

Hochschule für Angewandte Wissenschaften  
Fakultät Life Sciences Studiengang Health  
Sciences

“Big Data” im deutschen Gesundheitswesen -  
Potenziale der Verwendung personenbezogener  
Daten und Beurteilung der Anwendbarkeit in  
Deutschland

## **Bachelorarbeit**

**Vorgelegt von:**

*Julien Kuhlmann*

*1981317*

*Bertuender Prüfer: Prof. Dr. York Zöllner*

*zweiter Prüfer: Prof. Dr. Boris Tolg*

Tag der Abgabe: 05.08.2013

## **Anmerkung**

Zur besseren Lesbarkeit wurde das generische Maskulinum verwendet, falls nicht anders vermerkt sind immer beide Geschlechter gemeint.

## Zusammenfassung

Durch die fortschreitende Digitalisierung der Welt, fallen jeden Tag große Mengen an Daten, verschiedenster Quellen an. Diese bieten großes Potenzial, teilweise werden Daten als das Öl der Zukunft bezeichnet. Die Nutzung dieser Datenmengen für entscheidungsrelevante Ergebnisse wird als „Big Data“ bezeichnet. Hierbei werden Daten aus verschiedensten Quellen miteinander kombiniert und in nahezu Echtzeit ausgewertet. Experten erhoffen sich neue Möglichkeiten durch „Big Data“ in allen Branchen. Auch der Gesundheitssektor wird als wichtiger Markt für „Big Data“ betrachtet. Allein in den USA sollen Kostenersparnisse von bis zu 450 Mrd US\$ jährlich möglich sein. Doch auch zur Untersuchung von Krankheiten, Anwendung personalisierter Medizin und für politische Entscheidungen können Daten genutzt werden. Bisherige Untersuchungen befassen sich hauptsächlich mit dem amerikanischen Markt. Eine Bewertung der Anwendbarkeit im deutschen Gesundheitswesen wurde bisher nicht durchgeführt. Diese Bachelorthesis untersucht die rechtliche Lage, die Verfügbarkeit von personenbezogenen medizinischen Daten und die Haltungen wichtiger Akteure des deutschen Gesundheitssystems zum Thema „Big Data“.

Um dies leisten zu können wurde eine Literaturrecherche durchgeführt und verschiedene Akteure befragt. Anschließend wurden Handlungsempfehlungen genannt um „Big Data“ im deutschen Gesundheitswesen nutzen zu können.

Es konnte in Erfahrung gebracht werden, das „Big Data“ in Deutschland bisher keine großangelegte Anwendung findet. Die nötigen Daten liegen jedoch heute oft bereits in digitaler Form und gut kodiert bei Gesundheitseinrichtungen vor. Eine Infrastruktur zum Austausch dieser Daten und eine zentrale Stelle an der die Daten gesammelt werden fehlen jedoch. Die geplante Einführung der elektronischen Gesundheitskarte und der Telematikinfrastruktur sind erste Schritte zur Nutzung von „Big Data“. Bislang ist jedoch nicht abzusehen wann dieses Projekt fertig gestellt wird und in welchem Umfang die Daten gespeichert und genutzt werden können. Das Vertrauen dem Staat gegenüber zur Handhabung mit den persönlichen Daten der Bürger ist in letzter Zeit stark gesunken, womit eine Sammlung von medizinischen Daten durch diesen derzeit auf wenig Akzeptanz stoßen würde. Der Datenschutz in Deutschland regelt sehr klar welche Daten wofür genutzt werden dürfen und schränkt vieles ein. Die Ärzteschaft zeigt eine tendenzielle Abneigung gegenüber IT und dem Austausch von Daten, während die Leistungsträger auf eine schnelle Verbreitung drängen.

# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>7</b>
<b>2. Zielstellung</b>	<b>8</b>
<b>3. Methode</b>	<b>8</b>
<b>4. „Big Data“</b>	<b>9</b>
4.1 Einsatzmöglichkeiten von „Big Data“ außerhalb des Gesundheitswesens	12
4.2 Einsatzmöglichkeiten von „Big Data“ im Gesundheitswesen	14
<b>5. Anwendbarkeit von „Big Data“ im deutschen Gesundheitswesen</b>	<b>20</b>
5.1 Rechtliche Grundlage zur Nutzung personenbezogener Daten	20
5.2 Verfügbarkeit relevanter Daten	24
5.2.1 Nutzerverhalten	24
5.2.2 Medizinische Daten	25
5.2.3 Administrative Daten	30
5.2.4 Daten der Pharmaindustrie	31
5.2.5 Open Data	32
<b>6. Standpunkte relevanter Akteure zur Datenweiterverarbeitung</b>	<b>33</b>
6.1 Leistungserbringer	35
6.2 Leistungsträger	35
6.3 Versicherte	36
<b>7. Ergebnisse</b>	<b>37</b>
<b>8. Notwendige Entwicklungen</b>	<b>39</b>
8.1 Rechtliche Rahmenbedingungen	39
8.2 Infrastruktur	41
8.3 Akzeptanz	42
<b>9. Diskussion</b>	<b>43</b>
<b>10. Fazit</b>	<b>45</b>
<b>Literaturverzeichnis</b>	<b>47</b>
<b>Eidesstattliche Erklärung</b>	<b>54</b>
<b>Anhang</b>	<b>55</b>

## Abbildungsverzeichnis

Abbildung 1	Die drei Dimensionen von „Big Data“	10
Abbildung 2	Theoretischer Aufbau einer „Big Data“ Struktur	12
Abbildung 3	Datenquellen für „Big Data“ im Gesundheitswesen	15

## Abkürzungsverzeichnis

BDSG	Bundesdatenschutzgesetz
CDC	Center for Disease Control and Prevention
DaTraV	Datentransparenzverordnung
DRG	Diagnosis Related Group
eGK	Elektronische Gesundheitskarte
ePA	Elektronische Patientenakte
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
iGeL	Individuelle Gesundheitsleistung
IFG	Informationsfreiheitsgesetz
GKV(en)	Gesetzliche Krankenversicherung(en)
HBA	Heilberufsausweis
KBV	Kassenärztliche Bundesvereinigung
OTC- Medikament	Over-the-Counter Medikament
SGB V	Sozialgesetzbuch fünftes Buch

## 1. Einleitung

Unternehmen, Wissenschaft und Staat verfügen über mehr Daten denn je. Jede Sekunde werden allein im Internet mehr Daten versendet, als dieses vor 20 Jahren fassen konnte. Ständig fallen Daten an, entweder durch Menschen erstellt oder Maschinen generiert. Durch immer günstiger werdende Speichermöglichkeiten und Computerchips sind zunehmend mehr Institutionen in der Lage, Daten zu speichern und auszuwerten. Die Nutzung dieser Datenmengen für entscheidungsrelevante Ergebnisse wird als „Big Data“ bezeichnet.

Laut eines Berichts des Havard Business Review wurde im Jahre 2012 jeden Tag weltweit ungefähr 2,5 Exabyte an Daten produziert und jede Sekunde mehr Daten im Internet versendet, als noch vor 20 Jahren im gesamten Internet gespeichert waren (Brynjolfsson & McAfee, 2012).

In der IT- Branche ist der Begriff „Big Data“ einer der größten Trends der kommenden Jahre. Die Datenmengen die tagtäglich produziert werden beflügeln die Phantasien einer ganzen Branche. Die Möglichkeiten der Auswertung für Forschung und Entwicklung scheinen unbegrenzt. Doch was genau verbirgt sich hinter dem Begriff Big Data? Und welche Relevanz hat dieser für Gesundheitssysteme?

Die Unternehmensberatung McKinsey&Company hat berechnet, dass durch die konsequente Nutzung von „Big Data“ im amerikanischen Gesundheitswesen jährlich 300-450 Milliarden US\$ eingespart werden könnten. Das entspricht 12-17% der derzeitigen Ausgaben. Es sind also enorme Einsparpotentiale im Gesundheitssystem möglich. Doch auch für innovative Behandlungsformen und Entscheidungsträger bieten sich Einsatzmöglichkeiten. Durch große Datenbanken und schnelle Analysen von Laborproben kann ein erster Schritt zur personalisierten Medizin gegangen werden. Entscheidungsträger können die Tragweite ihrer Beschlüsse schneller evaluieren und öffentliche Gesundheitsbehörden können große Datenmengen nutzen, um präzisere Vorhersagen über die Verbreitung von Krankheiten zu treffen.

Die bisher vorliegenden Veröffentlichungen zu dem Thema befassen sich fast ausschließlich mit dem amerikanischen Markt und der Anwendbarkeit in diesem Gesundheitswesen. Diese Arbeit hat den Anspruch einen ersten Überblick zu schaffen in welchem Maße „Big Data“ im deutschen Gesundheitswesen anwendbar ist und anschließend Handlungsempfehlungen zu geben welche Bedingungen geändert werden müssen.

## 2. Zielstellung

Gegenstand dieser Bachelorarbeit soll es sein, Potentiale der Nutzung von „Big Data“ im Gesundheitsbereich aufzuzeigen und die Anwendbarkeit in Deutschland zu bewerten. Demnach: Ist „Big Data“ im deutschen Gesundheitswesen möglich und welche Schritte sollten unternommen werden, um „Big Data“ (besser) nutzen zu können?

Um sich der Beantwortung der gestellten Frage zu nähern, wird im ersten Abschnitt zunächst versucht, den Begriff „Big Data“ zu definieren und durch Anwendungsbeispiele zu verbildlichen.

Anschließend wird in einem zweiten Abschnitt die Anwendbarkeit von „Big Data“ im deutschen Gesundheitssystem untersucht. Hierzu werden zunächst rechtliche Rahmenbedingungen zum Umgang mit personenbezogenen Daten aufgezeigt, um dann eine Bestandsaufnahme über die Verfügbarkeit und Beschaffenheit relevanter Daten folgen zu lassen. Weiterhin werden die Positionen wichtiger deutscher Akteure des Gesundheitssystems einbezogen. Anschließend werden die Ergebnisse noch einmal zusammen getragen.

Auf Grundlage der Ergebnisse werden im dritten Abschnitt Handlungsempfehlungen für die zukünftige Nutzung von Daten in Deutschland in Bezug auf Recht, Infrastruktur und Akzeptanz gegeben.

Nach der Aufarbeitung der relevanten Informationen erfolgt im vierten und letzten Abschnitt eine Diskussion, in der Limitationen der Ergebnisse erörtert werden. In einem abschließenden Fazit wird die Anwendbarkeit von „Big Data“ im deutschen Gesundheitssystem bewertet.

## 3. Methode

Um das oben genannte Ziel beantworten zu können, wurde eine Literaturrecherche durchgeführt. Primär wurde die Suchmaschine Google und Google Scholar genutzt, zusätzlich wurde der Campus Katalog der HAW Hamburg und Springer Link durchsucht. Die Suche fand sowohl in englischer als auch deutscher Sprache statt. Neben wissenschaftlichen Veröffentlichungen wurde auch auf graue Literatur zurückgegriffen. Die Sozialgesetzbücher, das Informationsfreiheitsgesetz und das Bundesdatenschutzgesetz dienen als Grundlage für die Untersuchung der rechtlichen Aspekte von „Big Data“.



Positionspapiere und Pressemitteilungen verschiedener Verbände wurden untersucht. Weiterhin wurden ausgewählte Verbände angeschrieben und gebeten zu dem Thema „Big Data“ und dem Teilen von medizinischen Daten Stellung zu nehmen. Um die Positionen der Leistungsträger und Leistungserbringer besser beschreiben zu können wurden verschiedene Akteure des deutschen Gesundheitssystems gebeten, einige Fragen zu dem Thema „Big Data“ zu beantworten (siehe Anhang). Folgende Institutionen wurden angeschrieben: Marburger Bund, Verband der privaten Krankenversicherungen e.V., Boehringer-Ingelheim GmbH, GlaxoSmithKline, Kassenärztlicher Vereinigung Bayern, Techniker Krankenkasse, Deutsche Krankenhausgesellschaft, Bundesärztekammer. Bis auf die Bundesärztekammer gaben alle Angeschriebenen eine Rückmeldung (Antworten siehe Anhang).

Zu dem Thema „Big Data“ liegen derzeit keine veröffentlichten Aussagen deutscher Akteure vor. Als Anhaltspunkt wurden die Positionen der Interessengruppen zum Thema Telematikinfrastruktur und eGK beleuchtet. Dies bietet einen ersten Überblick über die Akzeptanz von IT im Gesundheitsbereich und den Austausch medizinischer Daten.

## 4. „Big Data“

„Big Data“ steht für Datenmengen, die mit bisherigen Techniken nicht mehr effektiv zu handhaben sind. Durch die fortschreitende Digitalisierung der Welt in allen Lebensbereichen, werden mehr Daten denn je generiert. Waren es bisher eher Unternehmen, die ihre eigenen Daten wie Transaktionen erstellten und auswerteten, kommen heute zahlreiche neue Quellen hinzu. Soziale Netzwerke, elektronische Akten, Sensoren in mobilen Endgeräten wie Mobiltelefonen, Unterhaltungselektronik, Kundenkarten uvm.

Der Begriff Big Data wird in der Fachliteratur oft unterschiedlich definiert, wobei es einen Konsens auf eine Differenzierung in drei Eigenschaften gibt, um die Beschaffenheit vorliegender Daten zu beschreiben. Diese werden oft als die drei „Vs“ bezeichnet und stehen für „Volume“ (Volumen), „Velocity“ (Geschwindigkeit) und „Variety“ (Vielfalt). Als Daten werden alle Erzeugnisse bezeichnet, die von Menschen oder Maschinen erstellt werden.

### **„Volume“ (Volumen)**

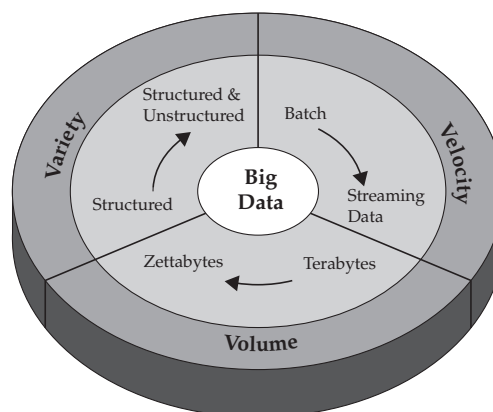
Wie der Name „Big Data“ schon verrät, handelt es sich hierbei um die Handhabung großer Datenmengen. Eine genaue Angabe, ab welcher Menge es sich um „Big Data“

handelt, gibt es jedoch nicht. Dies ist wohl auch darin begründet, dass zunehmend mehr Daten anfallen und heute gesetzte Grenzen bald schon wieder als veraltet gelten werden. Um zu verdeutlichen, wann und wie häufig Daten anfallen sind ein paar Beispiele zu nennen. Sobald Mobiltelefone angeschaltet werden, sich die Türen von Zügen öffnen, man am Flughafen eincheckt, online einkauft oder auch nur den Fernsehsender wechselt, jedes dieser Ereignisse erzeugt Daten.

Die Rechenleistung von Computern hat sich in den letzten Jahren rasant gesteigert. Entsprechend „Moore’s Law“ verdoppelt sich die Anzahl der Transistoren auf Mikrochips ungefähr alle zwei Jahre, womit auch die Leistung der Computer exponentiell wächst (Intel, 2013). Durch günstige Prozessoren und Sensoren werden zunehmend mehr Bereiche des Alltags mit Computerchips ausgestattet.

Zusätzlich zu schnellen Computern, die sich lokal gebunden in den Unternehmen befinden, ist es möglich, Rechenleistung auszulagern und auf Supercomputern enorme Datenmengen analysieren zu lassen (Ohlhorst, 2011).

Einer wachsenden Zahl an Unternehmen ist es nun möglich, große Mengen an Daten gleichzeitig zu analysieren. Die Daten müssen nicht nacheinander und voneinander getrennt in sogenannten „Silos“ gespeichert und bearbeitet werden, sondern können direkt miteinander verknüpft werden (O’Reilly Media, 2012, S. 13). „Silos“ werden hier definiert als getrennte Speicher(-orte) von Daten, die nicht miteinander kombiniert werden.



**Figure 1-1** IBM characterizes Big Data by its volume, velocity, and variety—or simply, V<sup>3</sup>.

**Abbildung 1** Die drei Dimensionen von “Big Data” (Zikopoulos, Eaton, deRoos, Deutsch, & Lapis, 2012)

## **„Velocity“ (Geschwindigkeit)**

Mit Geschwindigkeit sind zwei Dinge gemeint, Speicherung und Verarbeitung. Daten werden immer schneller produziert müssen schnell sortiert und in eine sinnvolle Reihenfolge gebracht werden (O'Reilly Media, 2012, S. 14). Häufig müssen Regeln festgelegt werden, welche Daten als relevant angesehen werden und auf welche es möglich ist zu verzichten, da nur begrenzte Ressourcen zur Verfügung stehen. Die verworfenen Daten könnten jedoch einen bisher unerkannten Nutzen haben.

Weiterhin müssen diese Daten möglichst schnell verarbeitet werden, um einen direkten Nutzen aus ihnen ziehen zu können. Dies ist wichtig, weil die Ergebnisse nach der Auswertung bereits veraltet sein können, da die Realität sehr dynamisch ist und sich stetig ändert.

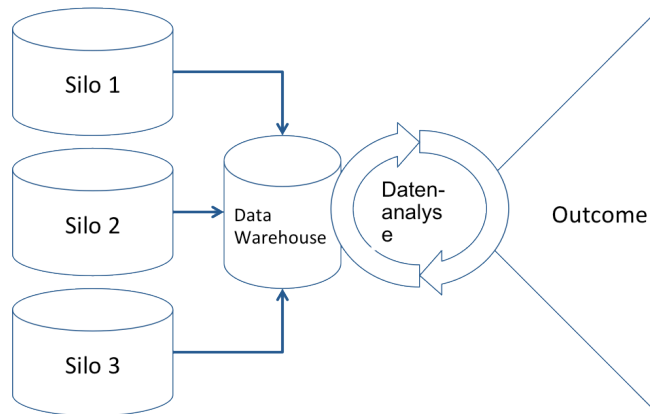
Um den Nutzen einer schnellen Verarbeitung und Analyse zu verdeutlichen, lässt sich das Bild einer Straßenkreuzung einer IBM Werbung verwenden. Dort wird gefragt, wie man die Straße überqueren solle, wenn alles was an Informationen vorliegt, lediglich ein Foto dieser Kreuzung von vor 15 Minuten ist (O'Reilly Media, 2012, S. 14). Es veranschaulicht, dass Entscheidungen bisher häufig auf Informationen beruhen die bereits veraltet sind, da sie nicht in (Nahezu-) Echtzeit betrachtet werden konnten.

## **„Variety“ (Vielfalt)**

Mit Vielfalt sind die Quellen und Formate gemeint, in denen Daten vorliegen. Entweder automatisch generiert oder durch Menschen erstellt. Sie können als Texte, Zahlen, Fotos oder anderen Formaten vorkommen. Daten können roh, strukturiert oder unstrukturiert vorliegen (Zikopoulos, Eaton, deRoos, Deutsch, & Lapis, 2012, S. 7).

Abbildung 2 zeigt den vereinfachten Aufbau einer „Big Data“ Infrastruktur. Wie beschrieben, werden Daten aus verschiedenen Bereichen oder auch verschiedenen Typs bisher in Silos gespeichert und getrennt voneinander analysiert. Bei „Big Data“ werden diese Daten jedoch kombiniert (hier als Data Warehouse bezeichnet). Im nächsten Schritt werden die neukombinierten Daten nun analysiert. Der Output kann ständig mit den ursprünglichen Daten verarbeitet oder verglichen werden, um neue Verhältnisse erkennen und reagieren zu können.

Es zeigt sich, dass „Big Data“ keine festen Dimensionen oder Definition hat, mit denen sich dieser Begriff beschreiben lässt. Es kann als Konzept verstanden werden, um die wachsende Datenmenge zu beschreiben und Ansatz diese zu nutzen.



**Abbildung 2 Theoretischer Aufbau einer „Big Data“ Struktur (The World Economic Forum, 2012)**

#### **4.1 Einsatzmöglichkeiten von „Big Data“ außerhalb des Gesundheitswesens**

Nachdem der Begriff „Big Data“ erklärt wurde, werden im folgenden Abschnitt erste Einsatzmöglichkeiten außerhalb des Gesundheitssektors erläutert. Ziel ist es, ein Verständnis für die Möglichkeiten von „Big Data“ zu entwickeln, um im späteren Verlauf dieses Wissen auf die Potentiale im Gesundheitswesen anwenden zu können.

Viele Institutionen verfügen zwar über große Mengen an Daten, können aus diesen jedoch keinen Nutzen ziehen. Derzeit beantworten Systeme Fragen oft erst ex post und geben Antworten über vergangene Vorgänge wie Verkäufe (O'Reilly Media, 2012, S. 60). Mit Big Data hingegen sollen Echtzeit-Aussagen und Voraussagen möglich sein, um bessere Entscheidungen treffen zu können. Informationen können genutzt werden, um bisherige Prozesse und Vorgänge zu analysieren und zu beschreiben, bei denen es vorher teuer war diese zu untersuchen.

Eine weitere Anwendungsmöglichkeit für Big Data ist die Entwicklung und Ermöglichung neuer Produkte. Durch die Analyse der zahlreichen Nutzerdaten und Interaktionen zwischen den Facebook-Freunden, kann Facebook beispielsweise jedem Nutzer einen eigenen persönlichen Nachrichtenstrom bieten und die angezeigte Werbung zielgenau positionieren (O'Reilly Media, 2012, S. 12). Während Werbung jahrelang Individuen in Zielgruppen einsortierte, kann heute jeder einzeln analysiert werden, um die beste Werbung zu platzieren.

Untersuchungen zeigen, dass Unternehmen, die ihre Entscheidungen heute schon auf Basis von Daten treffen und sich nicht auf andere Faktoren wie Intuition oder Erfahrungen

aus der Vergangenheit verlassen, 5% produktiver und 6% profitabler sind als ihre Wettbewerber (Brynjolfsson & McAfee, 2012, S. 64).

### **Ressourcenoptimierung**

Als Beispiel für eine intelligente Nutzung von „Big Data“ kann das Unternehmen „Otto“ angeführt werden. Das Ziel war es, die Absatzprognosen zu verbessern. Zum einen sollte die Lieferbereitschaft erhöht werden, also die Fähigkeit Waren zu liefern, wenn diese gefragt sind. Zum anderen sollte aber gleichzeitig vermieden werden, dass sich zu viele Waren im Lager befinden und somit die Lagerungskosten steigen würden. Durch die Einführung eines selbstlernenden Systems wurde der Einkauf unterstützt und konnte so bedarfsgerechter Waren bei den Zulieferern ordern. Wöchentlich fließen dafür 300 Millionen (Volumen) Datensätze ein, die Informationen wie Bewerbungsgrad des Artikels (online und offline), Artikeleigenschaften und Umfeldbedingungen beinhalten. Die gesamte Auswertung der Daten erfolgt in Echtzeit. So konnte das Unternehmen die Genauigkeit von Prognosen für einzelne Artikel um bis zu 40% steigern (Urbanski & Weber, 2012, S. 62).

### **Betrugsbekämpfung**

Fraud Detection<sup>1</sup> spielt im Versicherungs- und Kreditwesen eine große Rolle. Allein die weltweiten Kosten von Kreditkartenbetrug beliefen sich 2011 auf ca. 10 Milliarden Euro. Die Herausforderungen in diesem Bereich sind zum einen die hohe Anzahl an Transaktionen, die tagtäglich durchgeführt werden (Volumen). Zum anderen wird viel Aufwand in die Vertuschung dieser Betrügereien gesteckt, womit nicht immer klar ist, wonach gesucht werden muss (Vielfalt). Zusätzlich werden stets neue Strategien entwickelt, um erfolgreich zu betrügen. Eine, im Zusammenarbeit von der PAYMINT AG und dem Fraunhofer IAIS entwickelte, Software erstellt automatisch Regeln und durchsucht die anfallenden Daten nach Betrugsfällen. Die Software ist in der Lage, selbstständig Muster zu erkennen, die auf Betrüge schließen lassen. Durch die große Menge an Daten, auf die die Software zurückgreifen kann, können tausende Merkmale gebildet werden, auf die bei der Bewertung von Transaktionen Bezug genommen werden kann (Urbanski & Weber, 2012, S. 78-79).

### **Öffentliche Angelegenheiten**

Ein weiterer Bereich, der großes Potential für die Nutzung von „Big Data“ bietet, sind öffentlichen Angelegenheiten, die die Bevölkerung betreffen. Gesetzgebungsverfahren,

---

<sup>1</sup> Das Auffinden von Betrugsversuchen.

Stadtplanung, öffentliche Verkehrsmittel, Haushaltsplanung, Infrastruktur etc. (O'Reilly Media, 2012).

Regierungen verfügen über große Mengen von Daten in Form von Zensusdaten, Gesundheitsdaten der Krankenkassen, Steuerdaten und Ausgaben des öffentlichen Sektors (The World Economic Forum, 2012). Das königliche Technologie Institut Stockholm hat zusammen mit IBM ein System entwickelt, um das Verkehrsmanagement in Stockholm zu verbessern. Dafür wurden Verkehrsdaten, Wetterdaten und GPS-Daten analysiert. Es werden pro Sekunde 250.000 GPS-Daten (Volumen und Geschwindigkeit) und Daten aus weiteren Sensoren und Videokameras (Vielfalt), die im Verkehrsnetz eingebunden sind, ausgewertet, um in Echtzeit Entscheidungen treffen zu können. Als Resultat hat sich die Verkehrslage deutlich entspannt: 20% weniger Verkehr, 50% kürzere Fahrzeiten und 20% weniger Emissionen (Urbanski & Weber, 2012, S. 69).

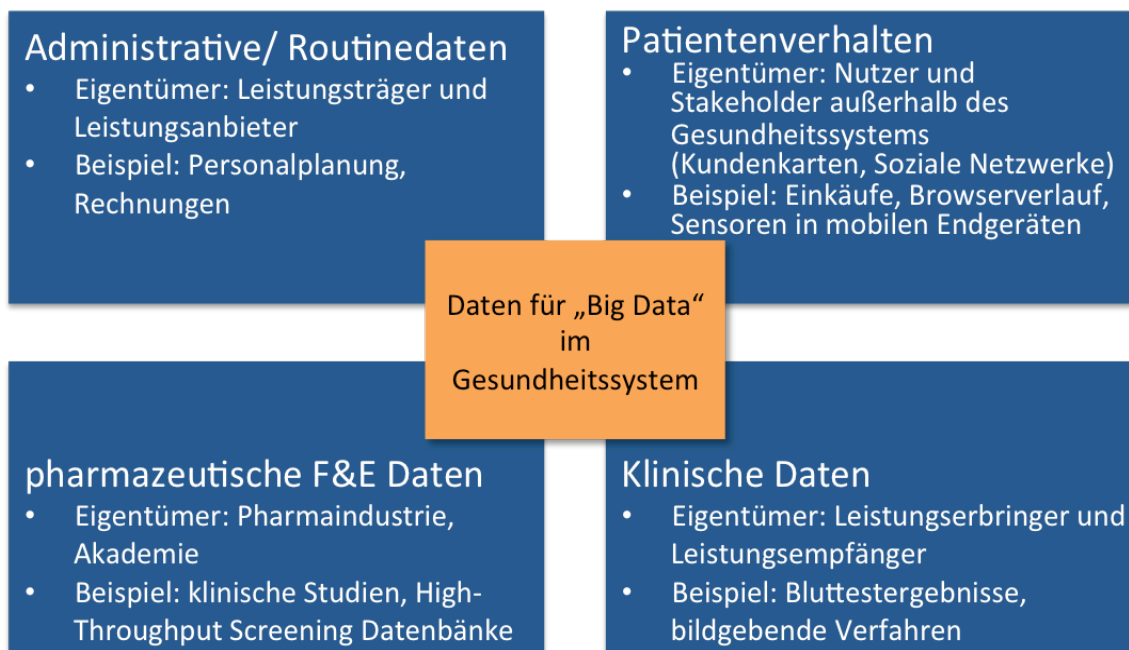
#### **4.2 Einsatzmöglichkeiten von „Big Data“ im Gesundheitswesen**

Nachdem im vorangegangenen Abschnitt Einsatzmöglichkeiten von „Big Data“ in Industrien außerhalb der Gesundheitsbranche erläutert wurden, befasst sich der folgende Abschnitt mit Einsatzmöglichkeiten innerhalb des Gesundheitssystems.

Es werden Potentiale der Nutzung von Daten in Hinblick auf Entscheidungsfindung, Leistungsoptimierung und innovative Behandlungsformen, ohne Anspruch auf Vollständigkeit, aufgezeigt.

Nutzbare Daten im Gesundheitssystem kommen primär aus vier Bereichen: Klinische Daten, Patientenverhalten, pharmazeutische Forschung und administrative/ Routinedaten (siehe Abbildung 3) (Groves, Kayyali, Knott, & Van Kuiken, 2013, S. 4).

Durch die genaue Dokumentation von Befunden und Leistungen im Gesundheitssystem liegen bereits viele Daten vor wie klinische Studien, GKV-Routinedaten und Krankenakten. Durch die zunehmende Digitalisierung im Gesundheitswesen, beispielsweise die elektronische Patientenakte, werden zukünftig weitere Datenquellen hinzu kommen, womit eine leichtere Verwertung möglich ist (O'Reilly Media, 2012). Die elektronische Patientenakte wird eine Schlüsselrolle einnehmen, da dort alle behandlungsrelevanten Daten gespeichert und institutionsübergreifend genutzt werden können. Weiterhin kann sie institutionsintern verwendet werden, um Prozesse zu optimieren oder als Datengrundlage für Entscheidungsunterstützungssysteme. Institutionsübergreifend kann sie von Gesundheitsbehörden und Forschung zur Untersuchung von Krankheitsverläufen, der Gesundheitsverteilung in der Bevölkerung, klinischen Outcomes, und Effektivitätsstudien genutzt werden. Auch als Grundlage für politische Entscheidungen sind die Daten denkbar (Hoffman & Podgurski, 2013).



**Abbildung 3 Datenquellen für “Big Data” im Gesundheitswesen (Groves, Kayyali, Knott, & Van Kuiken, 2013)**

### Entscheidungsfindung

Im Gesundheitswesen werden wie in jeder anderen Industrie ständig Entscheidungen getroffen die vielfältige Auswirkungen haben. Ärzte müssen Diagnosen fällen und die richtige Behandlung für jeden Patienten bestimmen, Leistungsträger entscheiden, welche Leistungen übernommen werden und Gesundheitsbehörden fällen Entscheidungen über die Verhütung von Krankheiten und Gesetzgebungsverfahren.

Entscheidungen werden bisher oft auf Grundlage von Erfahrung und Intuition getroffen, grade Ärzte vertrauen meist eher auf ihr eigenes Urteilsvermögen als auf computerunterstützte Systeme (Groves, Kayyali, Knott, & Van Kuiken, 2013, S. 10). Diese bieten jedoch die Möglichkeit auf standardisierte Pfade rückzugreifen und trotzdem situationsabhängig schnell Entscheidungen treffen zu können.

Bei der Wahl der passenden Medikation für Patienten wird auf die Wirksamkeit aus klinischen Studien und die Erfahrung der behandelnden Mediziner zurückgegriffen. Die Daten aus klinischen Studien sind jedoch Durchschnittswerte der gesamten Testpopulation und bilden Besonderheiten einzelner Patienten unzureichend ab. Die Erfahrung der Leistungserbringer ist auf die Fälle beschränkt, an denen sie beteiligt waren. Wird die Medikamentengabe und der Behandlungsverlauf in einer elektronischen Patientenakte dokumentiert entsteht ein verwertbarer Datensatz. Bei einer Vernetzung der Daten aus elektronischen Patientenakten mehrerer Patienten, können diese in einer Datenbank gesammelt und ausgewertet werden. Die so entstandene Datenbank kann genutzt werden, um Aussagen über die Effektivität von Medikamenten bei Subpopulationen zu treffen. Gleichzeitig können Aussagen zu Wechselwirkungen mit anderen Medikamenten getroffen werden. Es kann gezielt aufgezeigt werden, bei welchen Patienten die Medikamentengabe erfolgreich sein wird und bei welchen nicht (O'Reilly Media, 2012, S. 85-86). Dies sind erste Schritte zur personalisierten Medizin bei der einzelnen Patienten oder Patientengruppen, die ähnliche Merkmale teilen, spezifisch für sie geeignete Therapien erhalten.

Bereits heute stellen Clinical Decision Support Systems<sup>2</sup> eine Möglichkeiten zur Verbesserung der Behandlung von Patienten dar. Die Leistungserbringer werden bei der Entscheidungsfindung durch ein computergestütztes System unterstützt. Das System vergleicht die Eingaben der behandelnden Person mit bereits hinterlegten Daten aus der elektronischen Patientenakte, Medikamentendatenbanken und Literatur, um Behandlungsvorschläge zu bieten (Manyika, et al., 2011, S. 44). Studienergebnisse zeigen bereits, dass durch Verwendung dieser Systeme unerwünschte Nebenwirkungen durch Medikamentengabe signifikant abnehmen können (Potts, Barr, Gregory, Wright, & Patel, 2004). Durch Erweiterung dieser Systeme um die Erkenntnisse aus der zuvor beschriebenen Arzneimitteldatenbank, können die Ergebnisse verbessert werden.

Bei Vorhersagen über die Verbreitung von Infektionskrankheiten verlassen sich Gesundheitsbehörden auf Meldungen von Ärzten an zentrale Institutionen, die für die Auswertung der Daten zuständig sind. Diese Systeme sind häufig noch papierbasiert und

---

<sup>2</sup> Entscheidungsunterstützungssysteme.



über mehrere Behörden verteilt. Dies führt zu einer zeitlichen Verzögerung (Ginsberg, Mohebbi, Patel, Brammer, Smolinski, & Brilliant, 2009). Zeitlich nahe Reaktionen auf Ausbrüche von Krankheiten können Gesundheitsbehörden jedoch helfen, schneller Ressourcen bereitzustellen. Als Beispiel für eine Nutzung großer Datenmengen zur Identifikation der Verteilung von Krankheiten lässt sich „Google Flu“ nennen. Das Unternehmen verglich die historischen Suchanfragen in den USA aus seiner Suchmaschine aus den Jahren 2003-2008 mit den öffentlich zugänglichen Daten des CDC U.S. Influenza Sentinel Provider Surveillance Networks<sup>3</sup>. Die Daten des CDC beruhen auf positiven Laborbefunden aus 50 Bundesstaaten zu Typ und Subtyp des gefundenen Influenzaerregers. Das Ergebnis der Analyse beider Datensätze waren 45 Suchbegriffe, die am stärksten mit den Daten des CDC korrelieren. Wenn die Gesamtheit der „Google“-Nutzer also vermehrt nach diesen 45 Begriffen sucht, lassen sich daraus Muster erkennen, die auf grippeähnliche Symptome schließen lassen. Dadurch, dass zusätzlich die Standorte der Nutzer grob bekannt sind (durch die IP-Adresse), lässt sich auch eine geografische Verteilung erkennen. Als Resultat konnte „Google“ die Verbreitung von Influenza ähnlichen Symptomen<sup>4</sup> ein bis zwei Wochen vor den offiziellen Behörden berechnen (Ginsberg, Mohebbi, Patel, Brammer, Smolinski, & Brilliant, 2009). Das System wurde inzwischen auf 29 Länder ausgeweitet (Butler, 2013). Systeme wie dieses können Gesundheitsbehörden helfen, rechtzeitig Maßnahmen zur Bekämpfung übertragbarer Krankheiten zu ergreifen. Das Unternehmen selbst beschreibt das System nicht als Ersatz für bisherige Surveillancesysteme, deutet jedoch auf die Möglichkeiten hin, die Ergebnisse als Anhaltspunkt zu betrachten um bei vermuteten Ausbrüchen schneller reagieren zu können (Ginsberg, Mohebbi, Patel, Brammer, Smolinski, & Brilliant, 2009). Wie die Grippesaison 2012/2013 zeigte, ist das System noch anfällig für Störungen. „Google Flu“ hat die Größe des Ausbruchs um mehr als das doppelte der wahren Werte überschätzt. Als Grund wird die erhöhte Berichterstattung über die damalige Grippe angeführt, was vermutlich zu einem größeren Interesse der Internetnutzer zu diesem Thema führte und somit die Suchergebnisse beeinflusste (Butler, 2013).

Weitere Möglichkeiten zum Verständnis von Ausbreitungen von Krankheiten bieten Mobiltelefone. Caroline Buckee erforschte die Verbreitung von Malaria (Buckee, Wesolowski, Eagle, & Snow, 2013). In ihrer Studie nutzte sie die Bewegungsprofile der Mobiltelefone, die den Mobilfunkanbietern vorliegen. Die Position jedes Mobiltelefons

---

<sup>3</sup> CDC: Center for Disease Control and Prevention. Behörde in den USA zum Zweck des Schutzes der öffentlichen Gesundheit. <http://www.cdc.gov/flu/weekly/>.

<sup>4</sup> Influenza like Illness (ILI): Klinische Symptome, die auf eine Grippe hinweisen können. Diese werden anschließend mit den Laborwerten einer Stichprobe verglichen (Butler, 2013).

kann bestimmt werden, indem beobachtet wird, in welcher Funkzelle es sich befindet. Erstmals war es möglich, mit einfachen Mitteln die realen Bewegungsmuster ganzer Populationen darzustellen. Dies war bisher nur durch Hochrechnungen und Modelle annäherungsweise machbar. Die Erkenntnisse können in Verbindung mit Malariainfektionskarten genutzt werden, um Ansteckungswege zu erkennen und effektive Bekämpfungsmaßnahmen zu ergreifen (Talbot, 2013). So kann beispielsweise genau analysiert werden, welche Gewässer mit Insektiziden behandelt werden sollten, um den größtmöglichen Nutzen zu erzielen. Außerdem kann Personen, die sich einem Gebiet mit erhöhter Ansteckungsgefahr befinden, eine SMS geschickt werden, um sie darauf hinzuweisen Schutzmaßnahmen zu ergreifen.

### **Leistungsoptimierung**

Wie in dem Abschnitt „Entscheidungsfindung“ beschrieben, können durch die Nutzung von Patientendaten gezielter Therapien begonnen werden. Zusätzlich zu dem Nutzen für den Patienten, können die Ressourcen der Leistungserbringer, wie Zeit und Material, effektiver genutzt werden. Prozessdaten und Abrechnungsdaten in medizinischen Einrichtungen können analysiert und in Echtzeit auf Dashboards<sup>5</sup> dargestellt werden (Manyika, et al., 2011, S. 45). Die Bewegungen eines Patienten innerhalb einer Institution können durch moderne Krankenhausinformationssysteme bereits verfolgt werden. In Kombination mit ERP<sup>6</sup> Systemen, Prozessdaten für Arbeitsabläufe und prospektiven Vergütungsformen wie DRG-Systemen, können die Kosten pro Patient in Echtzeit betrachtet werden. Ziel ist es, Engpässe zu identifizieren und Prozesse zu optimieren (Manyika, et al., 2011, S. 45).

Auch die gesundheitsökonomische Evaluation von Behandlungen kann mit Hilfe von „Big Data“ durchgeführt werden. Dabei werden die Kosten dem Outcome alternativer Therapien und Leistungen gegenüber gestellt. Dies kann entweder auf Basis von Primärdaten, also Daten aus Studien, oder durch die Analyse von Sekundärdaten (administrativen Daten) geschehen (Schreyögg & Stargardt, 2012, S. 1). Letzteres bietet eine höhere Stichprobengröße, da potentiell alle erbrachten Leistungen erfasst werden. Weiterhin können Behandlungswege durch Überweisungen und Krankheitsverläufe durch Arbeitsunfähigkeitsbescheinigungen verfolgt werden (Schreyögg & Stargardt, 2012, S. 1-2). In Deutschland eignen sich dafür besonders die GKV-Routinedaten. Potentiell wäre auch eine Verwendung der Daten aus elektronischen Patientenakten denkbar, da hier weitere Merkmale des Patienten erfasst werden, die nicht an den Leistungsträger

---

<sup>5</sup> Visualisierung von Informationen.

<sup>6</sup> Enterprise Resource Planning (dt. Unternehmensressourcenplanung).

übermittelt werden. Dies wäre ein Wandel hin zu „realitätsnahe[n] Studie[n] unter Alltagsbedingungen“, wie sie bereits in einigen Ländern gefordert werden (Schreyögg & Stargardt, 2012, S. 6). Die gesundheitsökonomische Evaluation bietet großes Potential für Kostenbeschränkung im Gesundheitswesen, da Behandlungen aufgrund ihres Outcomes beurteilt werden können.

Auch Abrechnungsbetrug von Leistungserbringern im Gesundheitssystem ließe sich bekämpfen, wie am Beispiel der Fraud Detection bereits beschrieben wurde.

### **Innovative Behandlungsformen**

Therapietreue gilt als wichtiger Faktor auf dem Weg zu einer erfolgreichen Behandlung von Patienten. Um die Therapietreue von Patienten zu erhöhen, können verschiedene Kommunikationswege genutzt werden. Vorstellbar sind SMS Erinnerungen, Gruppentreffen oder Chats mit medizinischem Personal. Durch Auswertung des Patientenverhaltens können Profile von Menschen erstellt werden. Das Unternehmen „Eliza“<sup>7</sup> stellt diese Profile her, um herauszufinden, welche Kommunikationsform bei welchen Patienten am besten wirkt (O'Reilly Media, 2012). Ähnlich wie in der Werbeindustrie kann so zielgerichtet an Personen herantreten werden. Mit Hilfe dieses Wissens, kann die Therapietreue der Patienten erhöht werden und Komplikationen oder Wiedereinweisungen können verringert werden.

Wie bereits beschrieben, ist „Big Data“ geeignet, um große Datenmengen zu analysieren und Muster zu erkennen. Dies kann genutzt werden, um Verhaltensänderungen bei Patienten zu erfassen, die auf geringe Therapietreue oder Symptome hinweisen. Das Unternehmen „Ginger.io“ stellt Gesundheitsinstitutionen Software und den Patienten eine App für Smartphones zur Verfügung. Die App läuft im Hintergrund auf dem Smartphone des Patienten und erfasst Metadaten wie Standorte, Telefonverbindungen und Textnachrichten. Dazu wird das „normale“ Verhalten des Patienten zunächst analysiert und im Weiteren stetig untersucht ob es Abweichungen von diesem „Normalzustand“ gibt, die auf Nichtbefolgen der Therapieanweisungen hinweisen. Diese Daten werden an die behandelnden Personen weitergeleitet und ermöglichen diesen, Einblick in der Verhalten der Patienten zu erhalten. Die behandelnde Person kann darüber hinaus reagieren und den Patienten aktiv kontaktieren (Simonite, 2013).

Auch bei der akuten Versorgung von Patienten kann „Big Data“ helfen. Als Beispiel sei die Versorgung von Frühgeborenen genannt. An der University of Ontario wurde eine

---

<sup>7</sup> <http://www.elizacorporation.com/>.

Software eingeführt, die tausende Sensordaten pro Sekunde von 16 unterschiedlichen Sensoren analysiert, die das Frühchen überwachen. Bisher werden diese Parameter durch medizinisches Personal bewertet. Durch die neue Software können kritische Ereignisse die die Gesundheit gefährden könnten, bis zu 24 Stunden vor Eintritt von sichtbaren Symptomen erkannt werden. Dadurch kann beispielsweise bei Infektionen bereits frühzeitig eine gezielte Medikation beginnen, die schwächer dosiert werden muss (Urbanski & Weber, 2012, S. 70). Zusätzlich wurde entdeckt, dass vor einer Infektion die Vitalparameter oft besonders regelmäßig sind (Mayer-Schönberger & Cukier, 2013, S. 60). Dies wurde erstmals festgestellt, womit nun neue Behandlungspfade entwickelt werden können, um die Frühgeborenen möglichst schonend zu behandeln.

## **5. Anwendbarkeit von „Big Data“ im deutschen Gesundheitswesen**

Im Folgenden wird der Status Quo beschrieben, wie Daten in Deutschland genutzt werden können. Zunächst wird die Gesetzeslage untersucht, wie (personenbezogene) Daten in Deutschland genutzt werden dürfen. Anschließend erfolgt eine Untersuchung welche Daten in welcher Form vorliegen. Es folgt ein Überblick über Aufbau und Zuständigkeiten im deutschen Gesundheitssystem, um abschließend die Standpunkte von Schlüsselakteuren des deutschen Gesundheitssystems in diesem Kontext zu betrachten.

### **5.1 Rechtliche Grundlage zur Nutzung personenbezogener Daten**

Der folgende Abschnitt beschäftigt sich mit den rechtlichen Vorgaben bezüglich der Nutzung personenbezogener Daten und den geltenden Datenschutzbestimmungen. Weiterhin wird auf die Rechtslage zum Thema Telematik im Gesundheitswesen eingegangen, um die Möglichkeiten einer elektronischen Patientenakten und der elektronischen Gesundheitskarte zu erläutern, die große Bedeutung für das Thema „Big Data“ haben. Da bei „Big Data“ auch Daten außerhalb des Gesundheitswesens verwendet werden können, wird auch der allgemeine Datenschutz in Deutschland betrachtet sowie Grundlagen für den Zugang zu öffentlichen, nicht-personenbezogenen Daten.

## **Datenschutz personenbezogener Daten**

Der „Schutz personenbezogener Daten“ ist in der „Charta der Grundrechte der Europäischen Union“ Artikel 8 festgeschrieben. Auch das Bundesverfassungsgericht entschied 1983 das jeder Mensch das „Recht auf informelle Selbstbestimmung“ hat (vgl. BVerfGE 65,1) (Bundesministerium des Inneren, 2013). Danach hat „jede Person das Recht auf Schutz der sie betreffenden personenbezogenen Daten“. Diese Daten dürfen nur nach Einwilligung der betroffenen Person oder auf Grundlage gesetzlicher Legitimation verwendet werden. Jede Person hat das Recht auf Auskunft über die erhobenen Daten und diese zu berichtigen.

Der allgemeine Umgang mit personenbezogenen Daten in Deutschland wird im Bundesdatenschutzgesetz (BDSG) geregelt. Das BDSG ist den bereichsspezifischen Datenschutzrechten nachrangig. Insgesamt gibt es in Deutschland 39 Spezialgesetze die den Datenschutz in einzelnen Bereichen präzisieren (Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit). Spezielle Gesetze für den Umgang mit Daten im Gesundheitssystem befinden sich im SGB V Kapitel 10 „Versicherungs- und Leistungsdaten, Datenschutz, Datentransparenz“.

Nach BDSG §3a ist die Erhebung, Verarbeitung und Nutzung personenbezogener Daten und Wahl der verwendeten Datenverarbeitungssysteme so zu bestimmen, dass möglichst wenig personenbezogene Daten erhoben werden (Datenvermeidung und Datensparsamkeit). Ferner sind die Daten zu anonymisieren oder zu pseudonymisieren, soweit dies nach dem Verwendungszweck möglich ist. Nach der Anonymisierung fallen die Daten nicht mehr unter das BDSG.

Datenerhebung, Verarbeitung und Nutzung sind nach BDSG §4 Absatz 1 nur zulässig, wenn diese durch Gesetze oder Vorschriften erlaubt sind und der Betroffene schriftlich eingewilligt hat (Datenverarbeitungsverbot mit Erlaubnisvorbehalt). Für die wissenschaftlich Forschung können Ausnahmen gelten. Die personenbezogenen Daten müssen direkt bei den Betroffenen erhoben werden, dies schließt die Verwendung der Daten durch Dritte aus (BDSG §4). Ausnahmen bilden jeweils Rechtsvorschriften oder der Verwaltungszweck erfordert die Erhebung der Daten bei Dritten. Erhobene personenbezogene Daten dürfen nach BDSG §14 nur für die Zwecke gespeichert, verändert oder genutzt werden, für die sie erhoben wurden. Für die wissenschaftliche Forschung können auch hier Ausnahmen gelten.

Verfahren automatisierter Verarbeitung sind entweder den zuständigen Aufsichtsbehörden oder dem Bundesbeauftragten für den Datenschutz und

Informationsfreiheit nennen (BDSG §4d Absatz 1). Automatisierte Verarbeitung sind alle Prozesse, bei denen kein Mensch aktiv eingreifen muss, um sie durchzuführen.

Entscheidungen, die für die betroffene Person eine rechtliche Folge nach sich ziehen, dürfen nach BDSG §6a nicht ausschließlich auf die automatische Verarbeitung von personenbezogenen Daten gestützt werden.. Dies betrifft im wesentlichen auch „Big Data“, da somit medizinische Entscheidungen, die Patienten betreffen, immer einer Kontrolle durch befähigte Personen unterliegen.

Das Erstellen von Profilen und die Weitergabe für werbemäßige Zwecke durch nicht-öffentliche Stellen ist erlaubt, solange der Betroffene die verantwortliche Stelle klar erkennen kann. Ferner ist das Erheben besonderer Arten personenbezogener Daten<sup>8</sup> durch nicht-öffentliche Stellen zulässig, wenn der Betroffene eingewilligt hat. Der Betroffene muss nach §28 Absatz 7 nicht gesondert einwilligen, wenn die Verwendung der Daten für die Gesundheitsvorsorge, medizinische Diagnostik, Gesundheitsversorgung oder -behandlung oder für die Verwaltung von Gesundheitsdiensten erforderlich ist und die Verarbeitung dieser Daten durch medizinisches Personal mit entsprechender Geheimhaltungspflicht erfolgt.

Wenn personenbezogene Daten für Forschungszwecke erhoben werden, dürfen diese nur zu diesem Zweck verwendet werden und sind zu anonymisieren (BDSG §40).

### **Datenschutz im Gesundheitswesen nach SGB V**

Personenbezogene Daten werden in den Sozialgesetzbüchern auch als Sozialdaten bezeichnet. SGB X §67 präzisiert, dass diese Daten von einer in SGB I §35 genannten Stelle erhoben, verarbeitet und genutzt werden dürfen. Darunter fallen Sozialleistungsträger und deren Verbände. Allgemein gesprochen dürfen Krankenkassen nach SGB V §284 Sozialdaten nur erheben und speichern, wenn diese für administrative Zwecke, Qualitätssicherung, Überprüfung der Rechtmäßigkeit und Wirtschaftlichkeit der erbrachten Leistungen der Leistungserbringer, die Berechnung des Risikostrukturausgleichs und Risikopools und der Identifizierung potentieller Nutzer von Programmen nach SGB V §137g (Disease-Management-Programme) genutzt werden. Es soll verhindert werden, dass durch das unkontrollierte Erfassen und Auswerten sensibler personenbezogener Daten, und insbesondere die Datenzusammenführung, "gläserne Menschen/ Versicherte" entstehen oder die Erstellung von bestimmten "(Gesundheits-) Profilen" ermöglicht wird.

---

<sup>8</sup> Gesundheitsdaten, Informationen über die rassische oder ethnische Herkunft, politische, religiöse, gewerkschaftliche oder sexuelle Orientierung.

SGB V § 287 ist von entscheidender Bedeutung für die Verwendung personenbezogener bzw. Sozialdaten, da dieser regelt, dass Krankenkassen und Kassenärztliche Vereinigungen mit Erlaubnis der Aufsichtsbehörden die Datenbestände für zeitlich befristete und im Umfang begrenzte Forschungsvorhaben verwenden dürfen. Als vorrangige Ziele werden epidemiologische Erkenntnisse, Erkenntnisse über Zusammenhänge zwischen Krankheit und Arbeitsbedingungen oder Erkenntnisse über örtliche Krankheitsschwerpunkte genannt. Die Sozialdaten sind dabei zu anonymisieren.

Die auf der elektronischen Gesundheitskarte gespeicherten Daten dürfen nur durch in SGB V §291a Absatz 4 genannte Personen<sup>9</sup> verändert, gespeichert oder verwendet werden. Das vorherige Einverständnis der Betroffenen ist einzuholen, wobei dieser vorher umfassend und allgemeinverständlich über Funktionsweise und -umfang informiert werden muss. Der Zugriff auf die Daten darf nur in Verbindung mit einem elektronischen Heilberufsausweis erfolgen, der über die Möglichkeit einer sicheren Authentifizierung und eine qualifizierte Signatur<sup>10</sup> verfügt.

Zum Zweck der Qualitätssicherung sind alle an der vertragsärztlichen Versorgung teilnehmenden Ärzte, zugelassene Krankenhäuser und sämtliche andere Leistungserbringer nach SGB §135a Absatz 2 und Krankenkassen verpflichtet, personen- oder einrichtungsbezogene Daten der Versicherten und Leistungserbringer zu erheben, zu verarbeiten und zu nutzen. Die Richtlinien hierfür beschließt der Gemeinsame Bundesausschuss (SGB V §299). Die erhobenen Daten dürfen in der Regel nur stichprobenartig verwendet werden, müssen pseudonymisiert und von einer unabhängigen Stelle ausgewertet werden (mit Ausnahme der Kassenärztlichen Vereinigungen).

Das Bundesversichertenamt übermittelt der Datenaufbereitungsstelle (SGB V §303a) jährlich die nach SGB V §268 stichprobenartig erhobenen Daten zur Umsetzung des Risikostrukturausgleichs. In SGB V §303e ist geregelt, welche Institutionen auf die bei der Datenaufbereitungsstelle vorliegenden Daten zugreifen dürfen. Dies sind 18 Institutionen, welche ausschließlich aus dem Gesundheitswesen kommen. Weiterhin ist in diesem Paragraphen geregelt, zu welchem Zweck die Daten verwendet werden dürfen. In der

---

<sup>9</sup> Ärzte, Zahnärzte, Apotheker, Apothekerassistenten, Pharmazieingenieure, Apothekerassistenten, Personen, die bei den zuvor genannten oder in einem Krankenhaus tätig sind und berufsmäßige Gehilfen oder zur Vorbereitung auf den Beruf tätig sind und der Zugriff unter Aufsicht der zuvor genannten Personen erfolgt.

<sup>10</sup> SigG §2, eine fortgeschrittene elektronische Signatur die zum Zeitpunkt ihrer Erzeugung auf einem qualifizierten Zertifikat beruht und mit einer sicheren Signatureinheit erzeugt wurde. Diese können ausschließlich dem Signaturschlüsselinhaber zugeordnet werden und ermöglichen eine Erkennung nachträglicher Veränderungen.

„Verordnung zur Umsetzung der Vorschriften über die Datentransparenz“ (DaTraV) ist geregelt, dass die der Datenaufbereitungsstelle vorliegenden Daten nur pseudonymisiert an die anfragenden Institutionen weitergegeben werden dürfen. Weiterhin muss nach DaTraV §5 der Verwendungszweck vorher angegeben werden und die Daten dürfen nur mit Daten aus externen Beständen kombiniert werden, wenn „die betroffenen Versicherten nicht wieder identifiziert werden können.“

## 5.2 Verfügbarkeit relevanter Daten

Der folgende Abschnitt untersucht, welche Daten in Deutschland verfügbar sind und in welcher Form. Es werden die in Abbildung 3 genannten Quellen untersucht und zusätzlich der Bereich Open Data in Deutschland betrachtet.

### 5.2.1 Nutzerverhalten

Daten über das Nutzerverhalten außerhalb des Gesundheitssystems sind einer der Hauptquellen für die Nutzung von „Big Data“. Diese ergeben sich aus der Nutzung von Internetdiensten, Kundenkarten, Verbindungsdaten, Teilnahmen an Umfragen oder Gewinnspielen und weiteren Angeboten.

Alleine durch die Nutzung des Internets wird eine breite Datenspur hinterlassen. Es werden Cookies auf den Rechnern der Nutzer gespeichert, um diese später wieder zu erkennen. Diese werden dann mit dem Nutzungsverhalten des Nutzers angereichert: Suchanfragen, angeschaute oder bestellte Artikel in Online-Shops oder das Klick-Verhalten der Nutzer werden gespeichert. So können gezielt Profile über Websitegrenzen hinweg erstellt werden.

Im Jahr 2012 nutzten in Deutschland 77% der Personen über 10 Jahren das Internet täglich oder fast täglich. Besonders Personen über 65 Jahre nutzen das Internet kaum, nur 24% der Frauen und 44% der Männer in diesem Altersbereich nutzen das Internet, bei den Personen zwischen 45 und 64 Jahren nutzen 82% der Männer und 78% der Frauen das Internet. Bei den Personen unter 45 Jahren sind nahezu alle online (10-24 Jahre: 97%, 25-44 Jahre: 96%) (Statistisches Bundesamt, 2012). 16% der deutschen Haushalte ohne Internetzugang geben an, dieses nicht zu nutzen, da sie Bedenken hinsichtlich des Datenschutzes und der Privatsphäre haben. 81% der befragten Personen gaben an, innerhalb der letzten zwölf Monate Online-Shops genutzt zu haben (Statistisches Bundesamt, 2012).



Nach letzten Zahlen des statistischen Bundesamtes nutzen 2011 51% der deutschen Internetnutzer soziale Netzwerke (Statistisches Bundesamt, 2012). Es ist anzunehmen, dass sich dieser Anteil in den letzten zwei Jahren erhöht hat, da alleine Facebook von Juli 2011 (20,11 Mio.) bis Juni 2013 (26 Mio.) ca. 30% mehr Nutzer in Deutschland zählt (allfacebook.de, 2013).

Auch durch die Nutzung von Kundenkarten und Bonusprogrammen werden Daten über die Konsumenten gesammelt und Profile über das Nutzungsverhalten, inkl. Konsumverhalten und Bewegungsprofilen, können erstellt werden. In Deutschland nahmen 2011 16,59 Mio. Personen an mindestens einem und 12,73 Mio. Menschen an mehreren dieser Programme teil (Institut für Demoskopie Allensbach, 2011).

Ebenso lässt sich aus der Nutzung von Mobiltelefonen eine Vielzahl an Daten entnehmen. 2012 verfügten 90,2% der deutschen Haushalte über Handys oder Smartphones und jeder Haushalt verfügt durchschnittlich über 1,6 Mobiltelefone (Statistisches Bundesamt, 2013). Smartphones besitzen oft vielfältige Sensoren, die stetig die Aktivität der Nutzer überwachen.

Derzeit herrscht ein zunehmender Trend zur Selbstvermessung, das heißt, durch mobile Computer die über Sensoren wie GPS, Blutdruckmessgerät oder Mikrofone verfügen, können sich Nutzer ständig selbst messen. Die Vitalparameter oder der Grad der körperlichen Aktivität können jederzeit analysiert werden. Der Fachbegriff für diese Geräte ist „Wearable Computing“, wobei hier nicht nur Innovationen gemeint sind, sondern beispielsweise auch Hörgeräte die es bereits seit Jahren gibt. Diese Daten bieten großes Potenzial zur Analyse, da sie den tatsächlichen Alltag inkl. Vitalparameter auch von Personen abbildet, die nicht aktiv Ärzte aufsuchen.

### **5.2.2 Medizinische Daten**

Wie in Abbildung 3 beschrieben, werden Routinedaten, Daten über das Verhalten des Patienten, Daten aus der pharmazeutischen Forschung und klinische Daten aus Patientenakten als wichtigste Quellen zur effektiven Nutzung von Big Data im Gesundheitswesen genannt. Im Folgenden wird untersucht, inwieweit diese Daten in Deutschland vorliegen und in welcher Form.

## **Austausch medizinischer Daten**

Die elektronische Gesundheitskarte (eGK) und die elektronische Patientenakte (ePA) können einen großen Beitrag zur Nutzung von „Big Data“ im deutschen Gesundheitssystem leisten. Nachfolgend werden diese beiden Systeme beleuchtet.

In Deutschland wird bereits an einer Infrastruktur zur Erhebung, Speicherung und Verarbeitung der im Gesundheitswesen anfallenden Daten gearbeitet. Der Gesetzgeber beauftragt den Spitzenverbund der Krankenkassen, die Kassenärztliche Bundesvereinigung, die Kassenzahnärztliche Bundesvereinigung, die Bundesärztekammer, die Bundeszahnärztekammer, die Deutsche Krankenhausgesellschaft sowie die Spitzenorganisation der Apotheken damit, eine geeignete Telematikinfrastruktur aufzubauen. Diese gründen für diesen Zweck die Gesellschaft für Telematik, „gematik“ genannt (SGB V §291b). Aufgabe ist die Erstellung eines Sicherheitskonzeptes und die Festlegung der technischen Vorgaben. Inhalt und Struktur der Datensätze für die Einführung der elektronischen Gesundheitskarte und der elektronischen Patientenakte.

Die eGK wird mit dem Ziel eingeführt, den sicheren Austausch medizinischer Daten von Patienten institutionsübergreifend zu ermöglichen. Patientendaten sollen mit Zustimmung des Patienten für die jeweiligen Leistungserbringer zugänglich sein und dadurch eine institutionsübergreifende Patientenakte entstehen.

Die auf der eGK zu speichernden bzw. speicherbaren Daten sind in verpflichtende und freiwillige Daten unterteilt. Verpflichtende Daten sind nach SGB V §291a Absatz 2 ärztliche Verordnungen, also elektronische Rezepte und die europäische Krankenversicherungskarte. Weiterhin müssen alle Versichertenstammdaten, die auch auf der Krankenversicherungskarte nach SGB V §291 hinterlegt sind, gespeichert werden. Die eGK muss darüber hinaus folgende freiwillige Daten erheben, verarbeiten und nutzen können:

- Notfalldaten
- Elektronischer Arztbrief
- Daten zur Prüfung der Arzneimitteltherapiesicherheit
- Elektronische Patientenakte bzw. elektronische Fallakte<sup>11</sup>
- Durch Versicherte selbst zur Verfügung gestellte Daten

---

<sup>11</sup> von der gematik verwendeter Begriff.

- Daten über in Anspruch genommene Leistungen und deren vorläufige Kosten für die Versicherten
- Erklärungen zu Organspendebereitschaft
- Hinweise auf Vorhandensein und Aufbewahrungsorte von Vorsorgevollmachten und Patientenverfügungen

Derzeit (Stand: Juli 2013) hat die eGK keine weiteren verfügbaren Anwendungen als die alte Krankenkassenkarte. Es werden lediglich die Versichertenstammdaten und auf der Rückseite die europäische Krankenversicherungskarte unterstützt (gematik). Diese können bisher jedoch nicht online geprüft oder aktualisiert werden. Für die vorgesehenen Anwendungsmöglichkeiten wird derzeit eine zweite Generation der eGK, G2 genannt, von der gematik entwickelt. Die Online-Aktualisierung und -prüfung der Versichertenstammdaten wird erst nach dem Ende der ersten Stufe zur Einführung der eGK möglich sein. Die gematik beauftragt die Industrie, Angebote zum Aufbau der Telematikinfrastruktur, der Erprobung der Online-Aktualisierung der Versichertenstammdaten und der qualifizierten elektronischen Signatur, abzugeben (gematik, 2013). Für den Aufbau der Telematikinfrastruktur soll zunächst das zentrale Netz (Backbone) geschaffen werden, welches die sichere Anbindung der Krankenkassen und anderer Akteure sowie bereits bestehende Netze der Heilberufler ermöglichen soll. Voraussichtlich werden die Zuschläge Ende 2013 erteilt und anschließend von der gematik getestet. Weiterhin wurde die Industrie aufgefordert, technische und organisatorische Verfahren zur Herausgabe von Heilberufsausweisen (HBA) und Institutionskarten zu entwickeln. Mit diesen authentifizieren sich die Leistungserbringer, um Zugriff auf die auf der eGK gespeicherten Daten zu erhalten. Die letzte Angebotsrunde beinhaltet die Bereitstellung einer Public Key Infrastruktur (PKI) und der „Card Verifiable Certificates“. Diese ermöglichen eine gegenseitige Authentifizierung der eGK und des HBA, auch ohne Internetverbindung (gematik, 2013).

Spätestens zehn Monate nach der Erteilung des Zuschlages - also Ende 2014 - sollen die Systeme im Echtbetrieb in zwei Testregionen erprobt werden (gematik, 2013).

Gegenwärtig liegt kein öffentlicher Zeitplan der „gematik“ vor, aus dem hervorgeht, wann die zweite Stufe der Einführung der eGK begonnen werden soll. Diese beinhaltet das Notfalldatenmanagement, den elektronischen Arztbrief, die Arzneimittelsicherheitsprüfung und die elektronische Patientenakte/Fallakte. Somit ist bisher nicht abzusehen, wann flächendeckend eine institutionsübergreifende elektronische Patientenakte/Fallakte in Deutschland vorliegen wird. Auch nach erfolgreicher Einführung der eGK wird es voraussichtlich keine vollständige elektronische Patientenakte geben, da nach SGB V

§291a Absatz 3 Patienten zustimmen müssen ob und welche Daten auf der Karte gespeichert werden dürfen. Weiterhin dürfen nach Absatz 6 alle Daten die in §3 Absatz 1 beschrieben werden, auf Forderung des Patienten hin gelöscht werden. Somit besteht die Gefahr einer lückenhaften Protokollierung der Patientendaten, was durch die selektierten Daten zu einem verzerrten Bild führt. Ferner wird das Outcome einer Therapie in der Regel nicht erfasst, da Patienten dem Arzt häufig keine Rückmeldung geben, ob die Therapie angeschlagen hat.

Ungeachtet der langsamen Einführung der elektronische Patientenakte (ePA) auf Bundesebene, werden von einigen Leistungserbringern die Patientendaten bereits in elektronischen Patientenakten verarbeitet. Diese sind entweder institutionsinterne Systeme, also „Inhouse“, oder Online-Patientenakten, die institutionsübergreifend, also beispielsweise innerhalb eines Krankenhausverbundes oder der integrierten Versorgung, genutzt werden können (Uslu). In einer Studie aus dem Jahr 2012 (Hübner, Liebe, Egbert, & Frey, 2012, S. 44-46) gaben 22,6% der befragten Krankenhäuser (n=177) an, eine ePA vollfunktionsfähig in allen Einheiten zu nutzen, 10,2% nutzten eine ePA in mindestens einer Einheit ihres Krankenhauses. In den restlichen Krankenhäusern sind sie teilweise funktionsfähig (31,6%) oder in Planung (16,9%). Lediglich 13,1% gaben an, keine Implementierung einer elektronischen Patientenakte zu planen. Als Grund für die zurückhaltenden Investitionen in Systeme zur Nutzung elektronischer Patientenakten in Krankenhäusern kann das bisherige Fehlen von Standards genannt werden (Bundesverband Gesundheits-IT, 2013), welche die Interoperabilität der Systeme gewährleisten.

Fast alle niedergelassenen Ärzte verwenden einfache Computersysteme, um die Patientendaten zu erfassen. Diese sind jedoch oft recht rudimentär und bieten keine Schnittstellen zu anderen Systemen (Accenture, 2012, S. 151). Zusätzlich geben nur 12% an, elektronisch mit anderen Institutionen zu kommunizieren (Accenture, 2012, S. 151). Genannte Hauptgründe von Ärzten für das Nicht-Einführen von elektronischen Patientenakten und die fehlende Kommunikation mit anderen Institutionen sind: Datenschutzbedenken, Bedenken über Interoperabilität mit anderen Systemen, fehlendes Vertrauen der Institutionen untereinander, Kosten und die Sorge, dass die Datenerfassung zu viel Zeit in Anspruch nehmen würde. Ärzte nutzen IT in ihrem Klinikalltag erst, wenn sie eine spürbare Effizienzsteigerung feststellen (Accenture, 2012, S. 161).

## Prozessdaten

Da „Big Data“ auch Institutionen unterstützen kann Entscheidungen zu treffen und Prozesse zu optimieren, müssen Daten für diesen Zweck nicht mit anderen Institutionen ausgetauscht werden. Besonders wichtig ist hier eine möglichst genaue Erfassung der Abläufe innerhalb der Institution. In Krankenhäusern werden Krankenhausinformationssysteme (KIS) zur Erfassung, Bearbeitung und Weitergabe medizinischer und administrativer Aufgaben genutzt. Diese können entweder elektronisch oder aber auch in papierbasierter Form geführt werden.

In einer Studie zum Thema „IT im Krankenhaus“ gaben 22,5% der befragten Krankenhäuser (n=173) an, durch Nutzung betriebswirtschaftlicher Informationssysteme Prozesse besser abbilden und verbessern zu wollen. 16,8% wollen ein Data Warehouse / Business Intelligence System aufbauen (Hübner, Liebe, Egbert, & Frey, 2012, S. 56). Systeme zur Entscheidungsunterstützung werden bisher lediglich in ca. jedem zehnten Krankenhaus genutzt, in etwa einem Drittel der befragten Krankenhäuser ist eine Einführung nicht geplant. IT-Systeme zur Arzneimittelgabe (8,2%), Produktidentifikation (3,3%), Ortidentifikation der Patienten und des Personals (3,3%) verwenden nur wenige Krankenhäuser. Die Bestandsveränderungen können somit nicht den einzelnen Patienten zugeordnet werden. IT-Systeme zur Materialwirtschaft (75%), Essensbestellung (69,4%) und Apotheke (63,3%) werden jedoch genutzt (Hübner, Liebe, Egbert, & Frey, 2012, S. 31-38). Die Daten werden somit nur aggregiert erfasst, da keine Einzelerfassung der Prozesse stattfindet. Allerdings verwenden 73% der Krankenhäuser Kommunikationsserver, um die anfallenden Daten der verschiedenen Abteilungen und Systeme zusammenzufassen (Hübner, Liebe, Egbert, & Frey, 2012, S. 41).

Die innerhalb der Institutionen gesammelten Daten werden in Deutschland überwiegend nur für interne Prozesse weiterverwendet und nicht mit anderen Institutionen oder Gesundheitsbehörden geteilt. Der Anteil der Ärzte, die angeben ihre Daten mit anderen Institutionen und Gesundheitsbehörden zu teilen, liegt in Deutschland mit 28% unter dem internationalen Durchschnitt von 44% (Accenture, 2012, S. 34). Die erfassten Daten werden größtenteils unstrukturiert in die Systeme eingepflegt, lediglich ein Viertel der befragten Ärzte kodieren die erfassten Daten (Accenture, 2012, S. 155). Dies ist jedoch wichtig für eine Analyse und Weiterverarbeitung der Daten.

### 5.2.3 Administrative Daten

Die Gesetzlichen Krankenversicherungen (GKV) verfügen in Deutschland über große Mengen an Daten über die dort versicherte Bevölkerung, die es erlauben Patienten sektorübergreifend im Gesundheitssystem zu verfolgen. Diese werden Routinedaten genannt. Allein im Datawarehouse der Techniker Krankenkasse befinden sich ca. 5000 GB an Routinedaten (Horenkamp-Sonntag & Verheyen, 2012, S. 7). Es wird die gesamte GKV-Population erfasst, womit es keine Selektion und keinen Drop-Out gibt und die Daten vollständig dokumentiert sind.

In SGB V zehntes Kapitel ist festgelegt, dass alle administrativen Daten von den beteiligten Akteuren „im Wege elektronischer Datenübertragung oder maschinell verwertbar auf Datenträgern“ übermittelt werden müssen<sup>12</sup>. Dies ist wichtig, um vorhandene Daten analysieren zu können, ohne dass sie zunächst händisch übertragen werden müssen. Fehlerquellen durch Medienbrüche werden so reduziert.

Die den Krankenkassen vorliegenden Informationen dienen der Erfüllung der in SGB V §284 genannten Aufgaben. Folgende personenbezogene Daten liegen den gesetzlichen Krankenversicherungen vor (Schubert, 2012, S. 4):

- Versicherten Stammdaten: Wohnsitz, Alter, Geschlecht, Name, Versicherten-ID, Daten zum Leistungsträger, Laufzeit des Versichertenschutzes, Versichertengruppe und Art der Beschäftigung
- Ambulante Diagnosen und Leistungen
- Verordnungen
- Stationäre Aufenthalte
- Arbeitsunfähigkeiten
- Sachleistungen
- Leistung „Pflege“
- Patientendaten aus Disease-Management-Programmen
- Arzneimittelversorgung

Sämtliche Daten werden dem einzelnen Versicherten zugeordnet und pseudonymisiert. Durch das Gesetz zur Modernisierung der gesetzlichen Krankenversicherung im Jahr 2003, wurde jedem Versicherten der GKV eine eindeutig zuordenbare, lebenslang gültige Versichertennummer zugeteilt. Bei einem Wechsel der Krankenversicherung wird

---

<sup>12</sup> SGB V §295 Abrechnung ärztlicher Leistungen, §300 Abrechnung der Apotheken und weiterer Stellen, §301 Krankenhäuser, §301a Abrechnung der Hebammen und der von ihnen geleiteten Einrichtungen, §302 Abrechnung der sonstigen Leistungserbringer.

diese übernommen und alle dazu gespeicherten Daten an die neue Versicherung übertragen. Eine Verwechslung von Patientendaten oder eine Dopplungen von Patienten in Systemen durch die Vergabe unterschiedlicher Identifikationsnummern werden damit reduziert.

Im stationären Bereich ist die Kodierungsqualität der Daten seit der Einführung des DRG-Systems sehr hoch, was eine Analyse erleichtert. Außerdem werden die Daten sehr schnell an die GKVEn übertragen. Allerdings werden keine klinischen Parameter wie Laborwerte übermittelt. Ebenso bei den Daten aus der ambulanten Versorgung, diese werden zudem kumuliert (durch die Kassenärztlichen Vereinigungen) übertragen und sind frühestens neun Monate nach Quartalsende verfügbar (Horenkamp-Sonntag & Verheyen, 2012, S. 9). Ein weiterer Mangel ist die Dokumentation des Outcomes einer Intervention, da Patienten zum einen den Arzt häufig nicht wieder aufsuchen, um eine erfolgreiche Behandlung zu melden und zum anderen diese Daten nicht an die Krankenkassen gesendet würden.

Freiwillige Inanspruchnahme von Zusatzleistungen wie „IGeL“<sup>13</sup> und der Kauf bzw. die Nutzung von OTC-Medikamenten<sup>14</sup> wie bspw. ASS werden auch nicht erfasst (Horenkamp-Sonntag & Verheyen, 2012, S. 14).

#### 5.2.4 Daten der Pharmaindustrie

Derzeit werden die Ergebnisse aus Pharmastudien meist aus Wettbewerbsgründen unter Verschluss gehalten und nicht publiziert. Die „European Medical Agency“ fordert jedoch die proaktive Veröffentlichung der Pharmastudien, nachdem die Beurteilung über die Zulassung eines neuen Medikamentes abgeschlossen ist. Die Daten sollen in drei Kategorien unterteilt werden, welche jeweils andere Veröffentlichungsrichtlinien aufweisen (European Medicines Agency, 2013). Kategorie 1, „kommerziell vertrauliche<sup>15</sup> Daten“, beinhaltet sensible Daten, die Details über den Wirkstoff und einige bioanalytische Eigenschaften Auskunft geben, diese müssen nicht veröffentlicht werden. Kategorie 2, „öffentlicher Zugang“, beinhaltet alle Daten der Forschungsstudien, die keine personenbezogenen Daten enthalten und sind öffentlich zugänglich. Kategorie 3, „überwachter Zugriff“, beinhaltet personenbezogene Daten der teilnehmenden Patienten. Diese sollen vor Veröffentlichung anonymisiert werden. Die Daten werden nur auf Anfrage hin freigegeben und der Antragsteller muss bestimmte Kriterien erfüllen.

---

<sup>13</sup> individuelle Gesundheitsleistungen.

<sup>14</sup> Over-the-Counter, also nicht verschreibungspflichtige aber apothekenpflichtige Arzneimittel.

<sup>15</sup> commercially confident.

Ein weiterer Vorschlag des europäischen Ausschusses für Umweltfragen, öffentliche Gesundheit und Lebensmittelsicherheit zur Reform des bisherigen Umgangs mit klinischen Studien enthält die Idee, dass klinische Studien allgemein nicht als „kommerziell vertraulich“ eingestuft werden sollten, sobald der Entscheidungsprozess über die Zulassung abgeschlossen ist (European Medicines Agency, 2013). Somit könnten sie vollständig veröffentlicht werden.

### 5.2.5 Open Data

Da für „Big Data“ nicht nur personenbezogene Daten verwendet werden, sondern auch öffentlich zugängliche Daten, ist die Verfügbarkeit dieser Daten von Bedeutung. Auf Bundesebene ermöglicht das Informationsfreiheitsgesetz (IFG) jeder natürlichen und juristischen Person in Deutschland, Zugang zu amtlichen Informationen. Behörden haben Verzeichnisse über die vorhandenen Daten zu führen (IFG §11). Die Form in der Daten vorliegen müssen, d.h. ob als Rohdaten oder aufbereitet, ist nicht definiert (Klessmann, Denker, Schieferdecker, & Schulz, 2012, S. 127). Es dürfen nach IFG §10 Gebühren für die Auskunft und Bereitstellung von Daten erhoben werden. Zusätzlich zu dem IFG auf Bundesebene gibt es Informationsfreiheitsgesetze in elf weiteren Bundesländern<sup>16</sup> (Klessmann, Denker, Schieferdecker, & Schulz, 2012, S. 105-106). Derzeit ist eine proaktive Bereitstellung von Daten öffentlicher Stellen, d.h. die Daten werden publiziert, selten und der Zugang zu den Daten muss beantragt werden (Klessmann, Denker, Schieferdecker, & Schulz, 2012, S. 127).

Als Begriff für den freien Zugriff auf öffentliche Daten der Regierung und Verwaltung wird „Open Government Data“ genutzt. Eine Definition des IT-Planungsrates ist „die Bereitstellung von Daten der öffentlichen Hand zur Nutzung, insbesondere durch Weiterverwendung und Weiterverbreitung“ (IT-Planungsrat, 2012). Laut Richtlinie 2003/98/EG wird „Weiterverwendung“ definiert als „die Nutzung von Dokumenten, die im Besitz öffentlicher Stellen sind, durch natürliche oder juristische Personen für kommerzielle oder nichtkommerzielle Zwecke, die sich von dem ursprünglichen Zweck im Rahmen des öffentlichen Auftrags, für den die Dokumente erstellt wurden, unterscheiden“. Weiterhin wird dort geregelt, dass öffentliche Stellen die vorhandenen Dokumente in allen vorhandenen Formaten und Sprachen und soweit möglich in elektronischer Form zur Verfügung stellen müssen.

---

<sup>16</sup> Berlin, Brandenburg, Bremen, Hamburg, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Rheinland-Pfalz, Saarland, Sachsen-Anhalt, Schleswig-Holstein und Thüringen.



Im Zuge einer Studie im Auftrag des Bundesministeriums für Inneres wurde eine Bestandsanalyse zur Verfügbarkeit existierender Daten öffentlicher Stellen in Deutschland durchgeführt (Klessmann, Denker, Schieferdecker, & Schulz, 2012, S. 389-414). Es wurde eine manuelle (200 Behörden und Gebietskörperschaften aller föderaler Ebenen) und teilautomatisierte (719 Webauftritte von Einrichtungen und Gebietskörperschaften auf Bundes-, Landes- und Kommunalebene) Bestandsaufnahme durchgeführt. Besonders die Daten des Bundesamtes für Statistik erfüllen die Kriterien an offenen Daten<sup>17</sup>, da diese lizenzfrei genutzt werden können und in maschineninterpretierbaren Formaten vorliegen. Als Hauptgrund für die Einordnung als nicht-offene Daten der anderen untersuchten Institutionen, wird die geringe Maschineninterpretierbarkeit, die Verfügbarkeit für lediglich private Zwecke und die Erhebung von Gebühren genannt (Klessmann, Denker, Schieferdecker, & Schulz, 2012, S. 413). Ein Großteil der veröffentlichten Daten (90%) liegt im PDF-Format vor und ist somit nicht maschineninterpretierbar. In eine stichprobenhaften Untersuchung dieser Dateien wurde jedoch festgestellt, dass sie als „strukturierte Informationen vorliegen“, also aus offenen Formaten in PDF umgewandelt wurden (Klessmann, Denker, Schieferdecker, & Schulz, 2012, S. 414). Insgesamt haben die untersuchten öffentlichen Stellen jedoch großes Potenzial, Datensätze nach Open Data Kriterien zur Verfügung zu stellen.

## **6. Standpunkte relevanter Akteure zur Datenweiterverarbeitung**

Nachfolgend wird zunächst der Aufbau und die Zuständigkeiten im deutschen Gesundheitswesen beschrieben. Dies dient dem besseren Verständnis, um Strukturen und Position von Interessengruppen innerhalb des deutschen Gesundheitssystems besser beurteilen zu können.

Das Eingreifen des Staates ins deutsche Gesundheitssystem beschränkt sich weitgehend auf die Rahmensetzung und Rechtsaufsicht. Die Ausführung der Gesetze wird auf die Organe der sogenannten Selbstverwaltung übertragen (Simon, 2010). Diese haben gewisse Entscheidungskompetenzen, die sich aber im Rahmen der gegebenen Gesetze abspielen. Dennoch hat der Staat die letztendliche Entscheidungsmacht. Der Staat ist in Deutschland bedingt durch den Föderalismus, jedoch keine einzelne Institution. Er ist in Bund und Länder unterteilt, wobei beide Gesetzgebungskompetenzen haben.

---

<sup>17</sup> Folgende Kriterien wurden gewählt: Zugang, Maschineninterpretierbarkeit, Auffindbarkeit, Nutzungsbedingungen, Nutzungskosten, Aggregation, Verfügbarkeit, offene Standards, Aktualisierungsrate (vgl. Klessmann, Denker, Schieferdecker, & Schulz, 2012 S. 390-392).

Länderspezifische Regelungen betreffen zum Beispiel das Landeskrankenhausgesetz und die Landespflegegesetze. Der Bundesrat, in dem Vertreter der Länder sitzen, muss in sogenannte Zustimmungsgesetze einwilligen, bevor sie verabschiedet werden können. Auch ein Teil der Gesundheitsgesetzgebung ist zustimmungspflichtig (Simon, 2010, S. 98). In der gemeinsamen Selbstverwaltung, in der Vertreter der GKVen und Leistungserbringer sitzen, werden zentrale Fragen der Leistungserbringung und Vergütung entschieden. Wichtigstes Gremium ist hier der Gemeinsame Bundesausschuss (G-BA). Er trifft Grundsatzentscheidungen über den GKV-Leistungskatalog und Einzelthemen der Versorgung. Die Entscheidungen sind bindend, Patientenvertreter haben in diesem Gremium lediglich eine beratende Tätigkeit und sind nicht abstimmfähig. Auch Verbände haben eine wichtige Rolle bei der Selbstverwaltung, sie sind mittelbare Staatsverwaltung und Körperschaften des öffentlichen Rechts. Ihre Hauptaufgabe ist die Interessenvertretung ihrer Mitglieder. Einer der wichtigsten Verbände sind die Kassenärztlichen Vereinigungen (KVen) und die Kassenärztliche Bundesvereinigung (KBV). Sie haben hoheitliche Staatsaufgaben zu übernehmen, aber gleichzeitig die Interessen ihrer Mitglieder zu wahren. Somit sollen sie auf der einen Seite als Körperschaft öffentlichen Rechts dem Allgemeinwohl dienen, auf der anderen Seite aber Partikularinteressen der Mitglieder vertreten. Sie werden in der Regel in alle wichtigen gesundheitspolitischen Entscheidungen einbezogen (Simon, 2010, S. 199). Die Verbände der Krankenkassen und auf Bundesebene der Spitzenverband Bund unterstützen die GKVen bei der Durchführung ihrer Aufgaben und vertreten ihre Interessen. Dritte wichtige Säule sind die Krankenhausgesellschaften in denen sich die Krankenhäuser freiwillig zusammenschließen. Ihre Aufgaben sind in den Landeskrankenhausgesetzen geregelt und umfassen die Vertragsverhandlungen zwischen GKVen und Krankenhäusern und die Mitwirkung bei der Krankenhausplanung des Landes. Die Interessenvertretungen der Leistungserbringer und Leistungsträger haben also großes Mitspracherecht bei der Ausgestaltung des Gesundheitssystems, wohingegen Patienten lediglich angehört werden.

Nachdem Zuständigkeiten wichtiger Akteure des deutschen Gesundheitssystems erläutert wurden, folgt nun eine Betrachtung relevanter Interessengruppen, die mit Patientendaten agieren. Betrachtet werden hierbei Leistungserbringer in Form von Ärzten und Krankenhäusern, Krankenversicherungen als Leistungsträger und Patienten als Leistungsempfänger und Nachfrager. Ziel ist es, deren Position zum Thema „Nutzung personenbezogener Daten“ hinsichtlich der Beurteilung der Anwendbarkeit von „Big Data“ in Deutschland herauszuarbeiten.

## 6.1 Leistungserbringer

In einer Pressemitteilung vom Mai 2013 spricht sich die Kassenärztliche Bundesvereinigung grundsätzlich für eine Erweiterung des Datenaustausches zwischen Leistungserbringern aus. Sie ist jedoch nicht mit den geplanten Strukturen, wie sie mit der eGK eingeführt werden sollen, einverstanden. Die momentane Planung sei zu kassenzentriert und verlagere Verwaltungsaufgaben in die Arztpraxen. Laut einer Umfrage der Techniker Krankenkasse (TK) begrüßten 2009 jedoch noch drei Viertel der befragten Ärzte die Online-Aktualisierung der Versichertenstammdaten (Gangl & Birkner, 2009). Als Konsequenz will die KBV parallel eigene Strukturen zum Austausch ausbauen (Kassenärztliche Bundesvereinigung, 2013). Gleichzeitig fordert sie die gesamte Infrastruktur in die Hände der ärztlichen Selbstverwaltung zu legen. Der Marburger Bund antwortete auf Anfrage, sie unterstützten die Position der Bundesärztekammer. Es war jedoch kein Mitglied der Bundesärztekammer bis zur Fertigstellung der Arbeit zu erreichen.

In der Studie der TK gaben zudem 86% der befragten Ärzte an, bezüglich der Speicherung von Patientendaten auf zentralen Servern Bedenken zu haben, da dies zu Datenmissbrauch führen könnte. 39% der Ärzte glaubten, dass diese Form der Datenaufbewahrung gegen die ärztliche Schweigepflicht verstößt.

Die deutschen Ärzte sind demnach prinzipiell für elektronische Patientenakten und den Austausch medizinischer Patientendaten, mit den derzeitigen Plänen für den Ausbau dieser Strukturen jedoch unzufrieden.

## 6.2 Leistungsträger

Der Spitzenverband Bund der gesetzlichen Krankenversicherungen drängt auf die Einführung einer Telematikinfrastruktur zum Austausch medizinischer und administrativer Daten im Gesundheitssystem. Er fordert den Gesetzgeber auf „[...] stringente gesetzliche Rahmen- und Organisationsbedingungen für eine nutzenorientierte Telematikstruktur zu schaffen [...]“, da einige Leistungserbringerorganisationen die Ziele immer wieder unterlaufen würden (GKV-Spitzenverband, 2013). Michael Leonhart, Sprecher der AOK Bayern, sieht Datenschutzbedenken der Ärzte und Patienten als übertrieben an. Er weist darauf hin, dass bei Privatversicherten der Arzt alle Befunde und Leistungen erfasst und diese dann an die Versicherungen, also private Unternehmen, geleitet werden. Bei diesem System „kräht kein Hahn [nach Datenschutz]“ (Stecher, 2013).

Der Verband der deutschen Krankenversicherungen e.V. wollte sich auf Anfrage hin nicht zu der eGK und dem Austausch medizinischer Daten äußern, da sie vor einiger Zeit auf dem Projekt der Telematikinfrastruktur ausgetreten sind. Zum Thema „Big Data“ liegt kein Positionspapier der privaten Krankenversicherungen vor.

### 6.3 Versicherte

Wie in Abschnitt „Nutzerverhalten“ gezeigt, nutzen bereits viele Menschen soziale Netzwerke, Mobiltelefone, Geräte zur Selbstvermessung oder Bonusprogramme ganz selbstverständlich und erlauben den Anbietern dieser Dienste in vielen Fällen ihre personenbezogenen Daten zu verwenden. Die Nutzer willigen dazu den Allgemeinen Geschäftsbedingungen der Unternehmen ein. Oft tun sie dies ohne diese gelesen zu haben oder zu verstehen, in welchem Umfang die Daten tatsächlich genutzt werden (UN Global Pulse, 2012). In einer Studie zum Thema Datenmissbrauch aus dem Jahr 2010 geben 79% der befragten Deutschen an, „etwas bis keine“ Sorgen zu haben, dass ihre persönlichen Daten missbraucht werden. Nur 19% hingeben an, „große Angst“ vor Datenmissbrauch zu haben (IfD Allensbach, 2010). Gleichzeitig glaubt aber die Mehrheit der Deutschen, dass es bedenklich ist, eigene Daten an private Unternehmen weiterzugeben (mit Ausnahme von Onlineshops). Bei öffentlichen Einrichtungen wie Krankenkassen, Behörden und der Polizei hingeben findet es die Mehrheit nicht bedenklich, dass diese ihre Daten haben (TNS Emnid, 2011). Es herrschte also ein großes Vertrauen in die öffentlichen Behörden, dass personenbezogene Daten nicht missbräuchlich genutzt werden. Zum Zeitpunkt der Recherche für diese Arbeit, enthüllte der ehemalige NSA-Mitarbeiter Edward Snowden, dass US-Sicherheitsdienste in großem Umfang Daten deutscher Bürger abgreifen und speichern (Poitras, Rosenbach, & Stark, 2013). Ende Juli 2013 führte der Branchenverband Bitkom<sup>18</sup> eine Umfrage zum Vertrauen deutscher Internetnutzer zum Umgang mit persönlichen Daten durch. Die Zahl der Internetnutzer die sich sicher im Internet fühlten sank und die Personen die sich „völlig unsicher“ fühlten stieg von 12% (2011) auf 27% (2013). Auch das Vertrauen in den Staat in Bezug auf den Umgang mit ihren Daten vertrauen weit weniger Menschen. Waren es 2011 noch 52% die dem Staat „sehr stark“ oder „stark“ vertrauten, sind es 2013 nur noch 34% (Kempf, 2013).

Einer Umfrage zur Meinung TK-Versicherter zum Thema Datenaustausch zufolge, wünschen sich 12% der Befragten einen besseren Austausch medizinischer Daten zwischen Ärzten. Viele Versicherte sind der Möglichkeiten zur Nutzung ihrer

---

<sup>18</sup> Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.

medizinischen Daten gegenüber jedoch aufgeschlossen und wünschen sich mehr Informationen zu diesen Möglichkeiten. Die Mehrheit der Befragten glaubt, dass EU-weites Teilen von medizinischen Daten die Qualität der Behandlung steigern könnte (Dobrick, 2012, S. 36).

## 7. Ergebnisse

Es konnte gezeigt werden, dass „Big Data“ bereits vielfältige Einsatzmöglichkeiten in gesundheitsfremden Industrien hat. Auch im Gesundheitswesen gibt es erste Schritte zur Nutzung großer Datenmengen, um valide und schnelle Entscheidungen treffen zu können, Prozesse in Institutionen zu optimieren und die Versorgung von Patienten zu verbessern. In Deutschland gibt es einige Limitationen, die die Verwendung von „Big Data“ im Gesundheitswesen beschränken.

Die sehr weitreichenden Datenschutzbestimmungen erschweren die Nutzung medizinischer Daten. Prinzipiell gilt, dass in Deutschland personenbezogene Daten nur nach vorheriger schriftlicher Einwilligung der Betroffenen erhoben und genutzt werden dürfen. Ausnahmen zu dieser Regelung gelten meist nur für wissenschaftliche Zwecke. Die erhobenen Daten dürfen nur für die in der Einwilligung genannten Zwecke genutzt werden. Die Einverständniserklärungen zur Nutzung dieser Daten muss somit viele Eventualitäten abdecken, um sie für die vielfältigen Einsatzmöglichkeiten von „Big Data“ nutzen zu können. Weiterhin kann es hinderlich sein, dass die Verfahren, mit dem die Daten verarbeitet werden, zu melden sind und diese erst nach vorheriger Meldung geändert werden dürfen. Die Beschreibung der Nutzung darf nicht zu grob sein, es muss für den Betroffenen klar verständlich sein, in welchem Verfahren seine Daten genutzt werden. Eine Möglichkeit diese Beschränkungen zu umgehen, so dass das Datenschutzgesetz keine Anwendung mehr findet, ist die Anonymisierung der Daten. Daten, die im Gesundheitssystem anfallen, werden als besonders schützenswert betrachtet. Dementsprechend streng sind die Regelungen zu Erhebung, Speicherung und Verarbeitung dieser Daten. Neben den Krankenkassen ist es noch weiteren Organisationen rechtlich möglich, mit diesen Daten zu arbeiten. Für private Organisationen besteht derzeit keine Möglichkeit auf diese Daten zuzugreifen, wenn der Patient die Daten nicht aktiv bei den Gesundheitsinstitutionen einfordert und sie an sie weitergibt. Dies bedeutet eine Limitation für den Austausch der Daten und Nutzung für „Big Data“.

Bereits heute sind die Leistungsbringer und –träger jedoch bereits verpflichtet personenbezogene Daten zum Zweck der Qualitätssicherung zu erheben und zu nutzen. Die der Datenaufbereitungsstelle vorliegenden Daten dürfen in einem strengen Rahmen für die Forschung genutzt und sogar mit externen Daten, nach vorheriger Anonymisierung, kombiniert werden.

Der Großteil der deutschen Bürger nutzt regelmäßig aktiv Informations- und Telekommunikationstechnik, Bonusprogramme und soziale Netzwerke. Durch Open Data Bemühungen des Staates ist ein zunehmender Teil der öffentlichen Daten frei zugänglich und zur Weiterverarbeitung freigegeben. Forschungsergebnisse aus Pharmastudien sollen auf Drängen der EU zukünftig zumindest teilweise kostenlos öffentlich zugänglich sein. Durch die Routinedaten der GKVEn gibt es bereits heute eine sehr gut kodierte Datenbasis für die zukünftige Nutzung. Diese dürfen jedoch bisher nur von 18 festgelegten Institutionen genutzt werden. Weiterhin werden die Daten derzeit mit einer zeitlichen Verzögerung von bis zu neun Monaten an die Leistungsträger übermittelt.

Die elektronische Gesundheitskarte und die elektronische Patientenakte sind bisher nur geplant. Die Infrastruktur zum Austausch medizinischer Daten wird derzeit entwickelt, hat aber großes Potential, da sie allen beteiligten Akteure des deutschen Gesundheitssystems offen steht. Zusätzlich könnte sie administrative Vorgänge und den Zugriff auf Daten beschleunigen. Die Blockadehaltung der KBV erschwert jedoch eine rasche Umsetzung.

Fast alle Leistungserbringer verarbeiten ihre Daten mit IT-Systemen. Diese sind jedoch häufig noch sehr rudimentär und bieten keine Schnittstellen zur Kommunikation mit anderen Anbietern. Fehlende Standards, Investitionsstaus in Krankenhäusern und Datenschutzbedenken sind als Gründe zu nennen. Krankenhäuser verwenden im Vergleich zu niedergelassenen Ärzten leistungsfähigerer Systeme. Patientendaten werden nur sehr selten mit anderen Leistungserbringern geteilt. Im internationalen Vergleich hat Deutschland die geringste Quote bezüglich des Austauschs medizinischer Daten.

Die deutsche Ärzteschaft ist prinzipiell offen für den Austausch medizinischer Daten, aber unzufrieden mit der derzeitigen Planung der Telematikinfrastruktur in Deutschland. Sie fordern diese Systeme vollständig in der Hand der Ärzteschaft zu legen. Die Leistungsträger hingegen drängen auf eine schnelle Einführung der nötigen Infrastruktur und fordern die Ärzteschaft auf, die Blockadehaltung aufzugeben. Die Patienten haben seit den NSA-Enthüllungen ein stark verringertes Vertrauen in die Sicherheit ihrer Daten,

besonders dem Staat gegenüber. Vor diesem Ereignis hatte der Großteil nur geringe Sorge um Datenmissbrauch und die Menschen vertrauten staatlichen Institutionen. Nur ein geringer Teil der Patienten will einen besseren Austausch medizinischer Daten, gleichzeitig glaubt aber die Mehrheit, dass der Austausch die Qualität der Versorgung steigern würde.

## **8. Notwendige Entwicklungen**

Auf Grundlage der Ergebnisse der Untersuchung zur Anwendbarkeit von „Big Data“ werden im Folgenden notwendige Entwicklungen beschrieben, um „Big Data“ effektiv in Deutschland nutzen zu können. Es werden Lösungsvorschläge für die beschriebenen Defizite erörtert. Zunächst werden Empfehlungen gegeben, um die rechtlichen Rahmenbedingungen zu schaffen und eine technische Infrastruktur aufzubauen. Nachfolgend werden Möglichkeiten zur Erhöhung der Akzeptanz zum Teilen und zur Weiterverarbeitung medizinischer Daten genannt.

### **8.1 Rechtliche Rahmenbedingungen**

Nachdem bereits aufgezeigt wurde, dass potentiell viele verwertbare Daten vorliegen, gilt es nun das Problem zu lösen, wie diese Daten freigegeben werden und nicht nur innerhalb der Institutionen, sei es privat oder öffentlich, zugänglich sind. Gleichzeitig muss die Privatsphäre der Menschen, aber auch die Möglichkeit des Geschäftsgeheimnisses von Unternehmen gewahrt werden. Die Vereinten Nationen gründeten die Initiative „Global Pulse“, sie setzt sich dafür ein, „Big Data“ für das öffentliche Leben zu nutzen, um Entscheidungen besser treffen und evaluieren zu können. Sie verfolgen den Ansatz der „Data Philanthropy“. Dieser besagt, dass Unternehmen ihre Daten vollkommen anonymisieren und „sozialen Innovatoren“ bereitstellen, damit diese damit arbeiten können (UN Global Pulse, 2012, S. 25).

Die EU verfolgt einen ähnlichen Ansatz in ihrem Report „Redesigning health in Europe for 2020.“ Dort wird als ein wichtiger Baustein für die Zukunft von Gesundheitssystemen die Nutzung von Daten genannt. Die „Befreiung von Daten“ aus unterschiedlichsten Quellen sei entscheidend, um diesen Ziel zu erreichen. Regierungen sollten sicherstellen, dass Gesundheitsdaten robust sind, standardisiert gesammelt und anonymisiert werden und anschließend frei zugänglich zur Weiterverarbeitung durch Dritte sind (Europäische Kommission, 2012).

Auch Daten öffentlicher Stellen, die nicht personenbezogene Daten sind, können für „Big Data“ relevant sein. Wie schon für Gesundheitsdaten gefordert, sollte die Gesetzgebung zur Nutzung von Daten öffentlicher Stellen angepasst werden. Programmierschnittstellen zur Nutzung der vorhandenen Datenbestände, Integration von Daten Dritter, die diese öffentlich zugänglich machen wollen und Gesetze zur proaktiven Veröffentlichung von Daten sind notwendig (Klessmann, Denker, Schieferdecker, & Schulz, 2012, S. 458-460).

Wie im Abschnitt zur eGK beschrieben, ist die elektronische Patientenakte einer der freiwilligen Dienste, die der Patient in Anspruch nehmen kann. Weiterhin kann dieser selbst bestimmen, welche Daten auf der Karte gespeichert werden sollen. Hierdurch wird jedoch eine verfälschte Patientenakte erzeugt, auf die Ärzte sich im Zweifel nicht verlassen können und somit nicht mehr wert ist als eine institutsinterne Patientenakte (Müller, 2008). Eine gesetzliche Regelung, nach der alle Daten auf der Karte gespeichert werden, würde mehr Sicherheit für die Auswertung der Daten schaffen.

Derzeit ist es auf Grund von Datenschutzbestimmungen nicht möglich, personenbezogene Daten für einen anderen Zweck zu nutzen, als der Betroffene eingewilligt hat. ISO/TS 14265:2011 bietet Anhaltspunkte für Gesetzgeber und definiert klare Zwecke, für welche Gesundheitsdaten genutzt und verarbeitet werden können. Richtlinien und Gesetze zur Weiterverarbeitung von Daten können auf dieser Grundlage beschlossen werden (Geissbuhler, et al., 2013, S. 4).

Auch für die Anerkennung aus Ergebnissen aus „Big Data“ Projekten kann der Gesetzgeber Beiträge leisten. Im Bereich der gesundheitsökonomischen Evaluation herrscht beim Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) eine klare Rangfolge bevorzugter Studientypen vor: 1. randomisierte klinische Studien, 2. randomisierte Interventionsstudien, 3. prospektive Beobachtungsstudien, 4. retrospektive Beobachtungsstudien. Ergebnisse aus „Big Data“ Analysen fallen unter den vierten Punkt und ihnen wird ein geringerer Evidenzgrad beigemessen (Schreyögg & Stargardt, 2012, S. 673). In anderen Ländern sind Studien unter Alltagsbedingungen weitaus anerkannter und werden teilweise gegenüber RTCs bevorzugt (Schreyögg & Stargardt, 2012). Durch die im Versorgungsstrukturgesetz geplante Bereitstellung kassenübergreifender Routinedaten für wissenschaftliche Zwecke, wird ein Schritt zur Weiterverarbeitung vorhandener Daten getan (Schreyögg & Stargardt, 2012, S. 673). Rechtliche Vorgaben durch den G-BA oder das Bundesministerium für Gesundheit zur Nutzung von Alltagsstudien durch das IQWiG sind hier notwendig.



## 8.2 Infrastruktur

Um die Nutzungsrate von elektronischen Systemen zur Verarbeitung von Daten im Gesundheitssystem zu erhöhen, muss den Institutionen eine solide Finanzierung geboten werden (Accenture, 2012, S. 150). Wie im Abschnitt „Austausch medizinischer Daten“ gezeigt, zögern viele Krankenhäuser und Ärzte bisher, moderne IT-Systeme anzuschaffen, da sie das Investitionsrisiko als zu groß einschätzen.

Gemeinsam mit dem Gesetzgeber und der Industrie sollten Standards geschaffen werden, welche Dateiformate und Übertragungswege genutzt werden. Bisher fehlen Standards und Best Practice Beispiele, an denen sich Stakeholder orientieren können (Accenture, 2012, S. 160). Eine Orientierung an anderen Ländern kann hilfreich sein. So wurde in Österreich bereits die elektronische Gesundheitsakte eingeführt. Der Gesetzgeber hat die Standards festgelegt, so zum Beispiel die Verwendung von HL7<sup>19</sup>, einem international anerkannten Nachrichtenformats im Gesundheitswesen (bvitg - Bundesverband Gesundheits-IT, 2013). Auch auf EU-Ebene gibt es bereits Bestrebungen, Standards für elektronische Patientenakten zu schaffen. Im Rahmen der „Innovative Medicines Initiative“, einer großangelegten Initiative zur Steigerung der Wettbewerbsfähigkeit durch Private-Public-Partnerships, wird an dem EHR4CR<sup>20</sup> Projekt gearbeitet. Ziel ist die Entwicklung eines sicheren, interoperablen, skalierbaren, anpassbaren Systems, um Daten für die klinische Forschung zu nutzen (Geissbuhler, et al., 2013, S. 6). Zertifikate und Akkreditierungen für ePA-Systeme die in bestehende Systeme integrierbar sind, sind geplant.

Die in Deutschland geplante Telematikinfrastruktur ist bis heute (Stand Juli 2013) noch nicht errichtet. Bisher wurde die Industrie lediglich aufgefordert, Angebote zum Aufbau einer passenden Infrastruktur abzugeben (gematik, 2013). Um die Telematikinfrastruktur optimal einsetzen zu können, sollte sie offene Standards verwenden, skalierbar, flexibel, prozessorientiert, vertrauenswürdig sein und weiterhin Interoperabilität mit anderen verwendeten Systemen gewährleisten (Heidenreich & Blobel, 2009). Dies gilt auch für in einzelnen Institutionen verwendete IT-Systeme.

Ein großer Aspekt von „Big Data“ ist die Geschwindigkeit, mit der Daten anfallen und analysiert werden, dies geschieht in nahezu Echtzeit (Maier & Weber, 2013). Da Daten der Leistungserbringer bisher mit zeitlicher Verzögerung von bis zu neun Monaten übertragen werden (siehe Abschnitt 5.2.3) (Horenkamp-Sonntag & Verheyen, 2012, S. 9),

---

<sup>19</sup> <http://www.hl7.org/about/index.cfm?ref=nav>.

<sup>20</sup> Eletronic Health Record for Clinical Research, <http://www.ehr4cr.eu/>.

ist dieser Aspekt zumindest bei denen den Krankenkassen übermittelten administrativen Daten nicht gegeben. Durch die Telematikinfrastruktur und die elektronische Erfassung der von den Leistungserbringern erbrachten Leistungen und Befunde, würde die passende Infrastruktur zur zeitnahen Auswertung geschaffen werden.

Eine erfolgreiche Infrastruktur ist nur hilfreich, wenn die Anwender sie auch effektiv nutzen. Software und Serviceanbieter für IT-Systeme im Gesundheitswesen sollten auf eine einfache Anwendbarkeit und sinnvolle grafische Benutzeroberfläche achten und das System im Idealfall gemeinsam mit den Betroffenen entwickeln (Britnell & de Boer, 2012). Auch Trainings im Umgang mit den Systemen und das Aufzeigen der Vorteile durch die Nutzung sind wichtige Bausteine. Dies kann die Akzeptanz zur Nutzung dieser Systeme und somit die Datenqualität erhöhen.

### 8.3 Akzeptanz

(Bundesministerium für Justiz, 2013) (Bundesministerium der Justiz, 2013) (Bundesministerium der Justiz, 2013) Für die meisten Anwendungen von „Big Data“ ist es nicht notwendig, dass die Daten personenbezogen vorliegen. Sie können anonymisiert werden, um eine Rückverfolgung auf einzelne Personen auszuschließen. Dies ist jedoch durch die Vielfalt an Daten zunehmend schwerer, da durch deren Verknüpfung Profile erstellt werden können (Tucker, 2013). Pseudonymisieren einzelner Merkmale ist nicht ausreichend, neue Anonymisierungstechniken müssen entwickelt werden, um den nötigen Datenschutz zu gewährleisten (Machanavajhala & Reiter, 2012).

Bereits heute ist der Patient Besitzer seiner Daten, hat das Recht zu sehen welche Daten in Institutionen über ihn gespeichert werden (z.B. Patientenakte) und was mit ihnen geschieht. In der Praxis ist es jedoch schwer Einsicht in diese Daten zu erhalten. Eine elektronische Patientenakte auf die der Patient jederzeit zugreifen kann und prüfen kann wer auf die Daten zugreift, schafft Transparenz und erhöht die Akzeptanz zur digitalen Verarbeitung der Daten.

Doch auch auf der Seite der Leistungserbringer und Leistungsträger ist es notwendig Akzeptanz zu schaffen. Diese wurden in Deutschland gesetzlich dazu aufgefordert eine Telematikinfrastruktur zu entwickeln, die ein Austausch der zur Verfügung stehenden Daten ermöglichen soll. Dieser Top-Down Ansatz wird als einer der Punkte für mangelnde Akzeptanz der geplanten Features der eGK genannt (Accenture, 2012, S. 160). Wie in Abschnitt 6.1 beschrieben, haben viele Ärzte Bedenken gegenüber der Weitergabe von Patientendaten und führen Datenschutzprobleme und das besondere Arzt-Patienten-

Verhältnis an. Durch eine klar kommunizierte Data-Governance Strategie, die stetiger Evaluation unterzogen wird, kann eine größere Akzeptanz geschaffen werden (Accenture, 2012, S. 160). Weiterhin ist das Aufzeigen von Best Practices aus Ländern wie Dänemark und Österreich, die bereits Erfahrungen mit institutionsübergreifend bereitgestellten Patientendaten haben, eine Möglichkeit, um die Akzeptanz der Beteiligten zu erhöhen. Die von den Ärzten geforderte Effizienzsteigerung im Klinikalltag kann so aufgezeigt werden.

## 9. Diskussion

Dadurch, dass das Thema „Big Data“ ein relativ junges Thema ist, gibt es wenig kritische wissenschaftliche Publikationen hierzu. Ein großer Teil der vorhandenen Literatur wird von Unternehmensberatungen oder Technologieunternehmen veröffentlicht. Hierdurch herrscht eine Verzerrung, da diese Unternehmen mit dem Vertrieb neuer Technologien oder Dienstleistungen auf diesem Feld Umsatz generieren wollen. Es ist also zweifelhaft ob „Big Data“ in wirklich allen Bereichen die angepriesenen Potentiale erfüllen kann. Die Literatur ist aufgrund dieser Beeinflussung kritisch zu betrachten. Weiterhin liegt vorwiegend amerikanische Literatur vor, die auf den dort herrschenden Gegebenheiten beruht. Eine direkte Übertragung auf Deutschland ist nicht möglich. Signifikante Unterschiede in der Mentalität, Gesetzen, Finanzierung und dem Aufbau des Gesundheitssystems verhindern dies.

Die Datenschutzgesetze in Deutschland sind sehr komplex und beziehen sich häufig auf andere Gesetze. Zusätzlich hat jedes Bundesland eigene Gesetze und weitere 39 spezifische Gesetze bestehen. Eine erschöpfende Betrachtung der Gesetzeslage war aufgrund der Komplexität und dem eingeschränkten Umfang dieser Arbeit nicht möglich. Die betrachteten Gesetze bieten jedoch einen ersten Überblick über die herrschenden Einschränkungen. Diese sind jedoch nicht prinzipiell als schlecht einzusehen. Der starke Datenschutz ermöglicht auch ein Vertrauen der Bürger in die rechtmäßige Nutzung ihrer Daten und beugt einem Missbrauch vor.

Die Positionen der beteiligten Akteure in Deutschland sind schwer zu erfassen. Zum einen gibt es gezielt zu „Big Data“ keine Positionen oder Veröffentlichungen von deutschen Gesundheitseinrichtungen. Zum anderen antworteten die angeschriebenen Institutionen nicht zielführend. Entweder gab es keine Position zu dem Thema oder es wurde nicht gezielt auf die Fragen eingegangen. Eine ausführlichere Befragung deutscher Gesundheitseinrichtungen zum Austausch und der Nutzung personenbezogener Daten ist

sinnvoll. Auch bei den Patienten ist die Datenlage sehr schwach. Die zugrundeliegende Studie zum Austausch medizinischer Daten bezieht sich auf eine sehr kleine Stichprobengröße. Zusätzlich sind Patienten eine sehr heterogene Gruppe und haben keinen Spitzenverband der allgemeine Aussagen trifft. Die Verbände der Patienten verfolgen oft sehr starke Einzelinteressen.

Die allgemeine Position der deutschen Bevölkerung zum Umgang mit personenbezogenen Daten ist derzeit schwer zu untersuchen. So zeigte die Studie aus dem Jahr 2011 noch ein großes Vertrauen in den Staat. Nach den jüngsten Umfragen sank dieser Wert jedoch stark. Es bleibt abzuwarten, wie sich diese Meinung langfristig ändern wird und wie der Staat auf diesen Vertrauensverlust reagiert.

Trotz fehlender technischer Standards und Investitionsstaus investiert ein Großteil der Krankenhäuser in IT-Systeme zum Umgang mit Patienten- oder Prozessdaten. Es ist davon auszugehen, da vor allem private Kliniken und Klinikverbände diese Investitionen tätigen. Zum einen kann so ein Wettbewerbsvorteil entstehen und zum anderen können die Klinikverbände die einzelnen Kliniken besser miteinander vernetzen. Die Argumente der niedergelassenen Ärzte bezüglich der Nicht-Nutzung von IT und dem Austausch von Daten durch z.B. die eGK, sind kritisch zu betrachten. Sie geben Datenschutzbedenken an, obwohl in Krankenhäusern bereits gute Erfahrungen gesammelt wurden. Auch der Zusatznutzen ist in anderen Ländern bereits hinreichend belegt. Es sollte genauer untersucht werden, weshalb sie trotzdem eine eher abneigende Haltung haben.

Die von der EU und den UN vorgeschlagenen Ansätze zur Verfügbarkeit von Daten werden schwer umsetzbar sein. In einer weiteren Arbeit könnte untersucht werden, in wie fern es sowohl rechtlich als auch infrastrukturell möglich ist, das zentrale Stellen alle Daten von Unternehmen entgegen nehmen und aufbereiten.

Auch die eGK und die ePA sind trotz vorgesehener Möglichkeiten kritisch zu betrachten. Bisher konnte noch keine Funktion umgesetzt werden und die Kosten haben bereits ein Vielfaches des ursprünglich geplanten erreicht (Booz Allen Hamilton GmbH, 2006). Die vielen Akteure bei der Umsetzung des Projektes könnten durch Partikularinteressen weiter zu Verzögerungen und Verwässerungen des Projektes beitragen.

Aufgrund des gesundheitswissenschaftlichen Hintergrundes des Autors, behandelt die Arbeit weitestgehend keine technischen Details zu Software und Hardware.

## 10. Fazit

Ausgehend von den vorliegenden Informationen kann gesagt werden, dass im deutschen Gesundheitswesen „Big Data“ bisher kaum eine Rolle spielt und nur sehr begrenzt anwendbar ist. Es gibt rechtliche und strukturelle Beschränkungen, die zunächst überwunden werden müssen. Deutschland hinkt bei der Nutzung von Daten anderen Ländern bisher hinterher, ist aber auf dem richtigen Weg eine passende Infrastruktur aufzubauen. Rechtlich wird es jedoch schwer sein, diese Daten zu nutzen. Der Datenschutz ist in Deutschland ein sehr wichtiges Thema und die Sorge um Datenmissbrauch ist groß. Nur mit absolut sicheren Methoden der Verschlüsselung und Anonymisierung von personenbezogenen Daten wird eine Weiterverarbeitung dieser möglich sein.

Personenbezogenen und öffentliche Daten sind hingegen oft schon vielfältig vorhanden. Besonders medizinische Daten liegen den Leistungsträgern und Leistungserbringern oft sehr strukturiert vor. Durch das DRG-System werden Fälle in Krankenhäusern sehr genau kodiert. „Big Data“ hat in Deutschland also großes Potential wenn es um die Datengrundlage geht.

Dies ist jedoch nur begrenzt hilfreich, wenn auf der anderen Seite diese Daten nicht genutzt werden dürfen. Die rechtliche Grundlage lässt bisher keine Nutzung von medizinischen Daten, außer denen in SGB V vorgegeben, zu. Auch die Institutionen, die Zugriff auf diese Daten haben, sind festgelegt. Wie in den Handlungsempfehlungen beschrieben, sollte der Gesetzgeber neue Rahmenbedingungen zur Nutzung dieser Daten schaffen. Nach vollständiger Anonymisierung sollten sie frei zugänglich sein, da so die Grundlage für Innovationen geschaffen werden kann. Auch die Anerkennung von Ergebnissen aus „Big Data“ Projekten sollte gefördert werden.

Die Pläne zum Aufbau der Telematikinfrastruktur, der eGK und der ePA sind sehr ambitioniert und sind ein Schritt in die richtige Richtung. Deutschland kann mit einer erfolgreichen Umsetzung eine große Mengen an medizinischen Daten für die Weiterverarbeitung erfassen. Sie bieten die Grundlage für vielfältige Analysen und Studien, die dem Gesundheitssystem zugutekommen können. Gleichzeitig ist bisher jedoch noch keine der geplanten Funktionen einsatzbereit, womit eine langfristige Planung erschwert ist.

Das Sicherheitskonzept der eGK ist jedoch ein richtiger Schritt, um die Akzeptanz bei der Bevölkerung für den Datenaustausch zu erhöhen. Die Sorge der Menschen über die

Sicherheit ihrer Daten ist ein ernst zu nehmendes Thema. Ohne die Zustimmung der Betroffenen sollte keiner Zugriff auf die personenbezogenen medizinischen Daten möglich sein können. Nachdem das Vertrauen der Bürger in den Staat zum Umgang mit ihren persönlichen Daten in letzter Zeit stark gesunken ist, ist fraglich, inwiefern eine zentrale Stelle zur Speicherung aller medizinischer Daten Zustimmung finden würde.

Fragwürdig bleibt außerdem, ob es Institutionen erlaubt werden sollte so viele Daten über Menschen zusammenzuführen. Es wird wohl nie Techniken geben, mit denen Datensätze vollkommen anonymisiert werden können. Ein Missbrauch der Daten könnte weitreichende Folgen haben.

Es bleibt abzuwarten ob „Big Data“ die erhoffte Lösung für viele Probleme sein wird und wie sich die Bedingungen in Deutschland ändern werden.

## Literaturverzeichnis

Accenture. (2012). *Connected Health: The Drive to Integrated Healthcare Delivery*.

allfacebook.de. (06 2013). *Anzahl der aktiven Nutzer von Facebook in Deutschland von Januar 2010 bis Juni 2013 (in Millionen)*. Abgerufen am 25. 07 2013 von destatis: <http://de.statista.com/statistik/daten/studie/70189/umfrage/nutzer-von-facebook-in-deutschland-seit-2009/>

Amtsblatt der Europäischen Gemeinschaft. (23. 11 1995). Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr.

Booz Allen Hamilton GmbH. (2006). *Endbericht - zur Kosten-Nutzen-Analyse der Einrichtung einer Telematik- Infrastruktur im deutschen Gesundheitswesen* . Düsseldorf.

Britnell, M., & de Boer, J. (2012). *Accelerating innovation: the power of the crowd Global lessons in eHealth implementation* . (KPMG, Hrsg.)

Brynjolfsson, E., & McAfee, A. (Oktober 2012). Big Data: The Management Revolution. *Harvard Business Review* , 59-68.

Buckee, C., Wesolowski, A., Eagle, N., & Snow, R. (2013). Mobile phones and malaria: modeling human and parasite travel. *Travel Medicine and Infectious Disease* (11), 15-22.

Bundesministerium der Justiz. (2013). *Bundesdatenschutzgesetz - BDSG*.

Bundesministerium der Justiz. (2013). *Gesetz zur Regelung des Zugangs zu Informationen des Bundes - IFG*.

Bundesministerium des Inneren. (kein Datum). *Datenschutzrecht in der EU*. Abgerufen am 09. 07 2013 von Bundesministerium des Inneren : [http://www.bmi.bund.de/DE/Themen/Gesellschaft-Verfassung/Datenschutz/Datenschutzrecht-EU/datenschutzrecht-eu\\_node.html](http://www.bmi.bund.de/DE/Themen/Gesellschaft-Verfassung/Datenschutz/Datenschutzrecht-EU/datenschutzrecht-eu_node.html)

Bundesministerium des Inneren. (2013). *Der Schutz des Rechts auf informationelle Selbstbestimmung*. Abgerufen am 09. 07 2013 von Bundesministerium des Inneren: [http://www.bmi.bund.de/DE/Themen/Gesellschaft-Verfassung/Datenschutz/Informationelle-Selbstbestimmung/informationelle-selbstbestimmung\\_node.html](http://www.bmi.bund.de/DE/Themen/Gesellschaft-Verfassung/Datenschutz/Informationelle-Selbstbestimmung/informationelle-selbstbestimmung_node.html)

Bundesministerium für Justiz. (2013). *Sozialgesetzbuch fünftes Buch - SGB V*.

Bundesverband Gesundheits-IT. (27. 03 2013). *Interoperabilität: Gemeinsamer Standard für elektronische Patientenakten*. Abgerufen am 17. 07 2013 von [bvitg.de](http://www.bvitg.de):

<http://www.bvitg.de/pressemitteilung-conhit/items/interoperabilitaet-gemeinsamer-standard-fuer-elektronische-patientenakten.html>

Butler, D. (13. 02 2013). *When Google got flu wrong*. (N. P. Group, Herausgeber)

Abgerufen am 02. 07 2013 von Nature: <http://www.nature.com/news/when-google-got-flu-wrong-1.12413>

bvitg - Bundesverband Gesundheits-IT. (2013). *Gute Gesundheitsversorgung braucht mehr IT*. Positionspapier.

Charta der Grundrechte der Europäischen Union. (26. 10 2012). *Amtsblatt der Europäischen Union ( 2012/C 326/02)* . Amtsblatt der Europäischen Union.

Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit. (kein Datum).

*Spezialgesetze*. Abgerufen am 20. 07 2013 von Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit:

[http://www.bfdi.bund.de/DE/GesetzeUndRechtsprechung/Spezialgesetze/Spezialgesetze\\_node.html](http://www.bfdi.bund.de/DE/GesetzeUndRechtsprechung/Spezialgesetze/Spezialgesetze_node.html)

Dobrick, K. (23. 09 2012). *Attitudes and Experiences of TK Insurants Concerning EU-wide Electronic Exchange of Health Data*. Hamburg.

Europäische Kommission. (2012). *eHealth Task Force Report – Redesigning health in Europe for 2020*. Luxemburg: Publications Office of the European Union.

European Medicines Agency. (24. 06 2013). *European Medicines Agency releases for public consultation its draft policy on the publication and access to clinical-trial data*.

Abgerufen am 18. 07 2013 von European Medicines Agency:

[http://www.ema.europa.eu/ema/index.jsp?curl=pages/news\\_and\\_events/news/2013/06/news\\_detail\\_001825.jsp&mid=WC0b01ac058004d5c1](http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2013/06/news_detail_001825.jsp&mid=WC0b01ac058004d5c1)

Ärzteblatt. (05. 06 2007). *NICE: Krebsmedikament soll nur bezahlt werden, wenn es wirkt*.

Abgerufen am 05. 07 2013 von [aerzteblatt.de](http://www.aerzteblatt.de):

<http://www.aerzteblatt.de/nachrichten/28697>



Gangl, K., & Birkner, G. (2009). *Branchenbarometer E-Health*. (M. u. F.A.Z.-Institut für Management-, Hrsg.) Frankfurt am Main: Techniker Krankenkasse.

Geissbuhler, A., Safran, C., Buchan, I., Bellazzi, R., Labkoff, S., Eilenberg, K., et al. (2013). Trustworthy reuse of health data: A transnational perspective. *International Journal of Medical Informatics* (82), 1-9.

gematik. (11. 07 2013). *Freigabe für Aufforderung der Industrie zur Abgabe von verbindlichen Angeboten für Kartengeneration 2 erteilt*. Abgerufen am 17. 07 2013 von gematik - Gesellschaft für Telematikanwendungen der Gesundheitskarte mbH: [http://www.gematik.de/cms/de/header\\_navigation/presse/pressemitteilungen/pressemitteilungen\\_1.jsp](http://www.gematik.de/cms/de/header_navigation/presse/pressemitteilungen/pressemitteilungen_1.jsp)

gematik. (19. 06 2013). *Pressemitteilungen - Elmer: „Wichtiger Meilenstein für den Bau der sicheren Datenautobahn im Gesundheitswesen erreicht“*. Abgerufen am 17. 07 2013 von gematik - Gesellschaft für Telematikanwendungen der Gesundheitskarte mbH: [http://www.gematik.de/cms/de/header\\_navigation/presse/pressemitteilungen/pressemitteilungen\\_1.jsp](http://www.gematik.de/cms/de/header_navigation/presse/pressemitteilungen/pressemitteilungen_1.jsp)

gematik. (kein Datum). *Verfügbare Anwendungen*. Abgerufen am 17. 07 2013 von gematik - Gesellschaft für Telematikanwendungen der Gesundheitskarte mbH: [http://www.gematik.de/cms/de/egk\\_2/anwendungen/verfuegbare\\_anwendungen/verfuegbare\\_anwendungen\\_1.jsp](http://www.gematik.de/cms/de/egk_2/anwendungen/verfuegbare_anwendungen/verfuegbare_anwendungen_1.jsp)

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature* (457), 1012-1014.

GKV-Spitzenverband. (27. 06 2013). *GKV-Spitzenverband*. Abgerufen am 30. 07 2013 von Blockaden gegen Nutzen der elektronischen Gesundheitskarte beenden– Krankenkassen fordern gesetzliche Maßnahmen ein: [http://www.gkv-spitzenverband.de/presse/pressemitteilungen\\_und\\_statements/pressemitteilung\\_58816.jsp](http://www.gkv-spitzenverband.de/presse/pressemitteilungen_und_statements/pressemitteilung_58816.jsp)

Groves, P., Kayyali, B., Knott, D., & Van Kuiken, S. (2013). *The "big data" revolution in healthcare - Accelerating value and innovation*. McKinsey&Company. Center for US Health System Reform Business Technology Office.

Hübner, U., Liebe, J.-D., Egbert, N., & Frey, A. (2012). *IT-Report Gesundheitswesen - Schwerpunkt IT im Krankenhaus*. (A. u. Niedersächsisches Ministerium für Wirtschaft, Hrsg.) Hannover.

Hampe, K., & Puppe, M. (2012). *Vertrauen und Sicherheit im Netz*. (T. u. BITKOM - Bundesverband Informationswirtschaft, Hrsg.) Berlin.

Heidenreich, G., & Blobel, B. (2009). IT-Standards für telemedizinische Anwendungen. *Bundesgesundheitsblatt* (52), 316-323.

Hoffman, S., & Podgurski, A. (2013). *Big Bad Data: Law, Public Health, and Biomedical Databases*. *Journal of Law, Medicine & Ethics*.

Horenkamp-Sonntag, D., & Verheyen, F. (12. 06 2012). identifizierung geeigneter versichertengruppen für die integrierte versorgung. Berlin, Berlin, Deutschland.

IfD Allensbach. (09 2010). *Einmal ganz allgemein gefragt: Wie sehr beunruhigt Sie das Thema Datenmissbrauch, wie sehr machen Sie sich persönlich darüber Sorgen, dass persönliche Daten von Ihnen missbraucht werden könnten?* Abgerufen am 25. 07 2013 von destatis: <http://de.statista.com/statistik/daten/studie/164835/umfrage/angst-vor-datenmissbrauch/>

Institut für Demoskopie Allensbach. (2011). *Allensbacher Markt- und Werbeträger-Analyse*. Allensbach: Institut für Demoskopie Allensbach.

Intel. (12. Juni 2013). *Moore`s Law Inspires Intel Innovation*. Abgerufen am 12. Juni 2013 von Intel: <http://www.intel.com/content/www/us/en/silicon-innovations/moores-law-technology.html>

IT-Planungsrat. (12. 09 2012). Offenes Regierungs- und Verwaltungshandeln (Open Government). *Eckpunkte zur Förderung von Transparenz, Teilhabe und Zusammenarbeit*

Kassenärztliche Bundesvereinigung. (27. 05 2013). *Pressemitteilungen 2013*. Abgerufen am 30. 07 2013 von Kassenärztliche Bundesvereinigung: <http://www.kbv.de/telematik/43562.html>

Kempf, D. (2013). *Sicherheit und Vertrauen im Netz*. Berlin: BITKOM.

Klessmann, J., Denker, P., Schieferdecker, I., & Schulz, S. (2012). *Open Government Data Deutschland - Eine Studie zu Open Government in Deutschland im Auftrag des Bundesministerium des Innern*. (B. d. Inneren, Hrsg.) Berlin, Deutschland.

Müller, J. (2008). *Elektronische Patientenakte: Schlüsselrolle für den Datenschutz*.  
Abgerufen am 23. 07 2013 von [aerzteblatt.de](http://www.aerzteblatt.de):  
<http://www.aerzteblatt.de/archiv/59362/Elektronische-Patientenakte-Schluessselrolle-fuer-den-Datenschutz>

Machanavajjhala, A., & Reiter, J. P. (2012). Big privacy: protecting confidentiality in big data. *XRDS: Crossroads, The ACM Magazine for Students - Big Data* (19), 20-23.

Maier, W., & Weber, M. (2013). *Management von Big-Data-Projekten*. (T. u. BITKOM - Bundesverband Informationswirtschaft, Hrsg.) Berlin: BITKOM.

Manyika, J., Chui, M., Brown, B., Burghin, J., Dobbs, R., Roxburgh, C., et al. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute. MyKinsey & Company.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data - A Revolution That Will Transform How We Live, Work and Think*. London: John Murray.

Ohlhorst, F. J. (30. Dezember 2011). *Amazon Elastic Cloud Computer Cuts Cost of Supercomputer Access*. Abgerufen am 12. Juni 2013 von eWeek:  
<http://www.eweek.com/c/a/Cloud-Computing/Amazon-Elastic-Cloud-Computer-Cuts-Cost-of-Supercomputer-Access-260846/>

O'Reilly Media. (2012). *Big Data Now: 2012 Edition* (1. Auflage Ausg.). Sebastopol, Kalifornien, USA: O'Reilly Media Inc.

Poitras, L., Rosenbach, M., & Stark, H. (30. 06 2013). *Partner and Target: NSA Snoops on 500 Million German Data Connections*. Abgerufen am 28. 07 2013 von Spiegel Online:  
<http://www.spiegel.de/international/germany/nsa-spies-on-500-million-german-data-connections-a-908648.html>

Potts, A. L., Barr, F. E., Gregory, D. F., Wright, L., & Patel, N. R. (2004). Computerized physician order entry and medication errors in a pediatric critical care unit, . *Pediatrics* (113), 59-63.

- Schreyögg, J., & Stargardt, T. (2012). Gesundheitsökonomische Evaluation auf Grundlage von GKV-Routinedaten. *Bundesgesundheitsblatt* (55), 668-676.
- Schubert, D. (07. 03 2012). Routinedaten im Gesundheitswesen: Stand und Perspektiven Aus der Sicht der Sekundärdatenanalyse. München, Bayern, Deutschland.
- Simon, M. (2010). *Das Gesundheitssystem in Deutschland - Eine Einführung in Struktur und Funktionsweise* (3. Auflage Ausg.). Bern: Verlag Hans Huber.
- Simonite, T. (2013). Smartphone Tracker Gives Doctors Remote Viewing Powers. In M. T. Review, & M. T. Review (Hrsg.), *Big Data Gets Personal* (S. 16). Cambridge, Massachusetts, USA.
- Statistisches Bundesamt. (16. 05 2012). *53 % der Internetnutzer sind in sozialen Netzwerken aktiv*. Abgerufen am 19. 07 2013 von Destatis:  
[https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2012/05/PD12\\_172\\_63931.html](https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2012/05/PD12_172_63931.html)
- Statistisches Bundesamt. (2012). *Private Haushalte in der Informationsgesellschaft – Nutzung von Informations- und Kommunikationstechnologien*. Wiesbaden.
- Statistisches Bundesamt. (2013). *Wirtschaftsrechnungen - Laufende Wirtschaftsrechnungen Ausstattung privater Haushalte mit ausgewählten Gebrauchsgütern*. Wiesbaden.
- Stecher, E. (08. 07 2013). *Neue Gesundheitskarte ist bislang ein Flop*. Abgerufen am 01. 08 2013 von Nürnberger Zeitung: <http://www.nordbayern.de/nuernberger-zeitung/nz-news/neue-gesundheitskarte-ist-bislang-ein-flop-1.3018215>
- Talbot, D. (2013). Big Data from Cheap Phones. In MIT Technology Review, *Big Data Gets Personal* (S. 17-22). Cambridge, Massachusetts, USA: MIT.
- The World Economic Forum. (2012). *Big Data, Big Impact: New Possibilities for International Development*. Genf, Schweiz: The World Economic Forum.
- TNS Emnid. (09 2011). *Schätzen Sie es bei den folgenden Unternehmen, Organisationen bedenklich ein, wenn diese mit Ihren Daten arbeiten bzw. diese sammeln?* Abgerufen am 25. 07 2013 von destatis:  
<http://de.statista.com/statistik/daten/studie/204937/umfrage/vertrauen-in-organisationen-unternehmen-im-bezug-auf-sammeln-persoenerlicher-daten/>

Tucker, P. (2013). Has Big Data Made Anonymity Impossible? In M. T. Review, *Big Data Gets Personal* (S. 2-4).

UN Global Pulse. (2012). *Big Data for Development: Challenges & Opportunities*. New York.

Urbanski, J., & Weber, T. (2012). *Big Data im Praxiseinsatz – Szenarien, Beispiele, Effekte*. (T. u. BITKOM - Bundesverband Informationswirtschaft, Hrsg.) Berlin: BITKOM.

Uslu, D. M. (kein Datum). *Unerwartete Eigendynamik der Elektronischen Patientenakten in Deutschland*. Abgerufen am 17. 07 2013 von Deutsche MedizinMesse.de:  
[http://www.deutsche-medizinmesse.de/iframe\\_lecture\\_details.php?lecture\\_id=26](http://www.deutsche-medizinmesse.de/iframe_lecture_details.php?lecture_id=26)

Zikopoulos, P., Eaton, C., deRoos, D., Deutsch, T., & Lapis, G. (2012). *Understanding Big Data - Analytics for Enterprise Class Hadoop and Streaming Data*. USA: McGraw-Hill.

## **Eidesstattliche Erklärung**

Ich versichere, dass ich vorliegende Arbeit ohne fremde Hilfe selbständig verfasst und nur die angegebenen Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quelle kenntlich gemacht.

Julien Kuhlmann

Hamburg, den 05.08.2013

## Anhang

### Fragen an ausgewählte Akteure

1. Kennen Sie den Begriff "Big Data"? Wenn ja, was verstehen Sie unter diesem Begriff?
2. Welche Potenziale sehen Sie in der eGK?  
2a Welche Bedenken haben Sie gegenüber der eGK?
3. Welche Potenziale sehen Sie in der elektronischen Patientenakte?  
3a Welche Bedenken haben Sie gegenüber der elektronischen Patientenakte?
4. Wie sehen Sie die momentane Situation des Datenschutzes von personenbezogenen Daten?
5. Sollten personenbezogene Patientendaten (anonymisiert) von privatwirtschaftlichen und wissenschaftlichen Institutionen verwendet werden dürfen?

### Rückmeldungen der Angeschriebenen

#### Marburger Bund

Sehr geehrte Frau Kuhlmann,

was die elektronische Gesundheitskarte und die Entwicklung einer modernen Telematikinfrastruktur im Gesundheitswesen angeht, unterstützen wir die Haltung der Bundesärztekammer. Dort gibt es ein Projektbüro, das unmittelbar an der Entwicklung der eCard beteiligt ist. Bitte nehmen Sie direkt mit der Pressestelle der Bundesärztekammer Kontakt auf: 030/4004 56-700, [presse@baek.de](mailto:presse@baek.de).

Mit besten Grüßen

Hans-Jörg Freese

Leiter Verbandskommunikation / Pressesprecher

## **Verband der deutschen Krankenversicherungen**

Sehr geehrter Herr Kuhlmann,

vielen Dank für Ihre Anfrage und Ihr Interesse an der Privaten Krankenversicherung.

Sie haben sich ein sehr spannendes Thema ausgesucht. Zu meinem Bedauern gibt es im PKV-Verband keine Positionierung zum Thema „Big Data“, auch wenn der Begriff und dessen Bedeutung einigen Kollegen, besonders aus der IT-Abteilung, bekannt sein dürfte.

Zum Thema eGK, elektronischer Patientenakte und die Verwendung personenbezogener Daten möchten wir uns gegenwärtig nicht äußern. Wie Sie vielleicht mitbekommen haben, ist der PKV-Verband aus dem Projekt eGK vor einigen Jahren ausgestiegen und beteiligt sich folglich nicht mehr an deren Einführung.

Auch wenn wir uns nicht an Ihrer Studie beteiligen, wünschen wir Ihnen viel Erfolg für Ihre Bachelor-Thesis.

Mit freundlichen Grüßen

Dominik Heck

- Pressereferent -

## **Boehringer-Ingelheim**

Sehr geehrter Herr Kuhlmann,

vielen Dank für Ihre E-Mail. Wir müssen Ihnen hiermit mitteilen, dass wir generell an Befragungen für studentische Arbeiten nicht mehr teilnehmen. Die Anfragen wurden einfach zu viel.

Wir hoffen auf Ihr Verständnis und verbleiben

Mit freundlichen Grüßen / Kind regards, Gabriele Schiffner Boehringer

## **Kassenärztliche Vereinigung Bayerns**

Sehr geehrte Damen und Herren,

vielen Dank für Ihre E-Mail. Ihr Anliegen ist uns wichtig.

Wir haben Ihre Anfrage an die zuständige Stelle in unserem Haus weitergeleitet, diese wird sich schnellstmöglich mit Ihnen in Verbindung setzen.



Haben Sie noch Fragen? Bitte melden Sie sich einfach.

Freundliche Grüße

Sandra Ghulam

Fachreferentin Kommunikation

Mitgliederservice und Beratung

## **Techniker Krankenkasse**

Sehr geehrter Herr Kuhlmann,

vielen Dank für Ihre Anfrage vom 11. Juli 2013.

Der Begriff "Big Data" ist uns bekannt, wird innerhalb der TK jedoch nicht genutzt.

Die BITKOM hat zu folgenden Themen jeweils einen Leitfaden entwickelt, "Management von Big-Data-Projekten" zu finden unter: [http://www.bitkom.org/de/publikationen/38337\\_76511.aspx](http://www.bitkom.org/de/publikationen/38337_76511.aspx) und Big Data im Praxiseinsatz - Szenarien, Beispiele, Effekte zu finden unter: [http://www.bitkom.org/files/documents/BITKOM\\_LF\\_big\\_data\\_2012\\_online\(1\).pdf](http://www.bitkom.org/files/documents/BITKOM_LF_big_data_2012_online(1).pdf). Falls Ihnen diese Abhandlungen nicht bekannt sind, können Ihnen diese ggf. hilfreiche Informationen liefern.

Die Einführung der elektronischen Gesundheitskarte beruht auf die Umsetzung eines gesetzlichen Auftrags. In § 291 Sozialgesetzbuch Fünftes Buch (SGB V) befindet sich die Vorgabe, die Krankenversichertenkarte zur Gesundheitskarte auszubauen. Hiervon sind als Körperschaften des öffentlichen Rechts alle gesetzlichen Krankenversicherungen und folglich auch alle gesetzlich Versicherten betroffen. Eine Ausnahme ist nicht vorgesehen. Die elektronische Gesundheitskarte soll in erster Linie dazu dienen, eine Verbesserung in der Versorgung der Patientinnen und Patienten anzubieten. Die Qualität der medizinischen Behandlung hängt auch davon ab, ob die Ärztin oder der Arzt alle notwendigen medizinischen Informationen erhält. Mit der elektronischen Gesundheitskarte und dem "elektronischen Netz" (Telematikinfrastruktur), das im Gesundheitswesen aufgebaut wird, können Daten in Zukunft elektronisch sicher und schnell einem Arzt zur Verfügung gestellt werden. Mit ihr werden die technischen Voraussetzungen geschaffen, dass medizinische Daten in Zukunft auch auf elektronischem Wege sicher, schnell und einfach ausgetauscht und verarbeitet werden können.

Die jetzt ausgegebenen Gesundheitskarten sind für die Aufnahme weiterer Anwendungen vorbereitet. Der nächste Schritt ist der Aufbau der Telematikinfrastruktur. Durch die

geplante Online-Anbindung können die Karten aktualisiert und gesperrt werden. Die Kosten für den Kartenaustausch können somit eingeschränkt sowie der Missbrauch minimiert werden. Zu den weiteren zukünftigen Anwendungen, hierzu gehört auch die elektronisch Gesundheitsakte, können wir noch keine Aussage treffen, da die genauen Vorgaben für die Spezifikationen noch nicht vorliegen.

Weitere hilfreiche Informationen finden Sie auf folgenden Internetseiten:

Gematik: <http://www.gematik.de/cms/de/startseite/index.jsp> oder

Bundesministeriums für Gesundheit:  
<http://www.bmg.bund.de/krankenversicherung/elektronische-gesundheitskarte.html>

Freundliche Grüße

Susanne Kramer

### **Deutsche Krankenhausgesellschaft**

Sehr geehrter Herr Kuhlmann,

vielen Dank für Ihre Nachricht, zu der wir gern auf unsere Internetseite ([www.dkgev.de](http://www.dkgev.de)) verweisen.

Unter [Telematik](#) finden Sie die von uns dazu veröffentlichten Positionen. Ergänzend finden Sie [hier](#) unsere Information „Übersicht Gesundheitskarte“ zum Basis-Rollout der eGK.

Die aktuellsten Informationen zur Gesellschaft für Telematikanwendungen der Gesundheitskarte mbH (gematik) finden Sie unter [www.gematik.de](http://www.gematik.de)

Mit freundlichen Grüßen

### **GlaxoSmithKline**

Sehr geehrte Frau Kuhlmann,

vielen Dank für Ihre Anfrage. In der Regel unterstützen wir die Arbeiten von Studenten sehr gerne, leider muss ich Ihnen aber mitteilen, dass wir bei der Vielzahl der Anfragen nicht alle beantworten können. Deshalb werden wir bei dieser Befragung nicht teilnehmen.

Wir wünschen Ihnen für Ihre Bachelorarbeit alles Gute.