Fraunhofer
SIT

FRAUNHOFER-INSTITUT FÜR SICHERE INFORMATIONSTECHNOLOGIE

SIT TECHNICAL REPORTS

# WEB-TRACKING-REPORT 2014

02/2014

CASED

EC SPRIDE

FRAUNHOFER VERLAG


# Web-Tracking-Report 2014

Markus Schneider, Matthias Enzmann, Martin Stopczynski

Hrsg. Michael Waidner

SIT Technical Reports
SIT-TR-2014-01

Februar 2014

Fraunhofer-Institut für Sichere
Informationstechnologie SIT
Rheinstraße 75
64295 Darmstadt

FRAUNHOFER VERLAG

# IMPRESSUM

**Kontaktadresse:**
Fraunhofer-Institut für
Sichere Informationstechnologie SIT
Rheinstraße 75
64295 Darmstadt
Telefon 06151 869-213
Telefax 06151 869-224
E-Mail info@sit.fraunhofer.de
URL www.sit.fraunhofer.de

**Bibliografische Information der Deutschen Nationalbibliothek**
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Hrsg. Michael Waidner
SIT Technical Reports
SIT-TR-2014-01: Web-Tracking-Report 2014
Markus Schneider, Matthias Enzmann, Martin Stopczynski
ISBN: 978-3-8396-0700-8
ISSN: 2192-8169

Druck und Weiterverarbeitung:
IRB Mediendienstleistungen
Fraunhofer-Informationszentrum Raum und Bau IRB, Stuttgart

Für den Druck des Buches wurde chlor- und säurefreies Papier verwendet.


Fraunhofer-Informationszentrum Raum und Bau IRB
Postfach 800469, 70504 Stuttgart
Nobelstraße 12, 70569 Stuttgart
Telefon 0711 970-2500
Telefax 0711 970-2508
E-Mail verlag@fraunhofer.de
URL http://verlag.fraunhofer.de

# INHALTSVERZEICHNIS

# Web-Tracking-Report 2014

Markus Schneider, Matthias Enzmann, Martin Stopczynski
Fraunhofer-Institut für Sichere Informationstechnologie SIT, Darmstadt

---

Durch die Enthüllungen von Edward Snowden im vergangenen Jahr ist öffentlich geworden, in welchem Umfang Geheimdienste Verbraucher bei ihrer Internetnutzung und bei ihrer Kommunikation überwacht haben. Doch die heimliche Verfolgung von Verbrauchern im Internet ist nicht nur den Geheimdiensten vorbehalten: Viele Unternehmen sind fleißig damit beschäftigt, möglichst viele Daten über Verbraucher zu sammeln. Personenbezogene Daten und Informationen sind für sie ein wichtiger Rohstoff für viele neue Geschäftsmodelle, mit denen man hohe Umsätze erzielen kann.

Unternehmen fördern den Rohstoff *Daten* heute in großem Stil durch Web-Tracking. Mittels Web-Tracking können dritte Parteien nichtsahnende Verbraucher im Hintergrund bei deren Internetnutzung verfolgen und somit umfangreiche Einblicke in die Interessen, Wünsche, Probleme oder den Konsum von Verbrauchern gewinnen und Profile von Verbrauchern erstellen. Dabei können Tracker Verbraucher über Internetseiten verschiedener Anbieter hinweg verfolgen. Das so gewonnene Wissen über Verbraucher wird von Trackern heute hauptsächlich für gezielte Werbung verwendet. Es sind jedoch auch andere Verwertungsarten der durch Web-Tracking gewonnenen Daten möglich, die für einzelne Verbraucher oder auch die ganze Gesellschaft eine Reihe von Bedrohungen und Risiken mit sich bringen.

Dieser Report stellt verschiedene Aspekte des Themas *Web-Tracking* dar. Der Report erklärt, wie Web-Tracking, insbesondere Cross-Domain-Tracking, grundsätzlich funktioniert. Darüber hinaus werden Markt und Akteure im Zusammenhang mit dem Sammeln von Daten durch Web-Tracking beschrieben. Es werden ebenfalls die verschiedenen technischen Ansätze zur Implementierung von Trackingmethoden dargestellt. Neben Web-Tracking werden noch alternative Formen des Tracking beschrieben, insbesondere App-Tracking und Smart-TV-Tracking. Der Report beschreibt danach die Bedrohungen, Gefahren und Risiken für den einzelnen Verbraucher und für die Gesellschaft. Hierbei werden viele tatsächliche Ereignisse gesammelt und dargestellt, die in den vergangenen Jahren durch die Medien oder andere wissenschaftliche Arbeiten und Dokumentationen bekannt geworden sind. Danach wird kurz erklärt, wie man sich heute gegen Web-Tracking schützen kann. Darüber hinaus wurde eine quantitative Analyse auf Basis aktueller Datenerhebungen zu Web-Tracking durchgeführt, mittels derer dargestellt wird, in welchem Umfang man heute als Verbraucher aus Deutschland mit Web-Tracking zu rechnen hat, wenn man sich nicht dagegen schützt.

Der Report richtet sich sowohl an Experten im Bereich Informationssicherheit und Datenschutz als auch an normale Anwender. Für Experten, die bereits Hintergrundwissen zu Web-Tracking haben, stellt der Report Fakten zusammen und liefert eine umfangreiche Materialsammlung. Der Report sollte jedoch auch für normale Anwender ohne Informatikfachwissen verständlich sein. Mit dem Web-Tracking-Report können auch alle diejenigen, die das World Wide Web täglich nutzen, sich jedoch bisher weniger mit dem Thema *Web-Tracking* auseinandergesetzt haben, verstehen, was Web-Tracking für sie und ihre Lebenswirklichkeit bedeutet.

Key Words: Web-Tracking, Tracking, Privacy, Datenschutz, Trackingschutz, Do Not Track, Third-Party-Cookie, Flash-Cookie, Internetwerbung

---

Kontaktadresse: Fraunhofer-Institut für Sichere Informationstechnologie SIT, Rheinstrasse 75, D-64295 Darmstadt, Web: `http://www.sit.fraunhofer.de`, E-Mail: {vorname.nachname}@sit.fraunhofer.de

# 1. ZUSAMMENFASSUNG

**Einleitung**

Massenüberwachung und Spionage durch staatliche Organisationen wie Geheimdienste sind seit den Enthüllungen von Edward Snowden in der Öffentlichkeit intensiv diskutiert worden. Experten und Medien haben in diesem Zusammenhang versucht, zur Aufklärung und zum Verstehen dieser Vorgänge beizutragen. Jedoch sammeln nicht nur Geheimdienste Daten von Bürgern und Verbrauchern; auch Unternehmen sammeln Daten von Verbrauchern in großem Stil und dies, ohne dass Verbraucher davon etwas mitbekommen und ohne dass Verbraucher mit diesen Unternehmen bewusst in Kontakt treten.

Ein stark verbreitetes Mittel, um an diese Daten zu gelangen, ist Web-Tracking. Web-Tracking betrifft heute die ganze Gesellschaft: Zunächst natürlich diejenigen Personen, die häufig im Word Wide Web unterwegs sind, und in seinen Auswirkungen sogar auch solche Personen, die das World Wide Web gar nicht nutzen. Das Ziel des Web-Tracking-Reports lag darin, verschiedene inhaltliche Aspekte des Web-Tracking zu sammeln und Web-Tracking zu beschreiben. Der Schwerpunkt des Web-Tracking-Reports liegt in der Beschreibung dessen, was Web-Tracking für den Einzelnen und die Gesellschaft bedeutet und welche Bedrohungen und Risiken sich durch Web-Tracking ergeben. Hierfür wurden bekannt gewordene Fälle aus der ganzen Welt zusammengetragen, die zeigen, was mit den Daten von Verbrauchern geschieht bzw. geschehen kann und zu welchen Konsequenzen das führen kann. Darüber hinaus beschreibt der Web-Tracking-Report das heutige Ausmaß von Web-Tracking für die in Deutschland beliebtesten Webseiten.

**Was bedeutet Web-Tracking?**

Beim Web-Tracking geht es darum, dass Verbraucher von dritten Parteien —sogenannten Trackern— verfolgt werden, wenn sie im World Wide Web unterwegs sind. Tracker können dabei sehen, welche Webseiten Verbraucher aufrufen. Web-Tracking funktioniert als *Cross-Domain-Tracking* auch über Webseiten verschiedener Anbieter hinweg. All das geschieht im Hintergrund, ohne dass ein Verbraucher davon etwas mitbekommt.

Tracker tun dies, um Verbraucherdaten zu bekommen. Diese Daten sind ein wichtiger Rohstoff für viele Unternehmen. Wenn man weiß, welche Seiten ein Nutzer abruft, dann weiß man auch, für was er sich interessiert, was er gerne konsumiert, welche politische Einstellung er hat, wie alt er ungefähr ist, ob er allein lebt oder Familie hat, welchen Lebensstil er pflegt und vieles mehr. Diese Daten werden mit Web-Tracking eingesammelt. Solche Informationen werden beispielsweise für zielgerichtete Werbung benötigt.

**Der Markt rund um Web-Tracking**

Der Markt rund um Verbraucherdaten ist ein weltweiter Milliardenmarkt. In diesem Markt spielt Onlinewerbung eine herausragende Rolle. Bisher wurde Web-Tracking hauptsächlich durch die Werbeindustrie getrieben. Mittels Web-Tracking kann ganz zielgerichtet geworben werden. Mit einer geladenen Webseite bekommt jeder Verbraucher die individuell auf ihn zugeschnittene Werbung. Damit dies gelingen kann, müssen die Werbetreibenden möglichst viel über jeden einzelnen Verbraucher wissen. Werbeanzeigen orientieren sich damit nicht mehr an dem Inhalt einer geladenen Seite, sondern an dem, was der Werbetreibende über den Nutzer weiß. Je besser die Werbung auf den Konsumenten angepasst ist, desto höher ist die Erfolgsquote von Werbung und umso höher ist der Verdienst der Werbetreibenden.

Jedoch ist die Verwertung der durch Web-Tracking gewonnen Daten nicht auf Werbung beschränkt. Verbraucherdaten sind ein wichtiger Produktionsfaktor für viele Geschäftsmodelle, bei denen solche Daten erforderlich sind und verarbeitet werden. Auch wenn Daten ursprünglich für Werbung eingesammelt wurden, so können sie später auch für weitere Geschäftsmodelle verwendet werden. Es werden bereits einige Geschäftsmodelle im Zusammenhang mit im World Wide Web gesammelten Daten praktiziert, die über Werbung hinausgehen, wie z.B. Auskunftsdienste zu Personen im Internet, Scoring der Kreditwürdigkeit und die Abschätzung des Kundenrisikos für Krankenversicherungen.

Es gibt heute eine Reihe von Geschäftsmodellen, die mit verschiedenen Angeboten zur Erhebung von Verbraucherdaten einhergehen:

(1) Suchmaschinen
(2) Web2.0-Dienste
(3) E-Mail-Dienste
(4) Online-Textverarbeitung
(5) Cloud-Speicherdienste
(6) Content-Angebote
(7) Hosting-Dienste
(8) Web-Tracking bzw. Tracking allgemein[1]

Die unter (1)–(7) genannten Anwendungen unterscheiden sich von (8) Web-Tracking dadurch, dass sich Verbraucher bewusst für einen bestimmten Dienstanbieter entscheiden können und dabei auch abwägen können, ob sie dem jeweiligen Dienstanbieter bestimmte Daten geben möchten. Beim Web-Tracking ist dies anders, da Verbraucher nicht erkennen können, zu welchen dritten Parteien personenbeziehbare Informationen fließen. Die unter (1)–(7) genannten Angebote sind teilweise als Alternativen zur Datensammlung durch Web-Tracking zu verstehen, teilweise verstärken sie die Möglichkeiten des Web-Tracking auch, indem sie dazu beitragen, dass in bestimmten Fällen Tracker die durch Tracking gewonnenen Daten sehr einfach der korrekten Identität des Verbrauchers zuordnen können. Damit können die erhobenen Daten sogar außerhalb der Onlinewelt verwendet werden.

Zusätzlich können Unternehmensakquisitionen dazu beitragen, dass Personendaten aus verschiedenen Quellen miteinander kombiniert werden können. Somit können umfassende und genaue Profile von Personen entstehen.

Bisher hat die Politik den Markt rund um die Erhebung von Daten durch Web-Tracking relativ wenig reguliert und damit die Verbraucher vergleichsweise wenig geschützt. Eine EU-Richtlinie aus dem Jahr 2009, nach welcher Verbraucher gegen Web-Tracking besser zu schützen sind, wurde in Deutschland bisher nicht in nationales Recht umgesetzt. Ein neuer Vorstoß der EU mittels einer Verordnung wird vor den Europawahlen 2014 nicht verbindlich umgesetzt werden. Um die Entwicklung des rechtlichen Rahmens in den USA und in Europa mit steuern zu können, hat die Werbeindustrie an dem Do-Not-Track-Standard mitgearbeitet. Diese Lösung ist jedoch in mehrerer Hinsicht kein starker Schutz gegen Web-Tracking, wie bei den Schutzmaßnahmen gegen Web-Tracking beschrieben wird.

### Technische Betrachtung

Technisch betrachtet geht es den Trackern beim Web-Tracking hauptsächlich um zwei Ziele:

(1) Nach dem Aufruf einer Webseite durch den Verbraucher muss der Browser dazu gebracht werden, eine Verbindung zu dem Tracker aufzubauen, damit dieser erfährt, welche Webseite der Verbraucher geladen hat.
(2) Ein Tracker möchte verschiedene Webseitenaufrufe eines Verbrauchers diesem zuordnen können. Hierfür muss es möglich sein, den Verbraucher wiederzuerkennen, auch wenn eine bestimmte Zeit zwischen den Kontakten mit einem Tracker liegt.

[1] Wie später noch dargestellt wird, gibt es neben dem Web-Tracking noch weitere Formen des Tracking.

Damit dies gelingt, werden kleine Code-Elemente in Webseiten eingebettet. Lädt ein Verbraucher diese Webseiten, dann werden diese Code-Elemente mit den restlichen Webseiteninhalten heruntergeladen. Diese Code-Elemente sorgen dann dafür, dass im Hintergrund weitere Verbindungen aufgebaut werden. Über diese Verbindungen werden beispielsweise Bilder, Zählpixel, Landkarten, Videos, I-Frames oder aktive Inhalte wie JavaScript, Active X oder Flash übertragen. Mit jeder neuen Verbindung, die aufgebaut wird, sendet der Browser mit dem Referer-Parameter die URL der Webseite, die angezeigt wurde, bevor die neue Verbindung aufgebaut wurde. Über diesen Mechanismus fließen Informationen zur Webnutzung direkt zu den Trackern. Bei einigen Anbietern erfahren Tracker außer der Adresse der aufgerufenen Seite auch noch Informationen zur Benutzeridentität, die von einigen Anbietern in die URL der aufgerufenen Webseite eincodiert wird und somit mit dem Referer an den Tracker übertragen wird.

Um Verbraucher wiedererkennen zu können, gibt es für Tracker ein großes Spektrum von technischen Möglichkeiten, die man grob in zwei Klassen unterteilen kann:

- *Explizites Tracking*: Hier übertragen Tracker Daten auf den Computer eines Verbrauchers, wo sie gespeichert werden und bei einem späteren Kontakt zur Wiedererkennung wieder abgefragt werden. Dazu gehören
  - Third-Party-Cookies
  - Flash-Cookies
  - E-Tags
  - HTML5-Storage
  - Evercookies: Evercookies kombinieren verschiedene Trackingmethoden, z.B. die oben genannten. Löschen Nutzer einzelne Identifikatoren, die durch Trackingmethoden angelegt werden, dann stellen Evercookies die gelöschten Identifikatoren wieder her.
- *Schließendes Tracking*: Hierbei versuchen Tracker den Computer eines Verbrauchers anhand von Konfigurationsaspekten und Systemeigenschaften von den Computern anderer Verbraucher zu unterscheiden. Dazu gehören
  - IP-Adressen
  - Fingerprinting: Für Fingerprinting können Tracker Konfigurationsaspekte auf verschiedenen Ebenen abfragen:
    - Browsererweiterungen
    - Browsereinstellungen
    - Browserfamilie und Browserversion
    - Betriebssysteme und Anwendungen
    - Hardware

Tracker können mehrere Trackingmethoden gleichzeitig verwenden. Tracker sind stets damit beschäftigt, neue Trackingmethoden zu entwickeln, mit denen möglichst viele Verbraucher eindeutig, stabil und robust gegen Schutzmaßnahmen wiedererkannt werden können.

**Weitere Formen des Tracking**

Tracking ist nicht nur auf die Nutzung des World Wide Web beschränkt. Informations- und Kommunikationstechnologie wird zunehmend in Alltagsgegenständen und in der Unterhaltungselektronik verwendet. Dort werden Innovationen hauptsächlich durch Informations- und Kommunikationstechnologie getrieben. Die Internetfähigkeit dieser Geräte schafft die technische Voraussetzung für Tracking. Als Beispiele für Tracking außerhalb des World Wide Web werden in diesem Report insbesondere das App-Tracking auf Smartphones und das Smart-TV-Tracking für moderne internetfähige Fernsehgeräte erwähnt.

Mittels Smartphone-Apps können Verbraucherdaten zu dritten Parteien gelangen. Für Entwickler von Smartphone-Apps bestehen neben den Einnahmen aus dem App-Verkauf

weitere Einnahmemöglichkeiten im Zusammenhang mit Tracking, z.B. durch Werbung. Hierfür stehen sogenannte Tracking-Frameworks bereit, mittels derer die Programmierung von Trackingmethoden in Apps sehr effizient umgesetzt werden kann. Mit diesen Frameworks kann sehr einfach beschrieben werden, welche Daten und Aktionen der Benutzer erfasst und an Tracker gesendet werden sollen. Das Fraunhofer SIT hat die 400 populärsten Apps aus der Rubrik *Utilities* aus dem App Store von Apple auf das Enthaltensein von Tracking-Frameworks untersucht[2] und in 72% dieser Apps Tracking-Frameworks entdeckt. In 56% dieser Apps wurde sogar mehr als ein Tracking-Framework gefunden. Der Spitzenwert von Tracking-Frameworks je App lag bei 12 Stück.

Mit Smart-TV-Tracking können Daten zum Fernsehkonsum von Verbrauchern an Dritte versendet werden, ohne dass die Verbraucher hiervon etwas mitbekommen. Das Ziel dieser Technologie liegt in der Auswertungsmöglichkeit hinsichtlich Einschaltquote und in der Bereitstellung zielgerichteter Werbung. Wie Wissenschaftler der TU Darmstadt herausgefunden haben, versenden Smart-TV-Geräte Nutzungsdaten bereits vor der Aktivierung des Internetmodus durch den Nutzer. Insofern wird bereits Tracking durchgeführt, wenn Verbraucher davon ausgehen, dass sie noch nicht mit dem Internet verbunden sind.

Tracker versuchen, die über App-Tracking, Smart-TV-Tracking und Web-Tracking gewonnenen Daten über Geräte hinweg miteinander zu kombinieren. Dadurch lassen sich umfangreiche Profile mit personenbezogenen Daten anlegen.

**Bedrohungaspekte**

Es gibt viele Aspekte, welche Auswirkungen auf die Bedrohung von Verbrauchern durch Tracking haben. Diese werden im Folgenden aufgezählt:

(1) Bedrohungsaspekte im Zusammenhang mit der Erhebung von Daten
   (a) Internet der Dinge: Die Entwicklung des Internet der Dinge trägt dazu bei, dass weitere Datenquellen entstehen, über welche Daten erhoben werden, die mit den Daten aus dem Web-Tracking kombiniert werden können.
   (b) Internet der Dienste: Das Internet der Dienste bzw. Cloud-Computing trägt dazu bei, dass Dienstleister weitere Einblicke in Verbraucherdaten bekommen.
   (c) Eingeschränkte Abwehr und neue Trackingmethoden: Viele Verbraucher wissen nicht, wie sie sich gegen Tracking schützen sollen. Darüber hinaus werden neue Trackingmethoden entwickelt, die gegen bestimmte Abwehrmaßnahmen immun sind.
   (d) Ressourcenverbrauch: Durch Tracking werden im Hintergrund viele Verbindungen aufgebaut und Daten ausgetauscht, wodurch Ressourcen von Nutzern verbraucht werden, z.B. Zeit, Übertragungskosten und Energie.

(2) Bedrohungsaspekte im Zusammenhang mit der Verarbeitung von Daten
   (a) Kombinierbarkeit von Daten und Big Data: Daten aus verschiedenen Quellen können miteinander kombiniert werden, so dass umfangreiche Datensammlungen in Form von Personenprofilen entstehen. Wenn Big-Data-Algorithmen in der Menge der zu einer Person gesammelten Daten Ähnlichkeiten mit anderen Mustern erkennen, dann werden Verbraucher in entsprechende Kategorien einsortiert.
   (b) Verknüpfbarkeit mit echten Identitäten: Die über Tracking gewonnenen Daten können mit der echten Identität verbunden werden und sind somit auch außerhalb der Onlinewelt verwertbar. Agieren Online-Anbieter, bei denen Nutzer mit Klarnamen

[2]Diese Untersuchung wurde mit Appicaptor vorgenommen. Appicaptor ist ein am Fraunhofer SIT entwickeltes Werkzeug, mit dem Apps auf Sicherheitslücken hin untersucht werden können. (siehe `https://www.sit.fraunhofer.de/de/angebote/projekte/appicaptor/`)

registriert sind, auch als Tracker, dann können die Tracking-Daten mit der echten Identität in Zusammenhang gebracht werden. Es gibt auch einbettende Anbieter, die Tracker mit der echten Identität versorgen.

(c) Rückschlüsse aus sozialen Beziehungen: Tracker analysieren soziale Beziehungen zwischen Verbrauchern und verwerten diese in Profilen. Hier werden dann Erkenntnisse zu einer Person ggf. auf andere Personen übertragen, die mit dieser Person in Kontakt stehen.

(3) Bedrohungsaspekte im Zusammenhang mit der Verwertung von Daten und dem Markt

(a) Weitergabe von Daten und Datenhandel: Personenbezogene Daten sind ein wertvolles Wirtschaftsgut. Die Nachfrage nach Daten ist hoch und sie lassen sich gut vermarkten. Sie werden absichtlich weitergegeben oder auch gestohlen und können so bei Organisationen landen, welche von Verbrauchern abgelehnt werden.

(b) Reichweite von Trackern: Je stärker verbreitet Tracker sind, desto besser können sie Verbraucher verfolgen und desto geschlossener wird das Bild von Verbrauchern, welches Tracker von ihnen gewinnen.

(c) Neue Geschäftsmodelle: Die Sammlung von Daten ist nicht auf zielgerichtete Werbung beschränkt. Wenn Tracker erst einmal über den Rohstoff Verbraucherdaten verfügen, können sie neue Geschäftsmodelle zur Vermarktung entwickeln und einführen.

(d) Irreführung von Verbrauchern: Das nichtvorhandene Wissen auf Verbraucherseite wird ausgenutzt. Wenn Elemente von dritten Parteien in Webseiten eingebettet und bei jeder Verwendung von dort geladen werden, werden Verbraucher darüber im Unklaren gelassen, dass ihr Computer Verbindungen zu dritten Parteien aufbaut.

(4) Bedrohungsaspekte im Zusammenhang mit dem Verhalten von Verbrauchern

(a) Fehlendes Verbraucherwissen und Intransparenz: Viele Verbraucher wissen weder, was Tracking ist, noch, dass es stattfindet, und können auch die Auswirkungen nicht einschätzen. Daher setzen viele keine Werkzeuge zum Trackingschutz ein.

(b) Änderung von Vertrauensbeziehungen: Durch Unternehmensakquisitionen können Daten in die Hände anderer Unternehmen gelangen.

(c) Durchsetzung der Löschung von Daten: Für Verbraucher ist es praktisch unmöglich, die von Trackern gesammelten Daten löschen zu lassen.

**Bedrohungen und Risiken**

Durch Tracking entstehen echte Bedrohungen und Risiken für Werte von Einzelnen und der ganzen Gesellschaft. Im Folgenden werden solche Risiken kurz dargestellt:

(1) Digitale Analogien bekannter Risiken und Gefahren

(a) Stalking: Die Verfolgung von Verbrauchern im World Wide Web kann mit Stalking verglichen werden. Für Verbraucher stellt sich dies so dar, als würden sie bei jedem Schritt von einer ganzen Herde von Dritten verfolgt werden, die jeden dieser Schritte aufzeichnen und auswerten.

(b) Erfassung von Verbindungsdaten: Beim Tracking können unbekannte Dritte die Verbindungsdaten von Verbrauchern für Webseitenaufrufe erfassen.

(2) Gefahren und Risiken für ideelle Werte

(a) Die digitale Vermessung des Einzelnen: Beim Tracking erfassen unbekannte Dritte alle Facetten von Verbrauchern, die in irgendeiner Weise von wirtschaftlicher

Relevanz für den Tracker sein können. Hierzu gehören Eigenschaften, Gewohnheiten, Vorlieben, Abneigungen, soziale Beziehungen, Aufenthaltsorte, Gesundheitszustand, finanzielle Situation, Gemütsverfassung, Meinung, berufliche Situation, Konsumverhalten, Wünsche und Zukunftspläne von Verbrauchern.

(b) Auskunftsdienste für jedermann: Die Verwertung der durch Tracking gewonnenen Daten ist nicht auf zielgerichtete Werbung beschränkt. So gibt es mittlerweile Auskunftsdienste, welche die Privatsphäre von Personen erheblich verletzen. Bei einem Auskunftsdienst mit Schwerpunkt USA kann jeder über das Internet für wenige Cent zu einer gegebenen Person Informationen wie das Haushaltseinkommen, Familienstand, Ausbildung, Hobbys, Interessen und Wert des Immobilienbesitzes erfragen.

(c) Verlust der Freiheit: Tracking beeinträchtigt die Freiheit und die Selbstbestimmung von Personen. Wenn ein informierter Nutzer um die Risiken von Tracking weiß, wird er sich möglicherweise aus Angst vor den Konsequenzen anders verhalten, was ohne Tracking nicht notwendig wäre.

(d) Schwächung der Demokratie: Tracking beeinträchtigt die Pressefreiheit und den Pluralismus in unserer Demokratie. Mit der Verschiebung von Werbebudgets von Printanzeigen hin zu zielgerichteter Werbung im Internet entstehen für Zeitungsredaktionen finanzielle Probleme, die zum Zeitungssterben beitragen und somit den Meinungspluralismus gefährden. Darüber hinaus gefährdet Tracking die Meinungsfreiheit und Partizipation, da Meinungen und Einstellungen in die falschen Hände gelangen können, was Verbraucher daran hindert, sich frei Informationen zu beschaffen.

(e) Manipulation von Bürgern und Filterung von Information: Tracking schafft die Basis für die Manipulation von Bürgern und Verbrauchern. Abhängig von gespeicherten Daten werden die einem Verbraucher angezeigten Informationen gefiltert. Stellen zwei Personen dieselbe Suchanfrage, dann können sie unterschiedliche Ergebnisse bekommen.

(3) Gefahren und Risiken mit wirtschaftlichen Nachteilen für Verbraucher

(a) Diskriminierung im Online-Handel: Abhängig von der durch Tracking bekannten Zahlungsbereitschaft können verschiedene Verbraucher in Online-Shops dasselbe Produkt zu unterschiedlichen Preisen angeboten bekommen. Eine solche Preisdiskriminierung hilft dem Handel, aus seinen Kunden das finanziell maximal Machbare herauszuholen.

(b) Auswirkungen auf die Gesundheitsversorgung von Verbrauchern: Es sind bereits Fälle bekannt geworden, in denen Krankenversicherungen auf Sammlungen von Verbraucherdaten zurückgreifen. So wurde auf Basis von Verbraucherdaten das Risiko bei einzelnen Verbrauchern eingeschätzt, was für die Versicherung günstiger war als die Durchführung von medizinischen Tests.

(c) Auswirkungen auf Kreditvergabe: Die im World Wide Web gesammelten Daten eignen sich auch für die Berechnung des Scoring bei der Kreditvergabe. Von der Schufa ist bekannt, dass sie ein Projekt durchführen wollte, bei der für die Einschätzung der Kreditwürdigkeit Daten aus sozialen Netzen verwendet werden sollten. Andere Unternehmen, die intensiv im Bereich Online-Marketing aktiv sind, bieten ebenfalls Dienste für die Einschätzung der Kreditwürdigkeit an.

(d) Konsequenzen für Arbeitssuche und Arbeitsverhältnisse: Personalabteilungen nutzen das World Wide Web bereits als Quelle für Hintergrundinformationen zu Be-

werbern. Anstatt diese Information selbst organisieren zu müssen, sind Dienstleistungen auf Basis von Trackingdaten und anderen Datenquellen denkbar.

(e) Verfolgung von Kindern: Bestimmte Tracker haben sich auf die Verfolgung von Kindern spezialisiert, da Kinder eine hohe wirtschaftliche Relevanz haben und Kinder durch Werbung viel einfacher zu manipulieren sind als Erwachsene.

(4) Gefahren und Risiken im Zusammenhang mit strafbaren Handlungen

(a) Wirtschaftsspionage: Tracking eignet sich auch für Wirtschaftsspionage, da damit verfolgt werden kann, welche Webseiten von bestimmten Unternehmen aufgerufen werden, wodurch Strategien und Pläne aufgedeckt werden können. Darüber hinaus kann man die durch Tracking gewonnenen Daten dazu verwenden, um Schwachstellen für Korruptionsversuche zu identifizieren.

(b) Vorbereitung von IT-Angriffen: Tracking kann dazu verwendet werden, um Personen zu identifizieren, die sich besonders gut als Opfer für IT-Angriffe eignen. Bestimmte IT-Angriffe erfordern die Mitwirkung von Nutzern.

(c) Vorbereitung von Straftaten: Tracking kann auch zur Vorbereitung von Straftaten in der nichtdigitalen Welt ausgenutzt werden. Beispielsweise kann es Trackern möglich sein zu erkennen, ob man auf Reisen ist und sich fern der Heimat aufhält.

**Schutzmaßnahmen für Verbraucher**

Es gibt eine Reihe von Werkzeugen zum Schutz gegen Web-Tracking. Der Web-Tracking-Report kann das Thema *Schutzmaßnahmen* nur streifen. Eine ausführliche Behandlung mit Vergleich der verschiedenen Schutzmaßnahmen war im Rahmen dieser Arbeit nicht möglich.

Es gibt methoden- und ergebnisorientierte Schutzmaßnahmen. Die methodenorientierten Schutzmaßnahmen beziehen sich immer auf die Abwehr einer bestimmten Trackingmethode, gegen andere Trackingmethoden hingegen wirken sie nicht. Die ergebnisorientierten Schutzmaßnahmen abstrahieren von den einzelnen Trackingmethoden. Für Verbraucher, die keine Experten im Bereich Web-Tracking sind, ist die Verwendung von ergebnisorientierten Methoden von Vorteil.

Einige Werkzeuge zum Schutz gegen Tracking sind im Funktionsumfang von Browsern enthalten, andere müssen im Rahmen von Browsererweiterungen noch zusätzlich installiert werden, z.B. gibt es sehr viele Browser-Plugins für Firefox. Im Folgenden geben wir einen Überblick über Schutzmaßnahmen, die direkt mit den Bordmitteln von Browsern oder anderen Systemfunktionen verfügbar sind:

- *Methodenorientierte Schutzmaßnahmen*
  - Schutz gegen Third-Party-Cookies: Third-Party-Cookies sind das am weitesten verbreitete Mittel zur Wiedererkennung von Verbrauchern für Tracking. Verbraucher können die Verwendung von Third-Party-Cookies durch Browsereinstellungen verbieten. Jedoch ist hierbei darauf zu achten, dass ein Verbot von Third-Party-Cookies nicht notwendigerweise dazu führt, dass ein Browser keine Third-Party-Cookies sendet.
  - Schutz gegen Flash-Cookies: Ein Verbot von Flash-Cookies muss im Einstellungsmanager von Flash vorgenommen werden. Abhängig von der Version des Flash-Player ist der Einstellungsmanager an verschiedenen Stellen zugänglich.
  - Schutz gegen E-Tags: Um sich gegen das Tracking mit E-Tags zu schützen, muss man den Cache-Speicher des Browsers leeren. Man kann jedoch nicht verbieten, dass neue E-Tags in den Cache-Speicher geschrieben werden.

- Verwendung von Anonymisierungsproxys: Anonymisierungsproxys helfen, um die eigene IP-Adresse zu verbergen. Dieser Schutzmechanismus kann jedoch von Trackern umgangen werden, indem Tracker den Browser dazu bringen, eine direkte Verbindung mit dem Tracker aufzubauen.

— *Ergebnisorientierte Schutzmaßnahmen*

- Trackingschutzlisten: Eine Trackingschutzliste ist eine Sperrliste mit Einträgen von Trackern. Verfügt ein Browser über einen Sperrlistenmechanismus, dann werden sämtliche Verbindungsaufbauversuche zu Trackern blockiert, wenn ein Tracker in der Liste enthalten ist. Der Internet Explorer von Microsoft verfügt über einen solchen Sperrlistenmechanismus. Das Fraunhofer SIT generiert eine Trackingschutzliste, die allen Nutzern kostenfrei zur Verfügung steht.[3] Bei anderen Browsern kann ein solcher Sperrlistenmechanismus durch die Installation von Browsererweiterungen erreicht werden.
- Do Not Track: Dieser Standard bietet keinen harten Trackingschutz. Der Browser schickt die üblichen Daten an den Tracker, wobei mit einem Flag dem Tracker signalisiert wird, dass der Verbraucher kein Tracking wünscht. Insofern ist der Verbraucher auf die Kooperationsbereitschaft des Tracker angewiesen. Es ist jedoch auch bekannt, dass es Bibliotheken für Browser-Fingerprinting gibt, die das Do-Not-Track-Flag für die Verbesserung von Tracking ausnutzen.
- Private Mode: Der Private Mode eines Browsers ist kein Werkzeug, welches primär zum Trackingschutz entwickelt wurde. Der Schutz gegen Tracking ist lediglich ein Seiteneffekt des Private Mode. Beim Private Mode geht es zunächst darum, dass die Privatsphäre von Nutzern gegenüber anderen Personen geschützt wird, die denselben Computer nutzen. Für den Schutz gegen Tracking wird ausgenutzt, dass von Browsern im Private Mode bestimmte Identifikatoren nicht persistent gespeichert werden, so dass eine Wiedererkennung durch Tracker zwischen verschiedenen Browsersitzungen mittels dieser Identifikatoren unterbunden wird. Innerhalb einer Browsersitzung ist jedoch Tracking möglich.

### Web-Tracking quantitativ

Für die Generierung seiner Trackingschutzliste untersucht das Fraunhofer SIT regelmäßig die Webseiten vieler verschiedener Anbieter auf Tracking. Bei seiner Untersuchung berücksichtigt Fraunhofer SIT insbesondere solche Anbieter, deren Webseiten für den deutschen Sprachraum eine hohe Relevanz haben. In diesem Zusammenhang ist eine große Datensammlung zu Trackern und einbettenden Anbietern (Embedder) entstanden. Fraunhofer SIT untersucht seit November 2012 bis heute Webseiten von mehr als 1600 verschiedenen Anbietern auf Tracking. Die letzte Analyse von Webseiten, die in diesem Bericht Berücksichtigung fand, wurde zwischen dem 22. und dem 24. Januar 2014 durchgeführt.

Über dem kompletten Betrachtungszeitraum wurden insgesamt 1209 verschiedene Tracker auf den analysierten Webseiten gefunden. Da die Einbindung von Trackern Marktgesetzen folgt, zeigt sich hier eine gewisse Dynamik, so dass man zu unterschiedlichen Zeitpunkten verschiedene Tracker findet. Zwischen dem 22. und 24. Januar 2014 wurden auf den untersuchten Webseiten insgesamt 633 verschiedene Tracker gefunden.

Es zeigt sich, dass es Tracker gibt, die bei sehr vielen verschiedenen Anbietern eingebunden sind und somit eine gute Ausgangssituation haben, um Verbraucher auf Webseiten verschiedener Anbieter verfolgen zu können. Im Folgenden werden in verschiedenen

[3] siehe `http://www.sit.fraunhofer.de/tpl`

Betrachtungsweisen für die am stärksten verbreiteten Tracker die höchsten Anzahlen von Embeddern angegeben, bei denen derselbe Tracker gefunden wurde:

- Anzahlen Embedder je Tracker in der Messung 22.–24. Januar 2014 (siehe Kapitel 4.2.1): In dieser Messung wurde ein Tracker gefunden, der gleichzeitig bei 994 verschiedenen Anbietern eingebettet war. Die Tracker mit der zweit- und dritthäufigsten Verbreitung hatten Werte von 780 und 474 verschiedenen Anbietern.
- Höchste Anzahlen Embedder je Tracker von November 2012 bis 24. Januar 2014 (siehe Kapitel 4.2.3): Die drei höchsten gemessenen gleichzeitigen Verbreitungen von einzelnen Trackern im gesamten Messzeitraum lagen bei 1110, 971 und 741 Embedder.
- Durchschnittswerte Anzahlen Embedder je Tracker von November 2012 bis 24. Januar 2014 (siehe Kapitel 4.2.5): Die höchsten Durchschnittswerte von Trackern hinsichtlich Embedder-Anzahl über dem gesamten Betrachtungszeitraum lagen bei 762, 655 und 366.
- Gesamtanzahl Embedder je Tracker von November 2012 bis 24. Januar 2014 (siehe Kapitel 4.2.7): Die Tracker, die über dem gesamten Betrachtungszeitraum bei den meisten Anbietern eingebettet waren, wurden insgesamt bei 1463, 1278 und 1194 Anbietern gefunden.

Bei der Analyse wurde auch betrachtet, wie viele Tracker bei einzelnen Anbietern eingebunden sind. Hier hat sich gezeigt, dass es Anbieter gibt, die viele verschiedene Tracker einbetten, so dass Verbraucher beim Aufruf von deren Webseiten gleich von vielen unterschiedlichen Trackern verfolgt werden können. Im Folgenden werden in verschiedenen Betrachtungsweisen für diejenigen Embedder mit den meisten Trackern die Trackernanzahlen angegeben:

- Anzahlen Tracker je Embedder 22.–24. Januar 2014 (siehe Kapitel 4.2.2): Die Anbieter mit den meisten Trackern hatten 78, 74 und 71 Tracker gleichzeitig.
- Höchste Anzahlen Tracker je Embedder von November 2012 bis 24. Januar 2014 (siehe Kapitel 4.2.4): Im gesamten Betrachtungszeitraum wurden in Einzelmessungen Embedder gefunden, die gleichzeitig 113, 102 und 94 Tracker in ihre Seiten eingebettet hatten.
- Durchschnittswerte Anzahlen Tracker je Embedder von November 2012 bis 24. Januar 2014 (siehe Kapitel 4.2.6): Die höchsten Durchschnittswerte von eingebetteten Trackern lagen bei 54, 50 und 48 verschiedenen Trackern.
- Gesamtanzahl Tracker je Embedder von November 2012 bis 24. Januar 2014 (siehe Kapitel 4.2.8): Die Embedder mit den meisten verschiedenen Trackern über dem gesamten Betrachtungszeitraum hatten insgesamt 174, 170 und 170 verschiedene Tracker eingebettet.

**Fazit**

Die möglichen Auswirkungen von Web-Tracking mit allen seinen Bedrohungen und Risiken und das heutige Ausmaß von Web-Tracking werden wahrscheinlich vielen Verbrauchern Anlass zur Sorge geben. Den meisten Verbrauchern war bisher nicht bewusst, dass bzw. in welchem Umfang sie von Unternehmen im Internet verfolgt werden und was diese Unternehmen mit ihren Daten machen können. Für Verbraucher ist es wichtig, dass sie hinreichend informiert sind über das, was im Hintergrund mit ihren Daten geschieht. Die Konsequenz lautet nicht, Web-Tracking zu verbieten, jedoch sollten Verbraucher frei entscheiden können, ob sie Trackern ihre Daten geben möchten, statt nichtsahnend verfolgt zu werden. Sollten

sie sich gegen Tracking entscheiden, dann müssen Verbraucher wissen, welche Möglichkeiten sie hierzu haben.

Es wäre auch wünschenswert, wenn Anbieter von Webseiten, die bisher ohne zu zögern Code-Elemente von Trackern in ihre Webseiten eingebettet haben, ihre bisherige Praxis überdenken würden. Bei informierten Verbrauchern werden solche Anbieter immer stärker dem Vorwurf ausgesetzt sein, dass ihnen der finanzielle Vorteil durch Tracking wichtiger ist, als der Schutz ihrer Kunden.

Eine faire Informationspraxis wäre zum Schutz der Verbraucher wünschenswert. Werbeeinnahmen sind auch ohne Web-Tracking möglich. Wie sich die Praxis des Web-Tracking entwickeln wird, bestimmen neben den Anbietern auch die Verbraucher. Es ist die Aufgabe der Politik, den Rahmen zu gestalten und Regeln zu schaffen, damit auch uninformierte Verbraucher nicht länger der heutigen Praxis des Web-Tracking und Datensammelns schutzlos ausgeliefert sind.

# 2. EINLEITUNG

Spätestens seit den Enthüllungen rund um die Spionageprogramme *Prism* und *Tempora* wissen wir, dass Geheimdienste verschiedener Länder, wie z.B. der US-amerikanische Geheimdienst NSA, unsere Kommunikation und Daten zu unserer Internetnutzung aufzeichnen und analysieren. Moderne Informations- und Kommunikationstechnologien haben inzwischen sehr viele Bereiche unseres täglichen Lebens durchdrungen und deshalb eignen sie sich sehr gut dazu, umfassende Informationen über Personen zu bekommen; wir nutzen heute Informations- und Kommunikationstechnologien beruflich und auch privat zu vielen verschiedenen Zwecken. So kommunizieren wir mit Geschäftspartnern, Kollegen und Freunden im Internet, wir kaufen online ein, wickeln unsere Bankgeschäfte online ab, informieren uns im Internet über Dinge, die uns interessieren, holen uns Rat und Empfehlungen, oder wir nutzen das Internet, um uns von dem vielfältigen Angebot einfach nur unterhalten zu lassen. Wie durch die Enthüllungen von Edward Snowden nun bekannt wurde, werden praktisch alle Benutzer im Internet von staatlichen Organisationen wie dem US-amerikanischen Geheimdienst NSA überwacht und kontrolliert.

Doch das Ausspionieren und Analysieren von Personen bzw. deren Handlungen in großem Stil wird nicht nur von Geheimdiensten betrieben. Es gibt auch viele Unternehmen, die ein sehr großes Interesse an möglichst vielen Informationen zu Bürgern, Verbrauchern und potenziellen Kunden haben und diese in großem Umfang erfassen und verarbeiten. Diese Tatsache wird von der aktuellen journalistischen Berichterstattung weniger stark thematisiert als Prism und NSA. Eine relativ einfache Möglichkeit, umfangreiche Informationen über Personen zu gewinnen, besteht in dem sogenannten *Web-Tracking*, d.h. sie bei der Nutzung von verschiedenen Internetangeboten zu beobachten, Daten zu sammeln und diese auszuwerten. Aus den so gewonnenen Informationen lässt sich in verschiedener Hinsicht von Unternehmen wirtschaftlicher Nutzen ziehen. Je nach Internetnutzung lassen sich mit *Web-Tracking* detaillierte Einblicke in die Interessen, das Konsumverhalten, politische Einstellungen, Finanzsituation oder den aktuellen Gesundheitszustand gewinnen. Die gewonnenen Erkenntnisse lassen sich dann beispielsweise für speziell auf den Nutzer zugeschnittene Werbeangebote im Internet verwenden. Jedoch ist auch eine weitergehende Nutzung der Daten möglich, wie beispielsweise das Vorhaben der Schufa gezeigt hat, Analyseergebnisse von Web2.0-Inhalten bei der Ermittlung der Kreditwürdigkeit von Privatpersonen zu verwenden [Spi12]. Die Verwertungszwecke solcher Daten sind vielfältig und der Kreativität zur Nutzung solcher Daten sind keine engen Grenzen gesetzt, wie es das Schufa-Beispiel viele andere Beispiele in Kapitel 3 belegen.

Gemäß der publizierten Marktzahlen können Unternehmen mit den gesammelten Daten große Umsätze erzielen [AS10; AS11; BKMP12]. Die Branche der Datensammler ist in den vergangenen Jahren sehr stark gewachsen [Ang10] und es ist davon auszugehen, dass sich das Wachstum noch weiter fortsetzen wird. Es gibt eine sehr starke Nachfrage nach persönlichen Daten, auch wenn deren Sammlung in einigen Fällen in krimineller Weise geschieht, wie z.B. durch Einbruch in Webserver [AS10]. Gesammelte Verbraucherdaten werden gehandelt, unabhängig davon, wie sie erhoben wurden und ob dieser Handel ein Risiko für Verbraucher darstellt, wie z.B. der aktuell bekannt gewordene Handel mit Patientendaten gezeigt hat [Spi13c].

Ist ein Verbraucher heute im Internet unterwegs, dann fließen Daten zu einer Reihe von verschiedenen Unternehmen. Das sind zunächst die Unternehmen, deren Onlineangebote der Nutzer willentlich und wissentlich in Anspruch nimmt. Im Hintergrund fließen Daten jedoch

auch zu einer Vielzahl von anderen Unternehmen, was den Verbrauchern wahrscheinlich nicht bekannt ist.

Einem Verbraucher ist sicherlich bewusst, dass er bei der Nutzung des Internets wie dem Aufrufen von Webseiten oder nach der Registrierung bei Onlinediensten und deren aktiver Nutzung Informationen von sich selbst gegenüber einem Onlinedienst preisgibt. Die Entscheidung, dem Anbieter eines Onlinedienstes Daten zur eigenen Identität zu geben, wie z.B. beim elektronischen Einkaufen, trifft ein Nutzer willentlich und wissentlich. Auch ohne Verwendung der eigenen Identität oder auch ohne Nutzung eines eigenen Dienstkontos muss man als Verbraucher annehmen, dass ein Onlinedienst verschiedene Aktivitäten derselben Person zueinander in Beziehung setzen kann, z.B. um den Status eines digitalen Einkaufkorbs zu verwalten. Ohne dieses Inbeziehungsetzenkönnen von Aktivitäten sind bestimmte online angebotene Anwendungen schlicht und einfach nicht möglich. Um den Zustand einer Interaktionsabfolge ermitteln zu können bzw. um einzelne Anfragen einem Zustand zuordnen zu können, müssen also verschiedene Aktionen, die von dem selben Benutzer ausgehen, diesem Benutzer zugeordnet werden können. Hierfür muss ein Anbieter seine Nutzer innerhalb des eigenen Webangebots verfolgen und in diesem Zusammenhang bestimmte Zustände speichern. Es ist anzunehmen, dass ein Benutzer sich an dieser Form der Verfolgung wahrscheinlich nicht stören wird, sofern der Anbieter nur solche Daten erhebt, die für das Dienstangebot tatsächlich relevant sind und die erhobenen Daten nicht für weitere Zwecke verwendet oder an Andere weitergibt.

Die Realität sieht heute jedoch so aus, dass Daten von Nutzern zu Dritten gelangen können, ohne dass sie direkt von Onlinedienstanbietern an Dritte weitergegeben werden und ohne, dass Nutzer ihre Daten bewusst an solche Dritte geben: Ruft man Webseiten eines Anbieters ab, dann ist es heute üblich, dass ein Browser beim Aufbau der angefragten Webseite im Hintergrund zusätzlich weitere Internetverbindungen zu dritten Parteien initiiert. Über diese Verbindungen fließen dann Daten zu Dritten. Bei diesen dritten Parteien handelt es sich oftmals um Tracker, die in Erfahrung bringen möchten, welche Seiten man im Internet abruft, um darüber Information über den Benutzer zu bekommen. Diese Verbindungen zu dritten Parteien werden typischerweise von dem Benutzer gar nicht wahrgenommen; in der Adresszeile der Browser wird ausschließlich die URL angezeigt, zu welcher der Benutzer tatsächlich navigieren möchte. Über die im Hintergrund aufgebauten Verbindungen erfahren die jeweiligen Tracker, welche Webseiten Verbraucher abrufen. Sie können Verbraucher somit bei ihren Internetaktivitäten verfolgen. Der Tracker als Empfänger dieser Informationen ist dem Verbraucher nicht bekannt und er hat üblicherweise keine Zustimmung gegeben, dass Daten an einen Tracker geschickt werden.

Damit Tracking durch dritte Parteien überhaupt möglich ist, müssen die Anbieter von Webinhalten Code-Elemente mit entsprechenden Befehlssequenzen der Tracker in ihre Inhalte integrieren. Dies geschieht typischerweise nicht ohne Gegenleistung, z.B. gegen Bezahlung oder Möglichkeit zur kostenfreien Nutzung von Softwarediensten wie etwa Analytics. Gelingt es einem Tracker, dass seine Trackingfunktionen in die Inhalte verschiedener Webseitenanbieter –d.h. über viele Internetdomains– integriert werden, so kann er Verbraucher über mehrere Webseiten unterschiedlicher Domains verfolgen. Mit entsprechenden technischen Methoden kann ein Tracker die Aktivitäten eines Nutzers über verschiedene Domains hinweg in Beziehung zueinander bringen und diese somit auf einen Verbraucher zurückführen. Bei diesem sogenannten *Cross-Domain-Tracking* kann der Tracker ein sehr umfangreiches Bild von Verbrauchern gewinnen, da er in Erfahrung bringen kann, welche Webseiten Nutzer bei verschiedenen Anbietern abrufen. Insofern können Tracker einen sehr viel umfassenderen Einblick erhalten als dies Anbietern bei ihren eigenen Kunden möglich ist.

In bestimmten Konstellationen kann es Trackern sogar möglich sein, die über Tracking gewonnenen Informationen zu Benutzern mit deren tatsächlichem Namen in Verbindung zu bringen. Dies ist beispielsweise dann möglich, wenn Unternehmen sowohl als Anbieter von Webinhalten auftreten, bei denen sich Nutzer registrieren, als auch als Tracker agieren, die ihre Trackingmethoden in die Webseiten anderer Anbieter integrieren. Eine andere Möglichkeit bieten beispielsweise URLs von besuchten Seiten, die schützenswerte Informationen enthalten. Verwenden Anbieter von Webseiten beispielsweise die Identitäten ihrer registrierten Nutzer als Parameter in ihren URLs, was in der Realität kein seltener Fall ist [KNW11], dann bekommen Tracker die Identitäten ohne Zutun geliefert und Nutzer bleiben weder anonym noch pseudonym.

Die Daten der Nutzer haben sich in den vergangenen Jahren zu einer neuen Währung im Internet entwickelt. Nutzer bezahlen die Angebote im Internet mit ihren Daten, ohne dass sie es in dem Moment der Preisgabe ihrer Daten bemerken. Tracking geschieht im Hintergrund und wird von den Nutzern gar nicht erst wahrgenommen. Es ist anzunehmen, dass Nutzern das Ausmaß von Tracking heute nicht bewusst ist, d.h. wissen nicht, in welchem Umfang Unternehmen für wirtschaftliche Daten über sie sammeln und verarbeiten. Diese Annahme wird beispielsweise von [May09; Mit10; Kau11] bestätigt.

Tracking bedeutet, dass Nutzer Kontrolle über ihre Daten verlieren. Daten, die im Kontext der Internutzung entstehen, gelangen so zu vielen unterschiedlichen Unternehmen, die diese Daten für ihre Zwecke nutzen können. So können die Daten analysiert, kombiniert oder zur Nutzung durch Dritte weitergegeben bzw. verkauft werden. Für Nutzer ist es nicht nachvollziehbar, zu welchen und wie vielen Akteuren solche personenbeziehbaren Daten dadurch gelangen. Preisgegebene Informationen lassen sich in der Regel nicht zurückholen, zumal nicht klar ist, wem man sie gegeben hat. Tracking beeinträchtigt die Freiheit von Nutzern, wenn das Wissen um Tracking das eigene Verhalten und ihre Selbstbestimmung beeinflusst.

Wie eine von der TU Darmstadt im Jahr 2012 durchgeführte Studie ergeben hat, ist dem überwiegenden Teil der Nutzer die Kontrolle über die eigenen Daten sehr wichtig [BGW12]. Da Nutzer die Kontrolle über die durch Web-Tracking erhobenen Daten verlieren, stellt sich die Frage, wie sie Web-Tracking für sich bewerten. Das Anliegen dieses Reports besteht jedoch nicht darin, zu ermitteln, wie Nutzer Web-Tracking einschätzen und wie sie auf Web-Tracking reagieren. Für eine informierte Entscheidung von Nutzern ist ein gewisses Hintergrundwissen erforderlich. Das Ziel des Reports besteht darin, Nutzern dieses Hintergrundwissen zur Verfügung zu stellen.

Der Web-Tracking-Report bietet seinen Lesern Informationen zum aktuellen Lagebild des Web-Tracking, insbesondere für Nutzer aus dem deutschen Sprachraum. Eine persönliche Einschätzung von Web-Tracking und eine Entscheidung, wie man als Nutzer mit Web-Tracking umgehen möchte, ist nur dann sinnvoll möglich, wenn Hintergründe und die damit im Zusammenhang stehenden Implikationen hinreichend bekannt sind. Der Report beschreibt Web-Tracking sowohl qualitativ wie auch quantitativ. So wird dargestellt, wie Web-Tracking funktioniert und welche Technologien und Methoden Tracker heute verwenden, um Nutzer im Internet zu verfolgen. Darüber hinaus werden Risiken des Web-Tracking für Verbraucher beschrieben, indem mögliche Konsequenzen der Verwertung von den durch Web-Tracking erhobenen Daten dargestellt werden. Solche Risiken sollten einem Nutzer bekannt sein, bevor er entscheidet, wem er welche Daten gibt. Es liegt dann in der Entscheidung jedes Einzelnen, ob er diese Risiken für sich in Kauf nimmt oder ob er statt dessen Maßnahmen ergreift, sich gegen Web-Tracking zu schützen. Neben dem Web-Tracking werden weitere technische Möglichkeiten dargestellt, wie Nutzerdaten erhoben werden können.

Der Report stellt zusätzlich dar, in welchem Umfang heute Methoden des Web-Tracking angewendet werden. In diesem Rahmen wurde beispielsweise untersucht, mit wie vielen verschiedenen Trackern man als Verbraucher beim Besuch bestimmter Webangebote rechnen kann oder auf wie vielen Webangeboten man mit stark verbreiteten Cross-Domain-Trackern zu rechnen hat. Für diese quantitative Analyse wurden die Daten ausgewertet, die beim Fraunhofer-Institut für Sichere Informationstechnologie (SIT) in Darmstadt im Rahmen von Vorarbeiten zur Generierung von Trackingschutzlisten erfasst werden.[4]

Darüber hinaus kann der Web-Tracking-Report den gesellschaftlichen Diskurs zum Thema Web-Tracking unterstützen. Zu diesem Diskurs gehört die Klärung der Frage, welche Rechte ein Nutzer bzgl. der eigenen Daten im Zusammenhang mit Web-Tracking in der Zukunft haben sollte [Kri10a]. Für die Politik, in deren Verantwortung es liegt, die rechtlichen Rahmenbedingungen für Web-Tracking zu bestimmen, kann der Web-Tracking-Report zur eigenen Positionierung und zur Abwägung von Dringlichkeiten beitragen.

[4]Fraunhofer SIT hat ein System entwickelt, welches die Webseiten der für den deutschen Sprachraum wichtigsten Internetdomains, z.B. auf Basis des Deutschland-Rankings von Alexa (`http://www.alexa.com`), auf das Enthaltensein von Trackingmethoden für Cross-Domain-Tracking analysiert. Diejenigen Internetadressen, von welchen Trackingmethoden ausgehen, werden dann als Tracker bzw. potenzielle Tracker in eine Trackingschutzliste aufgenommen. Auf Basis dieser Trackingschutzliste werden alle Verbindungsaufbauversuche blockiert, über die im Hintergrund Daten zu den gefundenen Adressen fließen sollen. Somit werden auf den besuchten Seiten Trackingversuche durch Dritte unterbunden. Die Trackingschutzliste des Fraunhofer SIT ist unter `http:\www.sit.fraunhofer.de/tpl` kostenfrei verfügbar. Bisher wird die Trackingschutzliste des Fraunhofer SIT nur vom Internet Explorer (ab Version 9) umgesetzt. Durch die wiederholte Suche nach Trackingmethoden bei den für den deutschen Sprachraum wichtigsten Domains erfasst Fraunhofer SIT Daten zur Verbreitung von Trackern in Webseiten. Diese Daten wurden zur quantitativen Analyse in dieser Studie ausgewertet.

# 3. WEB-TRACKING

## 3.1 Der Begriff Web-Tracking

Ist man als Nutzer im World Wide Web unterwegs, so ist es nicht unüblich, dass man bei seinem Tun verfolgt wird. Hierbei muss man davon ausgehen, dass man von Akteuren im Hintergrund verfolgt wird, mit denen man weder wissentlich noch willentlich in Verbindung tritt und die einem als Nutzer meist gar nicht bekannt sind. Bei diesen handelt es sich nicht um Anbieter von Webseiten, die man als Nutzer bewusst aufgerufen hat. Selbstverständlich können Anbieter von Webseiten ihre Nutzer ebenfalls verfolgen, jedoch ist diese Verfolgung auf den eigenen Webauftritt beschränkt.[5] Bei der Verfolgung durch Dritte geht es meist darum, Daten über den Nutzer zu sammeln, zu speichern und diese für verschiedene Zwecke zu verwerten. Hierbei können ganze Profile über Nutzer angelegt werden. Die Verfolgung durch Dritte geschieht in Echtzeit. Echtzeit bedeutet in diesem Zusammenhang, dass eine Verfolgung von Nutzeraktivitäten im Web in dem Moment geschieht, in dem Nutzer Webseiten abrufen. Die Zwecke, für welche dritte Parteien Daten sammeln, werden den Verbrauchern in der Regel nicht genannt; auch werden Verbraucher dann nicht nach einer Einwilligung zur Erhebung und Verarbeitung dieser Daten gefragt.

Der Kontrollverlust von Nutzern über ihre eigenen Daten resultiert insbesondere aus der Verfolgung durch dritte Parteien. Vor diesem Hintergrund hat gerade die Verfolgung durch Dritte im World Wide Web eine besondere Brisanz. Diese Art der Verfolgung bezeichnen wir als *Web-Tracking*. Web-Tracking hat eine große Bedeutung für die Verteilung zielgerichteter Werbung im Internet.

Die technische Verfolgung von Benutzern ist keineswegs auf das World Wide Web beschränkt. Tracking gibt es heute auch bei Smartphone-Apps oder Smart TV. Diese weiteren Formen des Tracking werden in diesem Dokument jedoch nur am Rand betrachtet (siehe Kapitel 3.5).

Schematisch funktioniert Web-Tracking, wie in Abbildung 1 gezeigt wird. Wenn ein Nutzer Webseiten eines Betreibers abruft, dann empfängt er in der ausgelieferten Webseite eine Antwort, die für Tracking technisch präpariert ist. In der empfangenen Webseite sind Befehle eingebettet, welche die Interaktion mit dem Tracker auslösen. Hierfür stellt der Tracker dem Webseitenbetreiber —in dieser Rolle kurz als *Embedder* bezeichnet— die jeweiligen Passagen mit den Befehlen zur Einbettung in die eigenen Inhalte zur Verfügung. Der Browser führt die Befehle des empfangenen Inhalts aus und wird dadurch dazu gebracht, eine Verbindung zu einem Tracker aufzubauen, z.B. um ein Bild zu laden, das ggf. nur ein einziges Pixel enthält und für die Darstellung der eigentlichen Webseite gar keine Rolle spielt. Beim Aufbau der Verbindung mit dem Tracker teilt der Browser dem Tracker mit, welche Webseite der Nutzer gerade aufgerufen hat. Diese Information wird bei einer HTTP-Anfrage mittels der Referer-Angabe automatisch an den Tracker übertragen. Im Rahmen der Internetverbindung mit dem Tracker können dann verschiedene Trackingmethoden angewendet werden, die dem Tracker dazu dienen, den Nutzer wiederzuerkennen. Über eine funktionierende Wiedererkennung von Nutzern können Tracker dann Profile von Nutzern anlegen.

[5] Die Verfolgung von Benutzern durch Webseitenanbieter kann rein technische Gründe haben, wie z.B. zur technischen Umsetzung des Sitzungskonzepts bei der Kommunikation auf Basis von HTTP. Da HTTP zustandslos ist, müssen Sitzungen mittels technischer Methoden zur Verfolgung von Nutzern innerhalb des Angebots eines Webseitenanbieters implementiert werden. Beispielsweise müssen Onlineshops ihre Kunden innerhalb des Shops verfolgen, um ihnen nach jedem weiteren Klick den korrekten Warenkorb zuordnen zu können.

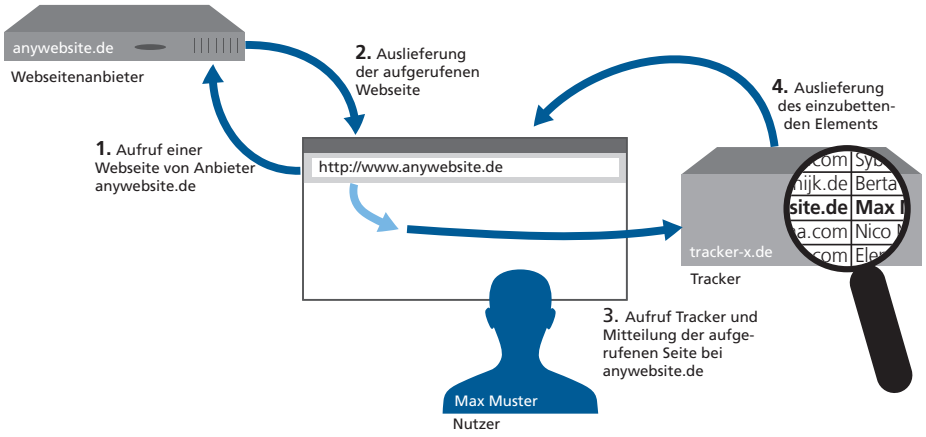


Abbildung 1. Schematische Darstellung Web-Tracking.

Die Wiedererkennung von Nutzern ist insbesondere dann sehr einfach für Tracker, wenn Anbieter von Webseiten in ihren URL benutzerspezifische Daten integrieren. Zu diesen benutzerspezifischen Daten können Informationen wie der vollständige Nutzername oder dessen Emailadresse gehören. Eine HTTP-Anfrage, bei der gegenüber Trackern personenbezogene Daten preisgegeben werden, könnte beispielsweise folgendermaßen aussehen: Angenommen, Max Mustermann ist bei `www.anbieterxy.com` als Nutzer registriert, wo er nach Informationen zur Behandlung irgendeiner Krankheit sucht. Befindet er sich auf einer Seite mit der fiktiven URL `www.anbieterxy.com/krankheitx.html?vorname=Max&name=Muster?email=mm@email.de`, von der aus Verbindungen zu einem Tracker `www.trackerxy.com` initiiert werden, dann enhält die HTTP-Anfrage typischerweise den Befehl `GET www.trackerxy.com` mit dem Referer-Eintrag im Header `www.anbieterxy.com/krankheitx.html?vorname=Max?name=Muster?email=mm@email.de`, woraus der Empfänger schließen kann, dass sich Max Muster für die Behandlung der Krankheit X interessiert.[6] Solche Fälle sind in der Praxis leider nicht unüblich, wie in [KNW11] gezeigt wurde.[7] Sind entsprechende Daten als Parameter in der URL der besuchten Webseite enthalten, dann werden diese mittels der Referer-Angabe an den Tracker übertragen. In [KNW11] werden eine Reihe von Beispielen gezeigt, bei denen Tracker in dieser einfachen Weise personenbezogene Daten von Nutzern bekommen können. Bei den Ergebnissen von [KNW11] ist bemerkenswert, dass gerade Webseiten von Vertretern der Gesundheitsbranche in diesem Zusammenhang besonders schlecht abgeschnitten haben.

Beim Web-Tracking können grundsätzlich sehr umfangreiche Informationen zu Personen gesammelt werden, zu denen beispielsweise Daten wie Vor- und Nachname, Emailadresse, Alter, Geschlecht, Wohnort, Einkommen, Familienstand, Gesundheitszustand, Einkäufe, bevorzugte Unterhaltungsmedien und Filme sowie allgemeine Interessen gehören [AM10]. Hinsichtlich der Interessen der Datensammler kann man heute keine klare Grenze erkennen. So werden beispielsweise auch Beiträge von Benutzern in Plattformen von sozialen Netzen

[6] Die Verantwortung bei der Verwendung von solchen personenbezogenen URL liegt beim Webseitenbetreiber, der personenbezogene Daten in die URL hineincodiert.

[7] In [KNW11] wird berichtet, dass 48% der untersuchten Webseitenanbieter in ihren URL Identifikatoren von ihren Nutzern verwenden, die dann von Trackern zur Wiedererkennung von Nutzern ausgewertet werden können.

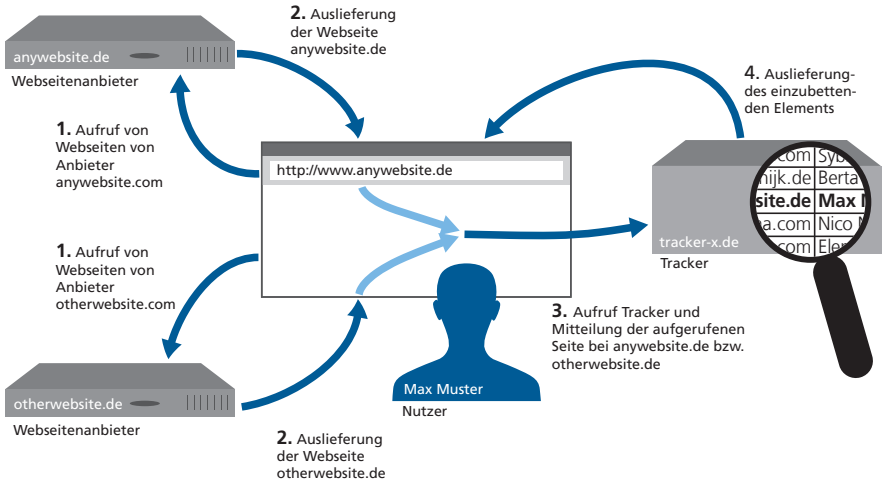


Abbildung 2. Schematische Darstellung Cross-Domain-Tracking.

analysiert und ausgewertet [AV10a]. Es scheint, dass alle Informationen gesammelt werden, die in irgendeiner Weise kommerziell verwertbar sein können.

Wenn im Folgenden verkürzt von *Tracking* gesprochen wird, dann ist damit immer *Web-Tracking* gemeint. Wenn von Tracking in einem anderen technischen Zusammenhang die Rede sein soll, dann wird ausdrücklich darauf hingewiesen.

## 3.2 Cross-Domain-Tracking

Für Tracker bietet Web-Tracking ein deutlich größeres Potenzial, wenn die trackerspezifischen Befehle nicht nur in den Inhalten einer Domain eingebettet sind, sondern in den Inhalten möglichst vieler Anbieter von Webseiten bzw. verschiedener Domains. Ist es einem Tracker möglich, Nutzer über viele Domains zu verfolgen, so kann er mehr Daten aus unterschiedlichen Zusammenhängen gewinnen, diese Daten kombinieren und somit Nutzerprofile besser verfeinern und fortentwickeln als wenn nur über einen einzelnen Webseitenbetreiber Zugang zu Nutzern besteht. Wenn ein Tracker Nutzer über die Angebote verschiedener Webseitenbetreiber verfolgen kann, dann bezeichnen wir dies als *Cross-Domain-Tracking*. *Cross-Domain-Tracking* erlaubt es Trackern, Erkenntnisse des Nutzerverhaltens über vielen unterschiedlichen Webangeboten aus verschiedenen Bereichen zu erfassen, zueinander in Beziehung zu setzen und Informationen zu Nutzern miteinander zu verknüpfen. In der Literatur wird für die Verfolgung von Nutzern über verschiedene Websites auch der Begriff *Retargeting* verwendet (siehe z.B. [HV10; TNB+10; Ell12; LT13; Gol13]). So lassen sich beispielsweise Erkenntnisse hinsichtlich persönlicher politischer Einstellungen, die sich u.U. durch Verfolgung eines Nutzers bei der Zeitungslektüre ergeben haben, durch Erkenntnisse bzgl. des Konsumverhaltens anreichern, wenn der Nutzer beim Stöbern durch Angebote in Online-Shops verfolgt wird.

Damit *Cross-Domain-Tracking* technisch umgesetzt werden kann, muss ein Tracker einen Nutzer über vielen Domains immer wieder erkennen können. Hierzu muss ein Tracker seine Trackingmethoden in den Webangeboten über allen Domains integrieren bzw. integrieren lassen, über denen er einen Nutzer verfolgen möchte. Besucht ein Nutzer diese Seiten, dann fließen hierzu Informationen zu dem Tracker, wie in Abbildung 2 schematisch dargestellt wird. Mit jeder ausgelieferten Webseite baut der Nutzerbrowser hierzu im Hintergrund Verbindungen auf.

Um *Cross-Domain-Tracking* betreiben zu können, haben Tracker eine hohe Motivation, ihre Trackingmethoden in den Webangeboten möglichst vieler Domains zu integrieren. Hier sind insbesondere die Webangebote mit hoher Nutzernachfrage für die Tracker interessant. So können in einer einzigen Webseite gleich eine Vielzahl von Trackern ihre Methoden integriert haben. Webseiten mit mehr als 50 verschiedenen Trackern sind heute keine Seltenheit mehr, wie wir in Kapitel 4 zeigen.

## 3.3 Markt und Akteure

### 3.3.1 Daten als Produktionsfaktor

Um das Sammeln, Verarbeiten und Verwerten von personenbeziehbaren Daten ist weltweit ein riesiger Markt entstanden. In diesem Markt werden jedes Jahr sehr viele Milliarden Euro umgesetzt, wobei uns der exakte Umfang nicht bekannt ist. Die genannte Größenordnung ist jedoch keineswegs unrealistisch, wenn man veröffentlichte Zahlen zugrunde legt, welche lediglich Teile dieses Marktes betreffen.[8] Daten von Personen sind mittlerweile ein sehr gefragtes Handelsgut geworden[9] und man kann mit ihnen viel Geld verdienen, so dass Kriminelle immer wieder versuchen, Daten aus Unternehmensdatenbanken zu stehlen [hei13; Spi13a; Spi13d].[10]

Meist geht die Nachfrage nach Daten von Organisationen aus, deren Interesse sich auf vorhandene oder potenzielle Kunden richtet, beispielsweise um neue Kunden zu gewinnen, um sich besser am Markt zu positionieren oder um Märkte besser verstehen zu können[Bau13; Car12; Spi13c]. Da sich angesichts der hohen Nachfrage nach personenbeziehbaren Daten viel Geld verdienen lässt, haben sich Unternehmen aufgemacht, für diese Nachfrage Angebote zu entwickeln (siehe z.B. [Spi13c]). Es gibt ein großes Spektrum an Angeboten von Unternehmen, die Daten sammeln, verarbeiten, analysieren oder kombinieren und die Ergebnisse in Form von Daten oder Dienstleistungen anbieten. Um diese herum haben sich wiederum zahlreiche Technologiehersteller etabliert, mit deren Produkten die Datensammler noch effektiver, effizienter, verlässlicher, genauer und robuster personenbeziehbare Daten erfassen können. Die Verwendung immer wieder neuer Technologien zum Datensammeln hat für Datensammler den Vorteil, dass hierfür keine oder wenige Schutzmaßnahmen verfügbar oder installiert sind, da Nutzer diese neuen Technologien u.U. noch nicht kennen [HSG+12].

[8] Allein für Onlinewerbung, für welche man Nutzerinformationen benötigt, wurden laut [War11] im Jahr 2011 weltweit 80,2 Milliarden US-Dollar ausgegeben. Für das Jahr 2015 wurde ein weltweites Umsatzwachstum auf 132 Milliarden US-Dollar vorhergesagt.

[9] Das US-amerikanische Unternehmen Rapleaf bietet jedem die Möglichkeit webbasiert zu einer gegebenen E-Mail-Adresse für einen Cent pro Merkmal Informationen wie Geschlecht, Haushaltseinkommen, Familienstand, Ausbildung, Kinder, Immobilienbesitz, Interessen oder Hobbys abzufragen [Kön13]. Die Quellen für die Datenbanken von Rapleaf sind unter anderem Einkaufshistorien, soziale Netzwerke, Grundbucheinträge oder Surfverhalten.

[10] Es ist in diesem Zusammenhang bemerkenswert, dass es Angreifer inzwischen gar nicht mehr nur auf Kreditkartendaten oder Zugangsdaten wie Passwörter abgesehen haben.

Dies treibt wiederum Datensammler und Technologiehersteller umso mehr an, hier weiter aktiv zu bleiben. Über all diesen Dingen wachen —mehr oder weniger wachsam— die Legislativen, Exekutiven und Jurisdiktionen der verschiedenen Länder, die mit ihren Datenschutzgesetzen die jeweiligen Regeln für die Sammlung und Verwertung von Daten aufstellen[11], die Einhaltung der Regeln überwachen und ggf. Verstöße sanktionieren [Kri10a; HSG+12; GT11b; GT11a].

### 3.3.2 Geschäftsmodelle zur Förderung von Daten als Rohstoff

Der gesamte Markt konnte sich in den vergangenen Jahren nur deshalb derart entwickeln, weil ihm genügend Daten als Rohstoff zur Verfügung standen. Die Durchdringung unseres Alltags mit Informationstechnologie und das hohe Ausmaß der Vernetzung durch das Internet schaffen die Voraussetzung, dass dieser Rohstoff immer wieder gefördert werden kann. Zu Verbrauchern lassen sich Profile anlegen und bei regelmäßiger Internetnutzung automatisch aktualisieren. Über das Internet laufen Daten zu Personen aus den verschiedensten Quellen zusammen. Durch die Verknüpfung von Daten aus diesen Quellen lassen sich Details von Personen zu Gesamtbildern kombinieren [Sch13b; Ste10b].

So wird praktisch alles, was von Datensammlern wirtschaftlich verwertet werden kann, erfasst und gespeichert, um dies in den eigenen Angeboten zu verwerten. Hierzu gehören Informationen wie Alter, Ausbildung, Beruf, Einkommen, Ethnie, Drogenkonsum, Familienstand, Geschlecht, Geschmack, Gesundheitszustand, Hobbies, Interessen, Konsumverhalten, Meinung, Namen, politische Einstellung, Religionszugehörigkeit, Risikobereitschaft, Schwangerschaft, sexuelle Orientierung, soziale Verflechtung, Sprachfertigkeit, Trinkgewohnheiten, Vorlieben, Wohnort [Eva09; Ste10b; May11; Sch13b]. Im Fokus der Sammelaktivitäten können alle Personen stehen, die das Internet als Verbraucher nutzen; es gibt jedoch auch Angebote, die sich auf bestimmte Gruppen spezialisiert haben, wie das US-amerikanische Unternehmen Lotame Solutions Inc. dies beispielsweise für sein Angebot *teeny bopper* gemacht hat, das sich insbesondere auf Kinder und Jugendliche konzentriert hat [Ste10a].

Um an diese Informationen gelangen zu können, müssen die Datensammler zunächst Daten erheben. Hierfür müssen sie entweder direkt oder indirekt mit den Nutzern über das Internet in Kontakt treten. Dies funktioniert oft über Dienstangebote, die Verbrauchern kostenfrei zur Verfügung gestellt werden. Zu den Quellen, über welche Daten gesammelt und personenbeziehbare Profile erstellt werden können, gehören:

(1) Suchmaschinen: Schickt ein Nutzer eine Anfrage an eine Suchmaschine, dann kann der Suchmaschinenbetreiber die Anfrage speichern und über den von einem Nutzer gestellten Suchanfragen ein Profil anlegen [KKMK09; Kau11].[12] Kann der Suchmaschinenbetreiber verschiedene Suchanfragen eines Nutzers derselben Person zuordnen, dann erfährt der Suchmaschinenbetreiber über die Inhalte der verknüpften Suchanfragen, für welche Themen und Inhalte sich der Nutzer interessiert. Suchmaschinenbetreiber nutzen diese Profile beispielsweise zur Personalisierung von Suchantworten, d.h. verschiedene Nutzer können unterschiedliche Antworten auf die gleiche Suchanfrage bekommen [GCF10; Kau11; HSK+13]. Eine andere wichtige Einnahmequelle ergibt sich aus der Online-Werbung, bei der die Werbeanzeige entsprechend der Suchanfragen ausgewählt

[11] Als Beispiel hierfür sei die europäische Richtlinie 2009/136/EG (ePrivacy-Richtlinie) der Europäischen Union angeführt [Eur09], die von den Mitgliedsstaaten teilweise in nationale Gesetze umgesetzt wurde bzw. noch umgesetzt werden muss.

[12] Es gibt mittlerweile auch Suchmaschinenangebote, die explizit erklären, dass sie die Daten aus den Suchanfragen nicht noch für weitere Zwecke auswerten. Ein Beispiel hierfür ist die Suchmaschine `duckduckgo.com`.

wird [Eva09; Kau11]. Suchmaschinenbetreiber können diese Profile über verschiedene Sitzungen speichern und über neue Suchanfragen sowie angeklickte Suchantworten weiterentwickeln. Ein Nutzer geht mit jeder Suchanfrage eine geschäftliche Beziehung mit dem Suchmaschinenbetreiber ein. Für die in der Regel kostenfreie Nutzung einer Suchmaschine gibt der Nutzer dem Suchmaschinenbetreiber Daten zu seinen Interessen und Bedürfnissen. Darüber hinaus wird Verbrauchern Werbung angezeigt, die mit der Suchanfrage in Zusammenhang steht. Für die Anzeige von Suchantworten, die möglichst am Anfang der Trefferliste erscheinen, lässt sich der Suchmaschinenbetreiber von den Werbenden und den Betreibern der priorisierten Webseiten bezahlen [GT11a; GM08; Vas10]. Da es sich bei dem Suchmaschinenbetreiber und dem Datensammler um dasselbe Unternehmen handelt, ist für den Nutzer zumindest ein gewisses Maß an Transparenz hergestellt. Seine Daten landen (zunächst) bei der Partei, zu welcher er eine Verbindung aufbaut. Deshalb wird in der englischsprachigen Literatur ein Suchmaschinenbetreiber vor dem Hintergrund des Datensammelns auch als *First Party* bezeichnet. Der informierte Nutzer kann entscheiden, welche Suchmaschine er nutzt und welchem Suchmaschinenbetreiber er somit seine Daten gibt.

(2) Web-2.0-Angebote: In den vergangenen Jahren haben Web-2.0-Angebote wie z.B. Facebook oder Twitter enorm an Nutzern hinzugewonnen. Nutzer können sich über diese Angebote vernetzen und tragen selbst zu den Inhalten von Webseiten bei. Die Betreiber dieser Angebote haben durch die von Nutzern generierten Inhalte die Möglichkeit, an Daten von Nutzern zu kommen. So können sie sehr einfach konkrete Informationen zu den Nutzern in Erfahrung bringen, wie z.B. Beruf, Interessen, Freunde, Bedürfnisse, die sie dann noch direkt mit den Klarnamen der Nutzer kombinieren können. Betreiber von Web-2.0-Plattformen bekommen diese Daten direkt von Nutzern geliefert und können darüber Profile konstruieren und weiterentwickeln [CKB12]. Die Anbieter dieser Dienste müssen zur Entwicklung dieser Profile strukturierte und unstrukturierte Daten verarbeiten und entsprechende Informationen extrahieren. Als Gegenleistung für die Plattformnutzung erhält der Plattformbetreiber die Daten und kann zusätzlich personalisierte Werbeinhalte schalten, wofür er von den Werbetreibenden bezahlt wird [GCF10; Tuc11]. Wenn Nutzer den Betreiber der verwendeten Web-2.0-Plattform bewusst auswählen, haben sie die Entscheidung selbst in der Hand, welchem Betreiber sie als First-Party ihre Daten geben.

(3) E-Mail-Dienste: Anbieter von E-Mail-Diensten haben Zugang zu der gesamten elektronischen Korrespondenz ihrer Kunden. Dies umfasst die von den Kunden versendeten und empfangenen Nachrichten. Machen Nutzer regen Gebrauch von E-Mails, dann können Anbieter von E-Mail-Diensten sehr umfangreiche Informationen von Nutzern erhalten. Anbieter von E-Mail-Diensten nutzen diese E-Mails als Datenquelle aus, um Nutzerprofile anzulegen [Kau11].[13] Ein bekanntes Beispiel hierfür ist Gmail, der E-Mail-Dienst von Google [Eva09; Hea13]. E-Mails enthalten zum großen Teil unstrukturierte Daten, wenn man von den Daten in den Kopfzeilen der E-Mails absieht. Um die unstrukturierten Inhalte von E-Mails automatisiert auswerten zu können, brauchen Betreiber Methoden des *Natural Language Processing*, wozu entweder eigene Lösungen angewendet oder entsprechende Dienstleistungen Dritter in Anspruch genommen werden. Als Gegenleistung für die Nutzung kostenfreier E-Mail-Dienste haben die Betreiber ein Geschäft auf Basis

[13] Anbieter von E-Mail-Diensten bekommen nicht nur Daten von ihren eigenen Kunden, sondern auch Daten von denjenigen Nutzern, mit denen der Kunde in Kontakt steht und E-Mails austauscht. Diese Nutzer müssen nicht notwendigerweise Kunden des E-Mail-Anbieters sein.

der Nutzerdaten entwickelt.[14] Innerhalb der Benutzerschnittstelle wird personalisierte Werbung angezeigt, für deren Anzeige sich der Anbieter des E-Mail-Dienstes von den Werbenden bezahlen lässt. Die Entscheidung eines Nutzers, welchen Dienstanbieter er als First-Party in Anspruch nehmen möchte und wer somit ggf. Zugang zu den eigenen Daten und Korrespondenzinhalten bekommen kann, liegt bei dem Benutzer.[15]

(4) Online-Textverarbeitungsdienste: Als Alternative zu den Textverarbeitungsprogrammen, die Nutzer auf ihrem Computer installieren können, gibt es Cloud-basierte Angebote, mittels denen man eine Textverarbeitungssoftware im Rahmen eines Dienstes nutzen kann. Bekannte Beispiele hierfür sind *Google Docs* (`http://docs.google.com`) und *Microsoft Office Web Apps* (`http://office.microsoft.com/de-de/web-apps/`). Mit einem solchen Dienst ersparen sich Nutzer Kauf und Installation von Textverarbeitungssoftware auf dem eigenen Computer und haben darüber hinaus immer direkten Zugriff auf die neueste Version der Software. Desweiteren erlauben diese Dienste die gemeinsame Erstellung von Dokumenten durch mehrere Nutzer. Durch die Verarbeitung und ggf. Speicherung von eigenen Dokumenten in der Cloud haben jedoch auch Dienstanbieter Zugang zu den verfassten Dokumenten und somit auch Einsichtmöglichkeiten für Analyse und Profilbildung. Ob bzw. in welchem Umfang die jeweiligen Anbieter von Online-Textverarbeitungsdiensten die erzeugten Dokumente für die Bildung von Profilen nutzen, ist uns nicht bekannt. Dies wurde im Rahmen des Reports weder untersucht noch recherchiert.[16] Google Docs betreffend ist jedoch bekannt, dass sich Google gemäß seiner Datenschutzrichtlinie, die sich auch auf Google Docs bezieht, die Option einräumt, die Inhalte von Google Docs diensteübergreifend zu verwenden. Damit besteht die Möglichkeit, Inhalte von Nutzerdokumenten für die Personalisierung von Suchanfragen zu verwenden und somit auch zur Anzeige von personalisierter Werbung [Mer12]. Ob bzw. welchen Anbieter eines Online-Textverarbeitungsdienstleisters ein Nutzer als First-Party in Anspruch nimmt, liegt in dessen freier Entscheidung. Auch hier besteht insofern Transparenz, dass er sich darüber im Klaren ist, wem er seine Daten gibt.

(5) Cloud-Speicherdienste: In den vergangenen Jahren haben sich eine Reihe von Cloud-Speicherdiensten am Markt etabliert. Diese bieten Nutzern die Möglichkeit, ihre Daten in Clouds zu speichern, was z.B. für Backup-Zwecke, hardwareunabhängige Verfügbarkeit oder für das Teilen von Dateien zwischen Personen von Vorteil sein kann. Beispiele für solche Cloud-Speicherdienste sind *Dropbox* (`https://www.dropbox.com/`), *Mozy* (`http://www.mozy.com/`) und *Google Cloud Storage* (`https://cloud.google.com/`). Einen Überblick über Cloud-Speicherdienste findet man in [BHH+12]. Die Betreiber der Cloud-Speicherdienste haben Zugriff auf alle Dateien, die ihre Kunden bei ihnen ablegen.[17] Abhängig von den Inhalten der in der Cloud gespeicherten Dateien können

[14] Ob es Anbieter von nicht kostenfreien E-Mail-Diensten gibt, die aufgrund des Bezahlangebots auf die Auswertung von Kundendaten verzichten, ist uns nicht bekannt.

[15] Dies gilt sowohl für den Nutzer, der Kunde des E-Mail-Anbieters ist, als auch für den Nutzer, der mit dem Kunden in E-Mail-Kontakt steht.

[16] In diesem Report geht es in erster Linie um Tracking. Die anderen am Markt verfügbaren technischen Möglichkeiten zum Sammeln von Daten werden hier lediglich zur Einordnung aufgelistet. Es geht in diesem Dokument nicht darum, andere Dienste zu untersuchen, die zum Sammeln von personenbezogenen Daten verwendet werden können.

[17] Viele Cloud-Anbieter werben damit, dass Daten verschlüsselt in die Cloud übertragen werden. Das bedeutet jedoch nicht, dass die Daten auch verschlüsselt in der Cloud gespeichert werden. Das wird von vielen Nutzern falsch verstanden. Die uns bekannten Cloud-Speicherdienste ermöglichen nur eine Verschlüsselung während des Datentransports. Wenn die Daten von dem Cloud-Speicherdienst empfangen werden, dann werden sie entschlüsselt und in Klartextform in der Cloud gespeichert. Möchte man als Nutzer die Daten

Anbieter umfangreiche Einblicke und Informationen zu ihren Kunden erhalten. Wie schon bei den Online-Textverarbeitungsdiensten ist es uns nicht bekannt, ob bzw. in welchem Umfang die Anbieter von Cloud-Speicherdiensten die Inhalte der bei ihnen gespeicherten Daten analysieren, auf Basis dieser Informationen Nutzerprofile anlegen und diese Daten an Dritte weitergeben. Wählt sich ein Kunde einen Cloud-Speicherdienst als vertrauenswürdige First-Party aus, so hat er zu berücksichtigen, dass noch weitere Anbieter von Cloud-Speicherdiensten involviert sein können. Bestimmte Cloud-Speicherdienste greifen nämlich im Hintergrund auf die Ressourcen anderer Cloud-Speicherdienste zurück und nutzen praktisch als Kunde deren Speicher. Als Beispiel sei hier der Cloud-Speicherdienst Dropbox genannt; Dropbox nutzt laut seiner Datenschutzerklärung den Speicherdienst *Amazon S3* [Dro13]. Somit können Dateien von Nutzern auch noch zu dritten Parteien gelangen, wobei nicht auszuschließen ist, dass diese die enthaltenen Informationen für ihre Zwecke nutzen können.

(6) Content-Angebote: Eine Möglichkeit, um mehr über Nutzer erfahren zu können, besteht darin, den Nutzern interessante Inhalte anzubieten. Das können Inhalte sein, die von anderen Nutzern beigetragen werden oder die von dem Content-Anbieter selbst produziert werden. Beispiele hierfür sind das Videoportal `www.youtube.com`, über welches Nutzer anderen Nutzern Videos zur Verfügung stellen können, das Photoportal `www.flickr.com` oder `books.google.de`, über welches Google Nutzern eingescannte Bücher anbietet, sowie `maps.google.de`, worüber man Landkarten und Routenplaner nutzen kann. Über die Bereitstellung attraktiver Inhalte geht es hier insbesondere darum, dass Nutzer möglichst regelmäßig die Webseiten des Content-Anbieters besuchen und von diesen Angeboten Gebrauch machen. Die anbietenden Unternehmen können auf verschiedenen Wegen wirtschaftlichen Nutzen aus diesem Angebot ziehen. Zunächst bekommen sie Daten von ihren Kunden, und zwar umso mehr, je häufiger diese von dem Angebot Gebrauch machen. Da bestimmte Funktionen in diesen Angeboten nur für registrierte Nutzer möglich sind, bekommt der Anbieter Identifikationsdaten von seinen Kunden. Darüber hinaus können inhalts- oder personenbezogene Werbeeinblendungen geschaltet werden, was dem Anbieter von Werbetreibenden finanziell entlohnt wird. Die Inhalte, welche Nutzern über die Webseiten des Content-Anbieters als First-Party zur Verfügung gestellt werden, können oftmals als Elemente alternativ in Webseiten anderer Anbieter eingebunden werden, wie z.B. für die Anzeige von Kartenausschnitten im Rahmen von Anfahrtsbeschreibungen. In diesem Fall agiert der Content-Anbieter dann als Third-Party; dieser Fall kann für Tracking ausgenutzt werden und wird unter Punkt (8) dieser Aufzählung beschrieben.

(7) Hosting-Dienste: Hosting-Dienste stellen Plattformen bereit, die es ihren Kunden erlauben, möglichst einfach ihre Inhalte ins *World Wide Web* bringen zu können. Diese Dienste müssen sich nicht allein auf Hosting beschränken: Es ist auch möglich, dass im Rahmen von komplexeren Angeboten zusätzlich zum Hosting Software angeboten wird, mittels derer Webauftritte gestaltet werden können. Ein Beispiel für ein solch komplexes Angebot ist Wordpress (siehe `www.wordpress.com`). Andere Varianten von Hosting-Angeboten bestehen in der Bereitsstellung von sogenannten *Content Delivery Networks*. Damit werden Inhalte nicht bei dem eigentlichen Anbieter der Inhalte,

vor den Augen der Speicherdienstbetreiber schützen, dann muss man die Dateien vor der Übertragung in die Cloud verschlüsseln. Eine Lösung hierfür ist das vom Fraunhofer-Institut für sichere Informationstechnologie entwickelte Produkt *OmniCloud*, welches im Jahr 2014 auf den Markt kommen wird (siehe `https://www.sit.fraunhofer.de/de/angebote/projekte/omnicloud/`.

sondern über den Betreiber des Content-Delivery-Network Nutzern zur Verfügung gestellt, der sich auf die Beschleunigung beim Laden von Webseiten und auf den Betrieb mit hoher Belastung spezialisiert hat. Ein Beispiel für einen solchen Anbieter ist Akamai (siehe `www.akamai.com`). Hosting-Dienste zeichnen sich dadurch aus, dass die Nutzeranfragen implizit bei dem Anbieter des Hosting-Dienstes landen und dieser somit automatisch Informationen über die Aktivitäten von Nutzern bekommt. Da ein Hosting-Dienst typischerweise Webauftritte von verschiedenen Webseitenbetreibern über seine Infrastruktur ermöglicht, ist er in der Lage, Benutzer über allen über ihn angebotenen Webseiten wiederzuerkennen und zu verfolgen. In diesem Rahmen besteht auch die Möglichkeit, Nutzerprofile anzulegen und diese beispielsweise für personenbezogene Werbung zu nutzen. Ein Hosting-Anbieter kann Nutzer ausschließlich über den über ihn zur Verfügung gestellten Angeboten verfolgen. Mit dieser bewussten Einschränkung könnte man auch bereits hier davon sprechen, dass zumindest die Möglichkeit für das Tracking von Nutzern besteht. Hosting-Anbieter und Webseitenanbieter als Embedder haben unterschiedliche Rahmenbedingungen, um an personenbezogene Daten von Nutzern kommen zu können.[18]

(8) Tracking bzw. Web-Tracking: Mit Tracking bzw. mit Cross-Domain-Tracking lassen sich sehr umfangreiche Einblicke zu Nutzern erlangen. Tracker müssen hierzu Betreiber von Webseiten gewinnen, so dass diese die für das Tracking erforderlichen Code-Elemente in ihre Webseiten einbetten. Diese Code-Elemente dienen Trackern dazu, dass Nutzer wiedererkannt werden können und somit in Erfahrung bringen können, welche Seiten ein Nutzer abruft — ggf. über viele Embedder hinweg. Tracker und die als Embedder agierenden Anbieter von Webseiten gehen hierzu eine geschäftliche Kooperation ein. Ein Tracker vergrößert mit jedem weiteren Embedder die Anzahl seiner neuen Datenquellen um die Menge der Nutzer, welche die Seiten des jeweiligen Embedders besuchen. Deshalb interessieren sich Tracker insbesondere für Webangebote mit hoher Reichweite. Für den Aufwand zur Einbettung von entsprechenden Code-Elementen in die eigenen Webseiten und entsprechender Auslieferung dieser Seiten werden Embedder typischerweise von Trackern entschädigt. Dies kann sowohl durch finanzielle Entschädigung geschehen oder auch durch andere Dienstleistungen, die direkt mit Tracking gekoppelt sind und die Embedder im Gegenzug kostenfrei nutzen können. Als Beispiel hierfür sind Analytics-Dienste (z.B. `www.google.com/analytics`) zu nennen, mit denen für Embedder die Nutzung ihres Webangebots erfasst, analysiert und aufbereitet werden, und Embedder dann Einblicke in die Analyseergebnisse erhalten können[19]. Weitere Beispiele sind Frontend-Dienste, über welche Embedder zur Integration bestimmter Funktionen in ihren Webseiten auf fertige JavaScript-Bibliotheken zugreifen können (z.B. `http://jquery.com/`), oder Angebote von Inhalten, die in Webseiten anderer integriert werden wie z.B. Landkarten, Bilder oder Videos. Tracker verwerten ihre Ergebnisse oft direkt für Werbeangebote z.B. über das Schalten von personalisierten Werbeanzeigen, über welche Tracker Einkünfte generieren können. Die Einbettung von Code-Elementen für Tracking wird in der Praxis häufig mit der Anzeige von Werbung kombiniert, d.h. der Embedder wird dann sowohl für die Einräumung einer Möglichkeit

[18] Eine Zusammenstellung von Trackingalternativen für Anbieter von Webseiten bietet WordPress seinen Kunden in [Wor13]

[19] Für Analytics werden in Webseiten Code-Elemente eingebettet, die zur Verfolgung von Nutzern innerhalb einer Website verwendet werden können. Werden die im Rahmen von Analytics gewonnenen Informationen über verschiedene Domains in Beziehung gesetzt, dann schafft Analytics selbst die Voraussetzung für Cross-Domain-Tracking.

zum Tracking als auch für das Angebot von Werbeflächen entschädigt. Aus der Perspektive eines Nutzers ist das Tracking nicht transparent. Ruft er Seiten eines Anbieters seiner Wahl als First-Party ab, dann können den Nutzer betreffende Daten zu einem oder gleich mehreren Trackern als Third-Party fließen. Selbst wenn der Nutzer im Rahmen einer bewussten Entscheidung bereit ist, seine Daten als Gegenleistung für Dienst oder Inhalt zu geben, dann ist ihm nicht bekannt, an wie viele Organisationen bzw. welche Organisationen genau seine Daten fließen.

Zu den über diese Wege erhaltenen Daten können Datensammler zusätzlich noch Daten aus Quellen anderer beziehen, z.B. durch Kauf oder Austausch, und diese mit den selbst erhobenen Daten kombinieren [GPS09; Dat13]. Über diesen Weg können vorhandene Profile weiter verbessert werden.

### 3.3.3 Implikationen des Rollenwechsels zwischen First-Party und Third-Party

In der Praxis gibt es Unternehmen, die zum Sammeln von Daten gleich mehrere der oben genannten Möglichkeiten zum Datensammeln (1) – (8) miteinander verbinden und damit Effekte erzielen können, die stärker sind, als die einzelnen Effekte der jeweiligen Ansätze zusammen. Durch Kombination von Informationen aus verschiedenen Quellen lassen sich für Datensammler oft neue wertvolle Informationen ableiten [FMSW11]. Gelingt es einem Unternehmen, gegenüber Nutzern die Rolle als First-Party mit der Rolle als Third-Party zu kombinieren, dann kann es seine Bedingungen für das Sammeln von Daten erheblich verbessern. Agiert ein Unternehmen gleichzeitig als First-Party, bei der sich Kunden registrieren, und auf Webseiten anderer Anbieter als Third-Party, dann können Nutzeraktivitäten auf Webseiten anderer Anbieter von Trackern ganz einfach der echten Identität eines Nutzers zugeordnet werden [KNW11]. Dies funktioniert auch dann, wenn der Nutzer zum Zeitpunkt seiner Verfolgung auf Webseiten anderer Anbieter nicht gleichzeitig bei dem First-Party-Angebot eingeloggt ist [CKB12]. Hierfür helfen den Trackern die technischen Möglichkeiten zur Wiedererkennung eines Nutzers, die unabhängig davon funktionieren, ob der Nutzer das Webangebot des Datensammlers nutzt oder andere Seiten besucht, auf denen die Code-Elemente eines Datensammlers für Tracking eingebettet sind, z.B. Social-Plugins. Um die Rahmenbedingungen für das Datensammeln bzw. das eigene Geschäft zu verbessern, ist es für Tracker also von Interesse, als First-Party attraktive Webangebote in Form von registrierungspflichtigen Inhalten oder Diensten zu schaffen und als Third-Party bei vielen anderen Betreibern von Webangeboten für Tracking geeignete Code-Elemente zu integrieren. Beispiele für Unternehmen, denen es erfolgreich gelungen ist, gleichzeitig in diesen beiden Rollen zu agieren, sind die Anbieter von Web-2.0-Angeboten, wie Facebook oder Twitter. Solche Unternehmen haben viele registrierte Kunden, welche die Plattformen regelmäßig nutzen. Mit der Einbettung von Social-Plugins, wie z.B. einem *Share Button*, in Webseiten anderer Anbieter wird bei jedem Aufruf dieser Seiten eine Verbindung zu der Web-2.0-Plattform aufgebaut, so dass die Aktionen eines Kunden verfolgt werden können. Da viele Unternehmen diese Social-Plugins auf ihren Webseiten zur besseren Vermarktung ihres eigenen Angebots intensiv nutzen, ist das Netz von entsprechend präparierten Webseiten mittlerweile sehr dicht, so dass Daten zu sehr vielen Nutzeraktivitäten erfasst werden können.

### 3.3.4 Die Werbeindustrie

Der Markt rund um die Sammlung und wirtschaftliche Verwertung von personenbeziehbaren Daten im Internet wurde in den vergangenen Jahren sehr stark durch die Werbeindustrie

getrieben.[20] Für die Werbeindustrie bekommen die zielgerichtete bzw. verhaltensabhängige Werbung eine immer größer werdende Bedeutung. Nutzer bekommen sowohl auf Basis der Inhalte von Webseiten, in denen die Werbung gezeigt wird, als auch auf Basis der über sie vorliegenden Daten angepasste Werbung angezeigt [GCF10]. Aus der Sicht von Werbetreibenden können Produkte somit verstärkt solchen Nutzern präsentiert werden, die zum adressierten Kundenkreis gehören. Das Internet bietet für die Analyse von Nutzern und die Verteilung nutzerspezifischer Werbung die besten Voraussetzungen. Das vorhandene und verwertbare Wissen über Nutzer und die Personalisierbarkeit von Inhalten im Internet machen Online-Werbung für die Werbeindustrie sehr attraktiv. Für zielgerichtete Anzeigen lassen sich um ein Vielfaches höhere Einnahmen erzielen als für konventionelle, ungerichtete Werbung. Der Trend zu mehr zielgerichteter Werbung wird durch verschiedene Interessen getrieben: Werbekunden sind bereit, für die höhere Erfolgswahrscheinlichkeit von zielgerichteter Werbung im Vergleich zu konventioneller Werbung höhere Preise zu zahlen [Har10; GT11b; LCC+12];[21] Anbieter von Webseiten können für die Einbettung von zielgerichteter Werbung auf ihren Webseiten höhere Preise erzielen; die Unternehmen, welche eine Werbeanzeige zwischen dem Werbekunden und einem Anbieter von Webseiten vermitteln, können ihre Umsätze mit zielgerichteter Werbung ebenfalls steigern. So hat sich ein Teil der Budgets von Werbekunden weg von konventioneller Werbung wie etwa in Zeitungen oder Zeitschriften hin zu Online-Werbung und dort insbesondere zu zielgerichteter kundenspezifischer Werbung verschoben.[22,23] Es ist anzunehmen, dass der Marktanteil von Online-Werbung und insbesondere personalisierter Werbung größer werden wird. Für das Jahr 2015 ist ein weltweiter Umsatz von 132 Milliarden US-Dollar prognostiziert, was einem Wachstum von knapp zwei Dritteln gegenüber dem Jahr 2011 entspricht [War11].

Dieses Wachstum ist im Wesentlichen getrieben durch die besondere Möglichkeit der auf den Nutzer zugeschnittenen Werbeeinblendungen. Mittels der durch Tracking gewonnenen

[20] Beispielsweise macht Google gemäß [Gar10] den mit weitem Abstand größten Teil seines Umsatzes mit Werbung. Auch viele andere der wichtigsten Websites generieren den größten Teil ihrer Umsätze durch Werbung [Eva09].

[21] Die Klickraten sind bei zielgerichteter Werbung gemäß [YLW+09] durchschnittlich um einen Faktor von 670% angestiegen; laut [Car12] konnten durch zielgerichtete Werbung Zuwächse bei Bestellungen von 15% gemessen werden. Die klassische Werbung in Zeitungen, Zeitschriften, Fernsehen und Radio ist statisch und erreicht somit auch viele Personen, die nicht als Kunde in Frage kommen [Eva09]. Zielgerichtete Werbung ist für den Geschäftserfolg des Werbenden viel effektiver. Darüber hinaus bietet klassische Werbung keine Mechanismen, mit denen man die direkte Wirksamkeit von verschiedenen Werbeaktivitäten erkennen und messen kann. Bei Online-Werbung hingegen lassen sich Klicks auf Anzeigen direkt auswerten. Damit lässt sich unmittelbar in Erfahrung bringen, welchen Nutzen einzelne Investitionen in Werbung bringen und wie gut welche Werbung beim Kunden ankommt.

[22] *Evans* geht in diesem Zusammenhang so weit, dass er in [Eva09] den von dem östereichischen Ökonomen Joseph Schumpeter geprägten Begriff der *schöpferischen Zerstörung* auf die Online-Werbung anwendet. Mit der Verschiebung von Werbebudgets zu Online-Werbung hin ändern sich Produktionsfaktoren in bestimmten Branchen, wie z.B. bei den Zeitungen. Durch das Ausbleiben von Werbeanzeigen hat sich ein deutlicher Umsatzrückgang für die meisten Zeitungen ergeben, was in den vergangenen Jahren für viele Zeitungen zu großen finanziellen Problemen und auch zu einem Zeitungssterben geführt hat. Traditionelle Medienunternehmen, die aufgrund anderer technischer Verteilmechanismen keine Möglichkeiten haben, Kundendaten zu sammeln, sehen sich schon seit längerer Zeit immer stärker im Nachteil im Wettbewerb um Werbeeinnahmen [Sto08; Chr11].

[23] Ein erfolgreicher Einsatz von kundenspezifischer zielgerichteter Werbung setzt voraus, dass genügend Nutzer mit entsprechenden Profilen die entsprechenden Webseiten besuchen, auf denen geworben werden soll. Ist diese Menge zu klein, dann sind auch größere Konversionsraten nicht hinreichend für deutliche Umsatzsteigerungen. In einem solchen Fall kann es die bessere Strategie sein, auf hohe Reichweiten mit niedrigeren Konversionsraten anstatt auf Zielfokussierung mit höheren Konversionsraten zu setzen.

Information werden die für den Nutzer passenden Werbeeinblendungen ausgesucht. Insofern kommt dem Tracking hier eine entscheidende Bedeutung zu. Mit Tracking kann praktisch auf aktuelle Änderungen von Bedürfnissen bei Nutzern reagiert werden. Holt beispielsweise eine Frau im Internet Informationen zu dem Begriff *Schwangerschaft* ein, dann können ihr unmittelbar danach —auch beim Besuch anderer Websites— Produkte für Babys, Schwangere und Mütter angezeigt werden. Die Auswertung von Personenprofilen kann u.U. so weit gehen, dass dem von dem Analysealgorithmus vermuteten Vater ebenfalls Anzeigen zu Babyprodukten angezeigt werden, ggf. sogar noch bevor dieser von seiner Vaterschaft Kenntnis hat [Duh12; Ram12].

Im Bereich der Online-Werbung hat sich ein ganzes Ökosystem gebildet. Moderne Online-Werbung geht davon aus, dass nutzerspezifische Daten vorhanden sind und somit auch Mechanismen, welche diese Daten liefern. Dieses Ökosystem ist durch technische Lösungen und auch spezielle Geschäftsmodelle geprägt und die Entwicklung neuer Technologien und Geschäftsmodelle scheint sich mit hoher Dynamik fortzusetzen [Ang10; AM10]. So gibt es beispielsweise seit vielen Jahren ganze Werbenetzwerke, die als Makler für Werbung agieren. Beispiele hierfür sind *Google AdSense*, *Yahoo!* oder auch das in Deutschland ansässige Unternehmen *Adtelligence*. Diese üben eine Rolle als Zwischenhändler zwischen Werbenden und Betreibern von Webseiten aus, die Werbeflächen anbieten. Solche Zwischenhändler können in Echtzeit die Schaltung einer Werbeanzeige vermitteln. Diese Vermittlung läuft computergesteuert ab. Diese Werbenetzwerke bieten ihre Dienste in der Regel gleichzeitig vielen Unternehmen an, die für ihre Produkte und Dienstleistungen Werbung schalten wollen. Auf der anderen Seite organisieren sie bei Webseitenbetreibern viele entsprechende Werbeflächen, über welche sie die Werbeanzeigen ausliefern. Liegen dem Zwischenhändler nun über Tracking gewonnene Informationen über einen Besucher einer Webseite vor und kann der Zwischenhändler mittels Trackingmethoden diesen Benutzer wiedererkennen, dann kann er aus dem Pool der ihm vorliegenden Werbeanzeigen diejenige auswählen, die am besten auf den Besucher der Seite passt. In einer Variation ist es auch möglich, dass der Zwischenhändler die Schaltung einer Anzeige für einen Besucher, der durch sein Profil als geeignet erkannt wird, unter den Werbenden in Echtzeit und meistbietend versteigert [RGF11]. Weitere übliche Modelle, nach denen die Bezahlung von Zwischenhändlern und Anbietern von Werbeflächen geregelt werden kann, sind *Pay per View* und *Pay per Click*.

### 3.3.5 Datenflüsse durch Unternehmensakquisitionen

Im Wettbewerb haben diejenigen Unternehmen einen Vorteil, die über umfangreiche Datensammlungen verfügen. Mit größeren Datensammlungen erhöht sich die Wahrscheinlichkeit dafür, dass für einen Besucher einer Webseite mit vermarktbarer Werbefläche Profildaten existieren. Neben dem Datensammeln durch Tracking ist es hierfür hilfreich, wenn sie Dienste oder Inhalte anbieten (z.B. solche, wie sie in der obigen Aufzählung unter (1) – (7) beschrieben wurden), deren Webangebote von möglichst vielen Nutzern regelmäßig besucht werden. In der Vergangenheit haben die großen Vertreter am Markt erfolgreiche kleinere Unternehmen durch Unternehmensakquisition übernommen, wie z.B.[24]

- AOL: `adsonar.com`, `advertising.com`, `tacoda.net`
- Facebook: `instagram.com`, `face.com`, `jibbigo.com`
- Google: `feedburner.com`, `youtube.com`, `picasa.com`, `doubleclick.net`

[24]Die hier gezeigten Beispiele von Unternehmensakquisitionen haben in Bezug auf die akquirierenden Unternehmen keinen Anspruch auf Vollständigkeit.

— Microsoft: `aquantive.com`, `skype.com`
— Omniture: `offermatica.com`, `hitbox.com`
— Valueclick: `mediaplex.com`, `fastclick.net`
— Yahoo!: `tumblr.com`, `rockmelt.com`, `xobni.com`, `qwiki.com`, `lexity.com`, `overture.com`, `iqengines.com`, `ztelic.com`, `adrevolver.com`, `rightmedia.com`, `yieldmanager.com`, `bluelithium.com`

Durch diese Akquisitionen konnten implizit auch bereits existierende Datenbestände übernommen werden und mit den Datenbanken der Akquisiteure kombiniert werden. Darüber hinaus sind vielversprechende Optionen zur Sammlung neuer Datenbestände entstanden [KW09]. Durch die Möglichkeit der Zusammenführung von zwei oder mehr Datenbanken können sich u.U. neue Informationen in Bezug auf die Nutzer ergeben. Durch Unternehmensakquisition gelangen Daten von Nutzern zu solchen Unternehmen, denen die Nutzer ihre Daten möglicherweise nicht hätten geben wollen. Es ist anzunehmen, dass die weiter erwarteten Wachstumspotenziale in der Werbebranche, der harte Wettbewerb und der Wunsch zur strategischen Positionierung von Unternehmen auch in den kommenden Jahren für weitere Unternehmenszukäufe sorgen werden.

### 3.3.6 Das Dilemma der Werbeindustrie

Die Interessen von Verbrauchern und Werbeindustrie hinsichtlich personenbezogener Werbung gehen weit auseinander [TKH+09]. Die Werbeindustrie und die Unternehmen, die mit der Werbeindustrie zusammen arbeiten, setzen auf zielgerichtete und nutzerspezifische Werbung im Internet, obwohl ihnen bekannt ist, dass der überwiegende Teil der Nutzer dies ablehnt. Gemäß einer in den USA im Jahr 2009 von Wissenschaftlern verschiedener Universitäten gemeinsam durchgeführten Umfrage lehnen zwischen 73% und 86% der Befragten verhaltensbasierte Werbung ab [TKH+09]. Trotz der Ablehnung dieser Methoden auf Nutzerseite halten die Marktakteure an ihrer Praxis fest. Hierfür lassen sich eine Reihe von Gründen finden. So ist der harte Wettbewerb in der Werbebranche zu nennen. Unternehmen aus der Werbebranche brauchen innovative Dienste und die entsprechende Technologie dafür, um sich von ihren Wettbewerbern abzugrenzen und gegen diese im Wettbewerb bestehen zu können. Das führt eher zum weiteren Aufrüsten für Tracking, so dass in neue Technologien für Tracking investiert wird. Mit Werbeangeboten, die auf Nutzer zugeschnitten sind, können Unternehmen —sowohl die Unternehmen, die Werbeflächen auf ihren Webseiten anbieten, als auch Unternehmen der Werbeindustrie— höhere Umsätze erzielen, auf die sie weder verzichten können noch wollen. Darüber hinaus gibt es eine entsprechende Nachfrage von Unternehmen, die ihre Produkte oder Dienste bewerben wollen, denen die Unternehmen aus der Werbebranche entsprechen möchten. Tut ein Unternehmen das nicht, dann riskiert es, Kunden an Wettbewerber zu verlieren. Insofern ist es für die Unternehmen der Werbebranche schwierig und riskant, auf die Wünsche der Nutzer einzugehen und das Sammeln von Daten sowie die Verwertung von Daten für Zwecke wie z.B. Werbung zu unterlassen. Es ist anzunehmen, dass ein allein durch den Markt geregelter Verzicht auf Tracking nicht möglich ist [HSG+12].

Damit eine Zurückhaltung beim Datensammeln nicht zu unerwünschten Umsatzverlusten führt, wäre es erforderlich, dass auch die Wettbewerber auf Tracking, Datensammeln und Verwertung dieser Daten für eigene wirtschaftliche Zwecke verzichten. Dies ist nur durch eine entsprechende Regulierung, eine effektive Kontrolle und Sanktionen bei Verstößen möglich. Hierbei ist eine entsprechende internationale Harmonisierung der rechtlichen Rahmenbedingungen wichtig.

### 3.3.7 Fragen zur Regelung des Marktes

Rechtliche Rahmenbedingungen zu schaffen ist Aufgabe der Politik. Vor einer Regulierung des Marktes ist es erforderlich, dass die Möglichkeiten zur Gestaltbarkeit dieses Rahmens ausgelotet werden. Auch wenn der Schutz von Privatsphäre und informationeller Selbstbestimmung in vielen Gesellschaften ein Grundrecht ist, muss es keineswegs das Ziel sein, Tracking sowie Sammlung von personenbeziehbaren Daten und deren wirtschaftliche Verwertung zu verbieten. Es sind hier auch Lösungen möglich, die Tracking erlauben und nicht im Widerspruch zu existierenden Grundrechten stehen und die einen Kompromiss zwischen den verschiedenen Interessen von Nutzern und der Werbeindustrie darstellen. In diesem Zusammenhang sind die folgenden grundsätzlichen Fragen zu beantworten [MM12]:

(1) Was sollen Verbraucher kontrollieren können? Dies bezieht eine weitere Frage mit ein, an welcher Stelle die Regelung ansetzen soll. Eine Möglichkeit besteht darin, dass Verbraucher entscheiden können, ob Daten auf der Basis von Tracking überhaupt zu dritten Parteien fließen dürfen. Eine andere Möglichkeit besteht in der Kontrolle, wie und zu welchen Zwecken die geflossenen Daten verwendet werden dürfen. Es sind auch Kombinationen beider Ansatzpunkte möglich: Wenn ein Nutzer dem Fluss von Daten durch Tracking zugestimmt hat, dann kann er in einer zweiten Stufe über die Verwertung der Daten entscheiden und ggf. bestimmte Optionen ausschließen.

(2) Wie soll die Voreinstellung aussehen? Dies Frage bezieht sich darauf, was für Tracker möglich sein soll, wenn ein Verbraucher Computer, Browser und Dienstangebote so nutzt, wie sie sich bei ihrer ersten Verwendung darstellen. Sollen dann die für einen Verbraucher kontrollierbaren Dinge so eingestellt sein, dass sie verboten oder erlaubt sind? Einigt man sich auf ein Verbot, dann funktioniert legales Tracking nur dann, wenn der Nutzer diese Einstellungen ändert. Im anderen Fall ist Tracking so lange möglich, bis der Nutzer das verbietet. In der Praxis muss man davon ausgehen, dass die meisten Verbraucher mit der gesamten Problematik noch nicht sonderlich vertraut sind. Vor diesem Hintergrund ist ein Verbot als Voreinstellung sehr viel verbraucherfreundlicher.

(3) Wo soll die Verbraucherentscheidung technisch umgesetzt werden? Grundsätzlich sind die Entscheidungen der Verbraucher an verschiedenen Stellen technisch umsetzbar. Die Beantwortung dieser Frage ist auch abhängig davon, wie die vorgenannte Frage (1) beantwortet werden wird. So kann die Entscheidung von Verbrauchern, ob sie Tracking erlauben möchten, beispielsweise im Browser vorgenommen werden, auf den Webseiten der Embedder oder beim Tracker selbst. Für Verbraucher ist es von Vorteil, wenn die technische Umsetzung zur Entscheidung möglichst einheitlich ist.

### 3.3.8 Strategie und Argumentationen der Werbeindustrie

Die Unternehmen der Werbeindustrie haben sich in Verbänden organisiert wie etwa die Network Advertising Initiative (NAI, `http://www.networkadvertising.org`), die Digital Advertising Alliance (DAA, `http://www.aboutads.info`), das Interactive Advertising Bureau (IAB, `www.iab.net`) und die European Advertising Standards Alliance (EASA, `http://www.easa-alliance.org`)), um gegenüber den Gesetzgebern gemeinsam die eigenen Interessen zu vertreten. Die US-amerikanische *Federal Trade Commission* (FTC) hatte eine Richtlinie mit Prinzipien hinsichtlich Transparenz, Offenlegung und Nutzerzustimmung herausgegeben, aus welchen die Werbeindustrie Vorschläge zur Selbstbeschränkung gemacht hat [LCC+12]. Aus der Perspektive der Werbeindustrie war es wichtig, eigene Vorschläge einzubringen, um ein Einlenken zu signalisieren und damit sehr restriktive Regelungen der

staatlichen Seite zu vermeiden. Im Zuge dieser Aktivitäten ist der Entwurf für den Do-Not-Track-Standard entstanden, der für Verbraucher eine Opt-Out-Möglichkeit vorsieht [MNS11; FS13]. Die Europäische Kommission unterstützt Do-Not-Track ebenfalls [Kro12]. Der Schutzumfang von Do-Not-Track ist sehr begrenzt; es ist also anzunehmen, dass die Interessen der Werbeindustrie durch Do-Not-Track nicht zu stark beeinträchtigt werden. Dennoch werden Lobbyisten nicht müde, weiter für die Interessen der Werbeindustrie zu kämpfen. So gibt es immer wieder Versuche, die Anwendung von Trackingmethoden zu rechtfertigen, indem sie behaupten, dass Verbraucher laut Umfragen personalisierte Werbung bevorzugen [Mar10; Gro13]. Auch wenn diese Ergebnisse den tatsächlichen Antworten der Befragten entsprechen, so ist unklar, ob den Befragten bei der Umfrage bewusst war, was personalisierte Werbung impliziert (z.B. Scannen von E-Mails, Verfolgung im Internet, Anlegung von Personenprofilen).

Die Europäische Union hat darüber hinaus in ihrer e-Privacy-Direktive beschlossen, dass die Anwendung von Trackingmethoden, wie z.B. die Anwendung von für diesen Zweck bestimmte Cookies, die Einwilligung eines Nutzers erfordert [Zui13]. Die e-Privacy-Direktive muss in den Mitgliedsstaaten in nationales Recht umgesetzt werden. Wie diese Umsetzung jeweils aussehen wird, ist noch unklar. Die Umsetzung sollte jedoch klare Regeln vorgeben, die für europäische Anbieter von Webseiten und Tracker verbindlich sind [Eur12b]. Sie sollte ebenfalls für Unternehmen gelten, die einen Nutzer an einem in Europa befindlichen Computer verfolgen wollen, die Daten jedoch in einem anderen Land verarbeiten [Eur12b]. Das europäische Parlament und der Rat haben im Zusammenhang mit der Umsetzung der Direktive auch bereits eine Verordnung vorgeschlagen [Eur12a]. Die Vorschläge der EU wurden massiv von den Lobbyisten der Werbeindustrie attackiert [Beu14]. Die Unternehmen fürchten, das Vorhaben könne ihre Geschäftsmodelle beeinträchtigen, weil sie die gesetzeskonforme Datenerhebung und Datenverarbeitung zu aufwendig machen würde [Beu14], bzw. weil sich damit wahrscheinlich weniger Daten gesetzeskonform sammeln lassen.

Zur Rechtfertigung ihres Tuns vertreten Tracker den Standpunkt, dass sie keine Datenschutzbestimmungen verletzen, da die gesammelten Daten keine Informationen zu dem Namen der betroffenen Personen enthalten [AM10]. Auch wenn in Fällen von Tracking keine Identifikationsdaten wie Klartextnamen erhoben werden, so können duch Tracking dennoch verschiedene Benutzer voneinander unterschieden werden. Da also Möglichkeiten zur Wiedererkennung für die Tracker bestehen (z.B. durch Third-Party-Cookies, IP-Adressen oder Browser-Fingerprinting), kann man diese Kommunikation nicht als *anonym* bezeichnen. Sie kann wegen der Wiedererkennbarkeit des Nutzers eher mit pseudonymer Kommunikation verglichen werden. Einer typischen pseudonymen Kommunikation würde jedoch eine Entscheidung eines Nutzers vorangehen, in der er beschließt, mit welcher Gegenseite er in welcher Form —ob anonym, pseudonym oder identifiziert— kommunizieren möchte. Beim Tracking ist es Verbrauchern in der heutigen Praxis meistens nicht transparent, dass im Hintergrund überhaupt eine Verbindung aufgebaut wird bzw. zu wem diese aufgebaut wird, so dass sich für den Verbraucher keine Frage hinsichtlich einer Entscheidung stellt, ob er eine Verbindung zu einer bestimmten Gegenseite aufbauen möchte, in welcher Form er diese Verbindung aufbauen möchte oder welche Daten zu welchem Zweck fließen. Selbst wenn Daten sogar anonym statt pseudonym fließen, so hat sich in der Vergangenheit gezeigt, dass auch nach Anonymisierung Beziehungen hergestellt werden konnten und Daten den ent-

sprechenden Personen korrekt zugeordnet werden konnten [And09; NS08; NS09; NS10].[25] Darüber hinaus gab es in der Praxis viele Fälle, bei denen Tracker den Klartextnamen von Nutzern erfahren konnten, da die Betreiber der einbettenden Webseiten die Namen in die URL eincodiert haben, so dass diese Information über die Referer an die Tracker geflossen ist [KNW11]. Darüber hinaus kann man in der Praxis auch nicht von einer Anonymisierung oder Pseudonymisierung ausgehen, wenn Unternehmen bei Angeboten sowohl in einer First-Party-Rolle (z.B. als Dienstanbieter wie etwa für E-Mail-Dienste oder als Betreiber einer Web-2.0-Plattform) als auch in einer Third-Party-Rolle als Tracker agieren. Wenn die so erfassten Daten zusammengeführt werden können, dann ist die pauschalisierte Behauptung, dass die durch Tracking gewonnenen Daten keinen Bezug zu Identifikationsdaten haben, nicht korrekt.

Die Marktakteure, die personalisierte Werbung befürworten und auch davon wirtschaftlich profitieren, setzen ihren Kritikern entgegen, dass die erzielten Werbeeinnahmen in den vergangenen Jahren die Entwicklung des Internets und insbesondere die Entwicklung des Dienstangebots im World Wide Web getrieben haben [VK10; Har10; TNB+10; GCF11; Zan11]. Nur durch die hohen Werbeumsätze konnten sich die Angebote so entwickeln, wie sie sich heute darstellen. Für die für Nutzer meist kostenfreien Angebote muss es Möglichkeiten zur Refinanzierung geben. Werbung und Tracking ermöglichen somit erst die zahlreichen kostenfreien Angebote [TNB+10]. Vertreter der Werbeindustrie mahnen, dass es das Internet in seiner heutigen Form nicht mehr geben wird, wenn man Tracking verbietet oder beeinträchtigt [Zan11]. Es ist jedoch anzunehmen, dass diese Ängste und Sorgen übertrieben sind: Denn auch ohne Tracking wird kontextbasierte Werbung immer noch möglich sein wie z.B. die Anzeige von Werbeinhalten, die sich alleine an den eingegebenen Suchinhalten orientieren [VK10].

## 3.4 Technische Betrachtung

### 3.4.1 Welche Daten eignen sich für Tracking?

Beim Web-Tracking geht es Trackern im Wesentlichen darum, in Erfahrung zu bringen, wer welche Webseiten aufruft. Die Daten hinsichtlich der aufgerufenen Webseiten werden bei HTTP standardmäßig geliefert. Die interessantere Frage bezieht sich hier auf das *wer*.

Zur Implementierung von Tracking können viele verschiedene technische Mechanismen verwendet werden. Potenzial für Tracking bieten alle technischen Ansätze, die dritten Parteien über längere Zeitintervalle Informationen liefern, anhand derer man Verbraucher bzw. die von ihnen verwendeten IT-Geräte wiedererkennen und voneinander unterscheiden kann [JBBM06]. Wiedererkennungsmöglichkeiten bestehen, wenn (1) Tracker Daten auf den Geräten der Verbraucher speichern und die gespeicherten Werte bei einem späteren Kontakt wieder auslesen (z.B. Third-Party-Cookies) oder wenn (2) Tracker verschiedene Verbraucher bzw. deren Geräte anhand von Wertebelegungen geeigneter Konfigurationsparameter voneinander unterscheiden können (z.B. IP-Adresse, Betriebssystem, Browsertyp und Browserversion). Sind einzelne Wertebelegungen zur Unterscheidung nicht hinreichend, dann können Tracker mit Kombinationen dieser Wertebelegungen arbeiten. Es ist auch eine Kombination von (1) und (2) möglich. In beiden Varianten überprüfen Tracker, ob die jeweils empfangenen Wertebelegungen bereits in der vorhandenen Datensammlung des Trackers

[25] Anonymisierung wird zunehmend schwieriger, da Vielfalt und Umfang von öffentlich verfügbaren Daten so stark zugenommen haben, so dass das Herstellen von Beziehungen bestimmter anonymisierter Daten zu Personen über Daten aus anderen Quellen wahrscheinlicher wird.

enthalten sind. Bei Übereinstimmung wird der jeweilige Nutzer dem entsprechenden Datensatz zugeordnet. Für die Wiederkennung nach (1) und (2) stehen im Bereich der Webtechnologien viele verschiedene technische Mechanismen zur Verfügung, die ursprünglich meist für andere Zwecke entwickelt wurden. Diese Mechanismen haben hinsichtlich ihrer Verwertung für Trackingzwecke unterschiedliedliche Eigenschaften und eignen sich deshalb besser oder schlechter für Tracking. So eignen sich beispielsweise solche Parameter besser, deren Wertebelegungen über entsprechende Zeiträume stabil bleiben oder bei denen von verschiedenen Verbrauchern genügend viele unterschiedliche Werte angenommen werden. Aus der Perspektive eines Trackers betrachtet sollten diese Zeiträume die Zeitintervalle zwischen verschiedenen Sitzungen umfassen, innerhalb welcher ein Computer ggf. einige Male heruntergefahren und wieder neu gestartet wird.

Für die Auswahl der für die Wiedererkennung zu verwendenden Parameter bzw. der damit einhergehenden technischen Mechanismen spielen die folgenden Aspekte eine Rolle:

- *Reichweite.* Für Tracker ist es von Vorteil, wenn die Wertebelegung bei möglichst vielen Nutzern abgefragt werden kann. Parameter, die möglicherweise nur bei einem eingeschränkten Kreis von Nutzern funktionieren, sind für Tracker weniger interessant.
- *Eindeutigkeit.* Wenn ein Parameter die Unterscheidbarkeit von verschiedenen Nutzern unterstützt, dann lässt sich das für Tracking ausnutzen. Je stärker sich die Wertebelegung zwischen verschiedenen Nutzern unterscheidet, desto besser ist dies für Tracking. Sollte die Wertebelegung sogar für alle Verbraucher unterschiedlich sein, dann können Tracker Nutzer daran eindeutig erkennen.
- *Stabilität.* Für Tracking ist es von Vorteil, wenn der Parameter seine Wertebelegung über möglichst großen Zeiträumen nicht verändert.
- *Robustheit.* Für Tracker sind Parameter interessant, die auch unter ungünstigen Bedingungen für Tracking noch zuverlässig funktionieren. Zu solchen ungünstigen Bedingungen zählen beispielsweise Gegenmaßnahmen von Nutzern, mit denen diese Tracking unterbinden wollen.

Tracker sind in der Praxis nicht auf die Verwendung einzelner technischer Parameter festgelegt. Sie können auch auf Kombinationen von solchen Parametern zurückgreifen. In diesem Fall ist es auch möglich, dass nicht alle verwendeten Parameter über den Wiedererkennungszeiträumen stabil bleiben müssen. Werden Tupel von Wertebelegungen über Kombinationen von Parametern verwendet, dann ist es auch denkbar, dass Nutzer auch dann wiedererkannt werden können, selbst wenn sich Werte von einzelnen Parametern ändern. Tracker können in diesem Fall ein neu gemessenes Tupel von Wertebelegungen dem Eintrag in ihrer Datenbank zuordnen, welcher zu dem vorliegenden Tupel von Wertebelegungen die geringste Distanz hat.

### 3.4.2 Prinzipien des Informationsflusses beim Tracking

Ein Browser hat beim Tracking Verbindungen zu Anbietern von Webseiten und zu Trackern. Die Daten zu den aufgerufenen Webseiten fließen dann über den Browser zu den Trackern. Für diesen Informationsfluss gibt es verschiedene Möglichkeiten. Um diese Möglichkeiten besser verstehen zu können, muss man wissen, wie Webseiten in Browsern aufgebaut werden, wenn Elemente einer Webseite von verschiedenen Anbietern geladen werden, und welche Regeln es für Informationsflüsse im Browser zwischen den Speicherbereichen verschiedener Anbieter gibt.

Wird eine neue Webseite in einem Browser aufgebaut, so legt der Browser für die Webseite einen neuen Ausführungskontext an. Zum Aufbau der Seite werden von dem angefragten Webseitenanbeiter die für die Webseite erforderlichen Elemente geladen wie z.B. I-Frames, JavaScripts, Stylesheets oder Bilder. Die für den Aufbau einer Webseite verwendeten Elemente können von verschiedenen Anbietern geladen werden. Für die von unterschiedlichen Anbietern geladenen Elemente einer Webseite können je nach Typ des geladenen Elements neue Ausführungskontexte angelegt werden. Die Speicherbereiche dieser Ausführungskontexte sind gegeneinander isoliert, so dass die Daten nicht über die Grenzen eines Ausführungskontextes in einen anderen Ausführungskontext gelangen können. Die Regeln für den Fluss von Informationen zwischen Ausführungskontexten folgen der sogenannten Same-Origin-Policy. So können beispielsweise Cookies nur von der Partei abgefragt werden, die die Cookies geschickt hat. Eine wichtige Ausnahme besteht in den Referer-Daten, die über die Grenzen der Ausführungskontexte hinweg ausgetauscht werden können. Darüber hinaus gelten für verschiedene Typen von Elementen in Webseiten unterschiedliche Regeln.

Wird beispielsweise in einer Webseite ein I-Frame von einer dritten Partei geladen, dann wird diesem I-Frame ein neuer Ausführungskontext zugeordnet und die Speicherbereiche sind voneinander isoliert. Wird hingegen in einer Webseite ein JavaScript von einer dritten Partei geladen, dann wird dieses JavaScript in dem Ausführungskontext des Embedder ausgeführt und das JavaScript kann auf die Daten im einbettenden Ausführungskontext zugreifen. Geschieht die Einbettung von Tracker-Elementen also derart, dass sie nicht gegen den Ausführungskontext des Embedder isoliert sind, dann können Tracker auch Daten in Erfahrung bringen, die über den Referer-Eintrag hinausgehen.

### 3.4.3 Varianten der Identifikation

Die von Trackern verwendeten Zustandsparameter als Identifikatoren von Nutzern bzw. deren Computern lassen sich dahingehend unterscheiden, wie die Wiedererkennung vorbereitet und durchgeführt wird. Man kann zwei Varianten unterscheiden [RKW12]:

— *Explizites Tracking.* Diese Art von Tracking verwendet eindeutige und von Trackern vorgegebene Identifikatoren. Die Identifikatoren werden von den Trackern entsprechend ausgewählt und auf den Computern von Verbrauchern gespeichert. Bei späteren Kontakten mit den Verbrauchern können Tracker diese dann anhand der Identifikatoren wiedererkannt werden. Ein Beispiel für einen solchen eindeutigen Identifikator sind HTTP-Cookies als Third-Party-Cookies. In der Literatur wird neben der Bezeichnung *explizit* (z.B. in [RKW12]) auch die Bezeichnung *zustandsbehaftet* (z.B. in [NKJ+13]) verwendet. Für Tracker sind solche Identifikatoren ein sehr effektives Mittel, um Benutzer zweifelsfrei wiederzuerkennen. Verwendet eine dritte Partei solche eindeutigen Identifikatoren für Verbraucher, dann ist dies ein starkes Indiz für Tracking.

— *Schließendes Tracking.* In diesem Fall werden von dem Tracker keine Identifikatoren vergeben. Bei dieser Art von Tracking versuchen Tracker, den Computer anhand von Konfigurationsaspekten zu erkennen und anhand von bekannten Wertebelegungen in der Konfiguration auf den Verbraucher zu schließen. Da bestimmte Konfigurationsaspekte bei mehreren Verbrauchern in ihren Wertebelegungen übereinstimmen können, müssen Tracker für ihr Schließen verschiedene Konfigurationsaspekte berücksichtigen. Dadurch wird die Wahrscheinlichkeit erhöht, dass ein Tracker einen Verbraucher eindeutig erkennen kann. Für die Erkennung werden Konfigurationsaspekte des Verbrauchercomputers ausgenutzt, die dieser dem Tracker mitteilt. Einige dieser Konfigurationsaspekte ge-

ben Computer immer von sich preis, andere geben sie nur dann heraus, wenn Tracker diese Konfigurationsaspekte explizit anfragen. Diese Art von Tracking wird auch als *zustandslos* bezeichnet (z.B. in [NKJ+13]). Die Erkennung von schließendem Tracking kann sehr schwierig sein.

Von außen betrachtet lässt sich nur beobachten, ob dritte Parteien durch die Anwendung von Trackingmethoden versuchen, Verbraucher wiederzuerkennen. Für welche Zwecke diese Wiedererkennungsmöglichkeit ausgenutzt wird, erkennt man aus der bloßen Erkenntnis, dass Trackingmethoden angewendet werden, nicht. Theoretisch besteht die Möglichkeit, dass die durch Trackingmethoden gewonnenen Identifikationsdaten nicht weiter verwertet werden. Natürlich besteht in einem solchen Fall die Frage, warum dieser Zusatzaufwand zur Wiedererkennung dann überhaupt betrieben wird. Um erkennen zu können, zu welchem Zweck der Aufwand betrieben wird, müsste man Einblicke in die interne Verarbeitung von Daten bei solchen dritten Parteien haben, es sei denn, das jeweilige Unternehmen stellt offen dar (z.B. in den Beschreibungen zum eigenen Geschäftsmodell), zu welchem Zweck versucht wird, Verbraucher wiederzuerkennen.

### 3.4.4 Abgrenzung klassisches Tracking und Ausnutzung von Sicherheitslücken

Ein Tracker setzt zur Erreichung seiner Ziele üblicherweise Funktionen und Methoden ein, die ihm die Web-Protokolle und die Browser bieten. Die für Tracking genutzten Funktionen und Möglichkeiten werden bewusst eingesetzt, auch wenn die ursprünglichen Ziele der technischen Mechanismen andere waren (siehe z.B. Cookies). Unternehmen, die für ihr Geschäft Tracking brauchen, haben die Funktionen und Methoden einfach zweckentfremdet. Diese Zweckentfremdung bedeutet jedoch nicht, dass Tracking mit Hacking gleichzusetzen ist. Beim typischen Tracking werden alle Funktionen genau so verwendet, wie sie im Desgin der Browser festgelegt wurden.

Es ist prinzipiell auch möglich, in Browsern oder Webservern vorhandene Sicherheitslücken auszunutzen, um Nutzer im Internet verfolgen zu können bzw. um andere Angriffe auszuführen. Hierfür kann beispielsweise Cross-Site-Scripting (XSS) als Methode verwendet werden (siehe z.B. [OWA13]). Solche Angriffe würde man jedoch eher der Kategorie *Hacking* und weniger *Tracking* zuordnen. Für XSS-Angriffe müssen Komponenten in IT-Systemen manipuliert werden. Darüber hinaus erfordern diese Angriffe nicht die Kooperation eines Embedder, wie das bei Tracking der Fall ist; die Anbieter von Webseiten sind hier vielmehr als weiteres Opfer der Angriffe zu sehen. Tracking über solche Angriffe ist jedoch aufwendiger als die Anwendung der anderen Trackingmethoden. Die Ausnutzung von Sicherheitslücken ist mglw. dann eine Alternative für Tracker, wenn sich keine Basis für eine Kooperation mit einem Embedder finden lässt. Auch wenn mittels solcher Angriffe Nutzer verfolgt und ausspioniert werden können, werden diese Angriffe typischerweise nicht zu den Methoden für Tracking gezählt.

### 3.4.5 Trackingmethoden

Im Folgenden werden technische Alternativen für die Durchführung von Tracking genannt und beschrieben. Die hier vorgenommene Aufzählung der Trackingmethoden hat keinen Anspruch auf Vollständigkeit. Die Vergangenheit hat gezeigt, dass Tracker bei der Erfindung und Entwicklung neuer Trackingmethoden sehr kreativ sind. Es werden immer wieder neue Methoden entwickelt und bestehende variiert oder verbessert, so dass eine Beschrei-

bung nicht vollständig sein kann. Für die angeführten Trackingmethoden kann zwischen unterschiedlichen zugrundeliegenden Modellen unterschieden werden.

### Modelle für Trackingmethoden

Für die aufgeführten Trackingmethoden sind zwei Modelle relevant. Die meisten Trackingmethoden basieren auf folgendem 3-Parteien-Modell: Das Modell beinhaltet Tracker, Embedder und Nutzer. Tracker und Embedder koopieren, indem Embedder Elemente von Trackern in ihre Webseiten integrieren. Immer wenn ein Nutzer diese Webseiten abruft, erfährt dies der Tracker; Tracking geschieht hier also synchron zu den Internetaktivitäten von Verbrauchern. Wenn der Tracker den Nutzer wiedererkennen kann, kann er den Webseitenabruf dem Nutzer zuordnen. In diesem Modell kann der Tracker den Nutzer ab dem Zeitpunkt verfolgen, ab dem die Webseiten des Embedder entsprechend präpariert sind.

In einigen praktischen Fällen ist jedoch ein 2-Parteien-Modell relevant: In diesem Modell gibt es nur Tracker und Nutzer. Der Nutzer tritt hier bewusst direkt mit dem Tracker in Kontakt; er geht jedoch nicht davon aus, dass es sich dabei um einen Tracker handelt. Der Tracker bietet eigene Webseiten an, die der Nutzer abruft. Zusammen mit der Webseite liefert der Tracker Code-Elemente an den Nutzer aus, mit denen er in Erfahrung bringen kann, welche anderen Webseitenanbieter der Nutzer im Vorfeld besucht hat. In diesem Modell kann der Tracker Aktivitäten des Nutzers verfolgen, die zeitlich vor dem Laden der präparierten Webseite liegen. Tracking geschieht hier also asynchron zu den Internetaktivitäten der Verbraucher, auf die sich die Verfolgung bezieht.

### Differenzierungskategorien für Trackingmethoden

Die hier angeführten Methoden können anhand der folgenden Fragen kategorisiert werden:

- Wie gelingt es Trackern, eine Verbindung zu einem Nutzer zu initiieren?
- Wie erhalten Tracker Informationen bzgl. der aufgerufenen Seiten?
- Wie können Tracker Nutzer identifizieren?

Je nach konkretem Vorgehen der Tracker können diese Aspekte abhängig oder unabhängig voneinander sein. Die Trackingmethoden werden anhand der Differenzierung, wie sie sich durch die vorgenannten Fragen ergibt aufgezählt.

### Trackingmethoden differenziert nach Verbindungsinitiierung

Die dargestellten Möglichkeiten zur Verbindungsinitiierung kann man nochmal hinsichtlich der verwendeten Technologie oder des Mediums unterscheiden. Die hier aufgeführten Methoden schließen sich nicht notwendigerweise gegenseitig aus. So kann beispielsweise zur Darstellung einer Landkarte (Medium) JavaScript (Technologie) verwendet werden.

Alle hier genannten Methoden zur Verbindungsinitiierung können sowohl für explizites als auch für schließendes Tracking verwendet werden.

*Medium: Bilder*

Damit ein Tracker in Erfahrung bringen kann, dass ein Nutzer eine Webseite aufruft, kann ein Bild in die Webseite eines Embedder integriert werden. Der von dem Embedder ausgelieferte Code für die Webseite veranlasst den Browser dazu, das zu integrierende Bild von einem Server des Tracker zu laden. Auch wenn ein Nutzer das Bild auf der aufgerufenen Webseite wahrnehmen kann, so ist es dem Nutzer meist nicht bewusst, dass das geladene

Bild nicht von der Partei zur Verfügung gestellt wird, deren Webseite er aufgerufen hat. Bei diesem Bild kann es sich beispielsweise um eine Werbeanzeige oder ein Social-Plugin eines Web-2.0-Anbieters handeln.

*Medium: Zählpixel*

Zählpixel werden auch Web-Bugs, Web-Beacons oder auch Web-Wanzen genannt. Bei Zählpixeln handelt es sich um Bilder, die von einem Tracker geladen werden und wie andere Bilder in Embedder-Webseiten integriert werden. Das Besondere an Zählpixeln ist, dass es sich um winzig kleine Bilder handelt, die ggf. nur aus einem oder sehr wenigen Pixeln bestehen und deren Inhalt farblos bzw. transparent sein kann. Zählpixel werden von Nutzern nicht wahrgenommen und haben für die Darstellung einer Webseite keine Bedeutung. Der einzige Zweck von Zählpixeln besteht darin, dass der Browser des Nutzers Kontakt mit dem Tracker aufnimmt.

*Medium: Landkarten*

Weitere beliebte Elemente, mit denen im Hintergrund Verbindungen zu Trackern initiiert werden können, sind Landkarten, die in Webseiten integriert werden. Die Einbindung geschieht hier in der Regel so, dass die Karte nicht von dem Webserver des Embedder, sondern von dem Anbieter des Kartenmaterials geladen wird. Die einbettende Seite enthält einen entsprechenden Verweis, dass die Karte von dem Kartenanbieter nachgeladen wird. Bei solchen Karten kann es sich um einfache Bilder handeln oder aber auch um interaktive Angebote, die beispielsweise mit JavaScript umgesetzt sind und die zusätzlich noch Funktionen wie einen Routenplaner anbieten. Nimmt ein Nutzer entsprechende Angebote eines Tracker in Anspruch, dann kann der Tracker auf diese Weise zusätzlich noch die Anschrift eines Nutzers in Erfahrung bringen.[26]

*Medium: Videos*

Die Einbindung von Videos in Webseiten erlaubt ebenfalls die Initiierung einer Verbindung zwischen einem Nutzer und einem Tracker. In Webseiten eingebundene Videos werden typischerweise von einem Videoportal geladen und nicht von dem Anbieter der einbettenden Webseite. Dieser integriert in seiner Seite lediglich den Code, damit das Startbild des Videos geladen und angezeigt wird. Dadurch werden durch die Einbindung von Videos Verbindungen von dem Nutzerbrowser aus initiiert, die der Betreiber eines Videoportals verwenden kann, um Nutzer zu tracken. Startet ein Nutzer ein Video, dann wird auch der Videostrom direkt vom Betreiber des Portals in den Browser geladen, wodurch der Betreiber des Portals zuätzlich noch eine weitere Rückmeldung bzgl. der Nutzerinteressen erhält.

*Medium: Neues Browserfenster (Pop-Up)*

Ein weitere Möglichkeit, mit der man den Nutzerbrowser dazu bringen kann, eine Verbindung zu einer dritten Partei aufzubauen, ist das Öffnen eines neuen Browserfensters (Pop-Up). Dies gelingt über entsprechende Befehle (z.B. JavaScript), die mit der abgerufenen Seite geladen werden. In dem Pop-Up-Fenster werden Inhalte dargestellt, die von dem Tracker geladen werden. Über das Pop-Up-Fenster wird der Anbieter der Inhalte nun zu

[26] Es ist davon auszugehen, dass nicht jedem Nutzer klar ist, dass die eingegebenen Daten direkt an den Anbieter der Routingfunktionalität und nicht an den Betreiber der geladenen Webseite geschickt werden.

einer First-Party, wenngleich der Nutzer weder bewusst noch willentlich eine Verbindung zu dieser Partei aufgebaut hat, da das Fenster ja automatisch geöffnet wurde.

*Technologie: I-Frames*

Bei I-Frames (synonym für *Inline-Frames*) handelt es sich um ein mit dem HTML4.0-Standard im Jahr 1999 eingeführtes Element, das der Strukturierung von Webseiten dient. Mittels I-Frames können Teile einer Webseite von verschiedenen Quellen parallel geladen werden. Für den Nutzer ist es üblicherweise nicht transparent, wenn die über I-Frames eingebundenen Teile einer Webseite von anderen Parteien geladen werden. Dem Nutzer wird beim Laden von I-Frames von anderen Quellen nur die URL der einbettenden Webseite in der Adresszeile des Browsers angezeigt. Durch das Laden von I-Frames können also im Hintergrund Verbindungen zu einem Tracker aufgebaut werden. Hierzu muss ein Embedder lediglich ein solches I-Frame in seine Webseiten integrieren.

*Technologie: Aktive Inhalte wie JavaScript, Active X, Flash, Silverlight, Java Applets*

Wenn aktive Inhalte in Webseiten eingebettet werden, dann stammen diese Inhalte oft nicht von dem Anbieter der geladenen Webseiten. Die geladenen Webseiten enthalten oft nur einen Verweis, so dass die aktiven Inhalte beim Aufbau der Seite direkt von einer dritten Partei geladen und in die restliche Webseite integriert werden. Das Einbindung aktiver Inhalte, die von Dritten zur Verfügung gestellt werden, ist für Nutzer meist nicht transparent. Als aktive Inhalte kommen Elemente wie beispielsweise JavaScript, Active X, Flash, Silverlight oder Java Applets in Frage. Aktive Inhalte eignen sich in besonderer Weise zur Wiedererkennung von Nutzern, wenn die Verbindung zwischen Nutzer und Tracker zustande gekommen ist. Wenn aktive Inhalte geladen sind und ausgeführt werden, dann können sie selbst weitere Verbindungen zu Trackern aufbauen und Daten senden, die zur Wiedererkennung von Benutzern verwendet werden können (siehe hierzu auch z.B. die Abschnitte *Flash Cookies* auf Seite 45 und *Fingerprinting* auf Seite 48).

**Trackingmethoden differenziert nach Informationsfluss bzgl. aufgerufener Seiten**

Im Folgenden werden verschiedene technische Möglichkeiten aufgeführt, mittels derer ein Tracker erfahren kann, welche Webseiten eines Embedder sich ein Nutzer anschaut.

*Referer*

Empfängt ein Webserver eine HTTP-Anfrage, dann kann diese Anfrage einen Referer enthalten.[27] Durch den Referer erfährt der Webserver die URL der Webseite, die im Browser angezeigt wurde, bevor die neue HTTP-Anfrage gestellt wurde [FGM+99]. Referer wurden eingeführt, um einem Webserver mitzuteilen, innerhalb welcher Webseite ein Nutzer einen Link angeklickt hat, um die neu aufgerufene Webseite zu laden. Referer werden jedoch auch in HTTP-Anfragen an Webserver geschickt, wenn Elemente, die in eine Webseite einzubinden sind, von einem anderen Anbieter geladen werden. Ein Beispiel für eine HTTP-Anfrage mit Referer kann folgendermaßen aussehen:

[27] Die korrekte englische Schreibweise lautet eigentlich "Referrer". In dem HTTP-Standard findet sich jedoch die falsche Schreibweise "Referer"[FGM+99]. Seitdem wird bei Bezugnahme im Zusammenhang mit HTTP oft bewusst die falsche Schreibweise verwendet.

```
GET /objects/trackingobject.js
Host: www.iamatracker.com
Referer: http://www.iamanembedder.com/articles/healthcare.htm
```

In dem Referer-Beispiel wird davon ausgegangen, dass in der einbettenden Webseite `http://www.iamanembedder.com/articles/healthcare.htm` das Skript `www.iamatracker.com/objects/trackingobject.js` geladen werden soll. Der Browser schickt in seiner HTTP-Anfrage zum Herunterladen des Skripts in dem Referer-Feld die Information `http://www.iamanembedder.com/articles/healthcare.htm` mit, so dass `www.iamatracker.com` erfahren kann, welche Seite der Nutzer geladen hat.

Die Verwendung von Referern bei HTTP-Anfragen ist nicht zwingend vorgeschrieben, jedoch ist die Option zur Versendung von Referern in den am stärksten verbreiteten Browsern voreingestellt. Werden Referer verschickt, dann kann dies unterschiedlichen Zwecken dienen. So können sie Anbietern von verlinkten Webseiten helfen zu erfahren, über welche Webseiten Nutzer auf den eigenen Seiten gelangen. Jedoch können sie auch für Tracking verwendet werden, da sie Trackern praktisch hinter dem Rücken von Nutzern mitteilen können, welche Webseite ein Nutzer gerade aufgerufen hat, und bieten somit ein einfaches Mittel, personenbeziehbare Daten in Erfahrung zu bringen [Kri10b]. Die Verwendung des Referer-Feldes ist mit jeder der genannten Methoden zur Verbindungsinitiierung kombinierbar.

*POST- und GET-Befehle*

Die Einbettung von POST- und GET-Befehlen in eine aufgerufene Seite bieten weitere Möglichkeiten, einer dritten Partei mitzuteilen, welche Webseite ein Nutzer gerade geladen hat. Da die Erklärungen für POST- und GET-Befehle identisch sind, beschränken wir uns in diesem Dokument auf POST-Befehle. Der POST-Befehl erlaubt es einem Browser, Daten zu einem Webserver zu übertragen [FGM+99]. Normalerweise verwendet man den POST-Befehl zur Übertragung von Formularinhalten, die ein Nutzer zuvor eingegeben hat. Man kann mit dem POST-Befehl jedoch auch die URL der gerade angezeigten Webseite an eine Adresse übertragen, deren Domainname sich von dem Domainnamen der angezeigten Seite unterscheidet. Durch Ausführung eines Skripts kann nach dem Herunterladen einer Webseite automatisch die URL der Webseite ermittelt und mit dem POST-Befehl an eine vorgegebene Internetadresse übertragen werden. Diese Übertragung erfordert keine weitere Interaktion des Nutzers. Ist die Webseite entsprechend präpariert, wird der Browser umgehend den entsprechenden Tracker über das Herunterladen der Webseite informieren, sobald diese geladen ist.

Die Übermittlung der Information per POST-Befehl an einen Tracker, welche Webseite ein Nutzer geladen hat, braucht keine zusätzliche Methode zur Verbindungsinitiierung. Ein POST-Befehl sorgt für eine Initiierung der Verbindung und schickt die Informationen zu aufgerufenen Webseiten über die initiierte Verbindung.

*History-Sniffing*

Beim History-Sniffing spioniert eine Partei einen Nutzer dahingehend aus, von welchen Webservern dieser Nutzer in der Vergangenheit Webseiten geladen hat, indem er sich Zugriff auf die Daten der Verbindungshistorie verschafft. Während die beiden vorangegangenen Methoden Informationen zu der aktuell geladenen Webseite gegenüber einem Dritten preisgeben, liefert das History-Sniffing Einblicke in die Surfhistorie eines Nutzers, die typischerweise von Browsern gespeichert wird [Ang11]. Ein weiterer Unterschied zwischen History-Sniffing

auf der einen Seite und Referern oder POST-/GET-Befehlen auf der anderen Seite besteht darin, dass Referer und POST-/GET-Befehle für Tracking nur Informationen zu entsprechend präparierten Webseiten liefern. Für History-Sniffing ist keine Kooperation zwischen dem Tracker und dem Embedder als Anbieter der Webseiten erforderlich. Arbeitet ein Tracker mit History-Sniffing, dann kann er Informationen zu besuchten Webseiten erhalten, welche der Tracker gar nicht für Tracking präparieren musste. Für History-Sniffing werden die entsprechenden Code-Elemente in die eigenen Webseiten eingebaut. Das Auslesen von Einträgen der Browserhistorie reicht dann zurück bis zu dem letzten Löschzeitpunkt. History-Sniffing funktioniert ohne Wissen und Zustimmung eines Nutzers.

Für History-Sniffing gibt es verschiedene Abstufungen. Zum Auslesen der kompletten Browserhistorie wurde eine Browserschwachstelle ausgenutzt [Clo02], die dann von den Browserherstellern behoben wurde. Eine Variante des History-Sniffing basiert darauf, dass Browser Links zu bereits besuchten Webseiten —d.h. Webseiten, deren URL in der Browserhistorie gespeichert sind— farblich kennzeichnen [JBBM06; WCJJ11; DDNP12]. Tracker können die Farbe dieser Links mittels Javscript in Erfahrung bringen und somit erkennen, ob die URL einer Webseite in der Browserhistorie gespeichert ist. Für History-Sniffing gibt es eine Reihe von Varianten [FS00; JO10; WCJJ11]. Gemäß der Untersuchung in [JJLS10] wird History-Sniffing auf vielen Webseiten angewendet.

**Trackingmethoden differenziert nach Identifikation des Nutzers**

*IP-Adressen*

IP-Adressen bieten eine sehr einfache technische Möglichkeit, um Benutzer wiederzuerkennen. Den meisten Benutzern sind jedoch IP-Adressen nicht statisch zugeordnet, d.h. Nutzer können über der Zeit unterschiedliche IP-Adressen zugeordnet bekommen, so dass IP-Adressen für Tracking gewisse Einschränkungen haben. Wird eine IP-Adresse einer anderen Person zugeordnet, so kann der Nutzer nicht mehr durch die bisher verwendete IP-Adresse verfolgt werden. Ein Tracker hat jedoch Möglichkeiten, anhand derer er erkennen kann, wenn IP-Adressen von anderen Nutzern verwendet werden, indem andere Daten mit der IP-Adresse kombiniert werden. Man kann einen anderen Nutzer mit hoher Wahrscheinlichkeit beispielsweise daran erkennen, wenn sich in einer Anfrage bestimmte Browserdaten geändert haben (z.B. Browsertyp, Versionsnummer des Browsers). Es ist jedoch davon auszugehen, dass IP-Adressen von Nutzern über kürzere Zeiträume stabil bleiben. Gemäß der Untersuchung in [CF07] haben 72% der Nutzer über einen Zeitraum von zwei Wochen eine IP-Adresse behalten können. Die Zeiträume für die Beibehaltung von IP-Adressen hängen von den Regeln der Internet-Provider ab und können stark variieren. Beispielsweise gibt es in Deutschland Provider, welche Computern alle 24 Stunden eine neue IP-Adresse zuweisen.

*Third-Party-Cookies*

Bei Third-Party-Cookies handelt es sich prinzipiell um gewöhnliche HTTP-Cookies, die zwischen Browser und Webserver ausgetauscht werden, wobei der Webserver jedoch nicht zu der Domain gehört, die in der Adresszeile des Browsers angezeigt wird. Die in der Adresszeile angezeigte URL enthält die primäre Adresse, zu welcher der Browser eine Verbindung aufbaut. Typischerweise setzt jede Domain, zu der ein Browser eine Verbindung aufbaut, ihre eigenen HTTP-Cookies. HTTP-Cookies werden später nur an Server solcher Domains

geschickt, von der die HTTP-Cookies empfangen wurden.[28] Third-Party-Cookies bieten die Möglichkeit für eindeutige Identifikatoren, die von einem Webserver in einem Browser gesetzt werden und dann später bei jeder Verbindung zwischen Browser und diesem Webserver ausgetauscht werden. Über diese eindeutigen Identifikatoren können Webserver Benutzer wiedererkennen und diese von anderen Benutzern unterscheiden. Mit Third-Party-Cookies ist eine solche Wiedererkennung auch für dritte Parteien im Hintergrund möglich, da die Verbindungen zum Austausch von Third-Party-Cookies typischerweise nicht von Benutzern wahrgenommen werden. Werden sie auf Benutzerseite gelöscht, dann setzt der Webserver ein neues Cookie. Das neue Cookie wird dann mit einem neuen Identifikationswert vergeben. Werden Cookies und Third-Party-Cookies gelöscht und danach neu gesetzt, dann kann ein Webserver die alten und die neuen Cookies nicht ohne weitere technische Hilfsmittel miteinander verknüpfen. Browser verfügen über Funktionen, um den Gebrauch von Third-Party-Cookies zu verbieten oder die gespeicherten Cookies zu löschen. Andererseits bieten Third-Party-Cookies für Tracker eine sehr beliebte Methode für Tracking: Sie sind einfach zu nutzen, sie werden von jedem Browser unterstützt, ohne dass man als Benutzer irgendwelche Erweiterungen installieren muss, die meisten Benutzer betreiben ihren Browser in einer Einstellung, in der Cookies akzeptiert werden und wenn Cookies einmal gespeichert sind, dann werden Cookies von den meisten Benutzern nur selten gelöscht. Insofern bieten Third-Party-Cookies geeignete Rahmenbedingungen, um bei vielen Benutzern mit sehr geringem Aufwand sehr effektives Tracking betreiben zu können. Löschen Benutzer ihre Cookies oder Third-Party-Cookies, so haben Tracker mittlerweile technische Tricks herausgefunden, wie sie aus der Ferne die gelöschten Cookies wieder rekonstruieren können und somit eine Verknüpfung zu dem alten Identifikator wiederherstellen können [Ang11; AASGH11].

*Flash-Cookies*

Der Begriff *Flash-Cookies* ist eine in der technischen Umgangssprache verwendete Bezeichnung für *Local Shared Objects* (LSO). Diese Bezeichnung wurde in Anlehnung an HTTP-Cookies gewählt. Flash-Cookies dienen einem vergleichbaren Zweck wie gewöhnliche Cookies oder Third-Party-Cookies: Sie erlauben Servern die Speicherung von Daten auf einem Client-Computer. Sie wurden im Jahr 2002 mit der Flash Version 6 eingeführt, um Einstellungen von Benutzern zu speichern, wie z.B. Lautstärkeeinstellungen bei der Flash-basierten Wiedergabe von Videos. Die Verarbeitung von Flash erfordert die Installation eines Flash-Plugin. Die Organisation der von einem Server geschriebenen Daten geschieht in der Regel außerhalb des Browsers. Flash-Cookies eignen sich sehr gut für Tracking [SCM+09; Sch09; Mit10; HSG+12], auch wenn dieser Verwertungszweck wahrscheinlich keine Zielsetzung bei der Einführung von Flash-Cookies war. Tracker können in Flash-Cookies eindeutige Identifikatoren schreiben, die bei einem erneuten Kontakt an die Server geschickt werden, welche die Flash-Cookies geschrieben haben. Anhand der eindeutigen Identifikatoren können Benutzer dann wiedererkannt werden. Flash-Cookies haben gegenüber HTTP-Cookies gerade für Tracking einige Vorteile. Ist der Umfang von HTTP-Cookies auf 4 KByte bechränkt, erlauben Flash-Cookies die Speicherung von 100 KByte [SCM+09; HSG+12]. Darüber hinaus haben Flash-Cookies kein Verfallsdatum [SCM+09; HSG+12]. Sie bleiben also für die Wiedererkennung so lange gespeichert und verwendbar, bis ein Benutzer sie eigenhändig löscht.

[28] Es gibt jedoch bereits Vorschläge zur Entwicklung einer neuen Art von Cookies, die von verschiedenen Domains geteilt werden können [Rab13]. Damit würden die Betreiber von Domains ihre Möglichkeiten zur Kombination von Datensammlungen verbessern.

Somit sind Flash-Cookies persistenter als HTTP-Cookies, was die Arbeit für Tracker erleichtert, d.h. sie sind langlebige, hartknäckige und beharrliche Elemente zur Identifikation von Benutzern und bieten einen stabilen Mechanismus zur Wiedererkennung von Benutzern. Flash-Cookies eignen sich im Gegensatz zu HTTP-Cookies für browserübergreifendes Tracking. Flash-Cookies werden im Gegensatz zu HTTP-Cookies außerhalb von Browsern gespeichert. Verschiedene Browser greifen auf dieselbe Flash-Installation mit denselben Flash-Cookies zurück. Sie können gesetzt und ausgelesen werden, ohne dass ein Benutzer hiervon etwas mitbekommt. Sie eignen sich auch deshalb gut für Tracking, weil sie der überwiegenden Anzahl von Nutzern nicht bekannt sind und diese somit keine Gegenmaßnahmen ergreifen [Sch09]. Diese Tatsache hängt auch damit zusammen, dass die Verwendung von Flash-Cookies für Tracking in der wissenschaftlichen Literatur erst seit dem Jahr 2009 thematisiert wird; die erste Arbeit zur Verwendung von Flash-Cookies für Tracking war [SCM+09]. Nicht alle Browser verfügen über eigene Bordmittel, um Flash-Cookies zu verwalten, zu verbieten oder zu löschen [SCM+09; HSG+12]. Einzige Möglichkeit zum Management und Löschen von Flash-Cookies bieten dann entweder zusätzlich installierte Browser-Plugins oder eigene Managementkomponenten des Flash-Player in der Systemsteuerung [hei11]. Da Flash-Cookies außerhalb des Browsers verwaltet werden, hinterlassen sie sogar Spuren, selbst wenn Benutzer den *Private Mode* ihres Browsers aktiviert haben [SCM+09]. Mittels Flash-Cookies ist es sogar möglich, gelöschte Third-Party-Cookies wiederherzustellen und somit das Löschen praktisch rückgängig zu machen [SCM+09; AASGH11; Ang10; Ang11]. Da viele Internetangebote nur mit installiertem und aktiviertem Flash-Plugin funktionieren (z.B. die in Videoportalen verwendeten Videoplayer basieren meistens auf Flash), verwenden viele Nutzer Flash. Deshalb bietet Flash eine gute Ausgangssituation für Tracker.

*E-Tags*

Im Jahr 2011 wurde entdeckt, dass E-Tags als Identifikatoren für Tracking eingesetzt werden [AASGH11]. Bei E-Tags handelt es sich um Cache-Cookies, die bereits im Jahr 1999 mit dem Hypertext Transfer Protocol – HTTP/1.1 eingeführt wurden [FGM+99]. Das Ziel des Caching besteht darin, dass Webseiteninhalte bei wiederholten Aufrufen aus Effizienzgründen von der Festplatte des Client-Computers geladen werden, wo sie nach der ersten Verwendung gespeichert wurden. Bei der ersten Verwendung eines Webseiteninhalts versendet ein Server mit dem ausgelieferten Inhalt ein Cache-Cookie, für welches der Server einen eindeutigen Identifikator vergeben kann. Wird bei einer späteren Anfrage derselbe Webseiteninhalt nochmal benötigt, dann sendet der Computer des Benutzers das E-Tag an den Server. Dieser wertet nun das E-Tag aus und muss zum Aufbau der Webseite diejenigen Inhalte nicht ausliefern, die bereits im Cache-Speicher auf der Festplatte liegen. Der Computer baut die Webseite dann mit den Daten auf, die er von seiner Festplätte lädt. Die Eindeutigkeit der E-Tags kann für Tracking ausgenutzt werden. Ein wichtiger Unterschied zwischen E-Tags und HTTP-Cookies für Tracking besteht darin, dass Nutzer normalerweise keine Möglichkeit haben, den Austausch von E-Tags zu verbieten. Bereits gespeicherte E-Tags werden durch das Löschen des Cache-Speichers des Browsers gelöscht [AASGH11]. Für den durchschnittlichen Nutzer ist dies nicht unbedingt offensichtlich; es gibt in Browsern keine Bedienelemente, die auf das Löschen von E-Tags hinweisen. Zusätzlich ist festzuhalten, dass E-Tags auch im *Private Mode* funktionieren. E-Tags werden auch für Evercookies verwendet [AASGH11; hei11; MM12] (siehe hierzu Abschnitt *Evercookies* auf Seite 48), d.h. sie können

zusammen mit Flash-Cookies dazu verwendet werden, um andere Identifikatoren nach ihrer Löschung zu rekonstruieren.

*HTML5-Storage*

HTML5 ist eine Sprache zur Strukturierung und zur semantischen Beschreibung von Inhalten wie Texten, Bildern und Hyperlinks in Dokumenten. Die mit dieser Sprache —oder einer ihrer Vorgängersprachen— beschriebenen Dokumente sind die Grundlage des World Wide Web; um Webseiten darstellen zu können, führen Browser die Anweisungen der Sprache aus. Der HTML5-Standard befindet sich noch in der Entwicklung, jedoch liegen bereits Entwürfe und Implementierungen vor. HTML5 soll die Nachfolge von HTML4 antreten.

HTML5 bietet eine weitere Möglichkeit, um Verbraucher mittels eindeutiger Identifikatoren zu verfolgen. Die Entwicklung dieses neuen Standards für HTML wurde hauptsächlich durch Rahmenbedingungen getrieben, welche in der Praxis echte Probleme darstellen: Webseiten werden von Verbrauchern mit verschiedenen Browsern auf verschiedenen Geräten mit unterschiedlichen Ressourcen aufgerufen (z.B. PCs und Smartphones mit unterschiedlicher Bandbreite, Bildschirmauflösung, Verarbeitungsgeschwindigkeit) [Ant12; Mat13]. Das bedeutet, dass die Anbieter von Webseiten ihre Inhalte für alle diese unterschiedlichen Typen von Empfängern gesondert vorhalten müssen, was mit bisherigen technischen Mitteln sehr aufwendig ist. HTML5 soll zur Lösung dieser Probleme beitragen. Abgesehen davon kann HTML5, insbesondere der HTML5-Web-Storage[29], jedoch auch zum Tracking verwendet werden [AASGH11; MM12; HSG+12; RKW12; SV11; WP12; Val10a; Mit10]. HTML5-Web-Storage stellt einen persistenten Speicher auf den Endgeräten von Verbrauchern dar, auf den Server über eine Schnittstelle zugreifen können, um dort Daten zu schreiben oder zu lesen.[30] Vereinfacht gesagt stellt HTML5-Storage damit eine ähnliche Funktionalität wie HTTP-Cookies oder Flash-Cookies dar und kann ganz analog von dritten Parteien zum Tracking verwendet werden. HTML5-Storage hat jedoch gegenüber einfachen HTTP-Cookies oder Flash-Cookies für Tracking gewisse Vorteile und schafft bessere Voraussetzungen für Tracking. Die Größe eines HTTP-Cookies ist auf 4 KByte beschränkt, die eines Flash-Cookies auf 100 KByte. HTML5-Storage erlaubt hingegen die Speicherung von Daten bis zu einem Volumen von 5 MByte. Inhalte im HTML5-Storage haben kein Verfallsdatum und bleiben so lange gespeichert, bis der Nutzer diesen Speicher selbst leert. Im Vergleich zu Flash-Cookies bietet HTML5-Storage für Tracker den Vorteil, dass zur Ausführung von HTML5 keine Browsererweiterungen erforderlich sind, d.h. alle HTML5-fähigen Browser können unmittelbar für Tracking mittels HTML5-Storage verwendet werden. Ein anderer Grund, der HTML5-Storage für Tracker attraktiv macht, besteht darin, dass das Thema HTML5-Storage bei den meisten Benutzern wahrscheinlich noch nicht als Ausgangspunkt von Bedrohungen wahrgenommen wurde und hierzu noch wenig bekannt ist. Darüber hinaus kann HTML5-Storage für das Tracking mittels Evercookies (siehe hierzu Abschnitt Ever-

[29]Die Spezifikation des HTML5-Web-Storage geschah ursprünglich im Rahmen der Standardisierung von HTML5. Später wurde das, was man mit HTML5-Web-Storage beabsichtigt hat, aus der HTML5-Spezifikation herausgelöst und in eine eigene Spezifikation ausgegliedert [Hic13]. Es sind viele synonyme Begriffe im Umlauf, was zur Verwirrung beträgt, wie etwa *Local Storage*, *Web Storage* oder *DOM Storage* [Tri13]. Im Zusammenhang mit Tracking wird HTML-Storage auch gelegentlich als *Supercookie* bezeichnet. In diesem Dokument wird die Bezeichnung *HTML5-Storage* verwendet.

[30]Die Zugriffe auf diesen Speicherbereich erfolgen geschützt durch die Same-Origin-Policy, d.h. Server können nur auf solche Inhalte lesend zugreifen, die auch von ihnen dort abgespeichert wurden.

cookies auf Seite 48) verwendet werden, d.h. für die automatische Wiederherstellung von Cookies nach deren Löschung.

*Evercookies*

Evercookies sind ein sehr fortgeschrittenes Werkzeug für Tracking. Der Begriff *Evercookie* steht für eine Software und ein allgemeines Prinzip. So heißt die von Samy Kamkar entwickelte API Evercookie, mit der man an mehr als 10 Stellen auf dem Computer Identifikatoren speichern kann [Veg10].[31] Mit einer solchen redundanten Speicherung von Identifikatoren wird es Benutzern praktisch unmöglich gemacht, alle einzelnen Identifikatoren zu löschen, da den Benutzern gar nicht klar ist, wo überall auf ihrem Computer Identifikatoren geschrieben werden. Die Evercookie-API verwendet unter anderem HTTP-Cookies, Flash-Cookies, Silverlight-Cookies, E-Tags und HTML5-Storage als Identifikatoren, mit denen jeweils unterschiedliche Speicherplätze verbunden sind. Neben dieser Software bezeichnet man mit Evercookie das allgemeine Prinzip, bei welchem viele verschiedene Identifikatoren redundant verwendet werden. Die hierbei verwertbaren Identifikatoren sind nicht auf die Identifikatoren beschränkt, die von der Evercookie-API verwendet werden. Eine allgemeine Zusammenstellung von Identifikatoren findet man in [MM12]. Es gibt bei Evercookies für Benutzer keine Möglichkeit, durch den Aufruf einer einzigen Funktion alle an verschiedenen Stellen gespeicherten Identifikatoren zu löschen. Selbst wenn ein Nutzer glaubt, dass er alle Identifikatoren gelöscht hat, dann ist es möglich, dass noch weitere Identifikatoren gesetzt sind, so dass weiterhin Tracking möglich ist [SV11]. Da ein durchschnittlicher Nutzer wahrscheinlich nicht dazu in der Lage ist, alle Identifikatoren zu löschen und somit immer irgendwelche Identifikatoren zurückbleiben, wurde dieser Mechanismus Evercookie genannt.[32] Es gibt jedoch noch spezielle Weiterentwicklungen von Evercookies, die sich neben der Redundanz von Identifikatoren dadurch auszeichnen, dass gelöschte Identifikatoren automatisch wiederhergestellt werden, wenn nur Teile der Identifikatoren gelöscht werden. Vor diesem Hintergrund bekommt die Bezeichnung Evercookie nochmal eine ganz andere Rechtfertigung: Selbst wenn Nutzer Evercookies löschen, werden diese wieder hergestellt. Die Verwendung von Evercookies für Tracking ist erst seit wenigen Jahren bekannt. Im Zusammenhang mit Evercookies sind insbesondere die Fälle von QuantCast und KISSmetrics in den Medien bekannt geworden [AASGH11; hei11]. Laut [Val10b] wirbt das Unternehmen ThreatMetrix für Evercookies, um damit Kriminelle erkennen zu können. Wenn Kriminelle bewusst ihre Cookies löschen, um anonym bleiben zu können, dann sollen Evercookies bei ihrer Wiedererkennung helfen.

*Fingerprinting*

Fingerprinting ist eine sehr fortgeschrittene Technik für Tracking. Hierbei geht es Trackern darum, bei Verbindungen mit Nutzern verschiedene Konfigurationsaspekte der Computer von Nutzern auszulesen und die Verbraucher darüber wiederzuerkennen. Beispiele für solche Konfigurationsaspekte sind etwa der verwendete Browser oder auch das Betriebssystem des Verbrauchers. Jeder einzelne Konfigurationsaspekt ist hierbei als Identifikator zu verstehen, der zur Wiedererkennung dient. Jedoch sind diese Identifikatoren nicht trennscharf genug, da ein Identifikator bei vielen Nutzern die gleiche Wertebelegung liefern kann. In diesem Fall werden dann Kombinationen von Identifikatoren verwendet, da durch Kombi-

[31] siehe hierzu auch `http://samy.pl/evercookie/`.
[32] Gelegentlich werden Evercookies auch *Supercookies* genannt; siehe z.B. [MM12].

nation von Konfigurationsaspekten die Trennschärfe ansteigt. Ist die Kombination von verschiedenen Identifikatoren trennscharf genug, dann ergibt sie eine Information, die einem Fingerabdruck gleichkommt.[33] Bei den für Tracking auswertbaren Konfigurationsaspekten kann unterschieden werden, ob ein Computer diese standardmäßig preisgibt, ohne dass der Server eines Tracker den Computer des Verbrauchers hierzu explizit auffordert oder ob dies nur nach expliziter Aufforderung von Trackerseite geschieht. Fingerprinting nach standardmäßiger Preisgabe von Konfigurationsaspekten bezeichnet man als *passives Fingerprinting*, Fingerprinting mit zusätzlicher Aufforderung zur Preisgabe weiterer Konfigurationsaspekte nennt man *aktives Fingerprinting* [MM12]. Für aktives Fingerprinting schicken Tracker typischerweise Skripte (z.B. als JavaScript), um die nicht standardmäßig herausgegebenen Konfigurationsaspekte herauszufinden. Browser teilen Servern die bei Fingerprinting verwerteten Informationen mit, damit Server ihre Inhalte so bereitstellen, dass sie für die besonderen Rahmenbedingungen von Browsern angepasst und die Inhalte entsprechend gut dargestellt werden können. Die für passives und aktives Fingerprinting verwertbaren Konfigurationsaspekte lassen sich danach unterscheiden, auf welcher technischen Ebene sie ansetzen [NKJ+13; MM12]:

- Ebene der Browsererweiterungen: Hierzu gehören beispielsweise die Liste der Plugins oder der unterstützten MIME-Types. Sogar die Reihenfolge der Einträge in diesen Listen ist ein weiterer starker Identifikator, z.B. wenn die Reihenfolge dieser Liste die zeitliche Abfolge der Installation wiedergibt.
- Ebene der Browsereinstellungen: In dieser Kategorie sind Konfigurationsaspekte für Fingerprinting verwertbar wie die eingestellte Zeitzone, die bevorzugte Sprache, die Einstellungen zu HTTP-Cookies und Third-Party-Cookies, die Flash-Option im Browser, Datum, Zeit[34] und sogar die Einstellungen für den Do-Not-Track-Header.[35]
- Ebene von Browserfamilie und Version: Auf dieser Ebene können die browserbezogenen Daten zum verwendeten Browser (User Agent) für das Tracking ausgewertet werden. Hierzu gehören die Browserfamilie und die Browserversion oder auch die in einem Browser verwendeten Komponenten wie z.B. die verwendete AJAX-Implementierung.
- Ebene von Betriebssystem und Anwendungen: Auf der Ebene von Betriebssystemen kann beispielsweise das verwendete Betriebssystem mit Versionsnummer (als Teil des User Agent) oder die Kernel-Version des Betriebssystems abgefragt werden. Darüber hinaus kann die Liste der installierten Schriftarten in Erfahrung gebracht werden. Neben den Schriftarten selbst ist hier auch die Reihenfolge der Schriftarten in der Antwort an den Tracker ein sehr guter Identifikator für den Computer.
- Ebene von Hardware: Auf dieser Ebene werten Tracker Konfigurationsaspekte wie die Bildschirmauflösung, die IP-Adresse sowie die vorhandenen Treiber aus. Da die Antwort zu den Treibern in Form einer Liste geliefert wird, kann die Reihenfolge der Listeneinträge wieder als zusätzlicher Identifikator ausgewertet werden.

[33] Aus einem Fingerabdruck kann man eindeutig auf die zugehörige Person schließen, wenn der Fingerabdruck einer Person bekannt ist. Vergleichbares wird hier mit dem Computer eines Verbrauchers gemacht.

[34] Die von einem Browser preisgegebene Zeit und die Systemzeit des Tracker bilden eine über bestimmte Zeiträume relativ stabile Differenz. Aus der Perspektive eines Tracker können verschiedenen Verbrauchern verschiedene Zeitdifferenzen zugeordnet werden. Diese Zeitdifferenzen stellen wiederum Identifikatoren zur Wiedererkennung von Computern dar. Die Fingerprinting-Bibliothek von Iovation (`http://www.iovation.com`) unterstützt die Abfrage dieses Konfigurationsaspektes.

[35] Die Bibliothek von BlueCava `http://www.bluecava.com` verwendet den Do-Not-Track-Header für Fingerprinting. Damit wird eine Lösung, die eigentlich dafür entwickelt wurde, um Benutzer besser zu schützen, dafür verwendet, um das Fingerprinting für Tracking zu verbessern.

Aus der Kombination dieser Informationen können eindeutige Identifikatoren konstruiert werden. Die Konstruierbarkeit von solchen Fingerprints für Tracking wurde 2009 von Mayer in [May09] und von Eckersley in [Eck09; Eck10b; Eck10a] beschrieben. Die Electronic Frontier Foundation hat mit Panopticlick[36] ein Projekt durchgeführt, bei dem untersucht wurde, wie gut sich in der Praxis Browser bzw. Computer mittels Fingerprint unterscheiden lassen. Die in dem Projekt verwendete Open-Source-Bibliothek Panopticlick ist weitaus nicht so mächtig wie die heute verfügbaren kommerziellen Angebote für Fingerprinting, wie ein qualitativer Vergleich in [NKJ+13] ergeben hat. Mit den vergleichsweise einfachen Mitteln von Panopticlick konnten von mehr als 470.000 Browsern 83,6% eindeutig erkannt werden. Weitere 5,3% der Browser haben in den betrachteten Konfigurationsaspekten mit genau einem anderen Browser übereingestimmt. Betrachtet man von den 470.000 Browsern nur diejenigen, bei denen noch zusätzlich entweder Flash oder Java installiert war, dann hatten unter diesen 94,2% eindeutige Fingerprints; 4,8% der Browser hatten Fingerprints, die mit einem Browser auf einem anderen Computer übereingestimmt haben. Es ist anzunehmen, dass man mit kommerziellen Fingerprinting-Bibliotheken wie z.B. von BlueCava[37], Iovation[38] oder ThreatMetrix [39] nochmal deutlich bessere Trennschärfen erreichen kann und über deren Fingerprints sehr nahe an eindeutige Identifikatoren herankommt. Da sich die beim Fingerprinting betrachteten Konfigurationsaspekte nicht häufig ändern, sind die Identifikatoren stabil genug für Tracking. Aber selbst wenn sich Konfigurationen ändern sollten, dann können die Konfigurationen, die zu demselben Computer gehören, in den meisten Fällen korrekt zugeordnet werden. In [Eck09; Eck10b] konnte diese Zuordnung in mehr als 99% der Fälle korrekt vorgenommen werden.[40] Die Entdeckung von Fingerprinting kann in der Praxis sehr schwierig sein. Für passives Fingerprinting müssen Tracker keine Cookies mehr speichern und auslesen und es muss auch kein Tracking-Skript geladen werden. Für passives Fingerprinting genügt das reine Zustandekommen einer Verbindung zwischen einem Tracker und einem Verbraucher. Fingerprinting ist immer noch ein Gegenstand der Forschung. Um weitere Aussagen über Fingerprinting und Tracking treffen zu können, ist es erforderlich, die Kommunikation zwischen potenziellen Trackern und Computern zu analysieren und ein Reverse-Engineering vorzunehmen, um die neuesten Tricks der Tracker für Fingerprinting zu erfahren. Das ist in der Regel sehr aufwendig und kann auch sehr schwierig sein, insbesondere dann, wenn Tracker es darauf anlegen, dass man ihre Techniken nicht einfach enttarnen können soll. Wissenschaftler schlagen auch selbst neue Mechanismen für Seitenkanäle vor (z.B.[BMFGI11; MBYS11; MS12b]), die man zum Tracking ausnutzen kann, die dann wiederum beim Reverse-Engineering von Tracker-Interaktionen angewendet werden können. Fingerprinting von Computern kann durch Fingerprinting von Personen ergänzt werden. Arbeitet ein Benutzer an verschiedenen Geräten, z.B. auf denen Browser unterschiedlich konfiguriert sind, dann geht es darum, den Benutzer geräteübergreifend wiederzuerkennen. Hierfür verwendet man eine Fingerprinting-Technik für Benutzer [Sno14]. Bei dieser Technik geht man davon aus, dass verschiedene Verbraucher unterschiedliche Gewohnheiten in der Nutzung des World Wide Web haben, z.B. ein Benutzer ruft immer wieder Seiten derselben Anbieter auf. Betrachtet man die Menge der von einem Verbraucher bei verschiedenen Anbietern aufgerufenen Seiten, dann kann man darin ein für einen Benut-

[36] siehe `https://panopticlick.eff.org`
[37] siehe `http://www.bluecava.com`
[38] siehe `http://www.iovation.com`
[39] siehe `http://www.threatmetrix.com`
[40] Diese Erfolgswahrscheinlichkeit kann auch dadurch gesteigert werden, indem man Tracking durch Fingerprinting mit einer anderen Trackingmethode kombiniert wie z.B. Flash-Cookies.

zer typisches Muster —also einen Fingerprint für den Benutzer— erkennen, die unabhängig von den Geräten sind, die ein Verbraucher verwendet. Können für verschiedene Fingerprints von Geräten oder Browsern die gleichen Fingerprints von Benutzern erkannt werden, dann hat man ein starkes Indiz dafür, dass die Geräte von demselben Benutzer verwendet werden. Diese Technik lässt ein Inbeziehungsetzen von Benutzern über verschiedene Geräte zu und ergänzt das Fingerprinting von Computern für geräteübergreifendes Tracking. Mittlerweile haben eine Reihe von Unternehmen Tracking mittels Fingerprinting im Angebot wie 41st Parameter, Iovation, ThreatMetrix oder BlueCava. Bei den drei Erstgenannten handelt es sich um Unternehmen, die mit Tracking für ihre Technologie ein neues Geschäftsmodell gefunden haben. Ursprünglich wurde Fingerprinting nämlich zum Schutz entwickelt. Die Motivation zur Technologieentwicklung lag im Schutz von Verbrauchern gegen Missbrauch von Kreditkarten und im Schutz gegen Software- und Medienpiraterie [Mil09; AV10b].[41]

### 3.4.6 Weitere Verarbeitung der durch Tracking gewonnenen Daten

Bei den Trackingmethoden, wie sie in 3.4.5 beschrieben wurden, geht es um die Förderung von Daten zum Webnutzungsverhalten von Verbrauchern als Rohstoff. Neben der Rohstoffquelle Web-Tracking gibt es noch weitere Quellen, wie die Erfassung von nutzergenerierten Inhalten im Zusammenhang mit Web-2.0 und soziale Netzwerke, Tracking von Apps und Tracking im Zusammenhang mit Smart TV.[42] Neben der reinen Förderung dieser Daten als Rohstoff, geht es für Tracker irgendwann darum, diesen Rohstoff weiter zu verarbeiten.

Wenn die Daten als Rohstoff vorliegen, dann müssen sie für die spätere Verwertung aufbereitet werden. Insbesondere ist dies der Fall, wenn Daten unstrukturiert vorliegen wie bei nutzergenerierten Inhalten. In der Praxis werden hierzu Verfahren des *Natural Language Processing* verwendet, um ein noch genaueres Bild von Benutzern zu bekommen; das US-amerikanische Unternehmen *LOTAME*[43] ergänzt sein Tracking mit dieser Technologie laut [AV10a]. *LOTAME* greift hierfür auf Technologie und Dienstleistung des britischen Unternehmens *OpenAmplify*[44] zurück. Die von Benutzern in Freitextformularfeldern eingegebenen natürlichsprachlichen Kommentare oder Mitteilungen werden hierfür im Hintergrund an *OpenAmplify* weitergeleitet und analysiert. Die Analyse solcher unstrukturierter Daten ist für Computer immer noch schwierig. Laut Aussagen von *OpenAmplify* liefern die eingesetzten Algorithmen noch genauere Einblicke bzgl. der Verbraucher [AV10a]. Gerade in Plattformen für soziale Netze lassen sich mit dieser Technologie tiefe Einblicke in die Interessen und Denkweisen von Benutzern herausfinden.

Nach der Strukturierung von Daten können diese mit Daten aus anderen Quellen kombiniert werden. Der somit erhaltene Datenbestand wird dann für den jeweiligen Verwertungszweck analysiert. Die Ergebnisse reichen von der Werbeanzeige, die sich an der Klickhistorie eines Verbrauchers und an den Inhalten seiner geschriebenen und empfangenen E-Mails orientiert, bis hin zu Scoring-Werten für die Vergabe von Krediten, in den Ergebnisse von

[41] Die Techniken des Fingerprintings wurden ursprünglich dazu entwickelt, Verbraucher zu schützen. Die Idee für das Fingerprinting bestand darin, erkennen zu können, ob abgefangene oder geteilte Zugangsdaten möglicherweise an einem anderen Computer eingegeben wurden. Anbieter von zahlungspflichtigen Inhalten wenden diese Technik an, um zu erkennen, ob sich verschiedene Benutzer Zugänge zu ihren Inhalten teilen. Fingerprinting kann auch zur Erkennung von Kreditkartenmissbrauch verwendet werden, z.B. um zu erkennen, wenn die Kreditkartendaten an einem anderen Computer eingegeben wurden.

[42] Die beiden letztgenannten Alternativen zum Web-Tracking werden im Kapitel 3.5 behandelt.

[43] `http://www.lotame.com/`

[44] `http://www.openamplify.com/`

Big-Data-Algorithmen eingehen und die aufgrund von Ähnlichkeiten in der Verwendung des World Wide Web mit anderen Verbrauchern vorgeschlagen wurden.

## 3.5 Weitere Formen des Tracking

Tracking ist heute nicht nur auf das World Wide Web beschränkt. Datensammler haben längst weitere Technologien und Anwendungsbereiche für das Tracking ausfindig gemacht und nutzen diese bereits. Die Durchführung von Tracking geht einher mit der Verbreitung von sogenannten Cyber-physikalischen Systemen, dem Internet der Dinge und der Dienste. Dadurch dass Dinge des täglichen Lebens, die Personen zugeordnet werden können, immer intelligenter werden und selbst mit dem Internet verbunden sind, worüber sie Dienste nutzen und Daten austauschen, enstehen neue Ansatzpunkte für Tracking. Die Vernetzung von verschiedenen Alltagsgegenständen über das Internet bietet die für das Tracking so wichtige Möglichkeit zur bidirektionalen Kommunikation in Anwendungsbereichen, in denen es vorher entweder gar keinen Netzzugang oder lediglich unidirektionale Kommunikationsmöglichkeiten gab (z.B. Rundfunknetze mit unidirektionaler Verteilkommunikation). Somit lassen sich bekannte Mechanismen aus dem Web-Tracking auf andere Technologiebereiche übertragen und dort für Tracking anwenden, wie z.B. für beliebige kommunikationsfähige Geräte wie Smartphones, Fernseher, Set-Top-Boxen oder Spielekonsolen [AV10b]. Selbst wenn es für bestimmte dieser Geräte keine eigenen Tracking-Mechanismen gibt, dann tragen diese Geräte dennoch dazu bei, dass immer mehr Daten zu Verbrauchern gesammelt werden, z.B. indem die Daten für Verarbeitungs-, Analyse- und Darstellungszwecke zu Internetdiensten hochgeladen werden und dort von anderen in Erfahrung gebracht werden können.

Die denkbaren Datenquellen für Tracking sind heute eigentlich unüberschaubar; der Trend zur Verschmelzung von Unterhaltungselektronik auf der einen Seite und Informations- und Kommunikationstechnologie auf der anderen Seite ist noch nicht an seinem Ende angekommen, und auch bei den durch Informations- und Kommunikationstechnologie getriebenen Innovationen bei Alltagsgegenständen werden in den kommenden Jahren noch viele weitere Produkte entstehen, die zur Sammlung von Verbraucherdaten beitragen können. Die Erfassung von Daten geschieht mit Geräten, die im öffentlichen Raum installiert sind, und durch Geräte, die den Verbrauchern selbst gehören. Als Beispiel für die Sammlung von Daten im öffentlichen Raum sei die Installation von Kameras für die Überwachung im Straßenverkehr und von öffentlichen Plätzen genannt. Die Technologien für das Tracking von Nummernschildern oder zur Erkennung von Gesichtern sind mittlerweile sehr weit fortgeschritten [AV12]. Zur Behandlung von Tracking mit eigenen Geräten sollen hier nur Smartphones bzw. Apps und moderne Fernseher für das sogenannte Smart TV kurz betrachtet werden; siehe hierzu die Kapitel 3.5.1 und 3.5.2.

Wenn Daten aus vielen Quellen erhoben werden, dann erhalten Tracker nochmal einen weiteren Mehrwert, wenn sie Daten für neue Informationen miteinander kombinieren. Um ein möglichst geschlossenes Bild über Benutzer zu bekommen, haben Tracker längst damit begonnen Lösungen zu entwickeln, um die über verschiedene Wege gewonnenen Daten zueinander in Beziehung setzen zu können [AV10b]. Nutzt ein Benutzer mehrere Geräte oder Anwendungen, so sollen die von einem Benutzer über verschiedene Wege gewonnenen Profile zusammengeführt werden können. Unternehmen wie z.B. BlueCava[45] haben sich zum Ziel gesetzt, Benutzer über ihre verschiedenen Geräte hinweg verfolgen zu können. Das

[45] `http://bluecava.com`

Cross-Domain-Tracking gewinnt durch das Cross-Device-Tracking eine neue Dimension, mit der nahezu alle Bereiche des täglichen Lebens erfasst und in Profilen zusammen getragen werden können.

### 3.5.1 Tracking durch Smartphone-Apps

Smartphones sind die ständigen Begleiter von Verbrauchern. Mit ihnen wird telefoniert, getextet für SMS, E-Mails und Instant Messaging, fotografiert, Musik gehört, im Internet gesurft, Kontakte und Kalender verwaltet, Wege gesucht. Darüber hinaus sind in Smartphones Sensoren enthalten, die Apps Kontextinformationen für viele sinnvolle Funktionen zur Verfügung stellen, z.B. Positionsdaten von Verbrauchern (GPS). Smartphones und Smartphone-Apps haben Zugang zu vielen Informationen über Verbraucher [UHL12]. Leider behalten Smartphones und Apps diese Informationen nicht geheim, sondern geben diese Informationen an Dritte heraus [TK10]. Bei diesen Dritten kann es sich um die Entwickler der Smartphone-Apps handeln, jedoch auch um solche dritte Parteien, für welche die Entwickler Möglichkeiten anlegen, dass sie auf Daten von Verbrauchern zugreifen können.[46]

Um zu verstehen, warum Entwickler dritten Parteien über ihre Apps Zugriffsmöglichkeiten auf die Smartphones einräumen, muss man das Ökosystem des App-Markts verstehen, das sich in den vergangenen Jahren entwickelt hat. Der überwiegende Teil von Apps wird den Verbrauchern entweder kostenfrei oder zu einem sehr geringen Preis angeboten [PPP+13]. Entwickler haben für ihr Geschäftsmodell neue Einnahmequellen entdeckt. Hierzu gehört die Einnahmequelle durch die Integration von Werbung in Apps, insbesondere auch personenbezogene Werbung. Für diese braucht die Werbeindustrie Informationen zum Benutzer, die sie sich über den Zugriff auf das Smartphone verschaffen kann. Entwickler können hierzu fertige Tracking-Frameworks in ihre Apps integrieren.[47] Mittels dieser Bibliotheken können Programmierer relativ einfach vorgeben, welche Daten, Aktionen und Events der Benutzer von der App erfasst werden sollen und wann und an wen diese Informationen zu versenden sind. Wird die App später vom Verbraucher ausgeführt, dann fließen die über das Tracking-Framework spezifizierten Informationen zum Tracker. In diesem Geschäftsmodell der Entwickler ersetzt die Einnahmemöglichkeit durch Bereitstellung von Werbefläche oder von Datenzugriffen die Einnahmen, die sich traditionell durch den Verkauf von Software ergeben haben [LEPM12]. Somit kann es für Entwickler betriebswirtschaftlich interessant sein, Apps kostenfrei anzubieten.

Das Ökosystem rund um den App-Markt konnte sich so entwickeln, da die Betriebssysteme für Smartphones die entsprechenden technischen Voraussetzungen erfüllen. Wird eine App installiert, dann erhält sie umfangreiche Rechte auf dem System. Durch die Gewährung weitreichender Rechte kann eine App auf viele Ressourcen zugreifen, wodurch die Privatsphäre der Smartphone-Inhaber erheblich verletzt werden kann. Im Folgenden werden einige technische Eigenschaften beschrieben, aus denen hervorgeht, warum sich Smartphones für Tracking so gut eignen [HJW12]:

[46] Damit sind hier explizit Unternehmen gemeint und nicht die Geheimdienste, auch wenn durch die Enthüllungen von Edward Snowden bekannt wurde, dass Geheimdienste über Smartphones auf Verbraucherdaten zugreifen [Spi14a].

[47] Als Beispiele seien hier das Framework von hasoffers (siehe `http://www.hasoffers.com` und `http://mobileapptracking.com`), das Admob-Framework von Google (siehe `http://www.google.com/ads/admob/index.html`), das Framework von Mindlab (siehe `http://www.mindlab.de` und `http://www.mindlab.de/rapptor/netmind-rapptor-appliance/app-tracking-technologie/`) oder das Framework von Mobclix (siehe `http://www.mobclix.com`) genannt.

— Persistente oder quasipersistente Identifikatoren: Apps können auf Identifikatoren zugreifen, die sehr viel beständiger sind als HTTP-Cookies. Diese Identifikatoren haben die Eigenschaft, dass sie eindeutig sind und von Verbrauchern entweder gar nicht oder nur sehr umständlich gelöscht werden können. Darüber hinaus hat man als Verbraucher keine Möglichkeit, das Versenden dieser Identifikatoren zu verbieten. Zu diesen Identifikatoren gehören beispielsweise
  ◦ System IDs wie AndroidID und Unique Device Identifier (UDID) bis iOS5 einschließlich und Identifier for Advertising (IDFA) ab iOS6. Die AndroidID ist eine 64-bit Nummer, die beim ersten Start des Betriebssystems automatisch angelegt wird. Die Anzahl der möglichen Kombinationen ist eine 19-stellige Dezimalzahl, so dass davon auszugehen ist, dass jedes Gerät einen eindeutigen Identifikator hat. Apps können auf diese Identifikatoren zugreifen.
  ◦ International Mobile Station Equipment Identity (IMEI), eine eindeutige 15-stellige Nummer, anhand derer jedes GSM- oder UMTS-Endgerät eindeutig identifiziert werden kann. Für CDMA-Geräte gibt es den Mobile Equipment Identifier (MEID). Diese Identifikatoren sind eindeutig und bleiben auch dann erhalten, wenn das Smartphone komplett neu aufgesetzt wird. Sie können so lange einer Person zugeordnet werden, wie diese Person ihr Gerät behält. Unter Android kann eine App auf diese Identifikatoren zugreifen, wenn sie die Berechtigung `READ_PHONE_STATE` hat.
  ◦ SIM-Karten-Identifikatoren wie International Mobile Subscriber Identity (IMSI) oder die Integrated Circuit Card ID (ICCID), einer eindeutigen ID für die Chipkarte. Diese Identifikatoren bleiben erhalten, solange man dieselbe SIM-Karte nutzt; typische Nutzungszeiträume gehen hier über mehrere Jahre. Um diese Identifikatoren unter Android auslesen zu können, braucht die App ebenfalls die Berechtigung `READ_PHONE_STATE`.
— Keine Unterscheidung zwischen First-Party und Third-Party: Beim Web-Tracking auf PCs unterscheiden Browser zwischen First-Party und Third-Party und Verbraucher haben die Möglichkeit, für First-Party und Third-Party unterschiedliche Berechtigungen einzustellen. Smartphones bzw. Smartphone-Apps bieten diese Differenzierung nicht an. Das bedeutet, dass jede App, die auf das Netz zugreifen kann, Nachrichten zu jeder beliebigen Adresse schicken kann. Wenn Entwickler also Code von Tracking-Frameworks in ihre App integrieren, dann wird dieser Code mit den gleichen Berechtigungen ausgeführt wie der Code, der die Funktionen der App implementiert. Wenn also die App bestimmte Berechtigungen erhält, dann haben die Trackingfunktionen die gleichen Berechtigungen.
— Keine Kontrolle und Steuerung von Tracking: Beim Web-Tracking haben Verbraucher abhängig von den verwendeten Trackingmethoden gewisse Möglichkeiten zu erkennen, ob Trackingversuche unternommen wurden (z.B. durch Kontrolle der gespeicherten Cookies). Darüber hinaus können sie gewisse Gegenmaßnahmen ergreifen (z.B. indem sie Cookies löschen oder Third-Party-Cookies verbieten). Diese Möglichkeiten gibt es beim App-Tracking nicht. Die einzige Entscheidung, die ein Benutzer zu seinem Schutz treffen kann, ist die, ob er eine App installiert oder nicht [LEPM12].

Apps können derart implementiert werden, dass Daten aus verschiedenen Ressourcen des Smartphones übermittelt werden. Zu diesen potenziell zugreifbaren Ressourcen gehören [HHJ+11]:

— Aufenthaltsort des Verbrauchers

- Telefonnummer bzw. andere eindeutige Identifikatoren des Verbrauchers
- Kontaktliste mit Adressbuch
- Kamera
- Mikrophon
- Browserhistorie
- Bookmarks
- Loglisten
- SMS-Nachrichten
- Kalender
- abonnierte RSS Feeds

In den vergangenen Jahren ist in einigen Untersuchungen aufgedeckt worden, wie Apps in der Praxis zum Tracking eingesetzt wurden und welche Daten hierbei an dritte Parteien versendet wurden. Hier einige Beispiele:

- Die App *TextPlus 4* für das iPhone hat eindeutige Identifikatoren an acht Third-Partys verschickt. Zusätzlich hat sie die Vorwahl, Alter und Geschlecht an zwei Third-Partys versendet [TK10].
- Die Musik-App *Pandora* (für Android und iOS) hat Alter, Geschlecht, Aufenthaltsort und eindeutige Identifikatoren an verschiedene Third-Partys verschickt [TK10].
- Das Spiel *Paper Toss* (für Android und iOS) hat eindeutige Identifikatoren an mindestens fünf Third-Partys versendet [TK10].
- Die iPhone-App *Grindr*, eine Netzwerk-App für Homosexuelle, hat drei Third Partys Zugriff auf Geschlecht, Aufenthaltsort und eindeutige Identifikatoren gegeben [TK10].
- Sowohl die kostenfreie wie auch die kostenpflichtige Version des Spiels *Angry Birds* haben einer Third-Party eindeutige Identifier und auch Aufenthaltsort mitgeteilt [TK10].
- In [EGC+10] wurden 30 Android-Apps analysiert. Die Hälfte der untersuchten Apps hat Third-Partys Zugriff auf den Aufenthaltsort der Verbraucher gegeben. In 20 dieser 30 Apps wurden 68 Stellen für potenziellen Missbrauch von schützenswerten Verbraucherdaten gefunden.
- In [EKKV11] wurden insgesamt 1407 iOS-Apps untersucht. Bei 55% dieser Apps wurde festgestellt, dass Identifikatoren im Hintergrund an Third-Partys verschickt wurden. Ebenfalls relativ häufig weiter gegeben wurde der Aufenthaltsort.
- Fraunhofer SIT hat bei 72% der Top400 kostenlosen Apps der Kategorie *Utilities* aus dem AppStore von Apple eingebundene Tracking-Frameworks entdeckt. Das Ergebnis einer Untersuchung vom Februar 2014 zeigt Abbildung 3. Diese Abbildung zeigt auf der linken Seite, in welchem Anteil der Apps welches Tracking-Framework gefunden wurde. Im rechten Teil der Abbildung ist zu sehen, wie viele verschiedene Tracking-Frameworks in welchem Anteil von Apps gefunden wurden. In nur 28% der untersuchten Apps waren keine Tracking-Frameworks integriert. In 16% der untersuchten Apps war nur genau ein Tracking-Framework integriert. Insgesamt haben 56% der untersuchten Apps mehr als ein Tracker-Framework eingebunden. Fraunhofer SIT hat für die Analyse von iOS- und Android-Apps mit Appicaptor ein Werkzeug entwickelt, mit dem Apps auf Sicherheitsschwachstellen untersucht werden können.[48] Die hier genannten Zahlen zur Verwendung von Tracking-Frameworks in Apps wurden mit Appicaptor ermittelt.

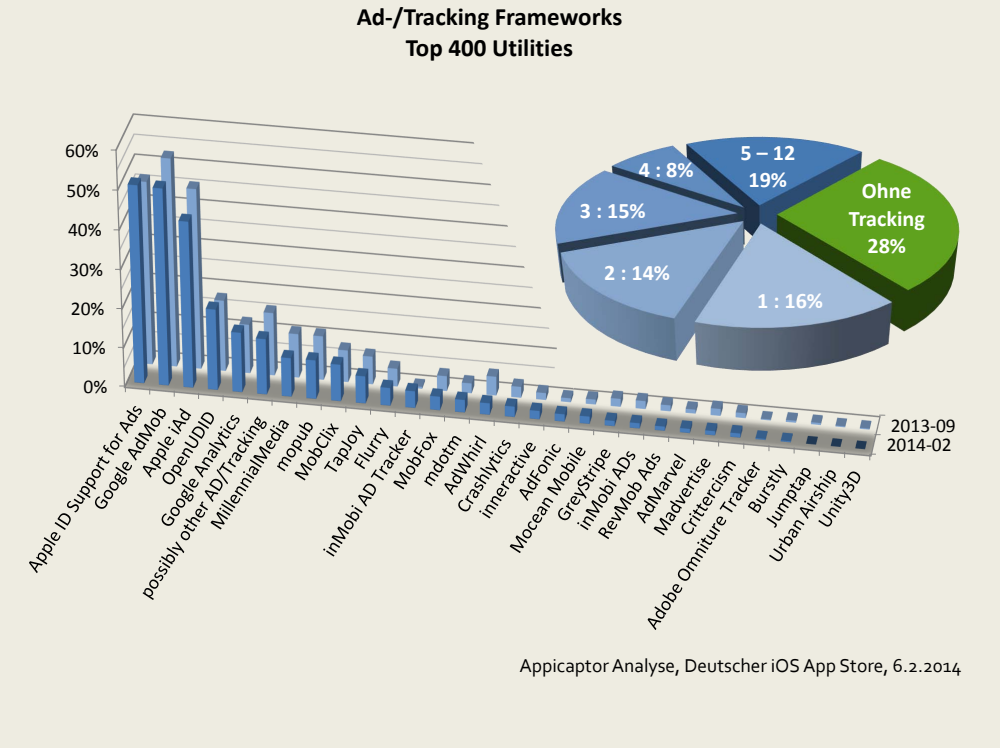


Abbildung 3. Die Verbreitung von Tracking-Frameworks in den Top400 kostenlosen Utility-Apps und Anteile von Apps ohne, einem oder mehreren integrierten Tracking-Frameworks.

Die über App-Tracking gewonnenen Daten lassen sich mit Daten, die über Web-Tracking gewonnen werden, kombinieren, woraus neue verwertbare Informationen für die Tracker entstehen. Beispielsweise können Tracker Erkenntnisse aus der Verfolgung von Verbrauchern durch das World Wide Web mit den Aufenthaltsorten von Verbrauchern kombinieren. Erfährt man durch App-Tracking, dass ein Verbraucher eine Arztpraxis besucht, dann lässt sich zusammen mit der durch Web-Tracking gewonnenen Information, dass dieser Verbraucher sich über Symptome und Verlauf einer bestimmten Krankheit im Internet erkundigt hat, schließen, dass der Verbraucher selbst erkrankt ist und kein Freund, Verwandter oder Bekannter des Verbrauchers.

Durch App-Tracking gewonnene Daten lassen sich auch sehr einfach mit der echten Identität und weiteren persönlichen Daten eines Verbrauchers in Zusammenhang bringen. Um an solche Daten gelangen zu können, können Anbieter von Apps beispielsweise vorschreiben, dass sich ein Verbraucher zur Nutzung mit seinem Facebook-Login anmelden muss. Ein prominentes Beispiel hierfür ist die offizielle App zu den Olympischen Winterspielen 2014 des DOSB [Mac14; Spi14c]. Damit kann der Anbieter alle Informationen aus dem öffentlichen Profil eines App-Nutzers in Erfahrung bringen.

### 3.5.2 Tracking durch Smart TV

Moderne Fernsehgeräte stellen eine weitere Datenquelle für Tracking dar. Früher konnten Fernsehgeräte nur Signale empfangen und keine versenden. Das ändert sich, wenn ein Fernsehgerät einen Internetanschluss bekommt. Über diesen Internetanschluss können Geräte

[48] siehe `https://www.sit.fraunhofer.de/de/angebote/projekte/appicaptor/`

eine Verbindung aufbauen und Daten versenden. Über diesen Anschluss werden Zuschauern dann eine Reihe von neuen Diensten parallel zum Fernsehempfang angeboten. Die neuen Geräte zusammen mit dem neuen Dienstangebot, welche die klassischen Verteilmedien mit dem Internet kombinieren, werden als *Smart TV* bezeichnet.

Ein technischer Standard hierfür ist HbbTV (Hybrid Broadcast Broadband TV). Dieser Standard wurde im Jahr 2012 eingeführt, man erwartet jedoch, dass bis zum Jahr 2016 bereits mehr als 20 Millionen der Haushalte in Deutschland mindestens einen internetfähigen Fernseher haben werden [Gol11].

Über diesen Internetanschluss möchten die Fernsehanstalten ihre Angebote interaktiver gestalten. Für Verbraucher ergibt sich dadurch die Möglichkeit des Abrufs von Sendungen aus der Mediathek oder der Abfrage von Zusatzinformationen, die über den Umfang des heutigen Videotexts hinausgehen. Der über den Internetanschluss vorhandene Rückkanal von HbbTV ist die technische Basis für eine Weiterentwicklung des sogenannten *Social Televisi-on*, bei dem die Bildschirme von Fernseher und PC miteinander verschmelzen. Während ein Verbraucher eine Fernsehsendung schaut, kann er auf demselben Bildschirm mit anderen Zuschauern diskutieren, sich an Abstimmungen beteiligen und vielleicht sogar live in das Fernsehgeschehen eingreifen.

Darüber hinaus soll die Kombination aus klassischem Verteilnetz und Internetanschluss auch für die Verteilung von Werbung genutzt werden können. Hier wird insbesondere an das Einspielen von lokaler oder personengerichteter Werbung gedacht. Die personengerichtete Werbung muss sich an Daten orientieren, die zu den Verbrauchern gesammelt und gespeichert werden, wofür Smart-TV-Tracking angewendet wird. In die dargestellten Inhalte können zusätzlich Inhalte einer Third-Party integriert werden. Die Sendeanstalten und Third-Partys können über Smart-TV-Tracking direkt das genaue Fernsehkonsumverhalten von Zuschauern in Erfahrung bringen.[49] Somit lassen sich Einschaltquoten exakt messen.[50] Smart TV erlaubt den Sendeanstalten eine Echtzeitmessung der Einschaltquote. Das kann dazu führen, dass Live-Sendungen in Abhängigkeit des Zuschauerverhaltens verkürzt oder verlängert werden.

Über Smart-TV-Tracking können für den Nutzer unbekannte dritte Parteien einen detaillierten Einblick in den Fernsehkonsum von Verbrauchern bekommen. Daraus lassen sich Informationen zu Interessen und Einstellungen ableiten, ohne dass der Benutzer dieser Weitergabe von personenbeziehbarer Information zugestimmt hat bzw. von diesem Informationsfluss Kenntnis hat.

In der Vergangenheit sind gerade in diesem Zusammenhang einige Sachverhalte aufgedeckt worden, die nicht im Interesse von Verbrauchern sind. So hat die Computerzeitschrift c't Smart-TV-fähige Fernsehgeräte verschiedener Hersteller analysiert und festgestellt, dass die Geräte Daten zum Fernsehnutzungsverhalten an den Gerätehersteller, an die Sendeanstalt und dritte Parteien schicken [hei14b; hei14c; Spi14b]. Wissenschaftler der Technischen Universität Darmstadt haben zusätzlich herausgefunden, dass Smart-TV-fähige Fernsehgeräte bereits Daten an Sendeanstalten und dritte Parteien schicken, bevor Verbraucher überhaupt den Modus zur zusätzlichen Internetnutzung neben dem klassischen Empfang über das Verteilnetz aktivieren [GOT13; GT14]. Die Geräte führen also bereits Smart-TV-

[49] Bei der ausschließlichen Nutzung traditioneller Verteilnetze konnten die Sendeanstalten nicht in Erfahrung bringen, welche Zuschauer welche Sendungen schauen und wann sie das Programm wechseln.

[50] Traditionell werden Einschaltquoten so gemessen, dass bei einer relativ kleine Gruppe von ca. 2000–6000 Zuschauern das Fernsehverhalten mit einer speziellen technischen Ausstattung aufgezeichnet und nachträglich ausgewertet wird. Über Hochrechnung ergibt sich dann die Einschaltquote für alle Zuschauer.

Tracking durch, wenn Verbraucher davon ausgehen, dass sie noch gar nicht mit dem Internet verbunden sind.

## 3.6 Bedrohung und Risiko für Verbraucher

Zur Anfangszeit der kommerziellen Nutzung des World Wide Web hat Peter Steiner im Jahr 1993 einen berühmten Cartoon veröffentlicht [Ste93]. Dieser Cartoon zeigt einen Hund, der an einem PC sitzend durch das World Wide Web surft mit der Aussage: „*On the Internet, nobody knows you're a dog.*“ Seit der Entwicklung von Web-Tracking gilt diese Aussage nicht mehr. Durch Tracking sind nicht diejenigen unwissend, denen man als Nutzer gegenübertritt, sondern vielmehr die Nutzer selbst, die nicht wissen, wer ihnen im Internet sprichwörtlich über die Schulter schaut. Entwickelt man als Nutzer eine kleine Vorstellung davon, was bei der heutigen Internetnutzung im Hintergrund alles geschieht, dann müsste man eigentlich Fragen stellen wie z.B. wer schaut zu, wer sammelt gerade Daten über mich, zu welchem Zweck sammelt er die Daten, was macht er mit den Daten, gibt er die Daten weiter und wenn ja, an wen und zu welchem Zweck, was bedeutet das für mich und welche Nachteile können dadurch für mich entstehen [Chr11].

Mit diesem Kapitel sollen Verbraucher verstehen können, was Tracking für sie selbst und für ihren Alltag bedeutet bzw. bedeuten kann. Hierzu wird dargestellt, welche Aspekte für Verbraucher im Zusammenhang mit Tracking Auswirkungen haben und welche Werte dadurch bedroht sind. In Kapitel 3.6.1 werden zunächst Aspekte genannt, aus welchen heraus sich Bedrohungen ergeben. In Kapitel 3.6.2 geht es dann um die Werte einzelner Verbraucher oder der Gesellschaft, die durch Tracking bedroht sind, indem Gefahren und Risiken im Hinblick auf diese Werte beschrieben werden.

Bisher sind keine Arbeiten bekannt, in welchen die verschiedenen Aspekte der Bedrohung von Tracking und ihre negativen Auswirkungen als konkrete Gefahren und Risiken umfänglich gesammelt und dargestellt wurden. In Arbeiten, in denen Implikationen von Tracking angesprochen werden, wird meist die Verletzung der Privatsphäre thematisiert. Dies ist zwar korrekt, auf dieser Ebene für Verbraucher jedoch sehr abstrakt; für das Verstehen von Gefahren und Risiken sind konkrete Darstellungen hilfreicher. Darüber hinaus gibt es verschiedene Risiken, bei denen zusätzlich zur Privatsphäre noch andere Werte Gefahren ausgesetzt sind.

Wenn Verbrauchern konkrete Gefahren und Risiken dargestellt wurden, dann haben sich diese hauptsächlich auf personalisierte Werbung bezogen. Allein schon deshalb lehnen viele Verbraucher Tracking ab. Nur vor dem Hintergrund von personalisierter Werbung waren 40% der in den USA befragten Personen bereit, ihr Online-Verhalten zu ändern, wenn sie wüssten, dass Daten zu ihren Online-Aktivitäten gesammelt würden [MC10]. Gemäß einer anderen Umfrage haben 66% der Befragten personalisierte Werbung und Tracking abgelehnt, ohne noch weitere Konsequenzen von Tracking zu betrachten [TKH+09].

Bedrohungen und Risiken, die im Zusammenhang mit der Verletzung der Privatsphäre entstehen, behandeln Aspekte wie etwa den Anspruch einer Person zur Kontrolle des Zugangs zu den eigenen Daten bzw. die eigene Information betreffende Daten und das Recht zur Verheimlichung von Informationen, die die eigene Person betreffen [Cha02]. Der Schutz der Privatsphäre sowie Bedrohungen und Risiken, die durch deren Verletzung entstehen, beschäftigt die Menschheit schon seit der Antike, wie etwa den griechischen Philosophen Sokrates [Hol09]. Die Schwelle, wo die Verletzung der Privatsphäre beginnt, variiert jedoch international und hängt ab von Einflüssen wie Kultur, Normen, Wertesystem und

jeweiligem politischen System [NH10]. So kann es sein, dass in einer Gesellschaft Dinge als Bedrohungen, Gefahren und Risiken wahrgenommen werden, die in einer Gesellschaft mit einem anderen Wertesystem nicht als kritisch eingeschätzt werden.

Die im Folgenden beschriebenen Bedrohungen und Risiken basieren auf Tracking allgemein und nicht nur auf Web-Tracking allein. Dennoch ist anzunehmen, dass Web-Tracking für einen erheblichen Anteil der Implikationen steht. Web-Tracking wird bereits lange angewendet und somit konnten über Web-Tracking bereits umfangreiche Datenbestände entstehen.

### 3.6.1 Aspekte der Bedrohung

Damit man als Verbraucher die Bedrohung durch Tracking besser verstehen kann, genügt es nicht zu wissen, dass Daten über das Verbrauchersurfverhalten in die Hände von unbekannten dritten Parteien gelangen können. In der Realität sind die Aspekte, die zur Bedrohung beitragen, komplexer. Um die eigene Situation als Verbraucher besser einschätzen zu können, ist es erforderlich, dass man die Zusammenhänge zwischen verschiedenen Sachverhalten und der Bedrohung versteht. Da die Bedrohung für verschiedene Werte relevant sein kann, ist es sinnvoll, zunächst von diesen Werten zu abstrahieren.

In diesem Kapitel werden verschiedene Aspekte genannt und es wird erklärt, wie diese zur Entstehung von Bedrohungen beitragen. Die Werte, welche durch Tracking beeinträchtigt werden, werden im anschließenden Kapitel betrachtet.

Die Aspekte der Bedrohung sind im Folgenden nach diesem Schema geordnet:

(1) Bedrohungsaspekte im Zusammenhang mit der Erhebung von Daten
   (a) Internet der Dinge
   (b) Internet der Dienste
   (c) Eingeschränkte Abwehr und neue Trackingmethoden
   (d) Ressourcenverbrauch
(2) Bedrohungsaspekte im Zusammenhang mit der Verarbeitung von Daten
   (a) Kombinierbarkeit von Daten und Big Data
   (b) Verknüpfbarkeit mit echten Identitäten
   (c) Rückschlüsse aus sozialen Beziehungen
(3) Bedrohungsaspekte im Zusammenhang mit der Verwertung von Daten und dem Markt
   (a) Weitergabe von Daten und Datenhandel
   (b) Reichweite von Trackern
   (c) Neue Geschäftsmodelle
   (d) Irreführung von Verbrauchern
(4) Bedrohungsaspekte im Zusammenhang mit dem Verhalten von Verbrauchern
   (a) Fehlendes Verbraucherwissen und Intransparenz
   (b) Änderung von Vertrauensbeziehungen
   (c) Durchsetzung der Löschung von Daten

#### Bedrohungsaspekte in Zusammenhang mit der Erhebung von Daten

*Internet der Dinge*

Alltagsgegenstände werden immer intelligenter; Informations- und Kommunikationstechnologie diffundiert in viele Geräte des täglichen Gebrauchs. Viele Produktinnovationen werden heute durch Informations- und Kommunikationstechnologie getrieben. Dies geschieht

insbesondere in Produktbereichen, bei denen es herkömmlich keinen Zusammenhang mit Computertechnik gab wie z.B. in der Energieversorgung und der Einführung von *Smart Grids* und intelligenter Stromverbraucher[51], den immer intelligenter werdenden Automobilen[52] oder der Entwicklung von Brillen mit integriertem Computer[53]. Diese Beispiele sind nur ein Anfang: In der Welt von morgen wird mittels sogenannter Cyber-physikalischer Systeme Computertechnik in sehr viele Geräte und Gegenstände Einzug halten, wo sie mittels Sensoren Daten erfassen und über eine Verbindung mit dem Internet mit anderen Systemen austauschen. Durch das so entstehende Internet der Dinge entsteht eine Infrastruktur, die perfekt dafür geschaffen ist, sämtliche Aktionen und Regungen von Personen zu erfassen. Diese Infrastruktur stellt neben ihrem Nutzen auch eine Bedrohung dar, z.B. Tracking im Web, Tracking der geographischen Position, Tracking der Gerätenutzung [KSS+12; Sch13b].

*Internet der Dienste*

Cloud-Computing bringt viele Vorteile mit sich: Statt der Verwendung eigener Ressourcen wie Hard- und Software nutzt man diese Ressourcen im Rahmen eines Dienstangebots. Viele Anwendungen werden dadurch kostengünstiger und schneller verfügbar. Es fallen je nach verwendetem Service- und Preismodell nur Kosten für die Ressourcen an, die man tatsächlich nutzt z.B. für Speicherplatz. Die Administration wird einfacher und die Nutzung von Informationstechnologie wird viel flexibler. Diesen Vorteilen des Cloud-Computing stehen jedoch auch gewisse Bedrohungen gegenüber [KSS+12; Wai13]. Durch die Dienstnutzung erhält der Betreiber Informationen des Verbrauchers und kann umfangreiche Daten in Erfahrung bringen. Im Rahmen heutiger Cloud-Angebote erfährt der Dienstanbieter die echte Identität des Nutzers und kann die Daten der Identität zuordnen. Über entsprechende Cloud-Angebote können Daten in Erfahrung gebracht werden, die aus vielen verschiedenen Anwendungen stammen.

*Eingeschränkte Abwehr und neue Trackingmethoden*

Der Großteil der Verbraucher hat nicht das erforderliche Hintergrundwissen, um mit Tracking umzugehen [LUB+11]. Verbraucher haben weder ausreichende Kenntnis über Tracking und Trackingmethoden noch über die technischen Möglichkeiten, Tracking abzuwehren. Viele Werkzeuge zur Abwehr sind methodenbezogen, d.h. sie wirken nur gegen eine bestimmte Trackingmethode und nicht gegen andere. Die Auswahl eines methodenspezifischen Werkzeugs setzt voraus, dass ein Benutzer von der entsprechenden Trackingmethode weiß. Wenn Verbraucher die Trackingmethoden nicht kennen, ist es schwierig für sie, sich mit diesen Werkzeugen davor zu schützen. Bei der Vielfalt von Trackingmethoden, die von Trackern parallel eingesetzt werden, ist es nicht realistisch, dass Verbraucher alle Trackingmethoden kennen. Insofern können sich Verbraucher mit methodenspezifischen Werkzeugen nicht komplett gegen Tracking wehren. Hinzu kommt, dass Tracker immer wieder neue Ideen für Trackingmethoden entwickeln und diese in die Praxis umsetzen [HSG+12]. Für neue Trackingmethoden gibt es in der Regel zunächst auch keine methodenspezifischen Werkzeuge zur Abwehr, die von Verbrauchern eingesetzt werden könnten. Dies gilt sowohl für Web-Tracking als auch für das Tracking mittels anderer Technologien wie z.B. Tracking mittels Smartphone-Apps. Für verschiedene Trackingmethoden, die von Smartphone-Apps ange-

[51] Siehe z.B. `http://www.spiegel.de/thema/smart_grid/`
[52] Siehe z.B. `http://www.simtd.de/index.dhtml/4952b17f6b687261372h/-/deDE/-/CS/-/`
[53] Siehe z.B. Google Glass `http://www.google.com/glass`

wendet werden, sind für den durchschnittlichen Verbraucher keine Werkzeuge verfügbar, mit denen man Tracking unterbinden kann, außer als komplett auf die Benutzung dieser Apps zu verzichten [HJW12].

*Ressourcenverbrauch*

Durch Tracking werden Ressourcen von Benutzern verbraucht. Das ist zwar nicht direkt als Bedrohung einzuordnen, jedoch handelt es sich dabei um einen Nachteil für Verbraucher. Der Ressourcenverbrauch hat mehrere Dimensionen: Geld, Zeit und Energie. Über die Internetverbindungen zu dritten Parteien, die für Trackingzwecke im Hintergrund aufgebaut werden, werden Daten übertragen. Die Datenübertragung geht zu Lasten der Übertragungskosten des Verbrauchers. Die Verbindungen zu Trackern verzögern den Aufbau der aufgerufenen Webseiten, insbesondere wenn viele Tracker in eine Seite eingebettet sind. Die zusätzlichen Datenübertragungen benötigen Energie auf der Seite des Verbrauchers. Der durch Tracking verursachte Ressourcenverbrauch macht sich insbesondere bei der Verwendung von Smartphones bemerkbar, da dort die Kommunikationskosten am höchsten sind, die verfügbare Bandbreite niedriger ist als im Festnetz und die im Akku gespeicherte Energie schneller verbraucht wird.

### Bedrohungsaspekte im Zusammenhang mit der Verarbeitung von Daten

*Kombinierbarkeit von Daten und Big Data*

Daten, die aus verschiedenen Quellen erhoben werden, können kombiniert werden [Kri13]. Durch die Kombination von diesen Daten können neue Informationen entstehen und somit tiefere Einblicke bzgl. des Nutzers erlauben [Ste10b]. Beispiele für solche Datenquellen sind Web-Tracking, Bewegungsprofile von Smartphones, Tracking von Apps über viele verschiedene Anwendungen hinweg, Verwendungsdaten von Kreditkarten und Kundenkarten, Mautsysteme, Videoüberwachung und Daten, die durch die Anwendungen im Zusammenhang mit dem Internet der Dinge und Dienste erhoben wurden. Unmengen solcher strukturierter und unstrukturierter Daten werden gespeichert und mit modernen Big-Data-Algorithmen analysiert. Web-Tracking spielt in diesem Zusammenhang eine besondere Rolle, da diese Art der Verfolgung für praktisch alle Personen relevant ist und diese Technik bereits seit längerer Zeit angewendet wird, so dass von großen Datenbeständen ausgegangen werden kann [KW09]. Darüber hinaus ist anzunehmen, dass Web-Tracking durch dritte Parteien Verbrauchern im Vergleich zu anderen Technologien zur Datensammlung am wenigsten bewusst ist. Aus der Kombination von solchen Daten lassen sich viele Bedrohungsszenarien ableiten: Die Kombination von Web-Tracking-Daten nach der Internetrecherche zu Symptomen einer bestimmten Krankheit und mit dem Bewegungsprofil via Smartphone-App, aus dem sich ein oder wiederkehrende Arztbesuche ableiten lassen, kann für den Abschluss einer neuen Krankenversicherung oder bei der Suche nach einer neuen Anstellung nachteilig sein, sofern Krankenversicherer oder potenzielle Arbeitgeber Zugang zu diesen Informationen bekommen können.

*Verknüpfbarkeit mit echten Identitäten*

Eine ernste Bedrohung durch Tracking ergibt sich, wenn die gesammelten Daten mit der echten Identität eines Verbrauchers verknüpft werden können [AS10]. Für eine solche Verknüpfung gibt es verschiedene Varianten. Anbieter von Webdiensten können ihre Kunden,

die ihnen mit ihrer echten Identität bekannt sind, beim Besuch von Webseiten anderer Anbieter als dritte Partei verfolgen [May11; MM12]. Das ist auch dann möglich, wenn ein Kunde nicht zeitgleich bei der entsprechenden First-Party angemeldet ist. Tracker können die echte Identität von Verbrauchern auch in Erfahrung bringen, wenn diese von einem Embedder entweder absichtlich oder unabsichtlich an Tracker weitergegeben werden. So haben bestimmte Embedder Identitätsdaten ihrer Kunden als Parameterwerte in die URL der von Kunden aufgerufenen Webseiten eincodiert, die dann mittels Referer an Tracker übertragen werden [KW09; KNW11; May11; MM12]. Es gibt viele Beispiele von Webseiten, welche Identitätsdaten an Dritte weitergegeben haben [SF10; May11]:

- Bei Facebook-Apps sind die Namen der Nutzer direkt an Werbetreibende geflossen. So sind IDs von Facebook-Kunden beispielsweise an das Unternehmen RapLeaf geflossen, von wo sie wiederum an weitere Unternehmen geflossen sind.
- Nach einem Klick auf eine Werbeeinblendung bei Home Depot wurden E-Mail-Adresse und Name des Nutzers im Hintergrund an 13 verschiedene dritte Parteien verschickt.
- Nach der Angabe eines falschen Passworts auf den Webseiten des Wall Street Journal wurde die E-Mail-Adresse des Nutzers im Hintergrund parallel an 7 Unternehmen verschickt.
- Nach der Registrierung bei NBC wurden Identitätsdaten automatisch an 7 andere Unternehmen gegeben.
- Nach der Registrierung bei Reuters wurden Identitätsdaten automatisch an 5 andere Unternehmen gegeben.

Diese Beispiele zeigen, dass Identitätsdaten durch Trackingmethoden an Dritte fließen. Darüber hinaus können Identitätsdaten dann von Trackern wiederum an weitere Dritte gelangen, z.B. durch Verkauf, Weitergabe, Hackerangriffe auf die Datenbanken der Tracker oder Fehler und Unachtsamkeiten von Trackern. Doch selbst wenn Daten von Verbrauchern ohne echte Identitäten vorliegen oder weitergegeben werden, dann sind Schlüsse auf die echte Identität möglich. Das Inbeziehungsetzen mit der echten Identität wird einfacher, je mehr Informationen vorliegen, auch wenn diese anonymisiert sind [NS08; NS10]. Ein Beispiel ist die Veröffentlichung von anonymisierten Suchanfragen durch AOL, bei der Suchanfragen echten Personen zugeordnet werden konnten [BZ06]. Ein anderes Beispiel beschreibt [And09], bei dem anonymisierte Gesundheitsdaten den korrekten Identitäten zugeordnet werden konnten. Selbst wenn Tracker also zunächst nur anonymisierte Daten von Verbrauchern erhalten, dann kann es durch die Menge der gesammelten Daten möglich sein, dass diese Daten den korrekten Identitäten zugeordnet werden können. Es ist zu erwarten, dass diese Zuordnungen mit der Entwicklung neuer Big-Data-Algorithmen über Unmengen von unstrukturierten Daten noch viel effektiver und effizienter werden.

*Rückschlüsse aus sozialen Beziehungen*

Tracker erfassen insbesondere auch die sozialen Beziehungen zwischen Verbrauchern. Über das Wissen über soziale Beziehungen lernen Tracker Personen besser kennen und diese Kenntnis kann ebenfalls wirtschaftlich verwertet werden. Die Ableitung von Wissen über Personen aus sozialen Beziehungen geht davon aus, dass Personen, zwischen denen es soziale Beziehungen gibt, in ihren Einstellungen, Bedürfnissen und Vorlieben einander ähnlich sind [CKB12]. In den Profilen von Benutzern können Erkenntnisse zu Personen aus dem eigenen sozialen Umfeld berücksichtigt und für verschiedene Zwecke angewendet werden. So können abhängig von den sozialen Beziehungen eines Benutzers Inhalte von Webangeboten unter-

schiedlich gestaltet werden. Das bedeutet, dass aktuelle Handlungen von anderen Personen, mit denen ein Benutzer sozial verflochten ist, Auswirkungen darauf haben können, welche Inhalte einem Benutzer in einem Webangebot dargestellt werden und welche nicht. Anbieter von Webseiten machen das dargestellte Angebot also davon abhängig, wie sich Freunde, Verwandte oder Bekannte eines Verbrauchers verhalten haben bzw. was über diese Personen bekannt geworden ist.

### Bedrohungsaspekte im Zusammenhang mit der Verwertung von Daten und dem Markt

*Weitergabe von Daten und Datenhandel*

Sind Daten in den Händen einer Partei, dann ist nicht auszuschließen, dass sie von dort aus zu weiteren Parteien gelangen. Personenbezogene Daten sind ein wertvolles Wirtschaftsgut, für das die Nachfrage hoch ist und das sich entsprechend gut vermarkten lässt. Neben der absichtlichen Weitergabe von Daten ist hier auch noch die unabsichtliche Weitergabe von Daten zu berücksichtigen, wie z.B. nach Hackerangriffen wie etwa dem Angriff auf Adobe, bei dem Daten von 38 Millionen Kunden in die Hände von Hackern gefallen sind [Spi13b]. Angriffe auf Datenbanken von Trackern oder anderen sind naheliegend, gerade weil personenbezogene Daten einen hohen Handelswert haben. Die Weitergabe von Daten ist schon seit langer Zeit als Problem erkannt und es wurde auch schon vor langer Zeit davon in den Medien berichtet; siehe z.B. [Sch01]. In [Sch01] wird zusätzlich von dem Fall berichtet, dass nach der Insolvenz des Unternehmens Toysmart versucht wurde, vorhandene Unternehmenswerte zu verkaufen und in diesem Zusammenhang die gesammelten Daten als Teil der Insolvenzmasse zum Verkauf angeboten wurde.

*Reichweite von Trackern*

Einige Unternehmen, die aktiv Daten von Verbrauchern sammeln, haben für Tracking inzwischen eine sehr große Reichweite. Das bedeutet, dass sie in sehr vielen wichtigen Webangeboten integriert sind und somit einen sehr großen Teil der Aktivitäten der meisten Verbraucher verfolgen können [HSG+12]. Laut [AM10] vereinigen die 50 wichtigsten Webangebote insgesamt 40% aller Webseitenaufrufe aus den USA. Es ist davon auszugehen, dass dies in anderen Ländern ähnlich ist, d.h. dass sich ein relativ großer Anteil von Seitenaufrufen auf relativ wenige Anbieter konzentriert. Kann ein Unternehmen also seine eigenen Tracking-Methoden größenordnungsmäßig noch bei einigen weiteren wichtigen Webangeboten für ein jeweiliges Land oder einen Sprachraum einbetten, dann können nahezu alle dortigen Verbraucher bei einem Großteil ihrer Internetaktivitäten verfolgt werden. Eine weitere Ursache für die große Reichweite ergibt sich aus der Akquisition von Unternehmen durch andere Unternehmen. Dadurch bekommen diese Zugang an die Datenbanken der akquirierten Unternehmen. Sind diese Unternehmen ebenfalls als Tracker aktiv, dann ergibt sich die gesamte Reichweite aus der Überdeckung der einzelnen Reichweiten der jeweiligen Tracker. Gemäß [Cle08] konnte allein das Google-Tochterunternehmen DoubleClick im Jahr 2008 mehr als 90% aller Verbraucher im Internet verfolgen.

*Neue Geschäftsmodelle*

Für Verbraucher ist es heute nicht klar, wie und im Rahmen welcher Geschäftsmodelle die eigenen, bis heute erhobenen Daten zukünftig verwertet werden. Es ist nicht auszuschließen,

dass es in der Zukunft Geschäftsmodelle und Verwertungen geben wird, die Verbraucherinteressen widersprechen und von Verbrauchern abgelehnt werden; dann möglicherweise sogar entschiedener als dies bei den heute bekannten Verwertungen der Fall ist. Das Problem besteht dann jedoch darin, dass die über Jahre gesammelten Daten dann bereits in den Händen des Unternehmens sind, welches sie verwertet. Allein alle aktuellen Geschäftsmodelle zu überschauen, die sich rund um Tracking und um die Verwertung der durch Tracking gewonnenen Daten ergeben haben, ist praktisch unmöglich. Um an Daten zu gelangen und aus diesen Profit zu generieren, geht die Branche so weit wie möglich [Jun13]. Für diese Profite wurde die Grenze des Erlaubten in einigen Fällen bereits massiv überschritten [AASGH11; HSG+12]. Die Opfer sind die Verbraucher als Einzelpersonen oder als gesamte Gesellschaft. Die Aussicht auf Umsätze und Verdienstmöglichkeiten macht Tracker und Datenverwerter kreativ, wodurch die technische Weiterentwicklung getrieben wird. Es ist anzunehmen, dass in der Zukunft noch viele neue Algorithmen zur Verwertung der Daten —und darauf aufbauend neue Geschäftsideen— entstehen werden. Bei technischer und wirtschaftlich vertretbarer Machbarkeit ist davon auszugehen, dass solche Ideen dann auch praktisch realisiert werden [FMSW11; AS11]. Mit der Sammlung der Daten werden die Rohstoffe für die zukünftige wirtschaftliche Verwertung auf Vorrat angelegt. Die konkrete Verwertung ist zum Zeitpunkt der Erhebung und Speicherung unter Umständen dem Tracker selbst noch gar nicht bekannt. Es ist beispielsweise nicht auszuschließen, dass Daten, die ursprünglich für Werbung gesammelt wurden, bei neuen Geschäftsideen für vollständig andere Zwecke eingesetzt werden. Die Bedrohung ist wegen noch nicht bekannter zukünftiger Verwertungen für Verbraucher unter Umständen noch zu abstrakt. Zukünftige Formen der Verwertung können hier noch nicht genannt werden. Es können lediglich Beispiele gegeben werden, welche solche Änderungen in der Verwertung rückwirkend betrachten: Mit der Entwicklung von Verfahren zur Vorhersage von Verbraucherverhalten wurden Möglichkeiten entdeckt, Daten, die ursprünglich für Marketingzwecke erhoben und gespeichert wurden, für die Versicherungsbranche zu verwerten [SM10], z.B. von Krankenversicherungen (siehe hierzu *Auswirkungen auf Gesundheitsversorgung von Verbrauchern* auf Seite 73).

*Irreführung von Verbrauchern*

Für Tracking wird das nichtvorhandene Wissen auf Verbraucherseite ausgenutzt. Darüber hinaus werden Verbraucher noch in die Irre geführt. Dieser Gedanke hat mehrere Facetten. Anbieter von Webdiensten erklären ihren Kunden, dass ihre Daten geschützt sind, z.B. Betreiber von Plattformen für soziale Netze [Tuc11]. Meist beziehen sie sich dabei ausschließlich auf den Schutz der Daten vor fremden Dritten wie z.B. Hackern. Dritte Parteien wie Tracker, die von den Anbietern in ihre Seiten eingebettet werden, werden hier außer Acht gelassen. Komponenten von dritten Parteien, die in Webseiten eingebunden werden und jeweils von fremden Anbietern geladen werden wie z.B. Widgets oder Video Player, werden nicht als fremde Komponenten betrachtet [MC10]. Einige Anbieter von Webdiensten erklären explizit, dass keine Daten ohne Zustimmung des Kunden an Dritte gegeben werden, was jedoch nicht den Tatsachen entspricht [May11]. Bekundungen zur Berücksichtigung des Privatsphärenschutzes und des Schutzes der Daten stehen auch in gewissem Konflikt mit der Tatsache, dass einige Webseitenanbieter ihre Kunden auch auf Webseiten anderer Anbieter verfolgen können, so dass sie weit mehr Informationen über einen Kunden bekommen können als Kunden glauben. Tracking wird gegenüber Verbrauchern nicht offen themati-

siert. Stattdessen wird es als Personalisierung zur Anpassung von Inhalten an Bedarfe und Wünsche von Verbrauchern dargestellt [Fer13; Str13].

### Bedrohungsaspekte im Zusammenhang mit dem Verhalten von Verbrauchern

*Fehlendes Verbraucherwissen und Intransparenz*

Nichtvorhandenes Verbraucherwissen in Bezug auf Tracking und die Intransparenz hinsichtlich des Agierens von Trackern stellen eine Bedrohung dar. Die meisten Verbraucher haben typischerweise kein Wissen darüber, dass unbekannte Dritte sie auf ihrem Weg durchs World Wide Web verfolgen und die gewonnenen Erkenntnisse verwerten. Ein Großteil der Verbraucher weiß nicht, was Tracking ist, was bei Tracking passiert und wie sich Tracking auf sie selbst auswirkt oder auswirken kann. Insofern ist diesen Verbrauchern auch nicht klar, zu welchem Zweck die Daten gesammelt und verwertet werden [May09]. Die Daten zu ihren Aktivitäten werden im Hintergrund ohne Erklärung und Anfrage bezüglich ihrer Zustimmung erhoben. Darüber hinaus haben Verbraucher keine Kenntnis, wie sie sich gegen Tracking zur Wehr setzen können [LUB+11]. In den meisten Fällen ist davon auszugehen, dass keine faire Informationspraxis der Tracker gegenüber den Verbrauchern vorliegt [MS12a]:

- Tracker holen im Allgemeinen keine Zustimmung bei den Verbrauchern ein, die Daten zu erfassen;
- Verbraucher haben keinen Zugriff auf ihre Daten zur Einsichtnahme und können somit nicht in Erfahrung bringen, was über sie gespeichert wird;
- die Daten werden ohne Zweckbindung erhoben, d.h. zukünftige Verwertungen bleiben damit offen, wie z.B. wenn Daten zunächst für Werbezwecke erhoben werden und später an Versicherungen zur Einschätzung des Risikos bei Kunden verkauft werden [SM10].

Darüber hinaus ist für Verbraucher nicht transparent, welche Auswirkungen die Verarbeitung ihrer Daten auf der Trackerseite hat. Die Verarbeitung mittels Big-Data-Algorithmen kann z.B. dazu führen, dass auf Basis von statistischen Auswertungen zukünftige Bedürfnisse von Verbrauchern vorhergesagt werden und den Verbrauchern Inhalte angezeigt werden, um sie entsprechend zu beeinflussen [RM13].

*Änderung von Vertrauensbeziehungen*

Hat ein Verbraucher ein hinreichendes Vertrauen zu einem Unternehmen und gibt diesem bereitwillig seine Daten, dann ist nicht auszuschließen, dass sich die Vertrauensbeziehung zu einem späteren Zeitpunkt ändert. Hierfür kann es verschiedene Gründe geben: Die Vertrauensbeziehung kann sich dadurch ändern, dass ein Verbraucher erkennt, dass das bisher entgegengebrachte Vertrauen nicht gerechtfertigt war und er deshalb seine Haltung ändert. Eine andere Ursache für die Änderung einer Vertrauensbeziehung kann darin bestehen, dass das Unternehmen von einer anderen Organisation übernommen wird, z.B. durch Unternehmensakquisition, und die von dem Unternehmen gesammelten Daten dadurch automatisch für ein anderes Unternehmen zugänglich sind. Es ist also festzuhalten, dass die Bedingungen, unter denen die Daten gesammelt wurden, nicht notwendigerweise stabil sind und sich jederzeit ändern können. Für Nutzer kann sich somit die Bedrohung hinsichtlich missbräuchlicher

Verwertung ihrer Daten ändern, nachdem sie ihre Daten bereitwillig zur Verfügung gestellt haben.

*Durchsetzung der Löschung von Daten*

Für Verbraucher ist es praktisch unmöglich, die über sie durch Tracking gesammelten Daten löschen zu lassen. Das Erwirken einer Datenlöschung scheitert bereits an einigen praktischen Dingen. Verbraucher wissen weder, dass dritte Parteien als Tracker Daten von ihnen speichern noch zu welchem Zweck ihre Daten erfasst und gespeichert werden [RL13]. Da Verbrauchern nicht bekannt ist, dass sie verfolgt werden und welche Tracker Informationen über sie speichern, ist ihnen nicht bekannt, bei welchen Unternehmen sie Einsicht in die über sie gespeicherten Daten nehmen sollen und gegebenenfalls eine Löschung der Daten fordern sollen. Selbst wenn eine Löschung der Daten praktisch möglich wäre, dann würde dies einen erheblichen Aufwand von einem Verbraucher verlangen. Verlangt ein Kunde, dass die unwissentlich gesammelten Daten bei Anbietern von Diensten (als First-Party), die auch als Tracker Daten sammeln (als Third-Party), gelöscht werden sollen, dann kann dies die Löschung aller Konten dieses Kunden bei diesem Anbieter erforderlich machen. Handelt es sich bei dem wissentlich genutzten Webangebot um einen E-Mail-Dienst, dann würde die Löschung der unwissentlich gesammelten Daten den Verlust der mit dem Verbraucherkonto einhergehenden E-Mail-Adresse bedeuten, was Verbraucher jedoch eher abschreckt, wenn das E-Mail-Konto intensiv genutzt wird.

### 3.6.2 Gefahren und Risiken

Die Nutzung und Verarbeitung von durch Tracking gewonnenen Daten kann zu Folgen führen, die für Verbraucher konkrete Nachteile mit sich bringen, nachdem sie ihre Daten nichtsahnend geliefert haben. Damit sind dann immer konkrete Werte der Verbraucher bedroht. Diese Werte können sowohl materieller als auch ideeller Art sein. Im Folgenden werden Gefahren und Risiken als Folgen von Tracking beschrieben. Für die angeführten Gefahren und Risiken wurden Fälle gesammelt, die in der Vergangenheit tatsächlich zu unerwünschten Folgen für Verbraucher geführt haben. Darüber hinaus werden ausgehend von bekannten Rahmenbedingungen denkbare Folgen beschrieben, zu denen uns bisher nicht bekannt ist, dass diese genau so eingetreten sind, die jedoch technisch ohne besondere Schwierigkeiten umgesetzt werden könnten. Einige Gefahren und Risiken werden zur Plausibilisierung direkt mit ihren nichtdigitalen Entsprechungen in Beziehung gesetzt.

Die Aufzählung der Gefahren und Risiken ist folgendermaßen strukturiert:

(1) Digitale Analogien bekannter Risiken und Gefahren
   (a) Stalking
   (b) Erfassung von Verbindungsdaten
(2) Gefahren und Risiken für ideelle Werte
   (a) Die digitale Vermessung des Einzelnen
   (b) Auskunftsdienste für jedermann
   (c) Verlust der Freiheit
   (d) Schwächung der Demokratie
   (e) Manipulation von Bürgern und Filterung von Information
(3) Gefahren und Risiken mit wirtschaftlichen Nachteilen für Verbraucher
   (a) Diskriminierung im Online-Handel
   (b) Auswirkungen auf die Gesundheitsversorgung von Verbrauchern

(c) Auswirkungen auf Kreditvergabe
(d) Konsequenzen für Arbeitssuche und Arbeitsverhältnisse
(e) Verfolgung von Kindern

(4) Gefahren und Risiken im Zusammenhang mit strafbaren Handlungen
(a) Wirtschaftsspionage
(b) Vorbereitung von IT-Angriffen
(c) Vorbereitung von Straftaten

### Digitale Analogien bekannter Risiken und Gefahren

*Stalking*

Für Verbraucher besteht durch Cross-Domain-Tracking das Risiko, dass sie bei der Nutzung vieler verschiedener Internetangebote von denselben Trackern verfolgt werden können. Würde Vergleichbares in der nichtdigitalen Welt stattfinden, dann würde man hierfür wahrscheinlich den Begriff *Stalking* verwenden. Tracking im Internet ist praktisch eine digitale Analogie zum Stalking in der nichtdigitalen Welt. Der Vergleich zwischen Tracking und Stalking wurde von anderen schon mehrfach hergestellt, siehe z.B. [YLL+12; Sch13c; Wya13]. In der digitalen und der nichtdigitalen Form von Stalking werden Personen willentlich und beharrlich verfolgt. Es gibt jedoch auch einige Unterschiede: Bei Stalking als nichtdigitale Verfolgung handelt es sich um einen Straftatsbestand (siehe §238 StGB Nachstellung), wenn sich daraus negative Folgen für das Opfer ergeben können; die digitale Verfolgung hingegen ist nicht strafbar, wenn keine weiteren Grenzen überschritten werden. Das nichtdigitale Verfolgen kann von Opfern in der Regel ohne technische Hilfsmittel erkannt werden und wird dann meistens als Belästigung oder gar als Bedrohung wahrgenommen. Diese Wahrnehmung des klassischen Stalkings besteht deshalb, weil den Opfern möglicherweise Stalker und potenzielle Folgen bekannt sind, z.B. weil sie von einem Stalker in der Vergangenheit schon massiv in ihrer Sicherheit bedroht worden sind und eine Wiederholung befürchten. Beim Tracken als digitalem Stalking bemerken die Verfolgten typischerweise nicht, dass sie verfolgt werden. Darüber hinaus sind den Verfolgten diejenigen, die sie verfolgen, nicht bekannt. Zusätzlich sind die potenziellen Folgen des Tracking den Verfolgten nicht bekannt; selbst wenn die Verfolgten eine Vorstellung von potenziellen Folgen haben, dann ist diese wahrscheinlich nur sehr abstrakt. Ein weiterer Unterschied zwischen beiden Formen des Stalkings besteht darin, dass Tracking nicht flüchtig ist, sondern jede erfasste Information gespeichert wird und bleibt. Dies würde in etwa der Analogie entsprechen, wenn Stalker ihre Opfer permanent mit einer Videokamera verfolgen würden, mit der sie sämtliche Aktivitäten der Opfer aufzeichnen würden, z.B. beim digitalen Einkaufsbummel, bei der Zeitungslektüre oder Knüpfen von Kontakten in sozialen Netzen. In der Realität ist die Wahrscheinlichkeit, dass irgendeine Person in der nichtdigitalen Welt Opfer von Stalking wird, eher niedrig. Hingegen gilt es als sicher, dass Tracker versuchen, jeden Internetnutzer zu verfolgen. Wird ein Internetnutzer verfolgt, dann kann man davon ausgehen, dass er gleich von mehreren bzw. vielen Trackern verfolgt wird. Das ist beim nichdigitalen Stalking eher unwahrscheinlich.

*Erfassung von Verbindungsdaten*

Durch Tracking können Tracker Verbindungsdaten von Verbrauchern in Erfahrung bringen. Mit jedem Aufruf einer Webseite, in die ein Trackermechanismus eingebettet ist, erfährt

der entsprechende Tracker, welche Webseite ein Verbraucher aufgerufen hat. Damit erhält er zwangsläufig die Verbindungsdaten für den Webseitenaufruf. Auch wenn in diesen Verbindungsdaten die Klarnamenidentität des Verbrauchers nicht enthalten sein sollte, so kann nicht garantiert werden, dass Tracker die Identität des Verbrauchers nicht in Erfahrung bringen können. So gab es bereits Beispiele, bei denen einbettende Seiten die Identität von Verbrauchern mit den Verbindungsdaten mitgeliefert haben [KW09; KNW11; May11; MM12]. Selbst wenn die Identität der Verbraucher nicht mit den Verbindungsdaten mitgeliefert wird, dann gibt es andere Möglichkeiten für den Tracker, diese in Erfahrung zu bringen z.B. wenn ein Verbraucher in einem anderen Zusammenhang eine First-Party-Beziehung zu diesem Tracker hat (siehe Kapitel *Verknüpfbarkeit mit echten Identitäten* auf Seite 61). Mit jeder Einbettung von Trackingmethoden in die Webseiten eines Anbieters wird die Voraussetzung dafür geschaffen, dass Verbindungsdaten über den Kommunikationsverkehr im World Wide Web im Hintergrund zu dritten Parteien fließen können. Diese Verbindungsdaten fließen immer dann, wenn ein Verbraucher keine Gegenmaßnahmen, wie z.B. Trackingschutz, ergriffen hat. Ob diese Erfassung von Verbindungsdaten durch Dritte eine Bedrohung für den Einzelnen darstellt, kann jeder Verbraucher nur für sich selbst beurteilen. Allein schon die Handhabung von Verbindungsdaten und deren Speicherung durch First-Party —also durch solche Unternehmen, mit denen Verbraucher ganz bewusst Kommunikations- bzw. Geschäftsbeziehungen eingehen— wird von den verschiedenen politischen Parteien in der Diskussion um die Vorratsdatenspeicherung sehr unterschiedlich bewertet [hei14a; Röt14]. Im Fall von Trackern fließen Verbindungsdaten jedoch zu Organisationen, bei denen davon auszugehen ist, dass die meisten Verbraucher nichts von diesem Datenfluss ahnen und dass den Verbrauchern die Empfänger der Verbindungsdaten nicht bekannt sind. Aufgrund der mittlerweile lang anhaltenden Diskussion um die NSA-Affäre, bei der es ebenfalls um das Ausspähen von Verbindungsdaten geht, ist anzunehmen, dass die meisten Verbraucher den Fluss von Verbindungsdaten an unbekannte Dritte ablehnen.

### Gefahren und Risiken für ideelle Werte

#### *Die digitale Vermessung des Einzelnen*

Ein weiteres Risiko des Tracking für Verbraucher besteht darin, dass Verbraucher regelrecht digital vermessen werden. Für diejenigen, die Zugang zu diesen Daten haben, wird der Verbraucher gläsern. Dank dieser Daten kennen Organisationen Eigenschaften, Gewohnheiten, Vorlieben, Abneigungen, soziale Beziehungen, Aufenthaltsorte, Gesundheitszustand, finanzielle Situation, Gemütsverfassung, Meinung, berufliche Situation, Konsumverhalten, Wünsche und Zukunftspläne von Verbrauchern. Diese Vermessung geschieht insbesondere mit dem Ziel, wirtschaftlichen Nutzen aus diesen Daten zu ziehen, z.B. für zielgerichtete Werbung. Weitere konkrete Verwendungen der Daten sind offen. Sie hängen von den Geschäftsideen der Unternehmen, die im Besitz dieser Daten sind, und der Kreativität und den Ideen zur Entwicklung von Big-Data-Algorithmen ab, die aus den Daten neue Informationen ableiten.[54] Inwieweit diese Verwertungszwecke dann ein Risiko für Verbraucher

[54]Die Ableitung von Information durch das Ziehen von Schlüssen aus vorhandenen Daten, die scheinbar ohne Zusammenhang sind, ist nicht zu unterschätzen [FMSW11]. Zur Zeit des kalten Kriegs, als die Welt noch deutlich weniger vernetzt und analoger war, hätte dem US-amerikanischen Militär beispielweise bereits eine Information über die Auslastung von Wäschereien in einer sowjetischen Hafenstadt geholfen, um einen Hinweis über das bevorstehende Auslaufen eines Marineschiffes zu haben [Sch13a]. Heute lassen sich aus den Daten, die aus verschiedenen Quellen stammen wie Web-Tracking, Suchanfragen, Bewegungsprofile

darstellen, ist schwierig zu beantworten, wenn die konkrete Verwertung der Daten noch gar nicht bekannt ist. Doch auch ohne genaue Kenntnis zukünftiger Verarbeitungszwecke dieser Daten stellt allein schon die digitale Vermessung einzelner Verbraucher an sich bereits ein Risiko dar. Auch ohne konkrete Vorstellung von zukünftigen Verwertungen klingen Aussagen wie *„Wir wissen, was Sie morgen tun werden"* vom Chef der Fair Isaac Corporation [Thu11] bedrohlich genug. Insofern sind Bedenken hinsichtlich der Risiken für Privatsphäre nicht übertrieben. Das gilt umso mehr in Ländern, in denen die politischen Systeme weniger stabil sind als in Deutschland und in denen heute nicht klar ist, wie die über Verbraucher gesammelten Daten und Profile zukünftig für andere Zwecke missbraucht werden können.

*Auskunftsdienste für jedermann*

Es gibt Menschen, die ihre Privatsphäre bereits bedroht sehen, wenn ihr Name mit Telefonnummer und ggf. ihre Anschrift in einem öffentlichen Telefonbuch abgedruckt ist. Tracking und das Sammeln von Daten auf anderen Wegen ermöglicht Auskunftsdienste, welche die Privatsphäre von Verbrauchern sehr viel stärker bedrohen und die man als Verbraucher weniger gut kontrollieren kann als die personenbezogenen Informationen, die in einem Telefonbuch abgedruckt werden. Mit der Sammlung, Speicherung und Analyse von Daten können Auskunftsdienste angeboten werden, die praktisch jedermann gegen Bezahlung detaillierte Informationen zu beliebigen Personen geben. Solche Auskunftsdienste sind keine düstere Zukunftsvision; es gibt sie bereits heute schon. Bei Rapleaf kann jeder über das Internet zu einer gegebenen Person Informationen wie Geschlecht, Haushaltseinkommen, Familienstand, Kinder, Ausbildung, Hobbys, persönliche Interessen und Wert der Immobilie abfragen [Kön13]. Pro Person und abgfragtem Merkmal berechnet Rapleaf einen US-Cent. Um für etwaige Anfragen gerüstet zu sein, speichert Rapleaf ebenfalls Daten auf Vorrat. Bei Rapleaf werden momentan (noch) vornehmlich Daten von Personen aus den USA gesammelt. Abfragen kann man diese Daten jedoch von überall und es ist nicht auszuschließen, dass zukünftig auch Daten von Personen aus anderen Ländern gesammelt und zum Abruf angeboten werden. Man kann hier annehmen, dass die Nachfrage das Angebot treiben wird. Mit solchen Angeboten werden Verbraucher nicht nur gläsern für bestimmte Organisationen, Details werden auch für jedermann zugänglich. Für Betroffene bestehen bisher keine Möglichkeiten, sich gegen solche Angebote zu wehren, insbesondere dann nicht, wenn die Angebote aus Ländern stammen, wo ein anderer rechtlicher Rahmen gilt. Das Risiko lässt sich praktisch nur dadurch eindämmen, indem man bewusster mit seinen eigenen Daten umgeht und Schnüfflern weniger Angriffsmöglichkeiten bietet.

*Verlust der Freiheit*

Tracking stellt ein Risiko für die Freiheit von Bürgern und Verbrauchern dar [Car10]. Ist einem Verbraucher bewusst, dass seine Handlungen aufgrund von Verfolgung und Beobachtung durch unbekannte Dritte Auswirkungen haben können, die für ihn von Nachteil sind, dann wird er konsequenterweise abwägen, ob er irgendetwas tut oder besser unterlässt. Wird sein Wissen um Tracking ihn dazu bringen, dass er Dinge unterlässt, die er ohne Tracking möglicherweise tun würde, dann wird seine Freiheit durch Tracking beeinträchtigt.

---

mittels Smartphone-Apps, Kontroll- und Steuerdaten im Zusammenhang mit dem Internet der Dinge viele Aussagen über Verbraucher ableiten [Stö14a; Kre14]. Aus Energieverbrauchsdaten von Smart Grids kann man herauslesen, welche Geräte ein Verbraucher hat [Sch13b] — und somit natürlich auch, welche er noch nicht hat.

Das führt zu einem Verlust der Selbstbestimmung bzw. der Autonomie des Einzelnen. Die Freiheit des Einzelnen hat einen hohen Wert in den westlichen Industrieländern. Sie ist eine gesellschaftliche Errungenschaft, die nicht selbstverständlich ist und für die in vielen Ländern lange Zeit gekämpft wurde; entsprechend ist sie in den Verfassungen vieler Staaten als hohes Individualrechtsgut geschützt. Der Verlust von Freiheit kann insbesondere dann starke Auswirkungen haben, wenn Menschen aufgrund bestimmter Eigenschaften von anderen diskriminiert werden und Nachteile zu befürchten haben, z.B. aufgrund von Religionsbekenntnissen oder sexueller Orientierung. Das Internet ist in dem Leben vieler Menschen allgegenwärtig: Sie nutzen es z.B. zum Einkaufen, als Bibliothek oder als Berater zu verschiedenen Fragen in diversen Lebenslagen. Hat ein aufgeklärter Internetnutzer die Befürchtung, dass er verfolgt wird und die gespeicherten Daten seiner Person zugeordnet werden können, dann wird er es möglicherweise unterlassen, sich im Internet Informationen und Rat zu bestimmten Krankheiten einzuholen oder sich online für bestimmte Dinge einzusetzen, da ihm daraus Nachteile entstehen können. Müssen sich Verbraucher bei jedem Klick darüber Gedanken machen, ob dieser Klick von Nachteil für sie sein kann, dann haben sie bereits einen Teil ihrer Freiheit verloren.

*Schwächung der Demokratie*

Tracking hat Auswirkungen auf wichtige Strukturen in unserer Demokratie und stellt somit eine gewisse Gefahr für die gesamte Gesellschaft dar. Das Risiko von Tracking für die Demokratie kann man folgendermaßen differenzieren.

- *Pressefreiheit und Pluralismus.* Eine wichtige Regel in unserer Demokratie ist die Pressefreiheit, die nach Artikel 5 des Grundgesetzes der Bundesrepublik Deutschland besteht. Artikel 5 des Grundgesetzes beinhaltet das Recht von Medien, Rundfunk und Presse auf die freie Ausübung ihrer Tätigkeit; dies betrifft insbesondere das unzensierte Veröffentlichen von Informationen und Meinungen. Die Presse kontrolliert in unserer Demokratie durch die Veröffentlichung von Informationen und Meinungen die demokratisch legitimierte Macht und wird daher oft als vierte Gewalt im Staat bezeichnet, die jedoch im Gegensatz zu den ersten drei Gewalten keine hoheitliche Verankerung hat. Bürger sollen auf die Informationen und Meinungen der Presse für ihren eigenen Meinungsbildungsprozess zugreifen können. Dies setzt jedoch neben der Informationsfreiheit auch einen Meinungspluralismus in den Medien voraus, der in der Regel durch viele verschiedene Medien gewährleistet wird. Tracking trägt jedoch dazu bei, dass die Voraussetzungen für diesen Meinungspluralismus weniger gut erfüllt sind. Tracking ist nämlich ein wichtiger Grund dafür, dass immer mehr Zeitungsredaktionen verschwinden [Eva09], die eine wesentliche Voraussetzung für diesen Pluralismus darstellen. Zeitungen haben sich in der Vergangenheit zu einem bestimmten Teil durch ihr Anzeigengeschäft finanziert. Werbetreibende haben jedoch einen großen Teil ihres Werbebudgets von Zeitungsanzeigen in die personenbezogene Werbung im Internet verschoben. Die personenbezogene Werbung wird durch Tracking ermöglicht. Welche Bedeutung diese Verschiebung von Werbebudgets für Zeitungen hat, wird offensichtlich, wenn man sich vergegenwärtigt, dass allein Google inzwischen mehr durch Werbung verdient als alle Zeitungen in den USA durch Werbung mit ihren Printausgaben zusammen [RL13]. Somit stellt Tracking ein Risiko für die vierte Gewalt in unserer Demokratie dar, indem Tracking die technische Voraussetzung zu ihrer Aushöhlung schafft.
- *Meinungsfreiheit und Partizipation.* Die Freiheit von Meinungen und Gedanken beinhaltet explizit die Freiheit der freien Informationsbeschaffung und der eigenen Ent-

scheidung, gegenüber wem man seine Meinung äußert und wem man diese zugänglich macht (siehe hierzu auch den Abschnitt *Verlust der Freiheit* auf Seite 69). Dies folgt auch aus der Allgemeinen Erklärung der Menschenrechte der Vereinten Nationen in Artikel 18 und Artikel 19.[55] Meinungsfreiheit und Partizipation können durch Tracking beeinträchtigt werden. Da Bürger bei der Beschaffung ihrer Information im Internet durch Tracking von ihnen unbekannten Dritten beobachtet werden, können Bürger gar nicht mehr frei darüber entscheiden, mit wem sie ihre Meinung teilen.[56] Tracking kann den informierten Bürger daran hindern, sich frei Informationen zu beschaffen, z.B. aus Angst vor negativen Konsequenzen. Das Internet ist jedoch heute zu einer sehr wichtigen Quelle für die Beschaffung aktueller Information geworden, z.B. beim arabischen Frühling, bei dem für die Organisation des Protests in der arabischen Welt das Internet eine sehr wichtige Rolle gespielt hat.

*Manipulation von Bürgern und Filterung von Information*

Tracking schafft die Grundlage, um Verbraucher im Internet manipulieren zu können bzw. um Informationen für Verbraucher gezielt zu filtern. Manipulation und Filterung stellen eine Gefahr für den Einzelnen und für die Gesellschaft dar. Wenn Filterung möglich ist, dann hat man damit die Grundlage, um bestimmten Verbrauchern gewisse Informationen vorzuenthalten oder sie bewusst mit einseitiger Information zu versorgen. Um zu wissen, für welche Verbraucher man welche Inhalte filtert, braucht man zunächst Informationen über die Verbraucher und man muss die Verbraucher, die im Internet Anfragen schicken, wiedererkennen können. Für beides braucht man Tracking. Die Filterung von Inhalten ist längst nichts mehr, was in ferner Zukunft liegt [Fer13]. Beispielsweise orientieren sich Antworten auf Suchanfragen an Daten, die zu der Person vorliegen, welche die Suchanfrage schickt [HSK+13]. Hierbei spielen Informationen wie z.B. Geschlecht, Alter oder Einkommen eine Rolle. Die im Hintergrund verwendeten Big-Data-Algorithmen entscheiden, welche Treffer auf eine Anfrage angezeigt und wie die einzelnen Treffer zu priorisieren sind. In [Fer13] geht die Argumentation so weit, dass die dargestellten Inhalte davon abhängig gemacht werden, zu welcher Gesellschaftsschicht ein Verbraucher gehört. Sicherlich kann Filterung gewisse Vorteile für Personalisierung haben und zur Steigerung des Verbrauchernutzens dienen. Doch Filterung und Personalisierung bringen zwangsläufig auch Möglichkeiten zur Manipulation mit sich [Car10]. Um die Erfolgswahrscheinlichkeit von Manipulationsversuchen zu erhöhen, ist es hilfreich, wenn viele Informationen zu einem Verbraucher vorliegen und ein Verbraucher wiedererkannt werden kann, so dass sich die Argumentationstaktik daran orientieren kann. Es ist offensichtlich, dass Tracking hierfür wichtig ist. Die Manipulation aufgrund umfangreicher Datensammlungen kann dabei vielen verschiedenen Zwecken dienen; sie ist nicht nur auf die Filterung und Vorenthaltung von bestimmten Inhalten im Netz beschränkt. Ein Beispiel für einen solchen Zweck war die Manipulation und Steuerung von Wählern im Wahlkampf bei der US-Präsidentschaftswahl im Jahr 2012. Das Wahlkampfteam von Barrack Obama hat wie es bisher wahrscheinlich in keinem anderen Wahlkampf weltweit der Fall war, Big-Data-Algorithmen angewendet [Moo13]. Die Wählerdaten sind aus einer Kombination aus verschiedenen Datenbanken entstanden, z.B. Wählerverzeichnisse, Konsumentendaten, Kreditkartenabrechnungen, Meinungsumfragen, Kundendatenbanken, soziale Netzwerke und Daten, die über Smartphone-Apps eingesammelt wurden. Laut eigenen

[55] siehe `http://www.un.org/depts/german/grunddok/ar217a3.html`

[56] Dieser Gedanke geht davon aus, dass aus der Kenntnis der Informationsquellen eines Bürgers mit hoher Wahrscheinlichkeit auf dessen Meinung geschlossen werden kann.

Einschätzungen des Obama-Teams lagen Daten zu fast jedem unentschlossenen Wähler in den Vereingten Staaten vor. Die Wähler wurden anhand von 80 Faktoren klassifiziert. Wurden Wähler früher eher aus der Vogelperspektive betrachtet, so konnten nun genau auf den Wähler zugeschnittene Botschaften geplant werden, die im Rahmen von Hausbesuchen oder Telefonanrufen an die Empfänger überbracht wurden. Darüber hinaus konnte man durch die Analyse der Daten von unentschlossenen Wählern herausfinden, zu welcher Zeit die meisten von ihnen welchen Fernsehsender schauen, um dann dort gezielt Wahlwerbespots zu schalten, so dass man mit geringstem Aufwand maximal viele Wechselwähler erreichen konnte [Rut13]. Ohne an dieser Stelle die heutige Politik von Präsident Obama bewerten zu wollen, ist es kritisch, wenn für den Ausgang von demokratischen Wahlen die Werkzeuge des Wahlkampfs wichtiger sind als die Inhalte der Parteien (siehe hierzu Abschnitt *Schwächung der Demokratie* auf Seite 70).

### Gefahren und Risiken mit wirtschaftlichen Nachteilen für Verbraucher

*Diskriminierung im Online-Handel*

Tracking kann Online-Shops bei der Preisdiskriminierung helfen. Bei Preisdiskriminierung bietet ein Verkäufer seinen Kunden das gleiche Produkt zu unterschiedlichen Preisen an.[57] Ein Kriterium für die Preisgestaltung kann das Wissen über die Zahlungsbereitschaft von Kunden sein. Ist von einem Kunden bekannt, dass er finanziell gut situiert ist und häufig teure Produkte kauft, dann ist es möglich, dass ihm ein Online-Shop Produkte zu höheren Preisen anbietet, als er dies anderen Kunden anbietet. Der Online-Handel eignet sich sehr gut für Preisdiskriminierung, da man als Kunde nicht erkennen kann, ob die angezeigten Preise individualisiert sind. Damit Preisdiskriminierung technisch gelingen kann, muss der Online-Shop den Kunden bereits wiedererkennen, bevor dieser sich im Online-Shop anmeldet. Hierfür kann Tracking ausgenutzt werden, indem ein Tracker einem Händler sein Wissen über den Kunden signalisiert. Preisdiskriminierung hilft dem Handel aus ihren Kunden das finanziell Machbare herauszuholen und ist somit für Händler betriebswirtschaftlich interessant. Es gibt eine Reihe von bekannt gewordenen Fällen für Preisdiskriminierung, bei denen Händler als First-Party wahrscheinlich ohne Unterstützung von dritten Parteien wie Trackern ausgekommen sind [Odl03; Str13]. In diesem Fall musste jedoch schon einmal eine Geschäftsbeziehung zwischen dem Händler und dem Kunden bestanden haben. Durch die Einbeziehung von Trackern ist es jedoch denkbar, dass Händler auch dann Preisdiskriminierung anwenden können, wenn sie den Kunden noch nicht kennen. Die relevanten Informationen hierfür können sie von einem Tracker erhalten, dem u.U. umfangreiche Informationen über den Kunden vorliegen.[58] Es ist in diesem Zusammenhang sogar denkbar, dass ein Tracker einem Händler mitteilt, bei welchen Wettbewerbern sich ein Kunde bereits vorher nach dem Produkt erkundigt hat, was für die Preisgestaltung ebenfalls relevant sein kann. Für Verbraucher stellt Preisdiskriminierung ein Risiko dar, da sich durch das Zusatzwissen von Händlern Kräfte zu Ungunsten der Verbraucher verschieben können. Tracking

[57] Wenn hier von Preisdiskriminierung die Rede ist, dann sind stets solche Formen der Preisdiskriminierung gemeint, die für Verbraucher Nachteile mit sich bringen. Es gibt auch positive Formen der Preisdiskriminierung, wie z.B. Preisnachlässe für Verbraucher, die weniger gut situiert sind.

[58] Es ist hier erwähnenswert, dass Google in den vergangenen Jahren an Verfahren zur Preisdiskriminierung gearbeitet hat [Chi12]. Preisdiskriminierung könnte für Google eine weitere Geschäftsoption für die durch Tracking gewonnenen Daten sein, welche bisher hauptsächlich für personalisierte Werbung genutzt wurden.

und die Anwendung der dadurch gewonnenen Daten kann dazu beitragen, dass Verbraucher noch in verstärktem Maß einer Preisdiskriminierung ausgesetzt werden.

*Auswirkung auf Gesundheitsversorgung von Verbrauchern*

Tracking kann negative Auswirkungen auf die Gesundheitsversorgung von einzelnen Verbrauchern haben. Laut [SM10] besteht das Risiko, dass Anbieter von Krankenversicherungen die Konditionenen einzelner Verträge bzw. den Vertragsabschluss von dem Verhalten von Nutzern im Internet abhängig machen. Für Versicherungen geht es bei diesen Daten darum, das eigene Risiko zu kontrollieren, indem sie aus Daten wie z.B. über Lebenswandel und Konsumverhalten auf Gesundheit oder Krankheitsrisiken schließen. Für Verbraucher bedeutet dieses Vorgehen eine Diskriminierung. Für Entwickler von Big-Data-Algorithmen und Unternehmen, die über große Datensammlungen verfügen, kann diese Verwertung von Daten ein interessantes Geschäft sein. Solche Unternehmen suchen nach Geschäftsideen, um aus ihren Daten und Kompetenzen Gewinne zu generieren: So hat die Fair Isaac Corporation einen *Medication Adherence Score* entwickelt, mit dem sie auf Basis von vorhandenen Daten zum Lebensstil vorherzusagen versucht, ob eine Person die ihnen gegebenen Medikationen befolgen wird [Thu11]. Dieses Scoring kann Krankenversicherungen angeboten werden. Gemäß [SM10] argumentiert das Unternehmen Deloitte, das an der Verwertung von personenbezogenen Daten für Krankenversicherungen arbeitet, dass die Anwendung von diesen Daten für Versicherungen deutlich günstiger ist als die Durchführung von medizinischen Tests, mit denen das Krankheits- und Versicherungsrisiko festgestellt werden soll. Auch wenn die Einführung von solchen Methoden dem betriebswirtschaftlichen Interesse von Versicherungsunternehmen dienen mag, so hat dies möglicherweise negative Auswirkungen auf Verbraucher: Klicken sich Verbraucher arglos durch das Internet oder informieren sie sich über Suchtkrankheiten —z.B. im Rahmen einer Weiterbildung als Mitarbeiter einer Personalabteilung—, dann ändert sich dadurch ihr Risikowert und sie bekommen Versicherungsangebote mit möglicherweise deutlich schlechteren Konditionen.[59] Diese Verwertung von Daten geschieht im Hintergrund, d.h. ohne dass Verbraucher davon etwas mitbekommen und ohne dass sie wissen, auf Basis welcher Daten Entscheidungen getroffen werden. Es ist jedoch auch möglich, dass Versicherungen eine andere Strategie wählen und Tracking für Belohnungsmodelle einsetzen, d.h. dass noch viel mehr Daten mit Kenntnis von Verbrauchern zu den Versicherern gelangen. Es ist denkbar, dass nur diejenigen Kunden die besten Konditionen bekommen, die bereit sind, sich umfangreich überwachen zu lassen [Sch13b]. Die Technologien als Grundlage hierfür sind entweder bereits verfügbar oder sie werden entwickelt. Auf der CES 2014 in Las Vegas haben gleich mehrere Unternehmen sogenannte *Fitness Tracker* vorgestellt, mit denen Verbraucher ihre eigenen sportlichen Aktivitäten überwachen können [Hil14].[60] Die eigenen Bewegungsdaten können zu Cloud-Diensten hochgeladen und dort zur Visualisierung und für andere Auswertungszwecke gespeichert und verarbeitet werden. Verbraucher müssen sich bewusst darüber sein, dass solche Daten auch zu Zwecken verwendet werden können, die ihnen nicht bekannt sind. *Fitness Tracker* oder ähnliche Produkte, mit denen gesundheitsbezogene Daten erfasst und anderen Parteien zugänglich gemacht werden (z.B. Pulsuhren, Schrittzähler, Personen-

[59] Bedenken hinsichtlich Missbrauch von Tracking für solche Zwecke und damit verbundene Risiken wurden bereits vor mehr als 10 Jahren geäußert, z.B. in [Sch01]. Damals waren weder Trackingtechnologie noch Technologien zur Auswertung großer Datenmengen jedoch noch nicht so entwickelt wie heute.

[60] Die CES ist eine der weltweit wichtigsten Messen für Unterhaltungselektronik und Computer. Sie findet jährlich zu Jahresbeginn in Las Vegas statt und stellt die Techniktrends des bevorstehenden Jahres vor.

waagen, Körperfettmessgeräte) können grundsätzlich von Versicherungen dazu verwendet werden, um einen Verbraucher hinsichtlich seiner Attraktivität als Kunde zu kategorisieren.

*Auswirkungen auf Kreditvergabe*

Ähnlich wie bei der Gesundheitsvorsorge ist es denkbar, dass die durch Tracking gewonnenen Daten ein Risiko für Verbraucher bei der Kreditvergabe darstellen. Banken können zur Einschätzung des Risikos, das sie bei der Vergabe eines Kredites an einen bestimmten Kunden eingehen, auf Informationen zurückgreifen, die durch das heimliche Verfolgen des Kunden im Internet gewonnen wurden. Entsprechend kann es sein, dass Kunden, wenn sie die für ihr Kreditscoring falschen Webseiten aufrufen, mit schlechteren Konditionen bei der Kreditvergabe bestraft werden. Dass dieses Risiko nicht abwegig ist, zeigt das gemeinsame Projektvorhaben der Schufa und dem Hasso Plattner Institut (HPI) in Potsdam, das in Deutschland eine gewisse Aufmerksamkeit in den Medien erregt hat. Im Jahr 2012 hatte die Schufa angekündigt, gemeinsam mit dem HPI ein entsprechendes Projekt durchführen zu wollen [Spi12; HRS12; Plö12]. Zur Ermittlung der Kreditwürdigkeit sollte in diesem Projekt untersucht werden, wie über die heutigen Verfahren hinausgehend Daten aus sozialen Netzwerken, wie z.B. Facebook und anderen Internetquellen, angewendet werden können. Nach einer Welle der Entrüstung und Kritik von Medien und Datenschützern hat das HPI von dem Projektvorhaben Abstand genommen [Han12], so dass auch die Schufa das Projekt nicht in dem beabsichtigten Rahmen durchführen konnte. Auch wenn dieses Forschungsprojekt nicht nach ursprünglichem Plan durchgeführt werden konnte, so scheint die Verwertung von Trackingdaten für das Scoring der Kreditwürdigkeit von Verbrauchern längst Realität zu sein [Thu11]. Als Beispiel seien hier die Aktivitäten und das Angebot des Unternehmens Experian genannt.[61] Experian bietet Unternehmen Dienstleistungen für Marketing auf Basis des Online-Verhaltens von Nutzern an —also auf Basis von Trackingdaten—[62] und bietet darüber hinaus Scoringdienste zur Kreditwürdigkeit von Verbrauchern aus Großbritannien an.[63] Experian behauptet, dass sie von 49 Millionen Konsumenten aus Großbritannien Daten vorliegen haben. Über Verbraucher werden Daten gesammelt, die Werte für mehr als 300 verschiedene Variablen liefern, mittels derer die Verbraucher dann in 93 verschiedene Kategorien eingeteilt werden. Die Datensammlung ist laut Experian immer aktuell, wofür Tracking erforderlich ist. Experian vermeidet jedoch darzustellen, in welcher Weise die Daten zum Online-Verhalten der Nutzer gewonnen werden.

*Konsequenzen für Arbeitssuche und Arbeitsverhältnisse*

Es ist keine Neuigkeit, dass Personalabteilungen heute üblicherweise das Internet nutzen, um Hintergrundinformationen über Bewerber einzuholen. Hierfür suchen sie beispielsweise zusätzliche Informationen zu Bewerbern im Internet, die diese dort freiwillig einstellen, z.B. eigene Beiträge und Bilder in sozialen Netzen wie Facebook. Das Einstellen von exzessiven Partybildern wird wahrscheinlich dazu führen, dass man gar nicht erst in die engere

[61] siehe `www.experian.com`

[62] siehe `http://www.experian.com/marketing-services/online-display-advertising.html`, `http://www.experian.com/marketing-services/consumer-insights.html` und `http://www.experian.com/hitwise/index.html`

[63] siehe `http://www.experian.co.uk/credit/credit-data-and-scorecards.html` und die Werbeinformation *Financial Strategy Segments 2011 — Experian's market leading financial consumer classification* unter `http://www.experian.co.uk/assets/business-strategies/brochures/financial-strategy-segments-factsheet.pdf`.

Auswahl bei der Stellenvergabe kommt. Aber auch bei bestehenden Arbeitsverhältnissen können Beiträge von Mitarbeitern in sozialen Netzen eine Rolle spielen. Solche Beiträge mit negativen Äußerungen über Arbeitgeber, Vorgesetzte oder Kollegen haben bisher bereits einige Male dazu geführt, dass bestehende Arbeitsverhältnisse gekündigt wurden (siehe z.B. [Mar12; Voß12]). Dass man bei Arbeitgebern mit dem bewussten Einstellen bzw. dem Zugänglichmachen von Inhalten mit unvorteilhaften Darstellungen zur eigenen Person oder entsprechenden negativen Äußerungen bzgl. des Beschäftigungsverhältnisses keine positiven Reaktionen hervorruft, ist nicht verwunderlich. Doch diese Dinge geschehen bewusst und lassen sich daher von Benutzern steuern. Im Vergleich dazu ist es sehr viel riskanter für Verbraucher, wenn ihre privaten Aktivitäten im Netz erfasst würden, daraus Profile und Dossiers zu einzelnen Mitarbeitern zusammengestellt und diese dann dem Arbeitgeber zur Verfügung gestellt würden. Das wäre grob damit vergleichbar, als würde ein Arbeitgeber einen Privatdetektiv auf einen Mitarbeiter ansetzen, um dessen Privatleben auszuspionieren. Dies für mehrere oder für alle Angestellte oder Bewerber zu machen, ist viel zu teuer. Einen Dienst zu nutzen, der auf Trackingdaten basiert, wäre wahrscheinlich sehr viel kostengünstiger. Darüber hinaus decken Trackingdaten längere Zeiträume ab und sind somit vollständiger. Stand heute sind uns noch keine spezialisierten Dienstangebote bekannt, die personenbezogene Daten, die aus Tracking gewonnen werden, den Personalabteilungen oder Arbeitgebern anbieten. Es ist jedoch nicht auszuschließen, dass es diese gibt oder geben wird. Wahrscheinlich können solche spezialisierten Angebote leicht generiert werden. Allgemeine Auskunftsangebote zu Personen, die für solche Zwecke theoretisch auch genutzt werden könnten gibt es ja bereits (siehe z.B. den Abschnitt *Auskunftsdienste für jedermann* auf Seite 69). Sollten Arbeitgeber allgemeine oder spezialisierte Tracking-basierte Informationsdienste nutzen, dann wäre das sehr riskant für Arbeitnehmer. Durch unbedachte Klicks im Internet könnten möglicherweise rasch Karrierechancen zunichte gemacht werden. Werden Trackingdaten für Arbeitsverhältnisse verwertet, dann wird die Freiheit der Arbeitnehmer als Bürger und Verbraucher massiv eingeschränkt (siehe z.B. den Abschnitt *Verlust der Freiheit* auf Seite 69). Da das Internet viele Bereiche des täglichen Lebens abdeckt, wie z.B. Informationsversorgung, Konsum oder Freizeit, können umfangreiche Informationssammlungen entstehen, so dass ein aufgeklärter Verbraucher bestimmte Angebote nicht mehr nutzt aus Sorge um negative persönliche Konsequenzen. Verfahren, die aus dem Online-Verhalten Gesundheitsstatus und damit verbundene Risiken für Krankenkassen ableiten, könnten für Arbeitgeber angepasst werden, damit diese bereits zu erwartende Krankheitstage bei der Einstellung berücksichtigen können (siehe z.B. den Abschnitt *Auswirkungen auf die Gesundheitsversorgung* auf Seite 73). Die durch Tracking gewonnenen Daten können Arbeitgebern im Hintergrund Informationen liefern zu Fragen, die bei einem direkten Kontakt nicht zulässig sind, wie z.B. die Frage nach einer bestehenden Schwangerschaft in einem Bewerbungsgespräch. Durch Tracking und aufgesetztem Auskunftsdienst können Arbeitgeber solche Themen bereits diskret im Vorfeld klären, so dass entsprechende Personen bei der Stellenvergabe ggf. gar nicht erst in die engere Auswahl kommen würden.

*Verfolgung von Kindern*

In einer Untersuchung wurde festgestellt, dass in Webangeboten, die sich explizit an Kinder und Jugendliche richten, besonders viele Tracker eingebunden sind [Ste10a]. Im Rahmen dieser Untersuchung wurden einzelne Anbieter von Webseiten gefunden, die auf ihren Seiten mehr als 200 verschiedene Trackingwerkzeuge einbauen. Das große Interesse von Trackern an Webseiten für Kinder und Jugendliche hängt wahrscheinlich damit zusammen, dass diese

Nutzergruppe von großer wirtschaftlicher Relevanz ist. Kindern und Jugendlichen können mit guten Werbebotschaften wahrscheinlich viel einfacher und effektiver Bedürfnisse suggeriert werden als Erwachsenen, die Botschaften und Methoden der Werbung sehr viel besser durchschauen als Kinder. Deshalb sind Kinder für Tracker begehrte Ziele. Höhere Klickraten bei Werbeanzeigen bringen den Trackern als Werbedienstleister für zielgerichtete Werbeanzeigen entsprechend höhere Einnahmen. Darüber hinaus werden die zu Kindern gesammelten Daten von Trackern zum Verkauf angeboten und weitergegeben. Die Verfolgung von Kindern wird allgemein kritischer gesehen als die Verfolgung von Erwachsenen.[64] Die ungleichen Kräfteverhältnisse zwischen den Trackern auf der einen Seite und den Kindern auf der anderen Seite, die nicht einschätzen können, was mit ihnen und ihren Daten geschieht und welchen manipulativen Kräfte auf sie einwirken, stellen eine Gefahr für Kinder und Jugendliche als Verbraucher dar.

### Gefahren und Risiken im Zusammenhang mit strafbaren Handlungen

*Wirtschaftsspionage*

Tracking stellt nicht nur ein Risiko für den einzelnen Nutzer dar, sondern auch für Organisationen, in deren Auftrag dieser Nutzer handelt, wie z.B. Arbeitgeber. Tracking kann Wettbewerbern Informationen für Wirtschaftsspionage liefern und stellt somit eine Bedrohung für Unternehmen dar. Durch Tracking erfahren Dritte beispielsweise, welche neue Themen für ein Unternehmen interessant sind. Somit können Informationen zu vertraulichen strategischen Vorhaben nach außen dringen. Hat ein Unternehmen ein solches Vorhaben, dann braucht es typischerweise Informationen, welche Mitarbeiter, die an dem Vorhaben arbeiten, aus dem Internet herunterladen. Werden nun solche Mitarbeiter bei ihren Recherchearbeiten über verschiedene Webangebote von Trackern verfolgt, können die Tracker Informationen zu den besuchten Seiten speichern und versuchen, aus den einzelnen Daten wie aus Mosaiksteinen ein größeres Bild zusammenzusetzen [FMSW11]. Auf diese Weise können Dritte beispielsweise ausspionieren, welche neuen Geschäftsstrategien, Produktentwicklungen oder Unternehmensakquisitionen gerade vorbereitet werden. Wenn ein Unternehmen es als Risiko bewerten würde, dass ein Fremder die Arbeit der eigenen Mitarbeiter bei ihren Recherchearbeiten im Internet aufzeichnet (z.B. auf einem Video), dann sollte es auch Tracking als Risiko verstehen. Der Bedrohungsaspekt von Tracking hinsichtlich Wirtschaftsspionage hat noch eine andere Dimension: Durch Tracking lassen sich unter Umständen Angestellte eines Unternehmens identifizieren, die als potenzielle Schwachstellen für Wirtschaftsspionage ausgenutzt werden könnten. Aufgrund umfangreicher Datensammlungen zu einzelnen Personen sind Trackern Vorlieben und Wünsche von Personen auf der einen Seite wie auch deren finanzielle Situation und mögliche Unzufriedenheit auf der anderen Seite bekannt, woraus sich ableiten ließe, ob eine hinreichende Erfolgswahrscheinlichkeit für Korruptionsversuche bestehen könnte.

*Vorbereitung von IT-Angriffen*

Tracking kann auch dazu verwendet werden, um informationstechnische Angriffe vorzubereiten. Aus der Vielfalt von Daten, die Tracker zu Verbrauchern sammeln, kann man

[64] In den Vereinigten Staaten gibt es Initiativen zum Trackingschutz, die explizit auf den besonderen Schutz von Kindern und Jugendlichen abzielen. Siehe hierzu beispielsweise `http://www.commonsensemedia.org/privacy` und `http://epic.org/privacy/kids/default.html`.

ableiten, ob es sich bei einem Verbraucher um einen eher mehr oder weniger technikaffinen Nutzer handelt.[65] Somit kann Tracking auch dazu verwendet werden, um die Menge von Verbrauchern in solche Personengruppen einzuteilen, die sich besser und schlechter für informationstechnische Angriffe eignen. Bei einigen informationstechnischen Angriffen ist der Angreifer auf die Mithilfe des angegriffenen Benutzers angewiesen. Beispiele hierfür sind Phishing-Angriffe, bei denen es darum geht, dass ein Benutzer sein Passwort auf einer präparierten Seite eingibt oder Malware-Angriffe, bei denen ein Angreifer einen Benutzer zur Vorbereitung soweit bringen muss, dass dieser z.B. eine präparierte Datei öffnet, so dass der Schadcode ausgeführt werden kann. Typischerweise ist die Erfolsquote bei solchen Angriffen sehr niedrig, d.h. es müssen sehr viele E-Mails verschickt werden, damit ein Benutzer auf einen Phishing- oder Malware-Trick hereinfällt. Mit dem zusätzlichen Wissen über Nutzer durch Tracking können solche Angriffe effizienter angelegt werden. Hierzu müssen die entsprechenden Daten über die Verbraucher erst einmal in die Hände von Angreifern kommen. Wir unterstellen hiermit nicht, dass Tracker ihre gewonnenen Daten zur Vorbereitung von IT-Angriffen ausnutzen. Wenn die von Trackern gesammelten Daten jedoch in die falschen Hände gelangen, dann können sie hierzu missbraucht werden.

*Vorbereitung von Straftaten*

Tracking kann theoretisch auch für das Ausspionieren von Personen zur Vorbereitung von Straftaten verwendet werden [FMSW11]. Durch Tracking erfahren unbekannte Dritte Details über unsere Gewohnheiten. Hat man Zugang zu diesen Daten, dann können daraus Informationen abgeleitet werden, welche auch hinreichende Auskunft über das Verhalten von Verbrauchern in der nichtdigitalen Welt geben, z.B. für Wohnungseinbrüche. Allein schon durch die Abrufzeiten von Webseiten lassen sich Muster rekonstruieren, aus denen hervorgeht, an welchen Wochentagen und welchen Uhrzeiten Verbraucher zu Hause oder eher unterwegs sind. Gehen Verbraucher auf einer Fernreise mit ihren Mobilgeräten online, dann können Dritte ebenfalls erkennen, dass der Aufruf von einer Adresse erfolgt, die einem anderen Land oder Kontinent zugeordnet werden kann. Entsprechend lässt sich dann schließen, dass Personen nicht innerhalb bestimmter Zeitfenster zurück sein können. Technische Systeme zur Verbrauchssteuerung in intelligenten Häusern, z.B. zur Heizungssteuerung [Stö14b], lassen sich ebenfalls dazu verwenden, um herauszufinden, ob sich Personen zu Hause aufhalten. Andere Quellen, welche solche Daten liefern, sind intelligente TV-Geräte oder Smartphone-Apps, welche Zugriff auf die Positionsdaten des Verbrauchers haben und diese an dritte Parteien melden. Grundsätzlich ist hier zu unterschieden zwischen Fällen, bei denen Personen freiwillig und bewusst ihre aktuellen Aufenthaltsorte im Netz publizieren (z.B. über Twitter) oder ob ihre Aufenthaltsorte für Verbraucher nichtsahnend zu dritten Parteien versendet werden. Die Ableitung von Aufenthaltsorten aus Trackingdaten hat für Kriminelle den Vorteil, dass man Algorithmen nach leeren Wohnungen und Häusern suchen lassen kann, anstatt physisch nach entsprechenden Objekten suchen zu müssen. Es geht hier nicht darum, Trackern die Absicht zu unterstellen, dass sie aus den gesammelten Daten Aufenthaltsorte extrahieren, um damit Haus- und Wohnungseinbrüche vorzubereiten. Wenn sich aus den gesammelten Daten jedoch entsprechende Informationen extrahieren lassen, dann ist nicht auszuschließen, dass solche Informationen durch undichte Stellen bei den datensammelnden Unternehmen in die falschen Hände gelangen. Solche undichten Stellen

[65] Es ist zu vermuten, dass dies oftmals schon aus dem Geschlecht und dem Alter einer Person mit relativ guter Wahscheinlichkeit geschlossen werden kann. Diese Daten werden beim Tracking praktisch immer erfasst.

könnten beispielsweise unzufriedene Mitarbeiter sein oder Mitarbeiter mit finanziellen Problemen. Uns liegen keine Auswertungen vor, in welchem Umfang durchgeführte Straftaten wie Haus- oder Wohnungseinbrüche durch das Sammeln und Auswerten von Aufenthaltsdaten im Internet vorbereitet wurden. Es ist jedoch bekannt, dass die Anzahl der Einbrüche weiter ansteigt und die Kriminellen immer neue Methoden zur Vorbereitung ihrer Tat suchen [Pol14].

## 3.7 Schutzmaßnahmen für Verbraucher

Fühlt sich ein Verbraucher durch Tracking bedroht, dann kann er sich gegen Tracking schützen. Hierzu stehen ihm verschiedene Werkzeuge zur Verfügung. Wenn er Werkzeuge gegen Tracking einsetzt, dann ist es möglich, dass bestimmte Webseiten bei aktiviertem Trackingschutz anders dargestellt werden als ohne aktivierten Trackingschutz. In solchen Fällen liegt es dann in der Entscheidung des Verbrauchers, ob ihm der Schutz seiner Daten oder eine ansprechende Darstellung der Seite wichtiger ist; beides zusammen ist in bestimmten Fällen nicht möglich.

Für Verbraucher stehen verschiedene Alternativen zur Verfügung, um sich gegen Web-Tracking zu schützen. Grundsätzlich kann man die Ansätze für Web-Tracking in zwei Kategorien unterteilen:

- **Methodenorientierte Schutzmaßnahmen**: Bei dieser Kategorie von Maßnahmen geht es jeweils um den Schutz vor einer bestimmten Trackingmethode. Um sich gegen Tracking zu schützen, muss man deshalb die verschiedenen Methoden kennen, die für Tracking angewendet werden, und dann zum Schutz gegen die jeweilige Methode ein geeignetes Werkzeug auswählen.
- **Ergebnisorientierte Schutzmaßnahmen**: Bei dieser Maßnahmenkategorie orientiert man sich am Schutzziel, ohne die einzelnen Methoden berücksichtigen zu müssen. Eine ergebnisorientierte Schutzmaßnahme abstrahiert somit von methodenorientierten Schutzmaßnahmen. Der Schutz gegen Tracking mit einer ergebnisorientierten Schutzmaßnahme umfasst somit alle relevanten methodenorientierten Schutzmaßnahmen.

Ergebnisorientierte Schutzmaßnahmen sind insbesondere für solche Verbraucher geeignet, die nicht alle verschiedenen Trackingmethoden kennen. Solche Verbraucher gehen davon aus, dass sie durch die Anwendung von ergebnisorientierten Maßnahmen besser geschützt sind, als wenn sie selbst für alle relevanten Trackingmethoden die besten methodenorientierten Maßnahmen auswählen und diese jeweils auf dem aktuellen Stand halten können. Nimmt ein Verbraucher an, dass er durch eine eigene Zusammenstellung von methodenorientierten Maßnahmen ein besseres Sicherheitsniveau erreichen wird als durch ergebnisorientierte Maßnahmen, dann wird er seine Schutzmaßnahmen methodenabhängig selbst auswählen. Es ist offensichtlich, dass eine Absicherung durch methodenorientierten Schutz vom Verbraucher verlangt, dass dieser hinsichtlich neuer Trackingmethoden und der Eignung von Werkzeugen stets auf dem aktuellen Stand bleiben muss. Hierbei ist zu erwähnen, dass insbesondere bei den methodenorientierten Schutzmaßnahmen nicht nur auf Standardfunktionen von Browsern zurückgegriffen wird; stattdessen werden solche Schutzmaßnahmen häufig durch Installation von Browsererweiterungen hinzugefügt. Möchte sich ein Verbraucher nicht ständig mit dem Thema *Tracking*, den Weiterentwicklungen hinsichtlich neuer Trackingmethoden und der Leistungsfähigkeit einzelner Werkzeuge beschäftigen müssen, dann wird er ergebnisorientierte Maßnahmen bevorzugen. Grundsätzlich kann man sich als

Verbraucher auch für eine Kombination aus ergebnis- und methodenorientierten Maßnahmen entscheiden.

In diesem Kapitel werden einige Werkzeuge zum Schutz gegen Tracking dargestellt. Es sei explizit darauf hingewiesen, dass hier nur eine kleine Auswahl von Werkzeugen dargestellt wird. Der Schwerpunkt liegt hierbei auf solchen Schutzmaßnahmen, die direkt als Bordmittel von Browsern verfügbar sind. Es ist nicht das Ziel dieses Dokuments, eine Anleitung zur Handhabung von Schutzmaßnahmen zu geben. Hierzu müsste man detailliert auf die verschiedenen Browser eingehen, was im Rahmen dieser Arbeit nicht möglich war. Es wird deshalb dringend empfohlen, andere Quellen zu lesen, z.B. Browseranleitungen. Ebenfalls werden in diesem Dokument keine Browsererweiterungen zum Trackingschutz behandelt. Es sind sehr viele solcher Browsererweiterungen verfügbar; diese werden ständig weiterentwickelt und immer wieder kommen neue Browsererweiterungen hinzu. Für eine möglichst aktuelle Darstellung von Browsererweiterungen ist ein Report wie dieser wegen der hohen Dynamik kein geeignetes Medium.

### 3.7.1 Methodenorientierte Schutzmaßnahmen

Es sei ausdrücklich darauf hingewiesen, dass die folgende Aufzählung von Maßnahmen nicht vollständig ist, d.h. es werden nicht für alle bekannten Trackingmethoden Schutzmaßnahmen vorgestellt. Insbesondere ist es wichtig, dass Verbraucher verstehen, dass sie bei der Aktivierung einzelner Schutzmaßnahmen nicht vollständig gegen Tracking geschützt sind, sondern im besten Fall nur gegen einzelne Trackingmethoden. Tracker können andere Trackingmethoden anwenden, gegen welche Verbraucher keinen Schutz aktiviert haben. Es gibt auch Tracker, die simultan verschiedene Trackingmethoden anwenden.

#### Schutz gegen Third-Party-Cookies

Moderne Browser verfügen über eine Funktionen, mit der man Third-Party-Cookies verbieten kann. Schaltet man diesen Schutz ein, dann können Tracker keine Third-Party-Cookies mehr in dem Browser speichern. Es ist anzunehmen, dass ein durchschnittlicher Verbraucher davon ausgeht, dass er nach dem Verbieten von Third-Party-Cookies gegen entsprechendes Tracking geschützt ist. Das stimmt jedoch leider nur bedingt, da die verschiedenen Browserhersteller diesen Befehl unterschiedlich umsetzen. Bei Firefox bedeutet ein Verbieten von Third-Party-Cookies, dass der Browser weder die Speicherung von Third-Party-Cookies noch das Senden von bereits gespeicherten Third-Party-Cookies erlaubt. Chrome, Safari und Internet Explorer verbieten nur die Speicherung von Third-Party-Cookies, erlauben jedoch das Senden [RKW12]. Nun sollte man annehmen, dass ein Browser keine Third-Party-Cookies versenden kann, wenn vorher keine solchen gespeichert werden konnten. Hier muss man jedoch berücksichtigen, dass es vorher eine Phase gegeben haben mag, in welcher ein Verbraucher Third-Party-Cookies akzeptiert hat, d.h. in dieser Phase konnten Third-Party-Cookies gespeichert werden. Entscheidet sich ein Verbraucher dazu, Third-Party-Cookies zu verbieten, nachdem er sie vorher erlaubt hatte, und löscht nicht zusätzlich die in seinem Browser gespeicherten Third-Party-Cookies, dann ist weiterhin Tracking mit Third-Party-Cookies möglich, auch wenn der Verbraucher ausdrücklich Third-Party-Cookies verboten hat. Es ist anzunehmen, dass die meisten Verbraucher bei der Entscheidung Third-Party-Cookies zu verbieten, nicht wissen, dass sie damit nur die Speicherung von Third-Party-Cookies verbieten und nicht das Senden. Um sich gegen Tracking mit Third-Party-Cookies zu schützen, ist es also wichtig, sowohl die Option zum Verbot von Third-Party-Cookies auszuwählen als auch die bereits gespeicherten Third-Party-Cookies zusätzlich zu löschen.

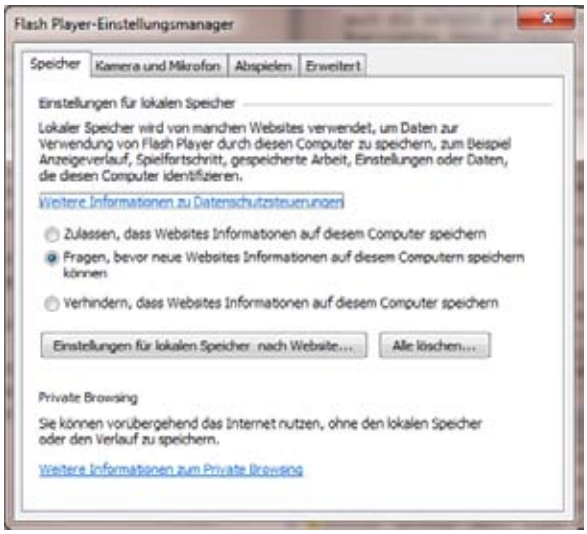


Abbildung 4. Der lokale Einstellungsmanager im Flash-Player.

Jedoch ist auch das Löschen nicht immer hinreichend. Spielt ein Tracker gegenüber einem Nutzer zusätzlich die Rolle einer First-Party, dann führt dies dazu, dass ein Cookie wieder gesetzt werden kann. Dieses Cookie kann dann später wieder im Rahmen eines Third-Party-Kontaktes an den Tracker gesendet werden. Um hier also gegen Tracking sicher zu sein, müssen Nutzer ihre Cookies häufiger löschen. Dennoch sei hier zur Vorsicht geraten: Mit der Verwendung von Evercookies können Tracker versuchen, die von Benutzern gelöschten Cookies wieder herzustellen [AM10; Ang11].

Third-Party-Cookies sind bei Trackern ein sehr häufig verwendetes Werkzeug. Wenn man sich gegen Tracking schützen möchte, sollte man Third-Party-Cookies auf jeden Fall verbieten. Zum Schutz gegen Tracking ist das Verbieten von Third-Party-Cookies jedoch nicht hinreichend. Man sollte diesen Mechanismus auf jeden Fall mit anderen Abwehrmaßnahmen kombinieren.

### Flash-Cookies blockieren und löschen

Die Einstellungen zur Handhabung von Flash-Cookies werden außerhalb des Browsers im Einstellungsmanager von Flash vorgenommen. Für aktuelle Flash-Versionen ist ein lokaler Einstellungsmanager im Flash-Player enthalten (siehe Abbildung 4), für ältere Flash-Versionen müssen die Einstellungen online[66] vorgenommen werden (siehe Abbildung 5).

In beiden Varianten des Einstellungsmanagers können Verbraucher festlegen, wie mit Flash-Cookies umgegangen werden soll: Verbraucher können Flash-Cookies entweder verbieten oder zulassen oder verlangen, dass sie dies im Einzelfall auf Rückfrage entscheiden möchten. In der lokalen Variante des Einstellungsmanagers kann man die bisher gespeicherten Flash-Cookies löschen, was mit dem Online-Einstellungsmanager nicht möglich ist. Hier ist nur ein manuelles Löschen der entsprechenden Dateien möglich. Verbraucher haben in der Online-Variante des Einstellungsmanagers jedoch die Möglichkeit, Flash-Cookies von Third-Partys zu verbieten, was in der lokalen Variante des Einstellungsmanagers nicht un-

[66] Siehe hierzu `http://www.macromedia.com/support/documentation/de/flashplayer/help/settings_manager03.html`.

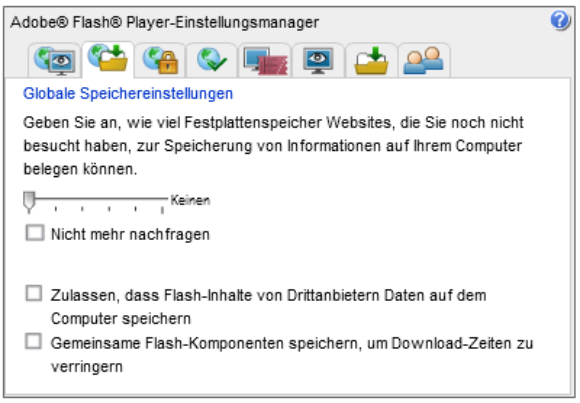


Abbildung 5. Der Online-Einstellungsmanager für Flash.

terstützt wird. Adobe, der Hersteller von Flash, gibt auf seinen Webseiten Informationen und Erklärungen zu den Managementfunktionen des Flash-Player.[67]

Die Verwendung des *Private Mode* eignet sich nur eingeschränkt zum Schutz gegen Flash-Cookies. Das automatische Löschen von Flash-Cookies nach Beendigung einer Browser-Sitzung funktioniert erst seit der Version 10.1 des Flash Player. Lädt ein Verbraucher innerhalb einer Browsersitzung Webseiten verschiedener Anbieter, dann ist im *Private Mode* Cross-Domain-Tracking durch eine Third-Party über diese Webseiten mittels Flash-Cookies möglich.

### Löschen von E-Tags

Zum Schutz gegen Tracking durch E-Tags gibt es leider keine methodenorientierte Schutzmaßnahme als direktes Gegenmittel; die Speicherung von E-Tags kann nicht unterbunden werden. Ebenso kann das Senden von E-Tags nicht verboten werden. Eine indirekte Möglichkeit, die Senden von E-Tags zu unterbinden, besteht darin, den Cache-Speicher des Browsers zu löschen. Durch das Löschen von E-Tags werden Identifikatoren für eine spätere Wiedererkennung beseitigt. Das kann zur Abwehr von Tracking jedoch nur dann effektiv sein, wenn Tracker alte und neue E-Tags nicht über andere Trackingmethoden zueinander in Beziehung setzen können.

### Verwendung von Anonymisierungsproxys

Das Ziel der Verwendung von Anonymitätsproxys besteht darin, die eigene IP-Adresse gegenüber anderen Parteien zu verbergen, so dass Tracking auf Basis der IP-Adresse unterbunden wird. Es sei jedoch vorab festgehalten, dass dies bei der Verwendung von Anonymisierungsproxys nur sehr eingeschränkt gegen Tracking schützen kann, da Tracker den Schutzmechanismus umgehen können. Anonymisierungsproxys agieren als Stellvertreter für Verbraucher, d.h. ein Stellvertreter nimmt eine Anfrage entgegen und leitet sie an denjenigen weiter, an den der Verbraucher seine Anfrage eigentlich stellen möchte. Die Anfrage empfängt der Anbieter von Webseiten dann von der IP-Adresse des Anonymisierungsproxy. Die Anbieter schicken ihre Inhalte als Antwort an den Anonymisierungsproxy, der die Antworten an den entsprechenden Verbraucher weiter leitet. Anonymisierungsproxys treten als Stellvertreter von mehreren Verbrauchern auf, so dass aus dem Nutzungsprofil des Stell-

[67] Siehe hierzu `http://help.adobe.com/de_DE/FlashPlayer/LSM/index.html`.

vertreters nicht auf ein Verbraucherprofil zurückgeschlossen werden kann. Bei der Nutzung dieser Technologie ist es jedoch sehr wichtig, dass der Betreiber eines Anonymisierungsproxy sehr vertrauenswürdig ist. Wird ein solcher Proxy von einem Verbraucher häufig verwendet, dann werden alle Verbraucheninformationen bei diesem Stellvertreter konzentriert. Er kann durch diese Konzentration vergleichbar viele Informationen zu einem Verbraucher gewinnen wie ein Tracker, der Verbraucher auf Webseiten verschiedener Anbieter verfolgen kann. Da ein Anonymisierungsproxy diese Daten theoretisch missbrauchen könnte, ist es sehr wichtig, dass er sich konform zu den Erwartungen der Verbraucher verhält.

Die Verwendung von Anonymisierungsproxys ist sehr einfach. Ein Benutzer muss hierzu seinen Browser nicht umkonfigurieren. Er lädt hierfür lediglich die Startseite des jeweiligen Anonymisierungsproxy und gibt in der Seite die Adresse der vom Anonymisierungsproxy aufzurufenden Webseite an. Bei der Verwendung von Anonymisierungsproxys ist jedoch Vorsicht geboten. Es besteht immer die Gefahr, dass besuchte Seiten oder Tracker auf anderem Weg versuchen, die IP-Adresse des Verbrauchers in Erfahrung zu bringen, z.B. mittels JavaScript.[68] Um sich gegen Tracking zu schützen, müssen Verbraucher also zusätzlich JavaScript deaktivieren, wodurch jedoch die Darstellung vieler Webseiten nicht mehr richtig funktioniert. Beispiele für Proxys sind:

— Anonymouse `http://anonymouse.org/anonwww_de.html`
— Hide My Ass `http://hidemyass.com/`
— KProxy `http://kproxy.com/`
— Webproxy.ca `http://www.webproxy.ca/`
— Zend2 `https://zend2.com/`

In diesem Dokument geht es nur um solche Angebote von Anonymisierungsproxys, die man zusammen mit einem der Standardbrowser ohne zusätzliche Software-Installation nutzen kann.[69]

Anonymisierungsproxys als einzige Schutzmaßnahme sind jedoch nicht hinreichend, um sich gegen Tracking zu schützen. Sie müssen immer mit anderen Schutzmaßnahmen kombiniert werden.

### 3.7.2 Ergebnisorientierte Schutzmaßnahmen

#### Trackingschutzlisten

Eine Trackingschutzliste kann man sich als eine Art Sperrliste oder Schwarzliste vorstellen. Sie enthält eine Sammlung von Trackern bzw. Adressen von Trackern. Betreibt man einen Browser mit einer Trackingschutzliste, dann überprüft der zugehörige Trackingschutzmechanismus im Browser vor dem Versuch, eine Verbindung zu einer Third-Party aufzubauen, ob die jeweilige Third-Party in der Liste als Tracker angeführt ist. Wird diese Third-Party in der Liste als Tracker geführt, dann unterbindet der Browser den Verbindungsaufbau zu

[68] Tracker können hierzu ein Skript einbetten, welches eine direkte Verbindung zu dem Tracker aufbaut und dem Tracker die tatsächliche IP-Adresse des Verbrauchers mitteilt.

[69] Lösungen, bei denen nicht nur einzelne Anonymisierungsproxys, sondern ganze Anonymitätsnetze verwendet werden können (siehe z.B. TOR `https://www.torproject.org/projects/torbrowser.html.en` oder Jondonym `http://www.anonym-surfen.de/jondofox.html`), erfordern die zusätzliche Installation von Software. Werden solche Netze von unterschiedlichen Organisationen betrieben, dann können die Eintrittspunkte in das Anonymitätsnetz variieren, so dass nicht alle Informationen bzgl. einer Person auf einen Anbieter konzentriert sind.

diesem Tracker. Da es gar nicht erst zu einer Verbindung mit dem Tracker kommt, können auch keine Daten versendet werden, anhand denen der Tracker den Verbraucher wiedererkennen kann. Trackingschutzlisten wirken somit komplett ergebnisorientiert, sie sind nicht auf bestimmte Trackingmethoden beschränkt.

Die Qualität des Schutzes durch Trackingschutzlisten ist also unmittelbar von dem Umfang und der Aktualität einer Trackingschutzliste abhängig. Es werden genau die Verbindungen blockiert, wenn die zugehörigen Adressaten in der Liste enthalten sind. Trackingversuche von bislang unbekannten Trackern werden mit Trackingschutzlisten nicht unterbunden. Um eine Trackingschutzliste auf dem Laufenden zu halten, muss diese deshalb kontinuierlich gepflegt werden. Hierzu untersuchen die Anbieter von Trackingschutzlisten die verschiedenen Webangebote auf das Enthaltensein von Trackern bzw. Trackingmethoden. Wird ein Tracker entdeckt, dann wird er in die Liste aufgenommen. Ist die Liste einmal im Browser konfiguriert, dann lädt der Browser die Aktualisierungen der Liste ohne weiteres manuelles Zutun des Verbrauchers nach.

Für Verbraucher ist es somit wichtig, Trackingschutzlisten nur von solchen Anbietern zu verwenden, die hinsichtlich der Generierung von Listen vertrauenswürdig sind, z.B. dürfen Anbieter in ihren Listen keine Tracker bewusst unterschlagen. Verbraucher können auch simultan mit Trackingschutzlisten verschiedener Anbieter arbeiten. Das Fraunhofer-Institut für Sichere Informationstechnologie (Fraunhofer SIT) in Darmstadt stellt eine solche Trackingschutzliste kostenfrei zur Verfügung.[70] Die Trackingschutzliste kann theoretisch von allen Verbrauchern genutzt werden. Unter den am weitesten verbreiteten Browsern kann heute nur der Internet Explorer (ab Version 9) ohne Weiteres mit Trackingschutzlisten umgehen; andere Browser brauchen zusätzliche Erweiterungen, um Trackingschutzlisten verarbeiten zu können.

Die Verwendung von Trackingschutzlisten eignet sich insbesondere für solche Verbraucher, die sich nicht ständig mit den Themen *Tracking*, *Trackingmethoden* und *Trackingschutz* beschäftigen möchten oder hierfür weder Zeit und Interesse aufbringen möchten noch über das notwendige Hintergrundwissen verfügen. Auf neue Erkenntnisse im Zusammenhang mit Tracking und deren Berücksichtigung zur Abwehr von Tracking können Anbieter von Trackingschutzlisten normalerweise viel besser reagieren als Verbraucher. Die Ergebnisorientierung von Trackingschutzlisten und ihrer Unabhängigkeit von den Trackingmethoden befreien Verbraucher davon, sich immer wieder mit der Effektivität verschiedener Abwehrmaßnahmen beschäftigen zu müssen.

Dennoch sollten Verbraucher nicht komplett auf andere Werkzeuge verzichten. Verwendet man eine Trackingschutzliste, dann kann man zusätzlich noch methodenorientierte Schutzmaßnahmen einsetzen, z.B. ein Verbot von Third-Party-Cookies. Ist ein Tracker in der Liste noch nicht enthalten, dann können somit Trackingversuche methodenorientiert unterbunden werden.

### Do Not Track

Der Do-Not-Track-Header bietet Verbrauchern eine Möglichkeit Trackern zu signalisieren, dass sie nicht verfolgt werden möchten. Hierzu können Verbraucher in ihren Browsern die Option *Opt-out* wählen. Entscheidet sich ein Nutzer für diese Option in seinem Browser, dann wird im Header einer jeden HTTP-Anfrage ein entsprechender Wert mitgeschickt.[71]

[70] Weitere Informationen zu dieser Trackingschutzliste und eine Möglichkeit zum Herunterladen dieser Liste findet man unter `https://www.sit.fraunhofer.de/de/angebote/projekte/tracking-protection-list/`.
[71] Do-Not-Track sieht die folgenden Werte für den HTTP-Header vor:

Auch im Fall *Opt-out* werden Verbindungen zu Trackern aufgebaut, über welche diese Daten erfassen und auswerten können.

Für Verbraucher ist es wichtig zu verstehen, dass Do-Not-Track kein harter Trackingschutz ist. Bei Do-Not-Track kann ein Nutzer den Schutz nicht direkt durchsetzen, er ist bei der Auswahl von Opt-out vielmehr auf die Kooperation von Trackern angewiesen. Ob ein Tracker seinen Wunsch tatsächlich befolgt, kann ein Verbraucher nicht überprüfen. Do-Not-Track wird von den wichtigsten Browsern unterstützt, auch wenn die Standardisierung noch nicht abgeschlossen ist. Inwieweit Tracker den per Do-Not-Track signalisierten Opt-out-Wunsch tatsächlich nachkommen, ist nicht bekannt.

Der Do-Not-Track-Header verbessert die Ausgangssituation für das Tracking durch Fingerprinting und kann somit für Tracking missbraucht werden. Do-Not-Track hilft auf der Ebene der Browsereinstellungen, um verschiedene Nutzer voneinander unterscheiden und wiedererkennen zu können (siehe Abschnitt *Fingerprinting* auf Seite 48). In der Tat machen sich technische Angebote Do-Not-Track zu Nutze; ein Beispiel hierfür ist die Fingerprinting-Bibliothek von BlueCava [NKJ+13].

An der Entwicklung von Do-Not-Track war die Werbeindustrie maßgeblich beteiligt [AS11]. Diese Aktivitäten hatten jedoch in erster Linie strategischen Charakter. Nachdem in den Vereingten Staaten von staatlicher Seite angekündigt wurde, dass man die Verwertung von personenbeziehbaren Daten im Zusammenhang mit Tracking regulieren wolle, ist die Industrie vorangeprescht, um damit auf die Regulierer beschwichtigend einzuwirken und durch eigene Vorschläge die Rahmenbedingungen für die eigene Zukunft selbst gestalten zu können. Die Sorge der Werbeindustrie war zu groß, dies vollständig in fremde Hände zu geben.

### Private Mode

Der *Private Mode* von Browsern bietet einen ergebnisorientierten Mechanismus, der beim Schutz gegen Tracking helfen kann. Es ist jedoch darauf hinzuweisen, dass ein *Private Mode* nicht vollständig gegen Tracking schützen kann. Primäres Ziel des *Private Mode* ist auch nicht der Schutz gegen Tracking; der partielle Schutz gegen Tracking ist vielmehr ein Seiteneffekt des *Private Mode*.

Das Ziel dieser Funktion besteht darin, Informationen über besuchte Webseiten gegenüber anderen Nutzern des Computers vertraulich zu halten (siehe z.B. [ABJB10]). Bei Verwendung des *Private Mode* geht es primär darum, dass lokal keine Verläufe über die Nutzung eines Browsers aufgezeichnet und langfristig gespeichert werden, die später von einem anderen Nutzer ausgelesen werden können. Abhängig von dem jeweils verwendeten Browser kann der Schutzumfang des *Private Mode* auch dahingehend weiter greifen, so dass ein besuchter Webserver Aktivitäten, die derselbe Benutzer im *Public Mode* und im *Private Mode* ausführt, nicht miteinander in Beziehung setzen kann. Darüber hinaus können browserabhängig weitere Ziele darin bestehen, eine Session eines Benutzers im *Private Mode* nicht mit einer anderen Session desselben Benutzers im *Private Mode* in Beziehung setzen zu können bzw. überhaupt erkennen zu können, ob ein Benutzer gerade den *Private Mode* oder den *Public Mode* verwendet. Auch wenn der *Private Mode* die Privatsphäre von Verbrauchern schützt, so schützt er nicht vollumfänglich gegen Tracking. Der *Private Mode* verbietet keine Verbindungen zu Third-Partys. Innerhalb einer Browsersitzung können Daten im Browser gespeichert werden, die zur Wiedererkennung innerhalb der selben Brow-

— `DNT=1`: Opt-out, Nutzer möchte nicht verfolgt werden

— `DNT=0`: Opt-in, Nutzer möchte verfolgt werden

sersitzung verwertet werden können. Die Daten werden lediglich nicht über das Ende eine Browsersitzung im *Private Mode* gespeichert.

# 4. QUANTITATIVE ANALYSE

Für Verbraucher stellt sich die Frage, in welchem Umfang sie von Web-Tracking betroffen sind, wenn sie im World Wide Web unterwegs sind. Wird man von einem Tracker verfolgt, dann merkt man hiervon als Verbraucher erst einmal wenig, wenn man von den Werbeanzeigen absieht, die häufig zu eingegebenen Suchbegriffen oder zu Inhalten von besuchten Webseiten passen. Das vermittelt einem Verbraucher jedoch keinen Eindruck davon,

- wie viele Tracker ihn typischerweise auf den Webseiten desselben Anbieters beobachten, oder
- bei wie vielen verschiedenen Anbietern er von demselben Tracker verfolgt werden kann.

Um solche Fragen beantworten zu können, muss man die Seiten verschiedener Anbieter dahingehend untersuchen, ob in ihnen Trackingmethoden integriert sind.

In Untersuchungen anderer wurde berichtet, dass auf Webseiten bestimmter Anbieter, die insbesondere für den US-amerikanischen Raum relevant sind, teilweise mehr als 100 verschiedene Trackingwerkzeuge bzw. Tracker erkannt wurden [AM10]. Für die vorliegende Studie war es nun für uns wichtig zu erkennen, wie sich dies für Webseiten darstellt, die insbesondere für den deutschen Sprachraum interessant sind. Hierzu wurden die im deutschen Sprachraum populärsten Anbieter von Webseiten dahingehend analysiert, ob diese Trackingmethoden von Third-Partys integrieren.

Anhand von integrierten Trackingmethoden kann man von außen erkennen, dass damit Verbindungen zu einer Third-Party ohne Zustimmung des Verbrauchers aufgebaut werden sollen. Über diese Verbindungen können Daten zu der Third-Party fließen, mittels derer diese erkennen kann, welche Webseite der Verbraucher bei dem einbettenden Anbieter abgerufen hat, und mit denen sie den Verbraucher später oder in anderem Kontext wiedererkennen kann. Was die Third-Partys danach mit diesen gewonnenen Daten machen —also ob sie sie bspw. speichern, verarbeiten, auswerten, weitergeben oder ob sie sie vollständig ungenutzt lassen—, kann man von außen betrachtet nicht immer erkennen. Hierzu bräuchte man uneingeschränkte Einblicke in die Verwertung von Daten bei jedem dieser Unternehmen. Es bleibt also oft unklar, was genau eine Third-Party mit den Daten macht, die im Hintergrund zu ihr fließen.

Es stellt sich die Frage, ab wann man eine solche Third-Party als *Tracker* bezeichnet, wenn man das World Wide Web hinsichtlich eingebetteten Trackingmethoden und im Hintergrund ablaufenden Verbindungen analysiert. Nach unserem Verständnis wird eine Third-Party zum Tracker, wenn sie Verbraucher verfolgt, d.h., wenn sie in Erfahrung bringen kann, auf welchen Webseiten Verbraucher unterwegs sind, ohne dass sie von den Verbrauchern hierzu aufgefordert wurden bzw. ohne dass Verbraucher hiervon Kenntnis haben. Es ist hierbei unmaßgeblich, zu welchem Zweck eine Third-Party diese Verfolgung vornimmt. Wenn wir im Folgenden von einem Tracker sprechen, dann ist damit immer eine Organisation (repräsentiert durch einen Domainnamen) gemeint, bei welchem wir durch unsere Analysen herausgefunden haben, dass

(1) sie Trackingmethoden angewendet hat,
(2) von mehreren Webseiten eines Anbieters (First-Party) im Hintergrund Verbindungen zu ihr aufgebaut wurden und
(3) dieses bei mindestens zwei Anbietern (First-Partys) beobachtet werden konnte.

In Kapitel 4.1 beschreiben wir, wie wir bei der Messung vorgegangen sind, in Kapitel 4.2 stellen wir die Ergebnisse der Messung dar.

## 4.1 Rahmen der quantitativen Analyse

Die quantitavie Analyse basiert auf Daten, die das Fraunhofer SIT für die Erstellung seiner Trackingschutzliste erhebt.[72] Um die Trackingschutzliste aktuell zu halten, untersucht Fraunhofer SIT seit November 2012 immer im Abstand weniger Tage das Webangebot von für den deutschen Sprachraum wichtigen Anbietern auf das Enthaltensein von Trackingmethoden. Insgesamt wurden für die Untersuchung die Webangebote von 1634 Anbietern erfasst (Stand 24. Januar 2014). Die Menge der untersuchten Webseiten enthält die Top500-Liste von Webseiten für Deutschland von Alexa[73], die Top100-Liste von Netcraft für Deutschland[74] und der Liste von ivw zu Onlinenutzungsdaten[75].

Für jede Analyse wurde eine Stichprobe von 20 verschiedenen Webseiten eines Anbieters geladen. Diese Stichprobe wurde auf das Enthaltensein von Trackingmethoden, die auf Third-Partys zurückzuführen sind, untersucht.[76] Wurden Trackingmethoden gefunden, dann wurde überprüft, ob diese Methoden derselben Third-Party in mehreren dieser Webseiten enthalten waren. Wenn dies zutraf, wurde untersucht, ob dies noch für dieselbe Third-Party bei einem anderen Anbieter erfüllt war, d.h. ob Cross-Domain-Tracking vorliegt. Wenn auch dies erfüllt war, dann wurde die entsprechende Third-Party in die Trackingschutzliste aufgenommen.

Im Zusammenhang mit der Erzeugung der Trackingschutzliste des Fraunhofer SIT sind viele Daten zu Trackern und auch zu Anbietern als Embedder von Trackingmethoden angefallen. Für den vorliegenden Report haben wir diese Daten nun vor dem Hintergrund von verschiedenen Fragen analysiert. Hierbei wurden Daten berücksichtigt, die zwischen November 2012 und Januar 2014 angefallen sind.

## 4.2 Ergebnisse der quantitativen Analyse

Im Betrachtungszeitraum wurden insgesamt 1634 Anbieter als potenzielle Embedder untersucht. Dabei wurden insgesamt 1209 verschiedene Tracker gefunden. Im Folgenden stellen wir die Ergebnisse unserer quantitativen Analyse dar. Bei den Ergebnissen, die sich auf Embedder beziehen, werden wir nicht für jeden Embedder angeben, wie viele Tracker bei ihm gefunden wurden. Bei allen Ergebnissen stellen wir immer nur die 25 Embedder mit den meisten Trackern dar. Ähnlich werden wir auch keine vollständigen Ergebnislisten zu den 1209 gefundenen Trackern angeben. Auch hier beschränken wir uns auf die 25 Tracker, die in der jeweiligen Betrachtung die höchste Verbreitung hatten.

### 4.2.1 Anzahl Embedder je Tracker aus aktueller Messung

Hier stellen wir die Frage, auf wie vielen Webseiten verschiedener Anbieter zu einem Zeitpunkt Tracker eingebunden waren. Hierfür verwenden wir eine im Januar 2014 durchgeführte Messung des Fraunhofer SIT.[77] Messung und Untersuchung wurden so durchgeführt, wie dies in Kapitel 4.1 beschrieben wurde. Insgesamt wurden in dieser Messung 633 Tracker

[72] siehe `https://www.sit.fraunhofer.de/de/angebote/projekte/tracking-protection-list/`
[73] siehe `http://www.alexa.com/topsites/countries/DE`
[74] siehe `http://toolbar.netcraft.com/stats/topsites?c=DE`
[75] siehe `http://ausweisung.ivw-online.de/online/`
[76] Der Umfang der hierbei betrachteten Trackingmethoden, die bei der Analyse berücksichtigt wurden, ist ebenfalls angestiegen. Wurden neue Trackingmethoden bekannt, dann wurden diese für die Generierung der Trackingschutzliste sukzessiv berücksichtigt.
[77] Die Messung wurde durchgeführt zwischem dem 22. und dem 24. Januar 2014.

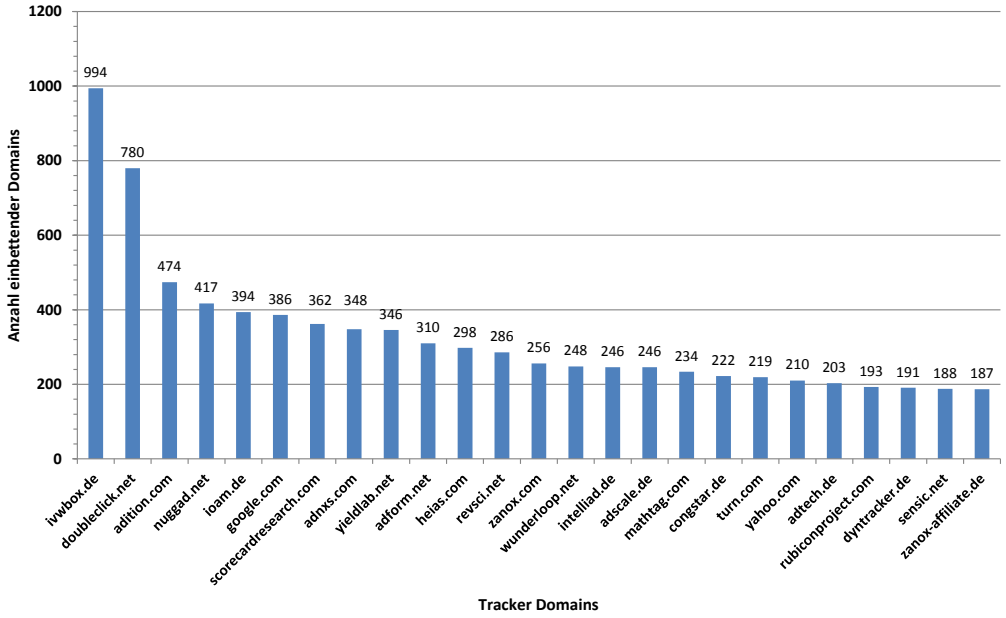


Abbildung 6. Die Rangliste der Tracker (Top25) nach der Anzahl von einbettenden Anbietern in der Messung vom 22.–24. Januar 2014.

gefunden.[78] Für jeden dieser 633 Tracker wurde ermittelt, bei wie vielen Anbietern er in der Messung erkannt wurde. Abbildung 6 zeigt die 25 Tracker, die zum Messzeitpunkt bei den meisten Anbietern von Webseiten eingebettet waren.

Bei dieser Analyse stellt sich heraus, dass viele Tracker bei sehr vielen Anbietern eingebettet sind. Die Beschränkung der Betrachtung auf die Situation hinsichtlich aktuell gefundener einbettender Webseitenanbieter stellt die Situation für den Verbraucher besser dar, als sie es tatsächlich ist. Cross-Domain-Tracking ist auch mit zeitlichem Versatz möglich. So ist es denkbar, dass ein Verbraucher in der Vergangenheit bei einem Anbieter A verfolgt wurde und dieser Verbraucher auf den Seiten eines anderen Anbieters B wiedererkannt wird, auch wenn zum Zeitpunkt der Wiedererkennung bei B Anbieter A diesen Tracker gar nicht mehr in seinem Webangebot einbettet. Die Wiedererkennungsmöglichkeiten bei Cross-Domain-Tracking gehen also weit über das gleichzeitige Eingebettetsein von Trackern hinaus, wie dies bei einer Betrachtung eines einzelnen Zeitpunkts suggeriert wird.

### 4.2.2 Anzahl Tracker je Embedder aus aktueller Messung

Bei dieser Betrachtung geht es um die Frage, wie viele Tracker einzelne Anbieter zum Messzeitpunkt in ihrem Webangebot integriert hatten. Es geht also um die Frage, von wie vielen Trackern ein Verbraucher zu diesem Zeitpunkt beim Aufruf von Webseiten bestimmter Anbieter gleichzeitig verfolgt werden konnte. Die Auffindbarkeit von Trackern ist bei dem in Kapitel 4.1 beschriebenen Rahmen auf die Stichprobe von 20 Webseiten je Anbieter

[78] Nach einer Messung werden neu gefundene Tracker jeweils in die bereits vorhandene Trackingschutzliste eingefügt. Insofern gibt der Umfang der Trackingschutzliste zu diesem Zeitpunkt nicht die Anzahl der in dieser Messung gefundenen Tracker wieder; die Liste ist wegen der in der Vergangenheit gefundenen Tracker, die in der aktuellen Messung nicht erkannt wurden, größer.

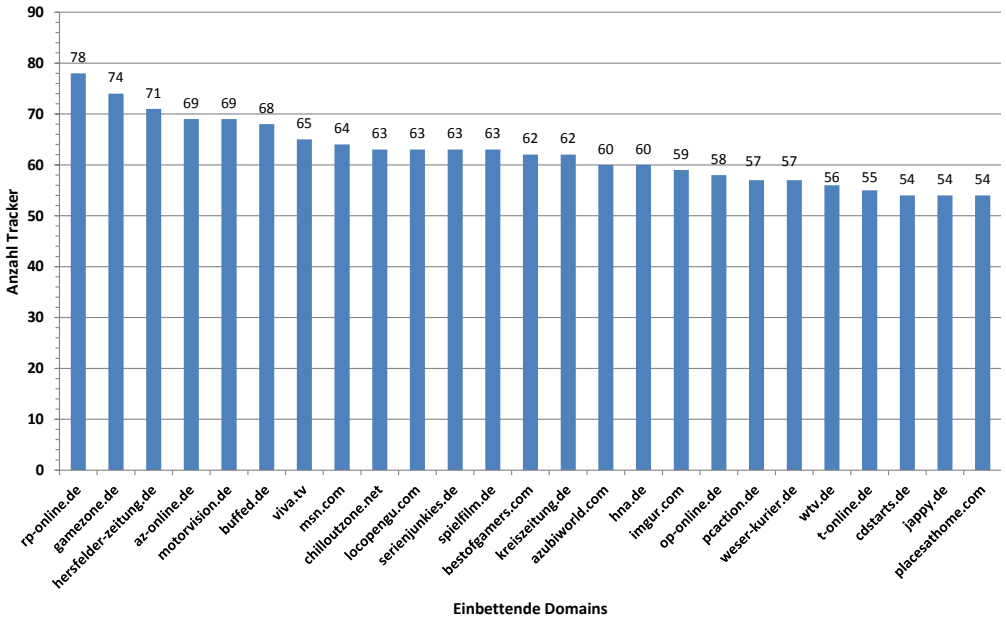


Abbildung 7. Die Rangliste der Embedder (Top25) nach der Anzahl von eingebetteten Trackern in der Messung vom 22.–24. Januar 2014.

beschränkt. Abbildung 6 zeigt die 25 einbettenden Anbieter (von den insgesamt 1634 Anbietern), die zum Messzeitpunkt die meisten Tracker in ihren Webseiten eingebettet hatten.

Wie man in Abbildung 6 erkennen kann, war zum Messzeitpunkt die gleichzeitige Einbettung von mehr als 50 Trackern keine Seltenheit. Ein Verbraucher, der zum Messzeitpunkt Webseiten dieser Anbieter aufgerufen hat, konnte von entsprechend vielen verschiedenen Trackern verfolgt werden.

### 4.2.3 Anzahl Embedder je Tracker aus Messungen zu anderen Zeitpunkten

Die Aussage von Kapitel 4.2.1 beschreibt lediglich das Ergebnis einer aktuellen Messung. Um diese Werte besser einordnen zu können, wollten wir wissen, wie hoch die stärksten Verbreitungen einzelner Tracker zu einzelnen Messzeitpunkten waren. Hierzu wurden für jeden Tracker diejenigen Einzelmessungen ausgesucht, bei denen Tracker ihre stärkste Verbreitung hatten. Das ist für einen gegebenen Tracker diejenige Messung, bei welcher der Tracker bei so vielen Anbietern eingebettet war, wie bei keiner anderen Messung. Die höchsten Werte für Tracker-Einbettungen bei Anbietern werden in Abbildung 8 dargestellt.

Vergleicht man die Werte von Abbildung 8 mit den Werten von Abbildung 6, dann fällt auf, dass die meisten Namen aus der aktuellen Messung auch in Abbildung 8 auftreten. Die Positionen in den beiden Ranglisten variieren leicht und die aktuellen Zahlenwerte in Abbildung 6 liegen unter den im gesamten Zeitraum gemessenen Maximalwerten in Abbildung 8, wie zu erwarten war. Jedoch unterscheiden sich die gemessenen Werte nicht um Größenordnungen.

### 4.2.4 Anzahl Tracker je Embedder aus Messungen zu anderen Zeitpunkten

Hier geht es nun analog zum vorangegangenen Kapitel 4.2.3 darum, die Ergebnisse der aktuellen Messung zu den einbettenden Anbietern aus Kapitel 4.2.2 einzuordnen. Dazu wurden

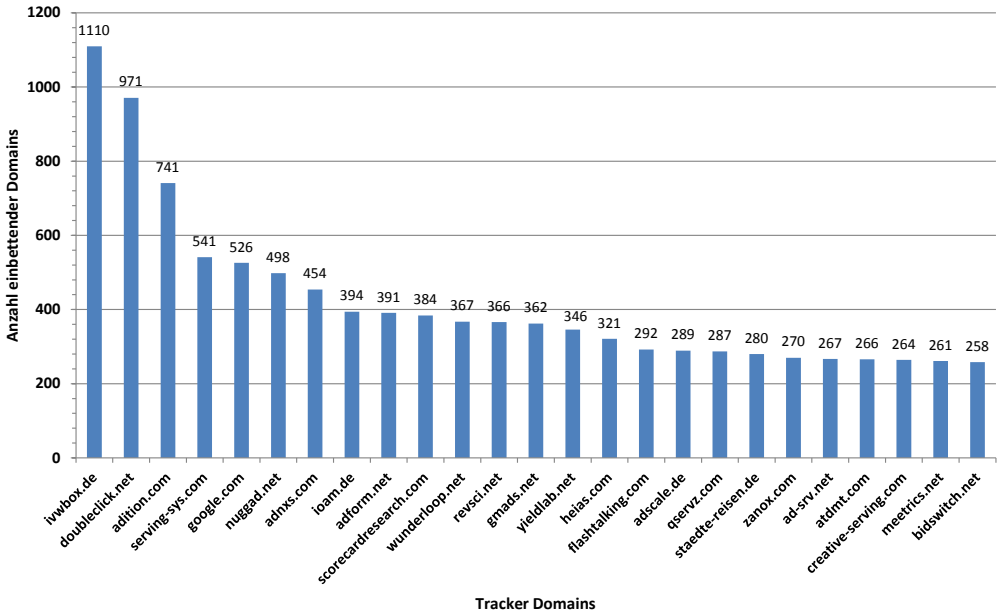


Abbildung 8. Die Rangliste der Tracker (Top25) nach der Anzahl von einbettenden Anbietern in Einzelmessungen zwischen November 2012 und Januar 2014.

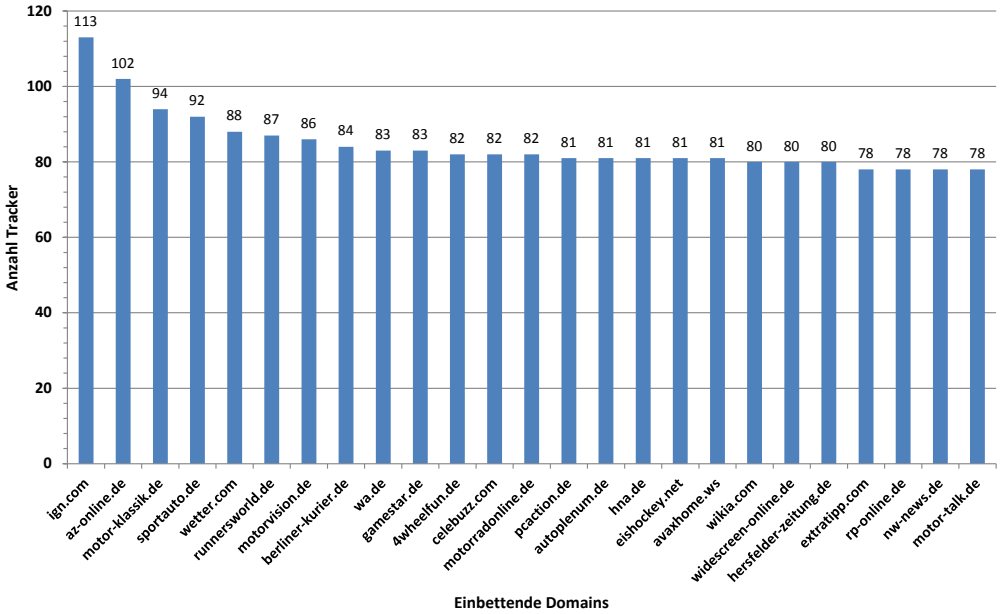


Abbildung 9. Die Rangliste der Embedder (Top25) nach der Anzahl von eingebetteten Trackern in Einzelmessungen zwischen November 2012 und Januar 2014.

für jeden Embedder diejenigen Einzelmessungen ausgesucht, bei denen ein Embedder die meisten Tracker integriert hat. Die höchsten Werte in Einzelmessungen für Einbettungen von Trackern bei Anbietern werden in Abbildung 9 dargestellt.

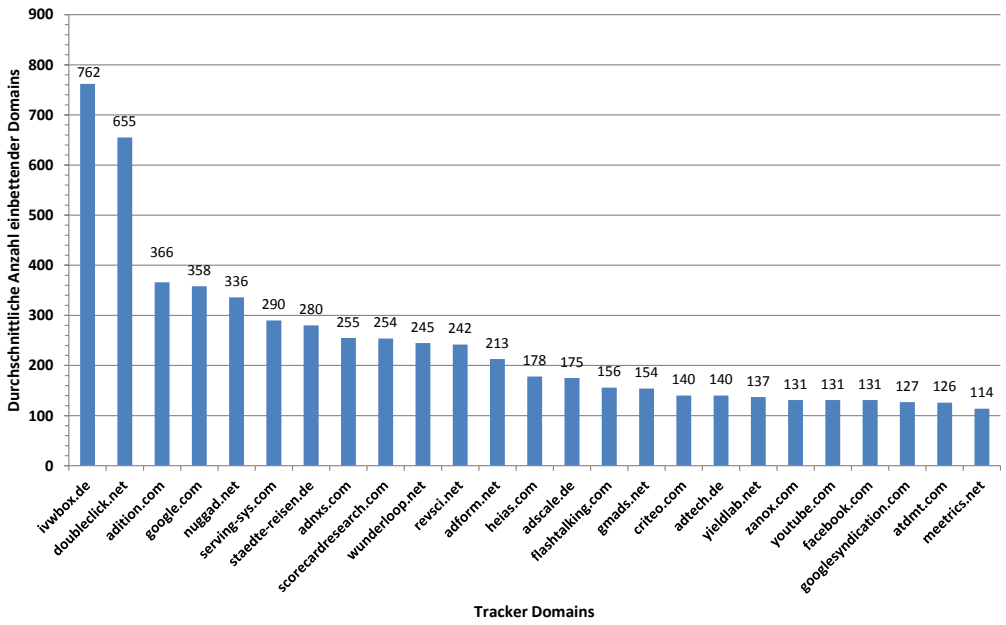


Abbildung 10. Die Rangliste der Tracker (Top25) nach der Anzahl von einbettenden Anbietern bei Einzelmessungen im Durchschnitt.

Vergleicht man die Werte von Abbildung 9 mit den Werten von Abbildung 7, dann fallen bei den Namen der Embedder schon größere Unterschiede auf. Die Anzahl der Tracker je Embedder scheint über der Zeit stärker zu variieren. Auch wenn die aktuelle Messung am 22.-24. Januar 2014 höchstens 67 Tracker für einen Embedder geliefert hat, dann ergibt die Messung in Abbildung 9, dass 80 Tracker und mehr je Embedder bei den betrachteten Domains keine Seltenheit sind.

### 4.2.5 Durchschnittliche Anzahl Embedder je Tracker

Einen anderen interessanten Einblick liefert die Antwort auf die Frage, bei wie vielen Anbietern im Durchschnitt ein Tracker eingebettet war. Hierfür wurde über allen Messungen, bei denen ein Tracker gefunden wurde, die Anzahl der Embedder ermittelt, bei denen der Tracker in einzelnen Messungen gefunden wurde, und danach der Durchschnitt über diesen Werten gebildet.[79] Die Top25 der Durchschnittswerte für Tracker hinsichtlich der Anzahl ihrer Einbettungen wird in Abbildung 10 dargestellt.

Vergleicht man die Aussage der Durchschnittswerte in Abbildung 10 mit den Werten in den Abbildungen 6 und 8, dann erkennt man, dass die durchschnittliche Verbreitung der Top25-Tracker sowohl den Werten in der aktuellen Messung als auch den Maximalwerten über allen Einzelmessungen ähnlich ist. Der überwiegende Anteil von Namen in Abbildung 10 findet sich auch in den Abbildungen 6 und 8 wieder. Auch die Positionierungen in allen Abbildungen sind ähnlich.

[79] Bei der Durchschnittsbildung wird die Anzahl der Messungen zugrunde gelegt, die seit der ersten Erkennung des jeweiligen Trackers durchgeführt wurden.

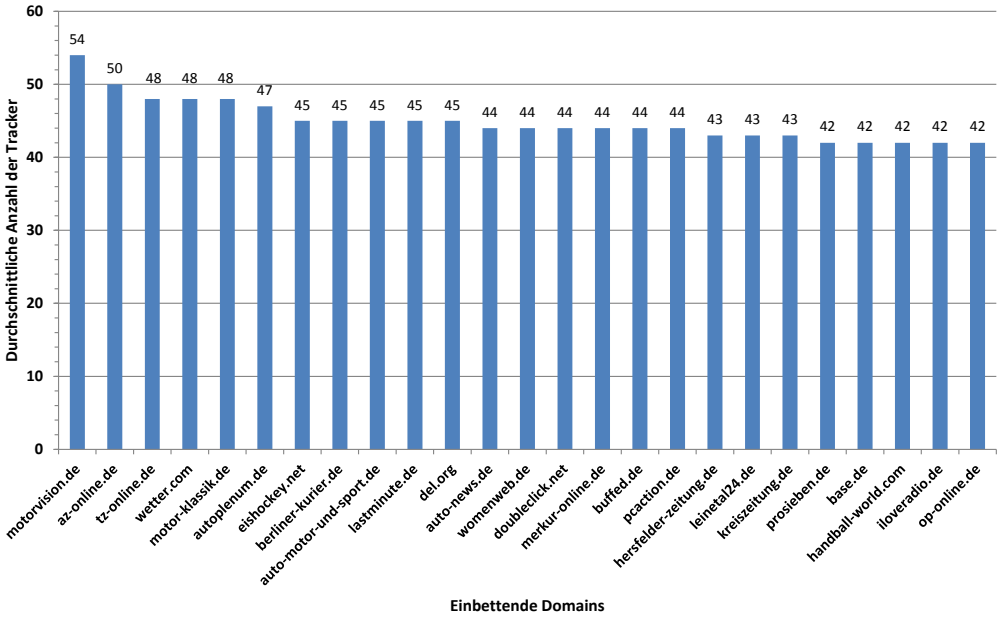


Abbildung 11. Die Rangliste der Embedder (Top25) nach der Anzahl von eingebetteten Trackern bei Einzelmessungen im Durchschnitt.

### 4.2.6 Durchschnittliche Anzahl Tracker je Embedder

Hier drehen wir die Durchschnittsbetrachtung aus Kapitel 4.2.5 herum und betrachten die Durchschnittswerte für die Anzahl von eingebetteten Trackern bei einzelnen Anbietern. Wir ermitteln die durchschnittliche Anzahl von Trackern, die bestimmte Anbieter einbetten. Hierbei gehen alle Messungen in die Durchschnittsbildung ein, seit der jeweilige Embedder in der Analyse berücksichtigt wurde. Das Ergebnis dieser Untersuchung wird für die Top25 Embedder in Abbildung 11 dargestellt.

In Abbildung 11 ist zu erkennen, dass die Top-Durchschnittswerte von eingebetteten Trackern für verschiedene Embedder sehr eng beeinander liegen. Das bedeutet, dass bei vielen Anbietern im Durchschnitt ähnlich viele Tracker zu finden sind. Es ist zu beachten, dass der Auswertung Einzelmessungen zugrunde liegen. Die Durchschnittswerte entsprechen also der Anzahl der Tracker, die man als Verbraucher bei einem Besuch der Anbieterseiten zu erwarten hat.

### 4.2.7 Gesamtanzahl Embedder je Tracker

Die Auswertungen zu den Trackern in den Kapitel 4.2.1, 4.2.3 und 4.2.5 beziehen sich immer auf Einzelmessungen. Damit wird jedoch der Aspekt nicht berücksichtigt, dass Tracker Verbraucher auch mit zeitlichem Versatz bei verschiedenen Embedder wiedererkennen können, wo sie möglicherweise gar nicht gleichzeitig eingebettet sind. Dennoch ist es möglich, dass ein Tracker einen Verbraucher durch eine Einbettung bei Embedder B wiedererkennt, den er bereits verfolgt hat, als der Tracker bei Embedder A eingebettet war, wo er jedoch mittlerweile nicht mehr eingebettet ist. Die Möglichkeit des Zusammenführens der zeitlich versetzt über Embedder A und Embedder B gewonnenen Daten hilft einem Tracker, ein entsprechend gutes Bild hinsichtlich des Verbrauchers zu bekommen, ohne dass ein Tracker simultan bei allen einbettenden Anbietern präsent ist.

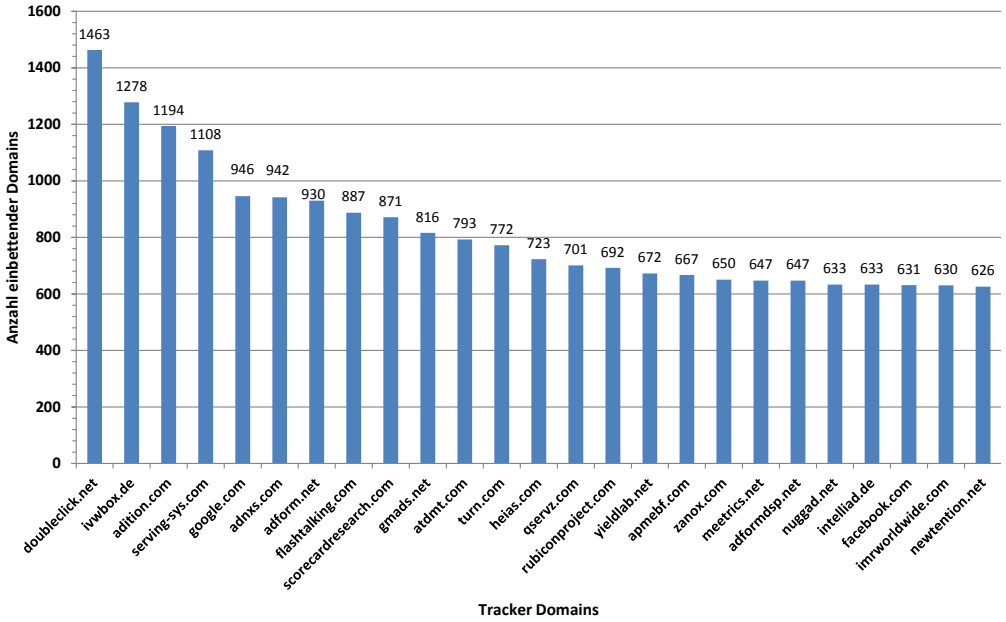


Abbildung 12. Die Rangliste der Tracker (Top25) nach der Gesamtanzahl von einbettenden Anbietern in allen Messungen zwischen November 2012 und Januar 2014.

Hier stellen wir also die Frage, bei wie vielen Anbietern wir insgesamt einen gegebenen Tracker bisher in unseren Analysen finden konnten. Die Antwort auf diese Frage ist für die Top25 Tracker in Abbildung 12 dargestellt.

Betrachtet man das Ergebnis in Abbildung 12, dann fällt auf, dass die Tracker mit der höchsten Verbreitung über dem gesamten Betrachtungszeitraum von November 2012 bis Januar 2014 bei einem sehr hohen Anteil der insgesamt 1634 verschiedenen Anbietern irgendwann einmal eingebettet waren. Doubleclick war in diesem Zeitraum bei rund 90% der betrachteten Anbieter eingebettet. Die Top10 der Tracker war in diesem Zeitraum praktisch bei der Hälfte aller betrachteten 1634 Anbieter eingebettet. Alle Top25 Tracker wurden in dem Betrachtungszeitraum bei mindestens 626 der insgesamt 1634 Anbieter entdeckt.

### 4.2.8 Gesamtanzahl Tracker je Embedder

Bei dieser Frage wollten wir erfahren, wie viele verschiedene Tracker insgesamt bestimmte Embedder über dem gesamten Betrachtungszeitraum eingebettet haben. Die Antwort auf diese Frage ist für die Top25 Embedder in Abbildung 13 dargestellt.

Auch hier fällt auf, dass sich die Werte für die einzelnen Embedder nicht stark voneinander unterscheiden. Darüber hinaus fällt auf, dass gerade Medienunternehmen über dem gesamten Zeitraum die meisten Tracker eingebettet haben. Die Top3-Einträge sind der Medienbranche zuzuordnen. Insgesamt gehören mehr als die Hälfte der in der Top25 genannten Embedder zur Medienbranche.

Vergleicht man die Ergebnisse in Abbildung 13 mit den Ergebnissen in den Abbildungen 7, 9 und 11, dann wird deutlich, dass viele Tracker über der Zeit variieren, d.h. es werden Tracker bei den Anbietern entfernt und dafür kommen wieder neue dazu. Ob es so etwas wie einen stabilen Kern von Trackern gibt, die über dem gesamten Zeitraum eingebettet waren, wurde im Rahmen dieser Analyse nicht untersucht.

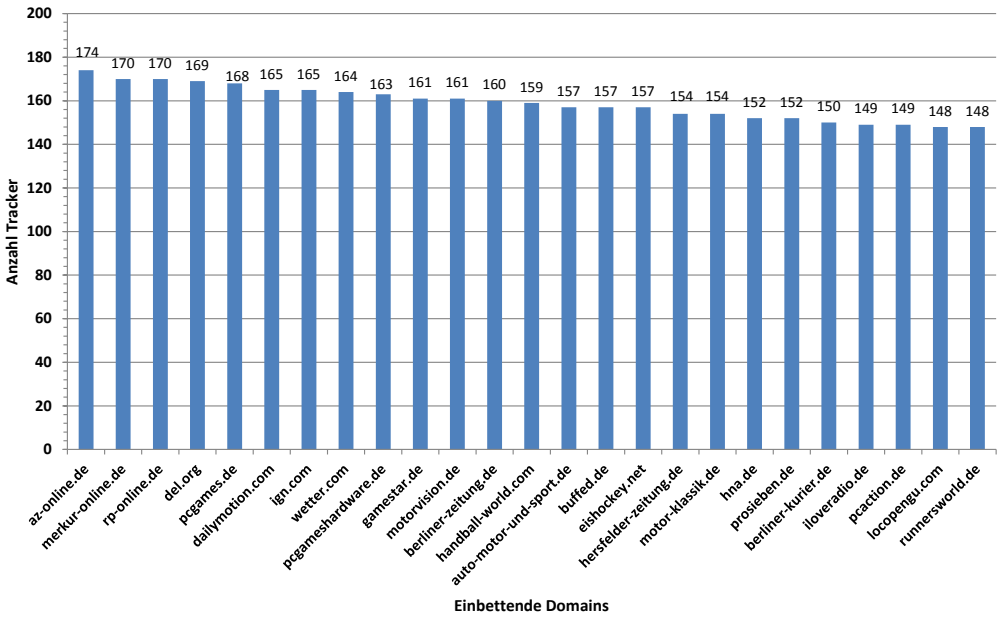


Abbildung 13. Die Rangliste der Embedder (Top25) nach der Gesamtanzahl von eingebetteten Trackern in allen Messungen zwischen November 2012 und Januar 2014.

# 5. SCHLUSSBEMERKUNG

Die Ergebnisse der vorangegangenen Kapitel haben gezeigt, wie sich die heutige Situation rund um Web-Tracking darstellt. Bei den Webseiten vieler Anbieter schauen den Verbrauchern Dutzende dritte Parteien über die Schulter, während die Verbraucher hiervon nichts ahnen. Das hat die quantitative Analyse in Kapitel 4 bestätigt. Die Ergebnisse haben auch gezeigt, in welchem Umfang ein Tracker Verbraucher über Seiten verschiedener Anbieter hinweg verfolgen kann. Viele Verbraucher wissen von dieser Verfolgung durch unbekannte Dritte beim Web-Browsen nichts, und wenn sie es wissen, dann ist ihnen oft nicht klar, was alles mit ihren Daten geschehen kann. Auch wenn die Verwertung der Daten für zielgerichtete Werbung bekannt sein mag, dann sind ihnen weitere Bedrohungen und Risiken durch Web-Tracking unbekannt. Wie im Vorangegangenen dargestellt wurde, ist die Verwertung der durch Web-Tracking gewonnenen Daten nicht auf Werbung beschränkt, selbst wenn sie heute für Werbezwecke eingesammelt werden. Doch die von Trackern gesammelten Daten werden nicht vergessen. Die Verwertung kann Lebensbereiche betreffen, von denen die Verbraucher heute noch nichts ahnen. Da sich in der Zukunft immer wieder neue Verwertungen ergeben werden, steht über den zukünftigen Risiken und Bedrohungen durch Web-Tracking zwangsläufig noch ein Fragezeichen. Selbst wenn heute ein Verbraucher der Sammlung von Daten für zielgerichtete Werbung zustimmen würde, dann würde er vielleicht einer Verwertung für die Berechnung des Risikos einer Bank bei der Kreditvergabe widersprechen.

Der Verbraucher ist heute das schwächste Glied in der Kette, insbesondere weil viele Verbraucher nicht hinreichend informiert sind, was hinter ihrem Rücken passiert, wenn sie im Internet unterwegs sind. Die Gelder, die hauptsächlich aus der Werbeindustrie stammen, treiben die Entwicklung von Tracking-Technologien hingegen weiter an. Insofern bestehen hier für den Verbraucher schlechte Aussichten. Es ist für die Zukunft nicht anzunehmen, dass die Online-Werbeunternehmen sich aus eigenem Antrieb in ihrem Tun beschränken und auf Tracking und zielgerichtete Werbung verzichten, erlaubt ihnen dies doch gerade die Vergrößerung ihrer Marktanteile zum Nachteil der Unternehmen, die mit klassischer Werbung ihr Geschäft machen (z.B. das klassische Printanzeigengeschäft). Auch von den Unternehmen, die für ihre Produkte Werbung machen, ist keine Änderung zu erwarten, da sie lieber in zielgerichtete Werbung investieren. So stellt sich noch die Frage, wie sich die Unternehmen verhalten, die zielgerichtete Werbung und Trackingmethoden in ihren Webseiten integrieren. Da sie für die Einbettung von Trackern in ihren Seiten eine Gegenleistung bekommen, bleibt die Einbettung natürlich weiter für sie interessant. Andererseits sagt die Einbettung aus, dass ein Anbieter ein größeres Interesse an der Gegenleistung hat als an dem Schutz seiner Kunden. Möglicherweise vertreten einige einbettende Anbieter den Standpunkt, dass der Wettbewerbszwang diesen Umgang mit Kunden unvermeidbar macht.

Eine Lösung hierfür könnte sein, dass sie die Entscheidung auf ihre Kunden übertragen und so lange keine Trackingmethoden anwenden, wie diese einer solchen Praxis nicht ausdrücklich zugestimmt haben. Jedoch besteht für Anbieter das Risiko einer solch fairen und transparenten Informationspraxis darin, dass sich zu viele Benutzer gegen Tracking entscheiden und somit auf potenzielle Einnahmen verzichtet werden muss. Eine andere Lösung könnte in einem Übereinkommen der mächtigsten Anbieter bestehen, indem alle auf Tracking und zielgerichtete Werbung verzichten und Werbung nur kontextorientiert schalten (z.B. Werbung für Bier, Autos, Baumärkte auf Sportseiten). Dadurch würde kein Anbieter unter Wettbewerbszugzwang gesetzt werden. Wahrscheinlich sind jedoch beide Vorschläge in der Realität nicht umsetzbar.

So lange sich in der Praxis nichts ändert, müssen Verbraucher selbst aktiv werden, um die negativen Konsequenzen von Web-Tracking zu vermeiden bzw. einzudämmen. Daten die Verbraucher bereits in der Vergangenheit an Tracker gegeben haben, können sie nicht mehr zurück holen — wie könnten sie auch, da sie ja gar nicht wissen, an wen ihre Daten geflossen sind. Wenn Verbraucher sich schützen wollen, dann stehen ihnen abhängig vom verwendeten Browser verschiedene Werkzeuge zur Verfügung.

Weitere Ansatzpunkte zur Verbesserung der Rahmenbedingungen hat die Politik. Sie kann Regeln schaffen und die Einhaltung dieser Regeln durchsetzen, so dass auch uninformierte Verbraucher nicht länger nichtsahnend und schutzlos der heutigen Praxis des Datensammelns durch Tracker ausgeliefert sind. Eine solche Änderung durch Rechtsgestaltung kann auch dazu führen, dass die Industrie einen Anreiz erkennt, neue datenschutzfreundliche Technologien zu entwickeln, die den Schutz des Verbrauchers auf der einen Seite und eine Möglichkeit zur zielgerichteten Werbung auf der anderen Seite miteinander verbindet.

Für die Interessen der Verbraucher wäre es wünschenswert, wenn die heutige Wildwestmanier des Datensammelns durch Tracking und der Verwertung dieser Daten möglichst bald beendet sein würde.

# LITERATUR


[AASGH11] Ayenson, Mika D.; Ashkan Soltani, Dietrich J. W.; Good, Nathaniel ; Hoofnagle, Chris Jay: *Flash Cookies and Privacy II: Now with HTML5 and ETag Respawning.* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1898390, July 29, 2011

[ABJB10] Aggarwal, Gaurav; Bursztein, Elie; Jackson, Collin ; Boneh, Dan: An Analysis of Private Browsing Modes in Modern Browsers. In: *19th USENIX Security Symposium, Washington, DC, USA, August 11-13, 2010, Proceedings*, 2010, S. 79–94

[AM10] Angwin, Julia; Mc Ginty, Tom: *Sites Feed Personal Details To New Tracking Industry.* The Wall Street Journal, http://online.wsj.com/article/SB10001424052748703977004575393173432219064.html, July 30, 2010

[And09] Anderson, Nate: *Änonymized"data really isn't – and here's why not.* http://arstechnica.com/tech-policy/2009/09/your-secrets-live-online-in-databases-of-ruin/, September 8, 2009

[Ang10] Angwin, Julia: *The Web's New Gold Mine: Your Secrets.* The Wall Street Journal, http://online.wsj.com/article/SB10001424052748703940904575395073512989404.html, July 30, 2010

[Ang11] Angwin, Julia: *Latest in Web Tracking: Stealthy 'Supercookies'.* The Wall Street Journal, http://online.wsj.com/article/SB10001424053111903480904576508382675931492.html, August 19, 2011

[Ant12] Anthes, Gary: HTML5 leads a web revolution. In: *Communications of the ACM* 55 (2012), Juli, Nr. 7, S. 16–17

[AS10] Angwin, Julia; Stecklow, Steve: *'Scrapers' Dig Deep for Data on Web.* The Wall Street Journal, http://online.wsj.com/article/SB10001424052748703358504575544381288117888.html, October 11, 2010

[AS11] Angwin, Julia; Steel, Emily: *Web's Hot New Commodity: Privacy.* The Wall Street Journal, http://online.wsj.com/article/SB10001424052748703529004576160764037920274.html, Februar 27, 2011

[AV10a] Angwin, Julia; Valentino-DeVries, Jennifer: *Analyzing What You Have Typed.* The Wall Street Journal, http://blogs.wsj.com/digits/2010/07/30/analyzing-what-you-have-typed, July 30, 2010

[AV10b] Angwin, Julia; Valentino-DeVries, Jennifer: *Race Is On to 'Fingerprint' Phones, PCs.* The Wall Street Journal, http://online.wsj.com/article/SB10001424052748704679204575646704100959546.html, November 30, 2010

[AV12] Angwin, Julia; Valentino-DeVries, Jennifer: *New Tracking Frontier: Your License Plates.* The Wall Street Journal, http://online.wsj.com/article/SB10000872396390443995604578004723603576296.html, September 29, 2012

[Bau13] Baumstieger, Moritz: *Die Datengräber.* Die Zeit, 28.2.2013, Nr.10, 2013

[Beu14] Beuth, Patrick: *Reform des EU-Datenschutzes ist erst einmal abgesagt.* Die Zeit, http://www.zeit.de/digital/datenschutz/2014-01/datenschutzreform-nicht-mehr-vor-europawahl/komplettansicht?print=true, 23. Januar, 2014

[BGW12] Buxmann, Peter; Gerlach, Jin ; Wenninger, Helena: *Der Preis des Kostenlosen – Ergebnisbericht zur Umfrage in Kooperation mit hr-iNFO.* http://blogs.hr-online.de/der-preis-des-kostenlosen/files/2013/05/Ergebnisbericht_Der_Preis_des_Kostenlosen.pdf, 2012

[BHH+12] Borgmann, Moritz; Hahn, Tobias; Herfert, Michael; Kunz, Thomas; Richter, Marcel; Viebeg, Ursula ; Vowé, Sven: *On the Security of Cloud Storage Services.* SIT Technical Report SIT-TR-2012-001, https://www.sit.fraunhofer.de/de/angebote/

projekte/cloud-studie/, März 2012

[BKMP12] Backes, Michael; Kate, Aniket; Maffei, Matteo ; Pecina, Kim: ObliviAd: Provably Secure and Practical Online Behavioral Advertising. In: *2012 IEEE Symposium on Security and Privacy*, 2012, S. 257–271

[BMFGI11] Boda, Károly; Máté Földes Ádám; Gulyás, Gábor G. ; Imre, Sándor: User Tracking on the Web via Cross-Browser Fingerprinting. In: *16th Nordic conference on Information Security Technology for Applications (NordSec 2011), Proceedings* Bd. 7161, Springer, 2011 (Lecture Notes in Computer Science), S. 31–46

[BZ06] Barbaro, Michael; Zeller, Tom: *A Face Is Exposed for AOL Searcher No. 4417749.* New York Times, http://www.nytimes.com/2006/08/09/technology/09aol.html?ex=1312776000&_r=0, August 9, 2006

[Car10] Carr, Nicholas: *Tracking Is an Assault on Liberty, With Real Dangers.* The Wall Street Journal, http://online.wsj.com/news/articles/SB10001424052748703748904575411682714389888, August 6, 2010

[Car12] Carroll, Robert: *The Alchemy of Behavioral Targeting.* http://www.destinationcrm.com/Articles/PrintArticle.aspx?ArticleID=81719, April 13, 2012

[CF07] Casado, Martin; Freedman, Michael J.: Peering through the Shroud: The Effect of Edge Opacity on IP-based Client Identification. In: *4th USENIX/ACM Symposium on Networked Systems Design and Implementation, Proceedings*, 2007, S. 173–186

[Cha02] Charters, Darren: Electronic Monitoring and Privacy Issues in Business-Marketing: The Ethics of the DoubleClick Experience. In: *Journal of Business Ethics* 35 (2002), Nr. 4, S. 243–254

[Chi12] Chirgwin, Richard: *Google snags patent on price discrimination – A sucker found in every cookie.* The Register, http://www.theregister.co.uk/2012/09/07/google_price_discrimination_patent/, 7th September, 2012

[Chr11] Christiansen, Linda: Personal privacy and Internet marketing: An impossible conflict or a marriage made in heaven? In: *Business Horizons* 54 (2011), November-December, Nr. 6, S. 509–514

[CKB12] Chaabane, Abdelberi; Kaafar, Mohamed Ali ; Boreli, Roksana: Big friend is watching you: analyzing online social networks tracking capabilities. In: *Proceedings of the 2012 ACM Workshop on online social networks (WOSN '12)*, 2012, S. 7–12

[Cle08] Cleland, Scott: *The Blind Eye to Privacy Law Arbitrage by Google – Broadly Threatens Respect for Privacy.* House Energy & Commerce Subcommittee on Internet Hearing, http://www.netcompetition.org/Written_Testimony_House_Privacy_071707.pdf, July, 17, 2008

[Clo02] Clover, Andrew: *CSS visited pages disclosure.* Bugtraq mailing list archives, http://seclists.org/bugtraq/2002/Feb/271, Februar 20, 2002

[Dat13] Datalogix: *Privacy.* http://www.datalogix.com/privacy, Version vom 16.4.2013, 2013

[DDNP12] De Groef, Willem; Devriese, Dominique; Nikiforakis, Nick ; Piessens, Frank: FlowFox: a web browser with flexible and precise information flow control. In: *ACM Conference on Computer and Communications Security (CCS 2012), Proceedings*, 2012, S. 748–759

[Dro13] Dropbox: *Dropbox Privacy Policy.* https://www.dropbox.com/privacy, Version April 10, 2013

[Duh12] Duhigg, Charles: *How Companies Learn Your Secrets.* The New York Times, http://www.nytimes.com/2012/02/19/magazine/shopping-habits.

html?pagewanted=print, February 16, 2012

[Eck09] Eckersley, Peter: *How Unique Is Your Web Browser?* Electronig Frontier Foundation, Technical Report, https://panopticlick.eff.org/browser-uniqueness.pdf, 2009

[Eck10a] Eckersley, Peter: *Browser Versions Carry 10.5 Bits of Identifying Information on Average.* Electronic Frontier Foundation, https://www.eff.org/deeplinks/2010/01/tracking-by-user-agent, January, 27, 2010

[Eck10b] Eckersley, Peter: How Unique is Your Web Browser? In: *Proceedings of the 10th International Conference on Privacy Enhancing Technologies (PETs 2010)* Bd. 6205, Springer, 2010 (Lecture Notes in Computer Science), S. 1–18

[EGC+10] Enck, William; Gilbert, Peter; Chun, Byung-Gon; Cox, Landon P.; Jung, Jaeyeon; McDaniel, Patrick ; Sheth, Anmol N.: TaintDroid: An Information-flow Tracking System for Realtime Privacy Monitoring on Smartphones. In: *9th USENIX Conference on Operating Systems Design and Implementation (OSDI'10), Proceedings*, 2010, S. 1–6

[EKKV11] Egele, Manuel; Kruegel, Christopher; Kirda, Engin ; Vigna, Giovanni: PiOS: Detecting Privacy Leaks in iOS Applications. In: *Network and Distributed System Security Symposium (NDSS 2011), Proceedings*, 2011

[Ell12] Ellis, Shelley: *The Future of Retargeting, Remarketing and Remessaging.* Marketing Land, http://marketingland.com/the-future-of-retargeting-remarketing-and-remessaging-7643, August 29, 2012

[Eur09] Europäische Union: *Richtline 2009/136/EG des Europäischen Parlaments und des Rates vom 25. November 2009 zur Änderung der Richtlinie 2002/22/EG über den Universaldienst und Nutzerrechte bei elektronischen Kommunikationsnetzen und -diensten, der Richtlinie 2002/58/EG über die Verarbeitung personenbezogener Daten und den Schutz der Privatsphäre in der elektronischen Kommunikation und der Verordnung (EG) Nr. 2006/2004 über die Zusammenarbeit im Verbraucherschutz.* Amtsblatt der Europäischen Union, L 337/11, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:337:0011:0036:De:PDF, 2009

[Eur12a] European Commission: *Vorschlag für Verordnung des Europäischen Parlaments und des Rates zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr (Datenschutz-Grundverordnung).* http://www.europarl.europa.eu/registre/docs_autres_institutions/commission_europeenne/com/2012/0011/COM_COM(2012)0011_DE.pdf, 2012

[Eur12b] European Commission: *Why do we need an EU data protection reform?* http://ec.europa.eu/justice/data-protection/document/review2012/factsheets/1_en.pdf, 2012

[Eva09] Evans, David S.: The Online Advertising Industry: Economics, Evolution, and Privacy. In: *Journal of Economic Perspectives* 23 (2009), Nr. 3, S. 37–60

[Fer13] Fertik, Michael: *The Rich See a Different Internet Than the Poor.* Scientific American, http://www.scientificamerican.com/article.cfm?id=rich-see-different-internet-than-the-poor, February, 18, 2013

[FGM+99] Fielding, R.; Gettys, J.; Mogul, J.; Frystyk, H.; Masinter, L.; Leach, P. ; Berners-Lee, T.: *Hypertext Transfer Protocol – HTTP/1.1.* RFC 2616, Standards Track, http://www.ietf.org/rfc/rfc2616.txt, June, 1999

[FMSW11] Friedland, Gerald; Maier, Gregor; Sommer, Robin ; Weaver, Nicholas: Sherlock holmes' evil twin: on the impact of global inference for online privacy. In: *New security paradigms workshop (NSPW 2011), Proceedings*, 2011, S. 105–114

[FS00] Felten, Edward W.; Schneider, Michael A.: Timing attacks on Web privacy. In: *7th ACM Conference on Computer and Communications Security (CCS 2000), Proceedings*, 2000, S. 25–32

[FS13] Fielding, Roy T.; Singer, David: *Tracking Preference Expression (DNT).* W3C Working Draft, `http://www.w3.org/TR/tracking-dnt`, September 12, 2013

[Gar10] Garfinkel, Simson: *Du sollst nicht ausspähen!* Technology Review, Oktober 2010

[GCF10] Guha, Saikat; Cheng, Bin ; Francis, Paul: Challenges in measuring online advertising systems. In: *10th ACM SIGCOMM Conference on Internet Measurement (IMC 2010), Proceedings*, 2010, S. 81–87

[GCF11] Guha, Saikat; Cheng, Bin ; Francis, Paul: Privad: Practical Privacy in Online Advertising. In: *8th USENIX Conferenve on Networked Systems Design and Implementation (NSDI 2011), Proceedings*, 2011

[GM08] Gunawardana, Asela; Meek, Christopher: Aggregators and Contextual Effects in Search Ad Markets. In: *WWW Workshop on Targeting and Ranking for Online Advertising*, 2008

[Gol11] Goldmedia: *Smart TV: Wer erringt die Portalhoheit auf dem Fernseher?* November, 2011

[Gol13] Goldfarb, Avi: What is Different About Online Advertising? In: *Review of Industrial Organization* 43 (2013), Nr. 1-2, S. 1–15

[GOT13] Ghiglieri, Marco; Oswald, Florian ; Tews, Erik: HbbTV - I Know What You Are Watching. In: *13. Deutscher IT-Sicherheitskongresses, Konferenzband*, 2013, S. ??

[GPS09] Gomez, Joshua; Pinnick, Travis ; Soltani, Ashkan: *KnowPrivacy.* UC Berkeley, School of Information, `http://www.knowprivacy.org/report/KnowPrivacy_Final_Report.pdf`, 2009

[Gro13] Gross, Grant: *Survey: Internet users like targeted ads, free content.* TechWorld, `http://www.techworld.com.au/article/print/459632/survey_internet_users_like_targeted_ads_free_content/`, April 19, 2013

[GT11a] Goldfarb, Avi; Tucker, Catherine E.: Online Advertising, Behavioral Targeting, and Privacy. In: *Communications of the ACM* 54 (2011), Nr. 5, S. 25–27

[GT11b] Goldfarb, Avi; Tucker, Catherine E.: Privacy Regulation and Online Advertising. In: *Management Science* 57 (2011), Nr. 1, S. 57–71

[GT14] Ghiglieri, Marco; Tews, Erik: A Privacy Protection System for HbbTV in Smart TVs. In: *IEEE 11th Consumer Communications and Networking Conference (CCNC 2014, Proceedings)*, 2014, S. ??

[Han12] Handelsblatt: *Institut springt ab – Schufa verliert Partner für Facebook-Projekt.* Handelsblatt, `http://www.handelsblatt.com/technologie/it-tk/it-internet/institut-springt-ab-schufa-verliert-partner-fuer-facebook-projekt/6726236.html`, 8. Juni, 2012

[Har10] Harper, Jim: *It's Modern Trade: Web Users Get as Much as They Give.* The Wall Street Journal, `http://online.wsj.com/article/SB10001424052748703748904575411530096840958.html#`, August 6, 2010

[Hea13] Healey, Jon: *Privacy advocates attack Gmail - again - for email scanning.* Los Angeles Times, August 15, 2013, `http://www.latimes.com/news/opinion/opinion-la/la-ol-google-gmail-privacy-reasonable-expectation-20130814,0,2662122.story`, 2013

[hei11] heise online: *Websites hebeln Anti-Cookie-Maßnahmen aus.* `http://www.heise.de/newsticker/meldung/Websites-hebeln-Anti-Cookie-Massnahmen-aus-1288914.html`, 31. Juli, 2011

[hei13] heise Security: *Hacker klauen Daten von 50 Millionen LivingSocial-Kunden*. http://www.heise.de/security/meldung/Hacker-klauen-Daten-von-50-Millionen-LivingSocial-Kunden-1851131.html, 27. April, 2013

[hei14a] heise online: *Schwarz-Rot: Vorratsdatenspeicherung bleibt weiter Zankapfel*. http://www.heise.de/newsticker/meldung/Schwarz-Rot-Vorratsdatenspeicherung-bleibt-weiter-Zankapfel-2084219.html, 13. Januar, 2014

[hei14b] heise Security: *Spion im Wohnzimmer: c't entdeckt Sicherheitslücken in zahlreichen Smart-TVs*. http://www.heise.de/security/meldung/Spion-im-Wohnzimmer-c-t-entdeckt-Sicherheitsluecken-in-zahlreichen-Smart-TVs-2097287.html, 27. Januar, 2014

[hei14c] heise Security: *Spion im Wohnzimmer: c't ertappt schnüffelnde Fernseher*. http://www.heise.de/security/meldung/Spion-im-Wohnzimmer-c-t-ertappt-schnueffelnde-Fernseher-2096578.html, 25. Januar, 2014

[HHJ+11] Hornyack, Peter; Han, Seungyeop; Jung, Jaeyeon; Schechter, Stuart ; Wetherall, David: These Aren't the Droids You're Looking for: Retrofitting Android to Protect Data from Imperious Applications. In: *Proceedings of the 18th ACM Conference on Computer and Communications Security (CCS 2011)*, 2011, S. 639–652

[Hic13] Hickson, Ian: *Web Storage*. W3C Recommendation, http://www.w3.org/TR/2013/REC-webstorage-20130730/, July 30, 2013

[Hil14] Hill, Simon: *Get in shape with 10 of the best fitness trackers from CES 2014*. Digital Trends, http://www.digitaltrends.com/mobile/ces-wants-get-shape-fitness-trackers-galore/, January, 11, 2014

[HJW12] Han, Seungyeop; Jung, Jaeyeon ; Wetherall, David: *A Study of Third-Party Tracking by Mobile Apps in the Wild*. University of Washington, Technical Report, UW-CSE-12-03-01, ftp://ftp.cs.washington.edu/tr/2012/03/UW-CSE-12-03-01.pdf, March, 2012

[Hol09] Holvast, Jan: History of Privacy. In: Matyáš, Vashek (Hrsg.); Fischer-Hübner, Simone (Hrsg.); Cvrček, Daniel (Hrsg.) ; Švenda, Petr (Hrsg.): *The Future of Identity in the Information Society* Bd. 298. Springer Berlin Heidelberg, 2009, S. 13–42

[HRS12] Horchert, Judith; Reißmann, Ole ; Stöcker, Christian: *Schnüffel-Plan der Schufa: Was Facebook über Sie verrät*. Spiegel Online, http://www.spiegel.de/netzwelt/netzpolitik/facebook-was-ein-profil-der-schufa-verraten-kann-a-837531.html, 7. Juni, 2012

[HSG+12] Hoofnagle, Chris Jay; Soltani, Ashkan; Good, Nathaniel; Wambach, Dietrich J. ; Ayenson, Mika D.: Behavioral Advertising: The Offer You Cannot Refuse. In: *Harvard Law & Policy Review* 6 (2012), Nr. 2, S. 273–296

[HSK+13] Hannak, Aniko; Sapieżyński, Piotr; Kakhki, Arash M.; Krishnamurthy, Balachander; Lazer, David; Mislove, Alan ; Wilson, Christo: Measuring Personalization of Web Search. In: *Proceedings of the Twenty-Second International World Wide Web Conference (WWW'13)*, 2013, S. 527–538

[HV10] Helft, Miguel; Vega, Tanzina: *Retargeting Ads Follow Surfers to Other Sites*. New York Times, http://www.nytimes.com/2010/08/30/technology/30adstalk.html?_r=0, August 29, 2010

[JBBM06] Jackson, Collin; Bortz, Andrew; Boneh, Dan ; Mitchell, John C.: Protecting browser state from web privacy attacks. In: *15th International Conference on World Wide Web (WWW 2006), Proceedings*, 2006, S. 737–744

[JJLS10] Jang, Dongseok; Jhala, Ranjit; Lerner, Sorin ; Shacham, Hovav: An empirical study of privacy-violating information flows in JavaScript web applications. In: *17th*

*ACM Conference on Computer and Communications Security (CCS 2010), Proceedings*, 2010, S. 270–283

[JO10] Janc, Artur; Olejnik, Lukasz: Web Browser History Detection as a Real-World Privacy Threat. In: *Computer Security - ESORICS 2010, 15th European Symposium on Research in Computer Security, Proceedings* Bd. 6345, Springer Verlag, 2010 (Lecture Notes in Computer Science), S. 215–231

[Jun13] Junge, Barbara: Wer hat meine Daten? In: Geiselberger, Heinrich (Hrsg.); Moorstedt, Tobias (Hrsg.): *Big Data – Das neue Versprechen der Allwissenheit.* edition unseld, Suhrkamp Verlag, 2013, S. 23–34

[Kau11] Kaufman, Lori M.: How Private Is the Internet? In: *IEEE Security & Privacy* 9 (2011), Nr. 1, S. 73–75

[KKMK09] Kosta, Eleni; Kalloniatis, Christos; Mitrou, Lilian ; Kavakli, Evangelia: Search Engines: Gateway to a New "Panopticon"? In: *Trust, Privacy and Security in Digital Business, 6th International Conference (TrustBus 2009), Proceedings* Bd. 5695, Springer Verlag, 2009 (Lecture Notes on Computer Science), S. 11–21

[Kön13] König, Tom: *US-Datenhändler Rapleaf: Jeder kann NSA.* Spiegel Online, `http://www.spiegel.de/netzwelt/web/datenhaendler-rapleaf-nsa-software-fuer-jedermann-a-925611.html`, 4. Oktober, 2013

[KNW11] Krishnamurthy, Balachander; Naryshkin, Konstantin ; Wills, Craig E.: Privacy leakage vs. Protection measures: the growing disconnect. In: *Web 2.0 Security and Privacy Workshop (W2SP 2011)*, 2011

[Kre14] Kremp, Matthias: *Heizungssteuerung per App — Wenn ich weg bin, wird es kalt.* Spiegel Online, `http://www.spiegel.de/netzwelt/gadgets/angefasst-heizungssteuerung-tado-im-test-a-943622.html`, 16. Januar, 2014

[Kri10a] Krishnamurthy, Bachalander: I know what you will do next summer. In: *ACM SIGCOMM Computer Communication Review* 40 (2010), Oktober, Nr. 5, S. 65–70

[Kri10b] Krishnamurthy, Bachalander: *Privacy leakage on the Internet.* 77th Internet Engineering Task Force, Proceedings, `http://www.ietf.org/proceedings/77/slides/plenaryt-5.pdf`, March 21-26, 2010

[Kri13] Krishnamurthy, Balachander: Privacy and online social networks: can colorless green ideas sleep furiously? In: *IEEE Security & Privacy* 11 (2013), Nr. 3, S. 14–20

[Kro12] Kroes, Neelie: *Online privacy and online business: An update on Do Not Track.* European Commission, Press Releases Database, `http://europa.eu/rapid/press-release_SPEECH-12-716_en.htm`, October 11, 2012

[KSS+12] Kelbert, Florian; Shirazi, Fatemeh; Simo, Hervais; Wüchner, Tobias; Buchmann, Johannes; Pretschner, Alexander ; Waidner, Michael: State of Online Privacy: A Technical Perspective. In: Buchmann, Johannes (Hrsg.): *Internet Privacy. Eine multidisziplinäre Bestandsaufnahme/ A multidisciplinary analysis (acatech STUDIE).* Springer Berlin Heidelberg, September 2012, S. 189–279

[KW09] Krishnamurthy, Balachander; Wills, Craig E.: Privacy diffusion on the web: a longitudinal perspective. In: *18th International Conference on World Wide Web (WWW 2009), Proceedings*, 2009, S. 541–550

[LCC+12] Leon, Pedro Giovanni; Cranshaw, Justin; Cranor, Lorrie Faith; Graves, Jim; Hastak, Manoj; Ur, Blase ; Xu, Guzi: What do online behavioral advertising privacy disclosures communicate to users? In: *2012 ACM workshop on Privacy in the electronic society (WPES 2012)*, 2012, S. 19–30

[LEPM12] Leontiadis, Ilias; Efstratiou, Christos; Picone, Marco ; Mascolo, Cecilia: Don't Kill My Ads!: Balancing Privacy in an Ad-supported Mobile Application Market. In:

*Twelfth Workshop on Mobile Computing Systems & Applications (HotMobile 2012), Proceedings*, 2012, S. 2:1–2:6

[LT13] Lambrecht, Anja; Tucker, Catherine: *When Does Retargeting Work? Information Specificity in Online Advertising.* Social Science Research Network, SSRN Working Paper Series, http://ssrn.com/abstract=1795105, Mai 2013

[LUB+11] Leon, Pedro G.; Ur, Blase; Balebako, Rebecca; Cranor, Lorrie Faith; Shay, Richard ; Wang, Yang: *Why Johnny Can't Opt Out: A Usability Evaluation of Tools to Limit Online Behavioral Advertising.* Technical Report, CyLab, Carnegie Mellon University, https://www.cylab.cmu.edu/files/pdfs/tech_reports/CMUCyLab11017.pdf, October, 31, 2011

[Mac14] Mack, Daniel: *Offizielle deutsche Olympia-App: Macht der DOSB Daten zu Gold?* http://danielmack.de/offizielle-deutsche-olympia-app-macht-der-dosb-daten-zu-gold/, 6. Februar, 2014

[Mar10] MarketingCharts: *In Right Context, Internet Users Like Targeted Ads.* http://www.marketingcharts.com/wp/uncategorized/in-right-context-internet-users-like-targeted-ads-13474/, July 8, 2010

[Mar12] Marwan, Peter: *Arbeitsrecht: Facebook-Postings als Kündigungsgrund.* ZDNet.de, http://www.zdnet.de/41559610/arbeitsrecht-facebook-postings-als-kuendigungsgrund, 25. Januar, 2012

[Mat13] Matsudaira, Kate: Making the mobile web faster. In: *Communications of the ACM* 56 (2013), März, Nr. 3, S. 56–61

[May09] Mayer, Jonathan R.: *Äny person... a pamphleteer- Internet Anonymity in the Age of Web 2.0.* Thesis, Woodrow Wilson School of Public and International Affairs, http://www.stanford.edu/~jmayer/papers_data/thesis09.pdf, 2009

[May11] Mayer, Jonathan: *Tracking the Trackers: Where Everybody Knows Your Username.* http://cyberlaw.stanford.edu/blog/2011/10/tracking-trackers-where-everybody-knows-your-username, October, 11, 2011

[MBYS11] Mowery, Keaton; Bogenreif, Dillon; Yilek, Scott ; Shacham, Hovav: Fingerprinting Information in JavaScript Implementations. In: *Web 2.0 Security and Privacy (W2SP 2011), Proceedings*, 2011

[MC10] McDonald, Aleecia M.; Cranor, Lorrie Faith: *Beliefs and Behaviors: Internet Users' Understanding of Behavioral Advertising.* Social Science Research Network, SSRN Working Paper Series, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1989092, August 2010

[Mer12] Meredith, Leslie: *How Google's New Privacy Policy Could Affect You.* Scientific American, http://www.scientificamerican.com/article.cfm?id=how-googles-new-privacy-p&print=true, January 27, 2012

[Mil09] Mills, Elinor: *Device identification in online banking is privacy threat, expert says.* CNET, http://news.cnet.com/8301-1009_3-10226742-83.html, April, 24, 2009

[Mit10] Mittal, Sonal: *User Privacy and the Evolution of Third-party Tracking Mechanisms on the World Wide Web.* Social Science Research Network, SSRN Working Paper Series, http://ssrn.com/abstract=2005252, 2010

[MM12] Mayer, Jonathan R.; Mitchell, John C.: Third-Party Web Tracking: Policy and Technology. In: *Proceedings of the 2012 IEEE Symposium on Security and Privacy*, 2012, S. 413–427

[MNS11] Mayer, J.; Narayanan, A. ; Stamm, S.: *Do Not Track: A Universal Third-Party Web Tracking Opt Out.* IETF Internet-Draft, http://tools.ietf.org/html/draft-mayer-do-not-track-00, March 7, 2011

[Moo13] Moorstedt, Tobias: Obamas Datenakrobaten. In: Geiselberger, Heinrich (Hrsg.); Moorstedt, Tobias (Hrsg.): *Big Data – Das neue Versprechen der Allwissenheit.* edition unseld, Suhrkamp Verlag, 2013, S. 35–54

[MS12a] Morton, Anthony; Sasse, M. Angela: Privacy is a Process, Not a PET: A Theory for Effective Privacy Practice. In: *Proceedings of the 2012 Workshop on New Security Paradigms (NSPW'12)*, 2012, S. 87–104

[MS12b] Mowery, Keaton; Shacham, Hovav: Pixel Perfect: Fingerprinting Canvas in HTML5. In: *Web 2.0 Security and Privacy (W2SP 2012), Proceedings*, 2012

[NH10] Nguyen, David H.; Hayes, Gillian R.: Information Privacy in Institutional and End-user Tracking and Recording Technologies. In: *Personal Ubiquitous Computing* 14 (2010), Januar, Nr. 1, S. 53–72

[NKJ+13] Nikiforakis, Nick; Kapravelos, Alexandros; Joosen, Wouter; Kruegel, Christopher; Piessens, Frank ; Vigna, Giovanni: Cookieless monster: Exploring the ecosystem of web-based device fingerprinting. In: *Proceedings of the 34th IEEE Symposium on Security & Privacy,*, 2013, S. 541–555

[NS08] Narayanan, Arvind; Shmatikov, Vitaly: Robust De-anonymization of Large Sparse Datasets. In: *29th IEEE Symposium on Security and Privacy (S&P 2008), Proceedings*, 2008, S. 111–125

[NS09] Narayanan, Arvind; Shmatikov, Vitaly: De-anonymizing Social Networks. In: *30th IEEE Symposium on Security and Privacy (S&P 2009), Proceedings*, 2009, S. 173–187

[NS10] Narayanan, Arvind; Shmatikov, Vitaly: Myths and fallacies of "Personally Identifiable Information". In: *Communications of the ACM* 53 (2010), Juni, Nr. 6, S. 24–26

[Odl03] Odlyzko, Andrew: Privacy, Economics, and Price Discrimination on the Internet. In: *Proceedings of the 5th International Conference on Electronic Commerce (ICEC'03)*, 2003, S. 355–366

[OWA13] OWASP: *Cross-site Scripting (XSS).* `https://www.owasp.org/index.php/Cross-site_Scripting_(XSS)`, September 13, 2013

[Plö12] Plöchinger, Stefan: *Was uns die Schnüffel-Schufa lehrt.* Süddeutsche, `http://www.sueddeutsche.de/digital/2.220/versuchte-facebook-twitter-analysen-was-uns-die-schnueffel-schufa-lehrt-1.1376581`, 14. Juni, 2012

[Pol14] Polizeipräsidium Mainz, Kriminaldirektion: *Wohnungseinbrüche – Polizei warnt vor neuer Masche!* Polizei Rheinland-Pfalz - Polizeipräsidium Mainz, `https://www.polizei.rlp.de/internet/nav/f6a/presse.jsp?uMen=f6a70d73-c9a2-b001-be59-2680a525fe06&sel_uCon=ef533022-35f0-341c-5ec3-f1f282c266d1&page=1&pagesize=10`, 3. Januar, 2014

[PPP+13] Petsas, Thanasis; Papadogiannakis, Antonis; Polychronakis, Michalis; Markatos, Evangelos P. ; Karagiannis, Thomas: Rise of the Planet of the Apps: A Systematic Study of the Mobile App Ecosystem. In: *Internet Measurement Conference (IMC '13), Proceedings*, 2013, S. 277–290

[Rab13] Rabinovich, Paul: Secure cross-domain cookies for HTTP. In: *Journal of Internet Services and Applications* 4 (2013), Nr. 1

[Ram12] Ramasastry, Anita: *Should Target Tell Your Loved Ones You Are Pregnant, Or Should You? — The Perils of Consumer Data Aggregation, Including Loss of Privacy.* Verdict, `http://verdict.justia.com/2012/02/28/should-target-tell-your-loved-ones-you-are-pregnant-or-should-you`, February 28, 2012

[RGF11] Reznichenko, Alexey; Guha, Saikat ; Francis, Paul: Auctions in do-not-track compliant internet advertising. In: *18th ACM Conference on Computer and Communications Security (CCS 2011), Proceedings*, 2011, S. 667–676

[RKW12] Roesner, Franziska; Kohno, Tadayoshi ; Wetherall, David: Detecting and defending against third-party tracking on the web. In: *9th USENIX conference on Networked Systems Design and Implementation (NSDI 2012), Proceedings*, 2012, S. 12–12

[RL13] Robert Levine, Martin K.: *Das Recht auf Vergessen.* Die Zeit, 10.10.2013, Nr.41, 2013

[RM13] Reißmann, Ole; Mayer-Schönberger, Viktor: *Ïch wünsche mir ein Recht auf Irrationalität".* Spiegel Online, http://www.spiegel.de/netzwelt/web/interview-mit-viktor-mayer-schoenberger-zu-big-data-a-936741.html, Dezember, 20, 2013

[Röt14] Rötzer, Florian: *Vorratsdatenspeicherung – erst mal abwarten.* Telepolis, http://www.heise.de/tp/blogs/8/155697, 17. Januar, 2014

[Rut13] Rutenberg, Jim: *Data You Can Believe In.* The New York Times, http://www.nytimes.com/2013/06/23/magazine/the-obama-campaigns-digital-masterminds-cash-in.html?_r=0&pagewanted=print, June, 20, 2013

[Sch01] Schwartz, John: *Giving the Web a Memory Cost Its Users Privacy.* The New York Times, http://www.nytimes.com/2001/09/04/business/giving-web-a-memory-cost-its-users-privacy.html, September, 4, 2001

[Sch09] Schoen, Seth: *New Cookie Technologies: Harder to See and Remove, Widely Used to Track You.* Electronic Frontier Foundation, https://www.eff.org/deeplinks/2009/09/new-cookie-technologies-harder-see-and-remove-wide, September, 14, 2009

[Sch13a] Schirrmacher, Frank: Der verwettete Mensch. In: Geiselberger, Heinrich (Hrsg.); Moorstedt, Tobias (Hrsg.): *Big Data – Das neue Versprechen der Allwissenheit.* edition unseld, Suhrkamp Verlag, 2013, S. 273–294

[Sch13b] Schmitt, Stefan: *Ein Bild aus tausend Spuren.* Die Zeit, 1.8.2013, Nr.32, 2013

[Sch13c] Schneier, Bruce: *SStalker Economy"Here to Stay.* https://www.schneier.com/essay-467.html, November, 20, 2013

[SCM+09] Soltani, Ashkan; Canty, Shannon; Mayo, Quentin; Thomas, Lauren ; Hoofnagle, Chris Jay: *Flash Cookies and Privacy.* SSRN, http://ssrn.com/abstract=1446862, August 10, 2009

[SF10] Steel, Emily; Fowler, Geoffrey A.: *Facebook in Privacy Breach.* The Wall Street Journal, http://online.wsj.com/news/articles/SB10001424052702304772804575558484075236968, October 17, 2010

[SM10] Scism, Leslie; Maremont, Mark: *Insurers Test Data Profiles to Identify Risky Clients.* The Wall Street Journal, http://online.wsj.com/news/articles/SB10001424052748704648604575620750998072986, November 18, 2010

[Sno14] Snowden, Edward: *Snowden exclusiv – Das Interview.* NDR, http://www.ndr.de/ratgeber/netzwelt/snowden263.html, 26. Januar, 2014

[Spi12] Spiegel Online: *Kreditwürdigkeit – Schufa will Facebook-Nutzer durchleuchten.* Spiegel Online, http://www.spiegel.de/netzwelt/web/schufa-will-kreditdaten-bei-facebook-sammeln-a-837454.html, 7. Juni, 2012

[Spi13a] Spiegel Online: *Hacker erbeutet Bankdaten von Millionen Vodafone-Kunden.* Spiegel Online, http://www.spiegel.de/netzwelt/netzpolitik/vodafone-hacker-stehlen-daten-von-millionen-kundena-921790.html, 12. September, 2013

[Spi13b] Spiegel Online: *Hackerangriff: Adobe korrigiert Zahl betroffener Kunden auf 38 Millionen.* Spiegel Online, http://www.spiegel.de/netzwelt/web/hacker-erbeuteten-daten-von-38-millionen-adobe-kunden-a-930855.html, 30. Oktober, 2013

[Spi13c] Spiegel Online: *Handel mit vertraulichen Daten – Millionen deutsche Patienten und Ärzte werden ausgespäht.* Spiegel Online, http://www.spiegel.de/netzwelt/

netzpolitik/patienten-apotheken-verkaufen-vertrauliche-daten-a-917118.html, 18. August, 2013

[Spi13d] Spiegel Online: *Kriminelle erbeuten Daten von 2,9 Millionen Adobe-Kunden.* Spiegel Online, http://www.spiegel.de/netzwelt/web/adobe-kriminelle-erlangen-kreditkartendaten-und-quellcode-a-926039.html, 4. Oktober, 2013

[Spi14a] Spiegel Online: *Geheimdienste greifen Daten von App-Nutzern ab.* http://www.spiegel.de/netzwelt/netzpolitik/angry-birds-nsa-und-gchq-zapfen-apps-an-a-945872.html, 27. Januar, 2014

[Spi14b] Spiegel Online: *Der Spion in deinem Wohnzimmer.* http://http://www.spiegel.de/netzwelt/gadgets/labortest-online-fernseher-ueberwachen-nutzungsgewohnheiten-a-945488.html, 25. Januar, 2014

[Spi14c] Spiegel Online: *Sportbund will Olympia-App überarbeiten.* http://www.spiegel.de/netzwelt/apps/sportbund-will-olympia-app-ueberarbeiten-a-951804.html, 6. Februar, 2014

[Stö14a] Stöcker, Christian: *Nest-Übernahme — Google will in Ihr Schlafzimmer.* Spiegel Online, http://www.spiegel.de/netzwelt/gadgets/nest-uebernahme-google-will-in-ihr-schlafzimmer-a-943406.html, 14. Januar, 2014

[Stö14b] Stöcker, Christian: *"Wir wissen, wann Sie zu Hause sind".* Spiegel Online, http://www.spiegel.de/netzwelt/gadgets/interview-mit-nest-labs-gruender-tony-fadell-ueber-googlekauf-a-945763.html, 27. Januar, 2014

[Ste93] Steiner, Peter: *On the Internet, nobody knows you're a dog.* The New Yorker (siehe auch http://en.wikipedia.org/wiki/On_the_Internet,_nobody_knows_you're_a_dog), July, 5, 1993

[Ste10a] Stecklow, Steve: *On the Web, Children Face Intensive Tracking.* The Wall Street Journal, http://online.wsj.com/article/SB10001424052748703904304575497903523187146.html, September 17, 2010

[Ste10b] Steel, Emily: *A Web Pioneer Profiles Users by Name.* The Wall Street Journal, http://online.wsj.com/article/SB10001424052702304410504575560243259416072.html, October 24, 2010

[Sto08] Story, Louise: *To Aim Ads, Web is Keeping Closer Eye On You.* New York Times, http://www.nytimes.com/2008/03/10/technology/10privacy.html?pagewanted=all&_r=0, March 10, 2008

[Str13] Stresing, Laura: *Ein individueller Preis für jeden.* Der Tagesspiegel, http://www.tagesspiegel.de/medien/digitale-welt/verbraucherschutz-ein-individueller-preis-fuer-jeden/8353500.html, 18. Juni, 2013

[SV11] Siddiqui, Mohd. Shadab; Verma, Deepanker: Evercookies : Extremely persistent cookies. In: *International Journal of Computer Science and Information Security* 9 (2011), Mai, Nr. 5, S. 165–167

[Thu11] Thurm, Scott: *Next Frontier in Credit Scores: Predicting Personal Behavior.* The Wall Street Journal, http://online.wsj.com/news/articles/SB10001424052970203687504576655182086300912, October 27, 2011

[TK10] Thurm, Scott; Kane, Yukari Iwatani: *Your Apps Are Watching You.* The Wall Street Journal, http://online.wsj.com/news/articles/SB10001424052748704694004576020083703574602, December, 17, 2010

[TKH+09] Turow, Joseph; King, Jennifer; Hoofnagle, Chris Jay; Bleakley, Amy ; Hennessy, Michael: *Americans Reject Tailored Advertising and Three Activities that Enable It.* Social Science Research Network, SSRN Working Paper Series, http://ssrn.com/abstract=1478214, 2009

[TNB+10] Toubiana, Vincent; Narayanan, Arvind; Boneh, Dan; Nissenbaum, Helen ; Barocas, Solon: Adnostic: Privacy Preserving Targeted Advertising. In: *Proceedings of the Network and Distributed System Security Symposium (NDSS 2010)*, 2010

[Tri13] Trilli, Kevin: *HTML5... I mean, Local... I mean, DOM... I mean, Web Storage!* TRUSTe Technology Blog, http://www.truste.com/developer/?p=297, March, 26, 2013

[Tuc11] Tucker, Catherine: *Social Networks, Personalized Advertising, and Privacy Controls.* The Tenth Workshop on Economics of Information Security (WEIS 2011), http://weis2011.econinfosec.org/papers/Social%20Networks,%20Personalized%20Advertising,%20and%20Privacy%20Cont.pdf, Juni 2011

[UHL12] Urban, Jennifer M.; Hoofnagle, Chris Jay ; Li, Su: *Mobile Phones and Privacy.* BCLT Research Paper Series, UC Berkeley Public Law Research Paper No. 2103405, http://ssrn.com/abstract=2103405, July, 10, 2012

[Val10a] Valentino-DeVries, Jennifer: *'Evercookies' and 'Fingerprinting': Are Anti-Fraud Tools Good for Ads?* The Wall Street Journal, http://blogs.wsj.com/digits/2010/12/01/evercookies-and-fingerprinting-finding-fraudsters-tracking-consumers/, December 1, 2010

[Val10b] Valentino-DeVries, Jennifer: *How to Prevent Device Fingerprinting.* The Wall Street Journal, http://blogs.wsj.com/digits/2010/11/30/how-to-prevent-device-fingerprinting/, November 30, 2010

[Vas10] Vascellaro, Jessica E.: *Google Agonizes on Privacy as Ad World Vaults Ahead.* The Wall Street Journal, http://online.wsj.com/article/SB10001424052748703309704575413553851854026.html, August 9, 2010

[Veg10] Vega, Tanzina: *New Web Code Draws Concern Over Privacy Risks.* New York Times, http://www.nytimes.com/2010/10/11/business/media/11privacy.html?pagewanted=print, October 10, 2010

[VK10] Vega, Tanzina; Kopytoff, Verne: *In Online Privacy Plan, the Opt-Out Question Looms.* New York Times, http://www.nytimes.com/2010/12/06/business/media/06privacy.html?_r=0&adxnnl=1&adxnnlx=1379593290-YH9GrcxJ8sOYKWNPn72bsw&pagewanted=print, December 5, 2010

[Voß12] Voß, Oliver: *Wann Facebook-Kündigungen unzulässig sind.* Wirtschaftswoche, http://www.wiwo.de/erfolg/beruf/streit-um-internet-eintrag-wann-facebook-kuendigungen-unzulaessig-sind/6775736.html, 20. Juni, 2012

[Wai13] Waidner, Michael: Informationssicherheit. In: Haupter, Ralph (Hrsg.): *Der digitale Dämon.* Redline Verlag, 2013, S. 35–48

[War11] Ward, Sam: *Web tracking has become a privacy time bomb.* USA Today, August, 3, 2011

[WCJJ11] Weinberg, Zachary; Chen, Eric Y.; Jayaraman, Pavithra Ramesh ; Jackson, Collin: I Still Know What You Visited Last Summer: Leaking Browsing History via User Interaction and Side Channel Attacks. In: *2011 IEEE Symposium on Security and Privacy, Proceedings*, 2011, S. 147–161

[Wor13] WordPress: *Plugin Directory – Tag: Tracking.* http://wordpress.org/plugins/tags/tracking, 2013

[WP12] West, William; Pulimood, S. Monisha: Analysis of Privacy and Security in HTML5 Web Storage. In: *Journal of Computing Sciences in Colleges* 27 (2012), Januar, Nr. 3, S. 80–87

[Wya13] Wyatt, Nelson: *Former U.S. vice-president Al Gore predicts lawmakers will rein in surveillance.* The Vancouver Sun, http://www.vancouversun.com/story_print.

html?id=9129866, November, 7, 2013

[YLL+12] Yang, Yuhao; Lutes, Jonathan; Li, Fengjun; Luo, Bo ; Liu, Peng: Stalking Online: On User Privacy in Social Networks. In: *Proceedings of the Second ACM Conference on Data and Application Security and Privacy (CODASPY '12)*, 2012, S. 37–48

[YLW+09] Yan, Jun; Liu, Ning; Wang, Gang; Zhang, Wen; Jiang, Yun ; Chen, Zheng: How much can behavioral targeting help online advertising? In: *18th international conference on World wide web (WWW '09), Proceedings*, 2009, S. 261–270

[Zan11] Zaneis, Michael: *'Do Not Track' Rules Would Put a Stop to the Internet As We Know It.* US News, http://www.usnews.com/opinion/articles/2011/01/03/do-not-track-rules-would-put-a-stop-to-the-internet-as-we-know-it_print.html, January 3, 2011

[Zui13] Zuiderveen Borgesius, Frederik: Behavioral Targeting: A European Legal Perspective. In: *IEEE Security & Privacy* 11 (2013), Nr. 1, S. 82–85

# DANKSAGUNG

Dieser Report konnte dank der CASED-Förderung des Hessischen Ministeriums für Wissenschaft und Kunst (HMWK), der Förderung des Bundesministeriums für Bildung und Forschung (BMBF) für EC SPRIDE und durch eine finanzielle Förderung von Microsoft entstehen. Die Autoren danken den Fördermittelgebern für ihre Unterstützung.

Besonderer Dank geht an Michael Waidner für viele hilfreiche Kommentare und Hinweise. Wir möchten uns ebenfalls bedanken bei der Forschungsabteilung *Test Lab Mobile Security* des Fraunhofer SIT, die uns für den Web-Tracking-Report die mit Appicaptor gewonnenen Auswertungen bzgl. App-Tracking zur Verfügung gestellt hat. Darüber hinaus danken wir Mona Bien für die Gestaltung des Umschlags.

# ANHANG A
# DAS FRAUNHOFER SIT

Das Fraunhofer-Institut für Sichere Informationstechnologie (Fraunhofer SIT) ist ein Institut der Fraunhofer-Gesellschaft (Fraunhofer). Fraunhofer ist die größte Forschungsorganisation für anwendungsorientierte Forschung in Europa. Die Forschungsfelder von Fraunhofer richten sich nach den Bedürfnissen der Menschen: Gesundheit, Sicherheit, Kommunikation, Mobilität, Energie und Umwelt. Deswegen hat die Arbeit von Fraunhofer großen Einfluss auf das zukünftige Leben der Menschen. Bei Fraunhofer arbeiten rund 23.000 Mitarbeiterinnen und Mitarbeiter in derzeit 67 Instituten und Forschungseinrichtungen.

Das Fraunhofer SIT ist eines dieser Institute; es hat seinen Standort in Darmstadt (`http://www.sit.fraunhofer.de`). Das Spezialgebiet des Fraunhofer SIT ist die Cyber-Sicherheit und der Schutz von Daten und Systemen. Das Fraunhofer SIT beschäftigt über 160 Mitarbeiter. Fraunhofer SIT betreibt mit seinen beiden Partnern, der Technischen Universität Darmstadt und der Hochschule Darmstadt, das *Center for Advanced Security Research Darmstadt*, kurz *CASED* (`http://www.cased.de`). CASED ist das größte Zentrum für IT-Sicherheitsforschung in Deutschland und gehört auch in Europa zu den größten Zentren für IT-Sicherheitsforschung. In CASED arbeiten rund 300 Wissenschaftler zusammen an Fragen und Problemen der IT-Sicherheit.

Damit sich Deutschland den großen Zukunftsfragen der Cyber-Sicherheit langfristig stellen kann, hat das Bundesministerium für Bildung und Forschung (BMBF) mit dem *European Center for Security and Privacy by Design*, kurz *EC SPRIDE* (siehe `http://www.ec-spride.de`), ein von den CASED-Partnern Technische Universität Darmstadt und Fraunhofer SIT getragenes Kompetenzzentrum, ausgewählt. EC SPRIDE ist das größte vom BMBF geförderte Kompetenzzentrum zum Thema Cyber-Sicherheit.

Das Fraunhofer SIT beschäftigt sich neben vielen anderen Fragen der IT-Sicherheit mit dem Schutz vor Web-Tracking. In diesem Zusammenhang hat Fraunhofer SIT eine Trackingschutzliste entwickelt, die Fraunhofer SIT in regelmäßigen Abständen aktualisiert. Diese Trackingschutzliste wurde hauptsächlich für den deutschen Sprachraum entwickelt, d.h. Fraunhofer SIT untersucht insbesondere die in Deutschland populärsten Internetangebote auf das Enthaltensein von Trackern. Die Trackingschutzliste wird im Internet zum kostenlosen Download zur Verfügung gestellt (siehe unter `https://www.sit.fraunhofer.de/tpl`).

Für Unternehmen, die sich oder die eigenen Mitarbeiter gegen Web-Tracking schützen möchten, erstellt Fraunhofer SIT auch Trackingschutzlisten, welche auf die für das jeweilige Unternehmen besonders relevanten Anbieter von Webseiten zugeschnitten sind. Damit können bisher von Unternehmen nicht wahrgenommene Informationslecks geschlossen werden. Warum sollten unbekannte Dritte auch wissen, zu welchen Themen in einem Unternehmen Informationen benötigt werden?

ISBN 978-3-8396-0700-8