

## Teil C – Big Data: Chancen und Risiken aus Sicht der Bürger

*Horst Müller-Peters*

### Summary

Ausgehend von den Arbeiten von Knorre sowie Wagner und Jost beschreibt diese Studie auf einer Basis von 1.000 Befragten das Verhältnis der deutschen Bevölkerung zu *digitaler Vernetzung, Big Data* und *Künstlicher Intelligenz*. Der Aufbau folgt der Logik

- a) vom „*Wissen*“ und „*Können*“ als Randbedingungen souveränen Entscheidens ...
- b) über das „*Wollen*“ in Form von Einstellungen und Bewertungen zu Technikfolgen, Anwendungsfeldern und Datenschutz ...
- c) zum tatsächlichen „*Handeln*“ in Onlinenutzung und Datenschutz.

Aufbauend darauf wird das Potenzial für einen möglichen Paradigmenwandel eruiert – in der Datenpolitik, in den medialen und gesellschaftlichen Narrativen und schließlich – beispielhaft für eine Branche – in der Versicherungswirtschaft.

Als Kernergebnisse lassen sich festhalten:

### 1. Wissen – Können – Wollen – Handeln

#### „*Wissen*“

Die digitale Welt ist für die Menschen in Deutschland kein „Neuland“ mehr. Grundbegriffe der Digitalisierung sind ins Allgemeinwissen eingegangen. Bestehende Wissensdefizite finden sich aber gerade rund um die digitale Vernetzung und Big Data, jene Themen also, die besonders den Umgang mit den eigenen Daten tangieren.

#### „*Können*“

Die Bürger fühlen sich aufgrund von Informationsüberlastung und monopolistischer Marktstrukturen nur sehr bedingt handlungsfrei oder souverän, und zwar quer durch alle Altersgruppen und Bildungsschichten. Regulierungen wie z.B. aus der Europäischen Datenschutzgrundverordnung, die vom Ideal eines mündigen und kritischen Bürgers ausgehen und die Interessen der Verbraucher daher über umfassende Information- und Zustimmungspflichten absichern wollen, drohen dadurch vielfach „ins Leere“ zu laufen.

#### „*Wollen*“

Der Schutz der eigenen Daten ist den Menschen unverändert wichtig. Dennoch ergibt sich eine ganz unterschiedliche Wahrnehmung von Big Data, je nach Blickwinkel, aus dem das Thema betrachtet wird:

Das *Fernbild* von Big Data:

- Sofern eher abstrakt nach möglichen Folgen gefragt wird, überwiegen aus Sicht der Bürger die Risiken.
- Nutzenpotenziale der neuen Techniken, wie ein Zugewinn an Sicherheit, Komfort oder Effizienz, werden vielfach nicht erkannt oder wiegen im Vergleich zu den möglichen Nachteilen weniger schwer.
- Auch bezüglich gesamtgesellschaftlicher und volkswirtschaftlicher Auswirkungen dominieren die Bedenken.
- In einer Klassifizierung nach Sorgen und Hoffnungen zeigen sich annähernd doppelt so viele Pessimisten (42% der Befragten) wie Optimisten (22% der Befragten). Die verbleibenden 36% geben sich bezüglich der Chancen und Risiken annähernd indifferent.

Das *Nahbild* von Big Data:

- Das negative Bild ändert sich, wenn anhand konkreter Anwendungsfelder der Nutzen von Big Data deutlich wird. Zahlreiche der getesteten Anwendungen in den drei Lebenswelten Mobilität, Wohnen und Gesundheit werden mehrheitlich begrüßt.
- Das gilt besonders, wenn die Sicherheit erhöht wird, wenn lästige Alltagstätigkeiten automatisiert werden oder wenn der Nutzer bei weitergehenden Tätigkeiten unterstützt wird, ohne dabei gleich die Kontrolle aus der Hand geben zu müssen.
- Ist der Nutzen der Anwendung ersichtlich, tritt selbst die Sorge vor Überwachung in den Hintergrund. Dies gilt sowohl für eine breite Anzahl von Anwendungen in den eigenen Lebenswelten als auch für eine (politisch derzeit höchst konträr diskutierte) polizeiliche Kriminalitätsbekämpfung mittels Big Data.
- Dennoch möchte der Bürger nicht zum „Objekt“ der Algorithmen degradiert werden: Klassifizierungen (sog. „Scorings“) oder gar automatisierte Entscheidungen werden mehrheitlich abgelehnt. Anwendungen im Marketing – z.B. durch „Mikrotargeting“, dynamische Preissetzung oder der Kundenbetreuung über Sprachcomputer, wird überwiegend mit Skepsis begegnet.

## „Handeln“

Die Bevölkerung ist im Netz, und zwar in doppeltem Sinne:

- Das Internet wird intensiv und über unterschiedlichste Geräte genutzt.
- Zugleich sind die Nutzer über zahlreiche Dienste und Geräte im „Netz“ der Anbieter und hinterlassen dort umfangreiche Datenspuren. Dabei dominieren – ungeachtet monopolistischer, gesamtwirtschaftlicher und datenschutzrechtlicher Bedenken – die bekannten Internetgiganten. Reichweite und Nutzenumfang der globalen Player überkompensieren „Nähe“ und höhere Datenschutzstandards nationaler Wettbewerber.
- Besorgte Bürger, aber sorglose Verbraucher – das Nutzer-Paradoxon findet demnach in unseren Daten weitgehende Bestätigung. Die Mehrzahl der Verbraucher ist sich dieses Konfliktes durchaus bewusst.
- Dennoch sind die Nutzer „im Kleinen“ nicht gänzlich datenschutz-passiv:
  - Datenschutzmaßnahmen an Computer und Smartphone werden – gerade von datenschutzsensiblen Bürgern – durchaus genutzt.
  - Die große Mehrheit verhält sich in Bezug auf Datenschutzmaßnahmen „pragmatisch“ (negativ formuliert ließe sich auch sagen: „leicht fahrlässig“). Jeder fünfte Nutzer zeigt sich „datenschutz-phlegmatisch“, nicht einmal jeder zehnte ist konsequent „datenschutz-aktiv“.

## 2. Neue Paradigmen für Politik, Wirtschaft und Gesellschaft?

Bürger befürworten ein *persönliches Eigentum* an den eigenen Daten, sind skeptisch gegenüber allen Arten *automatisierter Entscheidungen* und sind nur sehr begrenzt bereit, den Datenschutz *zugunsten der internationalen Wettbewerbsfähigkeit der Wirtschaft* hintenanzustellen.

Es herrscht breiter Konsens bezüglich der *Grundanforderungen* an datengestützte Systeme, nämlich *Fairness* und *Nachvollziehbarkeit*. Das impliziert *Transparenz*, *Vertrauen* oder *Kontrolle* sowie *Diskriminierungsfreiheit*, *Neutralität*, *Datenqualität* und *Datensicherheit*. (Die Forderungen nach *Datenqualität* und *Transparenz* stellen den Einsatz unstrukturierter Daten und selbstlernender Algorithmen aus technischer Sicht vor eine große Hürde, vgl. Wagner / Jost.)

Neue Paradigmen in der politischen Diskussion betonen *Datensouveränität* statt *Datenschutz*, freiwillige *Datenethik* der Unternehmen anstelle strikter *Regulierung*, sowie *Open-Data-Ansätze* anstelle oder ergänzend zum klassischen *Wettbewerbsrecht*.

- Die Konzepte von Datenschutz versus eigenverantwortlicher Datensouveränität des Einzelnen schließen sich aus Sicht der Befragten nicht aus, sondern werden gleichermaßen geteilt. Dennoch ist den meisten der Gedanke eher fremd, Daten im Sinne einer Ware zu „tauschen“ (auch wenn dies faktisch schon breite gesell-

schaftliche Realität ist) oder gar zu „verkaufen“ oder zu „spenden“. Ist der gesellschaftliche Nutzen aber offensichtlich (wie im Fall der medizinischen Forschung gegeben), erhöht sich die Bereitschaft zur Datenüberlassung ganz erheblich.

- Überwiegend fordern die Bürger jedoch eine enge staatliche Regulierung des Umgangs mit Daten. Der eigentlich gegenteilige Gedanke einer weitgehend freiwilligen Selbstverpflichtung der Unternehmen findet zwar ebenfalls eine recht breite Akzeptanz, dürfte aber im Detail stark vom Vertrauen in ebenjene abhängen.
- Ein Eingreifen des Staates gegen die Marktmacht der „Internetgiganten“ wird in Mehrheit unterstützt. Dabei findet auch das recht neue Modell der „Open Data“, bei dem der Wettbewerbsvorsprung von Monopolisten durch erzwungenes Data-Sharing beschnitten wird, um Innovation und Wettbewerb zu stärken, mehr Zustimmung als Ablehnung – und könnte bei intensiverer Kommunikation durchaus „mehrheitsfähig“ werden.

*Spion, Überwachungsapparat, Verführer, Droge und Diktator*: Wir haben die Befragten um Auskunft gebeten, wie weit sie positive und negative Analogien zum Einsatz von Big Data für zutreffend halten. Die fünf zuvor genannten Begriffe wurden mit Abstand meistgenannt und verdeutlichen die große Skepsis, die dem Thema aus der „Fernsicht“ entgegenschlägt. Positive Narrative, die den konkreten Nutzen neuer Anwendungen in die Vorstellungswelt der Menschen transportieren könnten, wie *Schutzengel, Wunscherfüller* oder *Heinzelmännchen*, folgen erst weit danach.

Die *Versicherungswirtschaft* könnte auf Basis von Daten und Vernetzung neue Rollen in der Gesellschaft übernehmen, indem sie sich vom *Kostenerstatter* wandelt zum „*Retter*“ oder „*Schadenmanager*“, zum „*Coach*“ oder „*Schadenverhüter*“ sowie zum „*Kontrollleur*“ oder „*Motivator*“.

- In der Summe zeigen sich die Befragten diesen neuen Rollen gegenüber durchaus aufgeschlossen.
- Deutlich zurückhaltender ist die persönliche Bereitschaft zur laufenden *Übermittlung der notwendigen Daten*, zum Beispiel durch Sensoren, Wearables oder auch durch den Arzt.
- Inwieweit ein Angebot – und auch die damit verbundene Weitergabe der Daten – im Einzelfall befürwortet wird, hängt mehr von dessen konkreter Ausgestaltung ab als von der dahinterstehenden Rolle. Gleiches gilt für die Frage, ob der Einsatz eher in der Lebenswelt *Mobilität, Wohnen* oder *Gesundheit* vorstellbar ist. Und selbst der Grad der Überwachung, der mit dem jeweiligen Angebot verbunden ist, tritt hinter die Attraktivität des jeweiligen Nutzenversprechens zurück.

In Summe lässt sich eine insgesamt hohe Bereitschaft der Bürger respektive Kunden erkennen, neue, erweiterte Rollen der Versicherer anzunehmen.

### 3. Fazit

Im Gesamtbild zeigt sich in den Ergebnissen ein doppeltes Paradoxon:

- Die Bürger fordern Datenschutz, beklagen übermächtige Konzerne und fühlen sich in ihrer Freiheit beschränkt – aber handeln nicht danach, sondern folgen „brav der Herde“.
- Die Bürger fürchten Big Data – aber sind gerne bereit, die aus der Vernetzung resultierenden Potenziale in ihre Lebenswelten einzulassen.

Die Antwort auf die Frage „Bürgerschreck oder Hoffnungsträger“ ist also weder das eine noch das andere, sondern muss lauten:

Big Data: Bürgerschreck und Hoffnungsträger!

Angesichts der weiter rasant fortschreitenden Vernetzung und des zunehmenden Einsatzes von künstlicher Intelligenz besteht die Gefahr, dass sich das doppelte Paradoxon nicht etwa auflöst, sondern weiter verfestigt. Dies kann in einer fatalistischen Haltung enden oder darin, dass auf Dauer eine Anpassung der Einstellungen an das eigene Verhalten erfolgt. Beides steht diametral zum Ideal eines „souveränen Datenbürgers“.

Ein politischer Paradigmenwechsel kann in Teilen dazu beitragen, Brücken zwischen den empfundenen Risiken und den erhofften Chancen zu bauen. Ebenso kann die Etablierung positiver gesellschaftlicher „Erzählungen“ helfen, das bisherige Schwarz-Weiß-Bild zu überwinden. Wesentliche Grundbedingungen dazu sind – neben dem eigentlichen Nutzen der Anwendungen – Fairness und Sicherheit der Systeme sowie Transparenz und Vertrauen in die Akteure. Und schließlich bleibt essenziell, die digitale Bildung der Bevölkerung auch in Bezug auf Big Data zu intensivieren, um Akzeptanz zu gewinnen und dem Ideal eines souverän handelnden Bürgers zumindest in Schritten näherzukommen.

## Gliederung

1. Aufbau und Methode	
2. Datenwissen	„Wissen“
3. Handlungsfreiheit	„Können“
4. Folgeabschätzung	
5. Anwendungsfelder	„Wollen“
6. Einstellungen	
7. Verhalten	„Handeln“
8. Datenpolitik und Datenethik	
9. Alte und neue Narrative	„Neue Paradigmen“
10. Neue Rollen – am Beispiel der Versicherungswirtschaft	



Vernetzung und Big Data, Künstliche Intelligenz und die damit verbundenen Algorithmen bestimmen zunehmend die Lebenswelt der Bürger, sind bisher aber erst wenig in den Köpfen angekommen. Bisherige Umfragen zeigen einen meist geringen Informationsgrad und eine tendenziell skeptische Haltung, aber zugleich eine umfassende, oft wenig bewusste Nutzung der Potenziale.

Aufbauend auf den vorliegenden Studien, theoretischen Grundüberlegungen sowie den Ausführungen von Knorre (Teil A) und Wagner / Jost (Teil B) soll im Folgenden das Verhältnis der deutschen Bevölkerung in Bezug auf Daten und Algorithmen erfasst und analysiert werden. Dabei werden einerseits „Wissen“ (Kap. 2) und „Können“ (Kap. 3) als Voraussetzung souveränen Handels beschrieben. Danach wird in Form von Folgeabschätzungen (Kap. 4), die Bewertung von Anwendungsfeldern (Kap. 5, mit besonderem Schwerpunkt auf den drei Lebenswelten „Mobilität“, „Wohnen“ und „Gesundheit“) sowie durch die Erfassung von Einstellungen zum Datenschutz (Kap. 6) das „Wollen“ der Bürger abgebildet.

Dem Wollen, Wissen und Können wird in Kapitel 7 das tatsächliche „Handeln“ gegenübergestellt. Wie verhalten sich die Bürger? In welchem Umfang werden Geräte und Dienste genutzt und welche aktiven Maßnahmen zum Schutz der persönlichen Daten werden tatsächlich ergriffen?

Anschließend werden die bei Knorre thematisierten Paradigmen zu Datenpolitik und Datenethik aus Sicht der Bürger beleuchtet (Kap. 8) sowie die Frage, inwieweit den zu beobachtenden und weitgehend negativ besetzten Narrativen rund um Big Data auch positive „Erzählungen“ gegenüberstehen (Kap. 9).

Abschließend werden in Kapitel 10 am Beispiel der Versicherungswirtschaft Erwartungen an neue Rollen von Marktteilnehmern dargestellt.

## Methode: 1.000 Befragte als „Multimode“-Erhebung

### 500 Telefoninterviews

bevölkerungsrepräsentative Zufallsstichprobe nach dem ADM- Mastersample

350 Interviews über das Festnetz, 150 über Mobilfunk,

Feldinstitut OmniQuest, Befragungsdauer 25 Minuten, November 2018

### 500 Onlineinterviews

bevölkerungsrepräsentativ quotierte Stichprobe im OmniQuest-Panel (mehrfach geschichtete Quotenstichprobe, quotiert nach Alter, Haushaltsgröße, Geschlecht, Region und Bildung)

Befragungsdauer 15 Minuten, November 2018

- Reduktion von Verzerrungen durch die Kombination zweier Erhebungsformen und Stichprobenverfahren.
- Kontrolle von Abweichungen durch Vergleich zwischen den Teilstichproben und Abgleich mit Bevölkerungsdaten
- Vermeidung von Reihenfolgeeffekten durch Rotation und von Ausstrahlungseffekten durch Stichprobensplits
- Weitestgehend „repräsentatives“ Abbild der deutschen Bevölkerung\*

Die Auswertung basiert auf einer Befragung von 1.000 Bürgern ab 18 Jahren. Diese Gesamtstichprobe wurde in zwei Teilstichproben mit unterschiedlicher Erhebungsform aufgeteilt. Dieser sogenannte Mixed-Mode-Ansatz hat den Vorteil, dass sowohl Personengruppen, die in Telefoninterviews schwer zu erreichen sind, als auch solche, die durch Online-Erhebungen nicht ansprechbar sind, in der Gesamtstichprobe ausreichend abgebildet sind, und dass sich auch Effekte der Interviewsituation in der Gesamtstichprobe ausgleichen.

Um Reihenfolgeeffekte zu vermeiden, wurden die Items innerhalb einer Frage nach dem Zufallsprinzip rotiert. Zudem wurde ein Teil der Fragen jeweils nur der (zufällig ausgewählten) Hälfte der Befragten gestellt um die Gesamtlänge des Fragebogens zu verkürzen („split half“)<sup>1</sup>.

Die realisierte Stichprobe wurde auf Basis der aktuellen Verbraucheranalyse „best 4 planning 2018 I“ mit der Struktur der deutschen Wohnbevölkerung abgeglichen. Auf eine nachträgliche Gewichtung der Antworten wurde aufgrund der insgesamt hohen Übereinstimmung verzichtet.

**Fazit:** Im Ergebnis liefert die Erhebung ein weitgehend „repräsentatives“ Abbild der deutschen Bevölkerung beziehungsweise – soweit die Fragen sich nur auf Internetnutzung beziehen, der online-aktiven Bevölkerung<sup>2</sup>.

---

<sup>1</sup> Bei der gegebenen Stichprobengröße von 1.000 Befragten ist bei Anteilswerten und einem Signifikanzniveau von 95% mit einer zufälligen Fehlerabweichung von ca. +/- 3% zu rechnen. Sofern Teilstichproben von ca. 500 Befragten vorliegen, erhöht sich der Fehlerbereich auf ca. +/- 4%. Die Stichprobengröße ist jeweils bei der Frage angegeben (n=...)

<sup>2</sup> Nichtsdestotrotz ist nicht auszuschließen, dass durch Selektionseffekte (Teilnahmebereitschaft an Telefoninterviews einerseits bzw. an einem Online-Panel andererseits) eine gewisse Verzerrung hin zu höherer Datenkompetenz und zu höherer Akzeptanz von Datenaustausch besteht. Zu den grundsätzlichen Grenzen der Repräsentativität von Befragungen vgl. Müller-Peters 2012 sowie Müller-Peters 2018.

## Gliederung

1.	Aufbau und Methode	
2.	<b>Datenwissen</b>	„ <i>Wissen</i> “
3.	<b>Handlungsfreiheit</b>	„ <i>Können</i> “
4.	Folgeabschätzung	
5.	Anwendungsfelder	„ <i>Wollen</i> “
6.	Einstellungen	
7.	Verhalten	„ <i>Handeln</i> “
8.	Datenpolitik und Datenethik	
9.	Alte und neue Narrative	„ <i>Neue</i> “
10.	Neue Rollen	<i>Paradigmen</i> “

Voraussetzung eines mündigen Umgangs mit den eigenen digitalen Daten ist zuerst einmal eine grundlegende *Kenntnis* des Themas. (Zur Relevanz von Fähigkeit und Motivation für überlegte Entscheidungen vergleiche z.B. Petty und Cacioppo 1986). Um das Wissen um Big Data und künstliche Intelligenz in der Bevölkerung zu erfassen, wurden eine Liste von Begriffen erstellt und die Interviewpartner befragt, inwieweit sie diese einem Bekannten erklären könnten. Die Stichworte umfassen sowohl Fachbegriffe als auch umgangssprachliche Formulierungen zu technischen Grundlagen und deren Anwendungen, zur Gesetzgebung und schließlich zu einigen Daten-Skandalen, die in den Medien breite Aufmerksamkeit erhalten haben.<sup>3</sup>

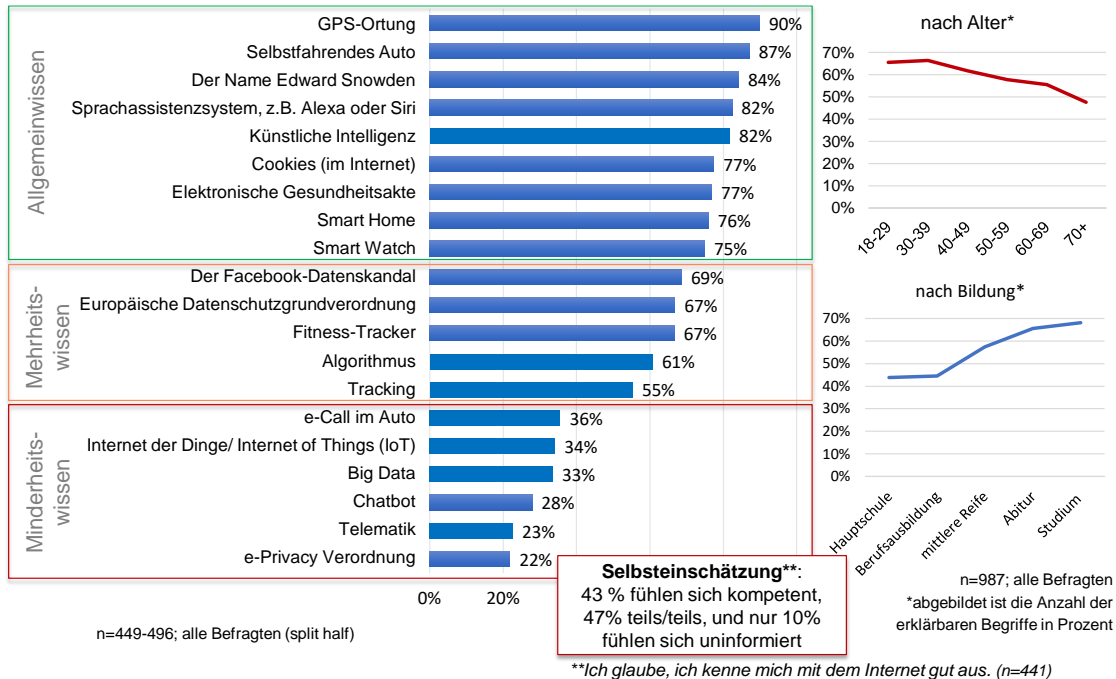
Neben dem Wissensstand ist auch die (gefühlte) *Wahlfreiheit* eine Bedingung für Souveränität im Umgang mit Daten. Einerseits haben auf der rechtlichen Seite die EU-Datenschutzgrundverordnung sowie in Folge das Bundesdatenschutzgesetz weitgehende Schutz- und Verfügungsrechte für die Bürger sowie umfassende Zustimmungspflichten installiert. Andererseits schränken Informationslücken, die Komplexität von Datenschutzerklärungen, digitale Ökosysteme und monopol-ähnliche Marktstrukturen die faktische Handlungsfreiheit in hohem Maße ein.

<sup>3</sup> Es liegen bereits eine Reihe aktueller empirischer Studien vor, in denen der Wissensstand zu Big Data, zu Algorithmen oder zu Künstlicher Intelligenz erfasst wurde, siehe zum Beispiel Bertelsmann-Stiftung 2018, GfK 2018, Marden 2017, PWC 2017 oder YouGov 2018, sowie im Überblick Knorre in diesem Band. Dabei wird aber in der Regel nur auf einzelne Begriffe abgezielt, während hier ein breiter definiertes „digitales Wissen“ der Bevölkerung als Voraussetzung souveränen Handelns gemessen werden sollte.



## Datenwissen als Bedingung für Souveränität

Könnten Sie einem Bekannten erklären, was mit den folgenden Stichworten gemeint ist?



Wie die Befragungsergebnisse zeigen, sind wesentliche Grundbegriffe der Digitalisierung bereits in das *Allgemeinwissen* eingegangen. Das betrifft vor allem weitverbreitete Anwendungen (wie GPS-Ortung oder Cookies) sowie plakative Begriffe aus der öffentlichen Diskussion (wie Selbstfahrendes Auto, Künstliche Intelligenz, aber auch der Name Edward Snowden). Bemerkenswert ist, wie schnell auch neue Angebote den Weg in die Köpfe gefunden haben (wie Smart Home, Smart Watch oder Alexa).

Weitere Begriffe sind zumindest einer *Mehrheit* bekannt, z.B. aktuelle politische Themen wie der Facebook-Datenskandal oder die Europäische Datenschutz-Grundverordnung, aber auch Termini wie Tracking oder Algorithmus (was nicht immer auch ein genaueres Verständnis beinhaltet).

Schließlich gibt es Begriffe, die nur einer *Minderheit* vertraut sind. Das sind vor allem Anwendungen, die relativ neu sind (wie Chatbot oder e-Call), aber auch grundlegende Begriffe digitaler Vernetzung, die eher im Hintergrund hinter sicht- und erlebbareren Anwendungen stehen (wie Big Data, Telematik, Internet of Things). Bezüglich des „Schlusslichtes“ *e-Privacy-Verordnung* ist anzumerken, dass sich diese politisch noch im Vorbereitungsstadium befindet.

Entsprechend der Prägung der Generationen durch digitale Techniken ist der Wissensstand der jüngeren Alterskohorten höher, wobei selbst die Gruppe 70+ noch annähernd jeden zweiten Begriff erklären kann. Ebenso steigt das Wissen mit dem Bildungsgrad und ist bei Männern im Durchschnitt etwas höher als bei Frauen.

**Fazit:** Die digitale Welt ist kein „Neuland“ mehr: Grundbegriffe der Digitalisierung sind angekommen, und ein Großteil der Bevölkerung fühlt sich zumindest in Teilen kompetent. Wissensdefizite bestehen aber gerade in Themen, die die Vernetzung von Daten behandeln und damit den souveränen Umgang mit den eigenen Daten besonders tangieren (z.B. Big Data, Telematik, IOT und e-Privacy).