Ausblick

Diese ersten Impulse für eine interdisziplinäre Evaluation bereits bekannter aber auch noch anvisierter Phänomene von „Smart Data“ und den datenschutzrechtlichen Schutzprinzipien zeigt einen umfangreichen Diskussions- und Forschungsbedarf. Die hier nur thesenhaft und teils noch verkürzt dargestellten Konfliktpunkte sollen Schwerpunkte der kommenden Arbeit der Begleitforschung „Smart Data – Innovationen aus Daten“ bilden. Neben der Analyse des aktuellen Rechtsrahmens wird es eine Auseinandersetzung mit den Phänomenen von Smart Data und den sich in der aktuellen rechtspolitischen Diskussion befindlichen Forderungen der Auflockerung einzelner Schutzprinzipien oder der Einführung eines „data ownership“ unvermeidbar machen, den Blick auch rechtsfortbildend auf Möglichkeiten einer grundlegenden Datenschutzreform zu lenken.

Das Spannungsverhältnis zwischen Smart Data und Datenschutz aufzulösen, wird eine zentrale Herausforderung der aktuellen Forschung bilden. Schon heute zeigt sich, dass bereits in Zwei-Personen-Verhältnissen aus Betroffenem und verantwortlicher Stelle u. a. die folgenden Kollisionslagen zu befürchten sind:

- Das Prinzip der Datensparsamkeit kann mit dem Anspruch von Smart Data in Konflikt geraten, wenn Erkenntnisse gerade aus großen Datenbeständen erlangt werden sollen.
- Die De-Personalisierung an der Datenquelle, steht Fällen entgegen, in denen Daten gerade mit Personenzug analysiert werden sollen.
- Bei der regelmäßigen Einbeziehung vielfältiger Sensorik und Datenquellen im privaten Umfeld ist zu befürchten, dass die Informiertheit des Betroffenen als Grundlage der Einwilligung nicht mehr ausreichend gewährleistet werden kann, wenngleich die Einwilligung als Legitimationstatbestand auch im europäischen Kontext weiterhin eine zentrale Rolle spielt.
- Informationen zu den Datenverwendungen und Transparenz zu Datenverwendern und Zwecken als Grundlage einer selbstbestimmten Entscheidung liegen z. T. zum Zeitpunkt der Datenerhebung noch nicht vor.
- Zwischen explorativen Analysen, deren Zweck erst nach der Auswertung feststeht, und dem Prinzip der Zweckbindung besteht ein Spannungsverhältnis.

Die Erwartungshaltung, dass gerade dedizierte Datenverarbeitung und -Analyse Impulse für neue Märkte mit Wissensdienstleistern generiert, führt zur Etablierung neuer Rollen und Marktakteure. In Mehrpersonenverhältnissen ist zu befürchten, dass sich die oben genannten Konflikte noch verstärken werden:

- Die Herausforderung in Mehrpersonenverhältnissen besteht in der Gewährleistung von Rechten sowohl der Betroffenen als auch der vorangegangenen verarbeitenden Stellen (u. a. auf Schutz personenbezogener Daten, Urheberrechten, etc.).
- Oftmals dürfte dem Betroffenen schon das „Ob“ einer (weiteren) Datenverarbeitung beim Dritten nicht bekannt sein. Informationen zum „Wie“ unterliegen z. T. als Betriebs- / Geschäftsgeheimnisse ebenfalls rechtlichem Schutz.
- Einer automatisierten Wissensgenerierung durch Korrelation sollte ein menschliches Korrektiv hinzugefügt werden, um zwischen Zufall und Kausalität zu unterscheiden und negative (insbesondere diskriminierende) Folgen zu vermeiden.
- Schließlich sind gesetzliche Legitimationstatbestände im Hinblick auf die Neuartigkeit der Geschäftsmodelle nur in geringer Zahl vorhanden oder bestehen nur in generalisierender Form.

Gleichzeitig ist anzumerken, dass selbst bei Verzicht auf die Verarbeitung personenbezogener Daten Wissen gewonnen werden kann, welches wiederum zu Entscheidungen führen kann, die ungünstig oder diskriminierend für einzelne Betroffene oder Gruppen von Betroffenen sind. Selbst wenn das Datenschutz-

recht in vielen Fällen nicht anwendbar ist, bleibt zu berücksichtigen, dass es sich um Eingriffe in die Rechte der Betroffenen handeln kann. Daher gelten die Forderungen nach Transparenz und Intervenierbarkeit auch bei solchen Datenverarbeitungen in ähnlicher Weise.

Vor dem Hintergrund besteht eine zentrale Herausforderung darin mit Hilfe eines durchdachten Zusammenwirkens von verschiedenen Gestaltungs- und Schutzmechanismen rechtlich zulässige und zugleich gesellschaftlich akzeptable Smart Data Einsätze zu erreichen. Die Methoden des „Privacy Preserving Data Mining“ können in diesem Zusammenwirken eine Komponente darstellen. Eine Anpassung des primär ordnungsrechtlich wirkenden Rechtsrahmens zur Lösung der vermeintlich unversöhnlich gegenüberstehender Prinzipien von Smart Data und Datenschutz allein erscheint nicht erfolgversprechend. Die rechtliche Stärkung neuer Mechanismen des technischen Datenschutzes ist als wesentliche Innovation z. B. in den Art. 23, 30 und 33 des Entwurfs einer Datenschutzgrundverordnung angelegt.

Durch die immer stärkere digitale Vernetzung auch persönlichster Lebensbereiche des Menschen ist ein umfassender Datenschutz notwendig. Dieser kann jedoch durch technische und organisatorische Maßnahmen, die allein in der Verantwortung von Marktakteuren stehen, nicht geboten werden. Insoweit bleibt der Gesetzgeber gefordert, den ordnungsrechtlichen Rahmen auszugestalten und vorhandene Instrumentarien auch aus Gründen der Klarheit und Rechtssicherheit an die technische Entwicklung anzupassen und nötigenfalls auszubauen.

Mitwirkung „Smart-Data – Smart Privacy?“

Fachgruppenmitglieder

Sebastian Bretthauer
Johann Wolfgang Goethe-Universität Frankfurt am Main, Smart Regio

Achim Klein
Universität Hohenheim, InnOPlan

Wolfgang Putz
Fraunhofer-Institut für Experimentelles Software Engineering (IESE), Pro Opt

Dr. Detlef Runde
Fraunhofer Heinrich Hertz Institut, sd kama

Prof. Dr. Indra Spiecker gen. Döhmann, LL.M. (Georgetown Univ.)
Johann Wolfgang Goethe-Universität Frankfurt am Main, Smart Regio

Jan Schallaböck
iRights.Law Rechtsanwälte, Smart Data Web

Prof. Dr. Beatrix Weber, MLE
Hochschule für Angewandte Wissenschaften Hof, Leiterin der Arbeitsgruppe „Daten als Wirtschaftsgut“

Gäste

Dr. Thilo Weichert
ehem. Datenschutzbeauftragter des Landes Schleswig-Holstein

PD Dr. habil. Feiyu Xu
DFKI Deutsches Forschungszentrum für Künstliche Intelligenz, SD4M

Mitwirkende

Dr. Jens Eckhardt
JUCONOMY Rechtsanwälte

Max von Grafenstein
Alexander von Humboldt Institut für Internet und Gesellschaft (HIIG)

Dr. Marc Hilber LL.M. (Illinois)
OPPENHOFF & PARTNER

Matthias Huber
FZI Forschungszentrum Informatik

Dr. Alexander Lenk
FZI Forschungszentrum Informatik, Leiter der Smart-Data-Begleitforschung

Prof. Dr. Jörn Müller-Quade
Karlsruher Institut für Technologie (KIT) / FZI Forschungszentrum Informatik Leiter der Fachgruppe Sicherheit

Dr. Oliver Raabe
Karlsruher Institut für Technologie (KIT) / FZI Forschungszentrum Informatik, Leiter der Fachgruppe Rechtsrahmen

Jochen Rill
FZI Forschungszentrum Informatik

Andreas Schlisske
Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein (ULD), iTesa

Manuela Wagner
Karlsruher Institut für Technologie (KIT)