

Tagung für Informatik und Recht – Eröffnungsreferat

Dr. Martin Dumermuth, Direktor des Bundesamtes für Justiz

Chancen und Risiken von Big Data

1 Anrede und Vorbemerkungen

Sehr geehrte Damen und Herren

Sehr geehrte Referierende

Herzlich willkommen!

Ich freue mich, Sie im Namen der vier Träger zur 12. Tagung für Informatik und Recht hier im Berner Rathaus begrüßen zu dürfen.

Die Tagung wird dieses Jahr getragen durch:

- die Berner Fachhochschule;
- die Universität Wien;
- der Verein „eJustice.CH“; und
- das Bundesamt für Justiz.

Vorab möchte ich danken:

- Zunächst der Regierung des Kantons Bern für die geleistete Unterstützung und das zur Verfügung stellen des Berner Rathauses.
- Zweitens danke ich den Sponsoren, welche die heutige Tagung überhaupt möglich machen:

Es sind dies:

- Fabasoft Schweiz AG, Bern
- Die Schweizerische Post
- Swisslex AG, Zürich
- Weblaw AG, Bern

Sie haben natürlich – geübt wie Sie sind im Erkennen von Gesetzmässigkeiten – die alphabetische Reihenfolge der Aufzählung sofort erkannt.

- Und drittens danke ich den Referierenden, die heute aktiv den wesentlichen Teil zum Erfolg dieser Tagung beitragen.
- Schliesslich möchte ich Ihnen allen, die sie hierher gekommen sind, für Ihr Interesse an diesem Thema danken.

«Big Data Governance»,

so lautet das Thema dieser Tagung, mit der wir die juristische und politische Diskussion in der Schweiz weiter anstossen möchten.

2 Einleitung: Drei Formen von Big Data

In einer ersten Annäherung können wir Big Data wie folgt umschreiben:

- Es geht darum, aus grossen Datenmengen geschäftskritische Informationen abzuleiten. Das heisst, Informationen, die in den Daten *implizit* vorhanden sind, können *explizit* gemacht werden.
- Typischerweise geht es dabei um sehr grosse *Datenmengen*, die aus unterschiedlichen Quellen zusammengeführt, in grosser Geschwindigkeit produziert und eventuell auch ebenso schnell verarbeitet werden.
- Dabei sind die Daten oft fehlerhaft, ihre tatsächliche Relevanz für das Geschäftsproblem ist unklar und sie sind oft auch nur für beschränkte Zeit gültig.

Man spricht deshalb im Englischen auch von den drei V: Volume, Variety und Velocity.

Bei aller Vielfalt kann man drei Kernformen von Big Data identifizieren:

- klassisches Big Data;
- Freestyle Big Data oder exploratives Big Data; und
- das sogenannte Big Data auf dem Graphen.

Beim **Klassischen Big Data** wird aus grossen Datenmengen abgeleitet, wie Eigenschaften respektive Attribute zusammenhängen, und diese Erkenntnis wird für das Einschätzen geschäftskritischer Eigenschaften genutzt.

Beim **Freestyle oder explorativen Big Data** ist wesentlich, dass man Zusammenhänge in Daten erkennt, von deren Existenz man zu vor nichts wusste.

Und das sogenannte „**Big Data auf dem Graphen**“ ermöglicht es schliesslich, die Zusammenhänge zwischen sehr unterschiedlichen Datensets auszuwerten und daraus Erkenntnisse abzuleiten.

Daneben gibt es noch viele andere Formen: Beispielsweise kann die menschliche Fähigkeit zur Mustererkennung dadurch unterstützt werden, dass ein Computer Daten visualisiert. Oder Daten können für die Simulation komplexer Systeme genutzt werden.

Lassen Sie mich nach dieser ersten Annäherung an das Thema nun die drei Kernformen etwas näher betrachten:

3 Standardnutzung von Big Data

3.1 Klassisches Big Data

3.1.1 Schwangerschaft

Beginnen wir mit dem klassischen Big Data.

Zunächst ein Beispiel:

Eine Supermarktkette möchte wissen, ob eine Kundin schwanger ist, um ihr entsprechende Sonderangebote zu machen, z.B. über eine App. Die Supermarktkette besitzt einerseits Daten darüber, wie sich das Einkaufsverhalten von Frauen ändert, wenn sie schwanger werden. Und sie hat über Kundenkarten Kenntnis des Einkaufsverhaltens ihrer Kundinnen.

- Ausgangspunkt ist also eine geschäftskritische Frage nach einer Eigenschaft (hier die Schwangerschaft);
- es bestehen ferner empirische Daten darüber, wie diese Eigenschaft mit anderen Eigenschaften korreliert (hier: Einkaufsverhalten); und
- schliesslich existieren Daten zu weiteren Eigenschaften (hier: das konkrete Einkaufsverhalten der Kundinnen)

Erkennt nun die Supermarktkette in ihren Daten eine Veränderung des Einkaufsverhaltens analog zum typischen Veränderungsmuster bei auftretender Schwangerschaft, dann kann sie daraus den Schluss ziehen, dass die betroffene Kundin wahrscheinlich schwanger ist.

3.1.2 Wechselwähler bei Obama

Ein anderes berühmtes Beispiel liefert Obamas zweiter Präsidentschaftswahlkampf.

Dem Wahlkampfteam war rasch klar, dass sich alle Anstrengungen auf die Wechselwähler in den umkämpften Staaten konzentrieren mussten. Die zentrale Frage lautete also: Wie identifiziert man Wechselwähler in den sogenannten Swing States?

Obamas Leute machten zuerst Telefonumfragen bei Wählern, über die grosse Datenmengen vorhanden waren – darunter Einkaufsdaten. Anschliessend wurde ermittelt, welche Datenkonfigurationen auf Wechselwähler hinwiesen.

Schliesslich wurde nach diesen Datenkonfigurationen gesucht, um potentielle Wechselwähler zu identifizieren und zu mobilisieren.

3.2 Freestyle Big Data

3.2.1 Definition

Anders ist die Ausgangslage beim sogenannten Freestyle Big Data oder dem explorativen Big Data. Hier geht es um eine Art des professionellen Fischens im Trüben. Man weiss nicht wonach man sucht, hält aber nach auffälligen Mustern Ausschau, die man anschliessend inhaltlich interpretiert. Daraus können interessante Schlüsse gezogen werden.

3.2.2 Zeitschrift

Man weiss beispielsweise aus Big Data Untersuchungen in einem unserer Nachbarländer, dass Angestellte in der ersten Arbeitsstunde oft nach Angeboten von Gratiszeitschriften surfen, statt zu arbeiten. Diese Erkenntnis ist gut für Gratiszeitschriften, weil die Publikumskontakte der Werbung verkauft werden können. Aus Sicht der Arbeitgeber liefert sie dagegen einen Hinweis auf ein unerwünschtes Phänomen.

3.2.3 Suchmaschine

In anderen Fällen ist das Erkennen von Mustern oder Anomalien schwieriger. Wenn Sie z.B. im Intranet eine Suchmaschine anbieten und in der Logdatei dieser Suchmaschine sehen, dass Abfragesessions in der Regel nur aus einer Abfrage bestehen, dann bedeutet das an sich noch nichts. Wenn sich aber gleichzeitig viele Mitarbeitende über das «ewige Nichtfinden im Intranet» beschweren, dann wird die Logdatei zu einer Anomalie. Denn «normal» wäre, dass Mitarbeitende *mehrmals* mit Synonymen versuchen würden, das Gesuchte zu finden. Es lohnt sich also, der Sache auf den Grund zu gehen und zu fragen, warum die Mitarbeitenden die Suche nach dem ersten Scheitern sofort aufgeben. Solche *neuen Fragen* sind das typische Ergebnis von Freestyle Big Data. Muster deuten auf kritische Phänomene hin, die man dann aber genau analysieren muss.

3.3 Big Data auf dem Graphen

Das sogenannte „Big Data auf dem Graphen“ schliesslich nutzt sogenannte Linked Data. Dabei werden Metadaten verwendet, um die Beziehungen zwischen Daten zu beschreiben. Mit Daten und Metadaten lässt sich dann ein Graph konstruieren, dessen Knoten die *Daten* repräsentieren und dessen Kanten *Beziehungen* zwischen Daten darstellen. So entsteht ein Wissensgraph, der die Wissensobjekte inhaltlich zueinander in Bezug setzt.

Auf diesem Graphen kann man nach Informationen suchen, die etwa mit herkömmlichen Suchmaschinen nicht gefunden werden. So findet man beispielsweise Informationen, die auf viele Texte verteilt sind, oder sogenannte "Informationslöcher", d.h. Bereiche, wo Informationen fehlen.

Mit „Big Data auf dem Graphen“ zeigen sich etwa auch Innovationsbereiche, die nicht durch Patente abgedeckt sind.

Darüber hinaus erlaubt die Linked Data Technologie die Integration von inhaltlich *ähnlichen Daten*, die aber nach *abweichenden Standards* erstellt worden sind. So kann man z.B. unterschiedlich dokumentierte Krankheitsgeschichten für die medizinische Forschung zusammenführen. Dieses Zusammenführen des Ähnlichen ist insbesondere dann entscheidend, wenn sich Datenstandards ändern.

4 Nutzungsbereiche

4.1 Marketing

Im Geschäftsleben findet Big Data vor allem im Marketing Anwendung. Seit vielen Jahren schon sorgen etwa im Private Banking Customer Relationship Management Systeme – kurz CRM-Systeme – dafür, dass sich neue Produktangebote nur an Kunden richten, die sich mit hoher Wahrscheinlichkeit auch dafür interessieren.

Aber auch beim Design von attraktiven Promotionspaketen – z.B. von Zusatzausstattungen in der Autoindustrie – findet Big Data Anwendung. Das Gleiche gilt bei der Weiterentwicklung von Produkten und Diensten, für die man aus den Kundendaten ein Bedürfnis abgeleitet hat. So kann man beispielsweise attraktive Informationsangebote oder spezifische Ausbildungslehrgänge zusammenstellen.

4.2 Politik

Auch in der **Politik** wird die Bedeutung von Big Data künftig zunehmen. Es geht nicht nur um die Personalisierung des Wahlkampfes, wie uns das Beispiel aus den USA gezeigt hat. Big Data kann auch etwa dazu verwendet werden, das eigene Politverhalten – etwa bei Abstimmungen im Parlament – optimal auf die Steigerung der Wiederwahlchancen auszurichten. Ob dabei auch die Qualität der Sachpolitik zunimmt, ist allerdings eine andere Frage.

Es lassen sich aber auch Gedankenspiele entwickeln, die viel revolutionärere Auswirkungen haben. So könnte man es beispielsweise Computern überlassen, repräsentative Entscheidungsgruppen, die Rückhalt bei den Stakeholdern haben, so zusammenzustellen, dass ein schneller Kompromiss wahrscheinlich ist.

4.3 Wissenschaft

In der **Wissenschaft** kommt Big Data bereits heute intensiv zur Anwendung und gilt in Zukunft wohl bald als unverzichtbar. Wir werden erleben, dass viel häufiger als bisher in Daten Zusammenhänge erkannt und Modelle abgeleitet werden, an die früher nie jemand gedacht hat. Damit wird Big Data zu einem mächtigen Forschungsinstrument, das aber mit grosser Vorsicht genutzt werden sollte, da ein möglicher Daten-Bias – ich komme darauf zurück – zu einem völlig falschen Modell führen kann.

Vor allem für die Hypothesenbildung schafft Big Data aber sehr nützliche Möglichkeiten.

Vorstellbar ist aber auch, dass in Zukunft Proposals für Forschungsprojekte eine Zusatzauswertung durch Computer erfahren, die aus der Projektskizze die Impact-Chancen im Fall einer Realisierung schätzen. Big Data kann so auch zur Wissenschaftssteuerung eingesetzt werden.

4.4 Medizin

Einer der wichtigsten Anwendungsbereiche wird in Zukunft die **Medizin** sein. Hier geht der Trend Richtung Personalisierung. Diagnosen werden zwar auch in Zukunft nach standardisierten Verfahren erstellt, die Therapie wird aber individualisiert und auf der Basis von Analogieschlüssen festgelegt. Dabei werden insbesondere die genetischen Daten und die Lebensumstände der Patientinnen und Patienten berücksichtigt.

Noch weiter geht die Vision, dass Gesundheitsdaten individualisiert *interpretiert* werden. Beispielsweise würde Bluthochdruck nicht nach Alterstabellen diagnostiziert, sondern es müsste *individuell* geprüft werden, ob der Blutdruck für den konkreten Patienten zu hoch ist und seine Gesundheit tatsächlich gefährdet. Der Schwellenwert würde also je nach individueller Disposition festgelegt, die sich aus Genom, den Lebensumständen und der bisheriger Krankheitsgeschichte ergäbe. Möglich macht so eine personalisierte Medizin der Wechsel vom Denken in *Kausalitäten* zum Denken in *Korrelationen*. Das ist ein grosser Paradigmenwechsel, mit Chancen und Risiken.

5 Praktische Herausforderungen

5.1 Einführung

So wunderbar der potentielle Nutzen von Big Data ist, so gross sind auch die **praktischen Herausforderungen**. Beim klassischen Big Data liegt die Crux beispielsweise darin, die geschäftlich interessierende Frage klar zu formulieren. Wir müssen genau wissen, was wir suchen. Dann brauchen wir ein mathematisches Modell, dessen Eigenschaften wir gut verstehen, damit wir einen passenden Algorithmus für die Datenverarbeitung definieren können, der verlässliche Resultate liefert.

Die praktischen Herausforderungen sind gross, erscheinen aber bewältigbar. Doch genau diese Machbarkeit von Big Data wirft grundsätzliche Probleme auf.

5.2 Zusammenführen unterschiedlicher Datenquellen

Zunächst wird es möglich, im Sinne eines Mix-and-Match ganz unterschiedliche Datenquellen zusammenzuführen, deren Datenmodelle überhaupt nicht zueinander passen. Das ist eine riesige Chance und überwindet viele Probleme der Vergangenheit, beispielsweise in der medizinischen Forschung. Aber sehr leicht kann dieses «Zusammenmischen» auch zu grobem Unfug führen.

5.3 Korrelationen statt Kausalitäten

Zum zweiten ist das Denken in Korrelationen eine grosse Herausforderung für uns Menschen, die nicht dadurch reduziert wird, dass der Computer uns die Rechenarbeit abnimmt. Nur weil etwas korreliert, muss es noch keinen tieferen Zusammenhang geben. Und selbst wenn es einen solchen gibt, ist die Interpretation einer Korrelation als kausaler Zusammenhang zwar auf den ersten Blick oft naheliegend aber gleichzeitig hochproblematisch.

5.4 Personalisierung

Zum dritten ermöglicht Big Data einen Grad an Personalisierung, den wir bisher nicht gekannt haben. Nicht nur die Produktwerbung sondern auch der Produktpreis können personalisiert werden und der Traum der Ökonomie von der totalen Preisdifferenzierung kann Wirklichkeit werden. Auch wenn es in Europa erst zaghafte Versuche gibt, wird dieser Trend aus den USA auch zu uns kommen.

Sie können dieses Phänomen selber praktisch testen. Suchen Sie im Internet ein Hotelzimmer. Dann löschen Sie Ihren Cache und wiederholen die Suche. Im Ergebnis wird der angebotene Preis nicht derselbe sein.

Standardisierte Preise stellten in der Vergangenheit in gewissen Bereichen einen sozialen Kontrakt dar, der durch Preisdifferenzierung verloren gehen wird. So wird beispielsweise der Solidaritätsgedanke im Versicherungswesen auf eine harte Probe gestellt, wenn aus den persönlichen Daten die Wahrscheinlichkeit von Versicherungsfällen vorhersagt werden kann, und die Preise differenziert werden.

5.5 Prognose

Zum vierten kann die Fähigkeit, Implizites explizit zu machen, auch unsere Flexibilität in der Zukunft einschränken.

Big Data suggeriert, man könne aus den Daten der Vergangenheit und Gegenwart die Zukunft voraussagen. Eine bessere Prognosefähigkeit mag zwar von grossem Nutzen sein. Sie wird aber dann problematisch, wenn wir die nötige Flexibilität verlieren, um auf unerwartete Szenarien sachgerecht reagieren zu können, die trotz geringer Wahrscheinlichkeit eintreten.

6 Governance-Probleme

6.1 Sechs Governance Felder

Die geschilderten Probleme werfen verschiedene kritische Governance-Fragen auf. Ich möchte in der Folge sechs Governance-Felder skizzieren.

6.2 Datenqualität

Das erste hängt mit der schwer zu beherrschenden Bias-Problematik und der Frage der Datenqualität zusammen. Die Ableitung von Modellen aus Daten ist mit Gefahren verbunden und darf nicht übereilt geschehen. Die Aussagekraft des resultierenden Modells muss mit Plausibilitätschecks überprüft werden, z.B. mit verwandten Daten aus einer anderen Quelle, oder mit zusätzlichen Datenerhebungen, die auf die Überprüfung des Modells ausgerichtet sind.

Die Beurteilung der Datenqualität ist schwierig. Dies gilt namentlich, wenn es sich um zusammengefügte oder bereits verarbeitete Daten handelt. Unterschiedliche Modelle können bei gleichen Daten zu konträren Resultaten führen, und ohne tiefes Modellverständnis ist die Beurteilung der Validität von Ergebnissen schwierig.

Umgekehrt können unterschiedlich selektierte Daten in Anwendung desselben Modells ebenfalls zu sehr unterschiedlichen Ergebnissen führen. Ferner ist es völlig unklar, wie man mit Ergebnissen umgehen, die sich auf kein Erklärungsmodell stützen, sondern aus undurchschaubaren Berechnungen entstehen.

Stellen Sie sich vor, eine Computerauswertung von Kameras im öffentlichen Raum führt zum Schluss, dass dramatische öffentliche Unruhen kurz bevor stehen und Wasserwerfer eingesetzt werden sollen. Niemand weiss aber, worum es geht und warum der Computer das empfiehlt.

Wollen wir wirklich diesen Computerempfehlungen folgen, auch wenn es sich nur um das zufällige Zusammentreffen mehrerer Polterabendgruppen handelt?

Zu diesem grossen Problemfeld braucht es in der Praxis klare Governance-Richtlinien. In erster Linie zum Nutzen des Unternehmens oder der Verwaltungsstelle, die auf Big Data setzt.

In manchen Bereichen ist eine solche Governance aber auch rechtlich relevant. Denken wir etwa an den Fall, dass Mitarbeiterevaluationen auf Big Data basieren und es aufgrund fehlerhafter Daten und Modelle zu systematischer Unfairness kommt.

6.3 Fachkompetenz

Der zweite grosse Problembereich ist die **benötigte Fachkompetenz**. Big Data benötigt sehr unterschiedliche Fachkompetenzen: Mathematik und Statistik, angewandte Modellierung, Informatik, bereichsspezifisches Fachwissen, Marktverständnis und natürlich Rechtsexpertise.

Diese Kompetenzen in einer Organisation intern verfügbar zu haben ist keine Selbstverständlichkeit. Zudem tun sich fachlich heterogene Teams häufig schwer mit der Zusammenarbeit. Und schliesslich ist es für eine Geschäftsleitung regelmässig eine grosse Herausforderung, die Qualität der Big Data-Ergebnisse sinnvoll einzuordnen.

Was bislang fehlt sind multidisziplinäre Schulungsprogramme, insbesondere adäquate Governance-Ausbildungen für die Geschäftsleitungsebene. Hier sind Fachhochschulen und Universitäten gefordert. Vor allem für Mitglieder von Geschäftsleitungen braucht es angemessene Weiterbildungsangebote.

6.4 Interpretation

Ein weiteres grosses Problem im Umgang mit Daten ist das **Verständnis ihrer Bedeutung**. Solange es um die eigene Datenbank geht, ist dieses Problem in der Regel überblickbar. Wenn aber zusätzlich auch Daten aus fremden Quellen benutzt werden, nehmen die Schwierigkeiten zu.

In den letzten Jahren wurde beispielsweise Open Data zu einem grossen Trend. Dabei geht es vor allem darum, dass die öffentliche Verwaltung ausgewählte Daten zur Sekundärnutzung durch Dritte in maschinenlesbarer Form zur Verfügung stellt. So sind beispielsweise viele nützliche Apps für Besucher und Bewohnerinnen von Städten entstanden, mit denen man z.B. Spielplätze oder Behindertenparkplätze finden kann.

Zunehmend wird aber gefordert – und immer mehr auch umgesetzt – dass politisch relevante Daten veröffentlicht werden und so Transparenz über das Verwaltungshandeln entsteht.

Ferner wird etwa postuliert, Firmen sollten ihre Daten veröffentlichen um Transparenz über ihr Verhalten zu schaffen. So würden sie für Kundinnen und Kunden sowie Arbeitnehmende attraktiv, wenn sie beispielsweise eine gute Nachhaltigkeitsbilanz veröffentlichen würden.

Das sind durchaus vielversprechende Trends, die aber auch Schattenseiten haben. Das Veröffentlichen von Rohdaten oder aufbereiteten Daten kann nur dann Transparenz schaffen, wenn sie auch verstanden werden. Dieses Verstehen ist leider oft Fiktion. In diesem Zusammenhang ist deshalb auch von **Transparenzfiktion** die Rede.

Was es braucht, sind **Beipackzettel für Datensätze**, die ihre Semantik und ihre Qualität erläutern. Sonst kann durch Pseudotransparenz die öffentliche Meinung manipuliert werden. Das ist vor allem bei gesellschaftlich oder politisch sensiblen Open Government Data wichtig.

6.5 De-Anonymisierungsrisiko

Ein weiteres grosses Governance Feld betrifft das **De-Anonymisierungsrisiko**.

Wenn unterschiedliche Daten über die gleiche anonyme Person aus verschiedenen Lebensbereichen verknüpft werden, kann das zur Identifizierung der betroffenen Person, d.h. zur De-Anonymisierung der Daten führen. Auch ohne dass Daten neu hinzukommen, kann allein schon der Technologiefortschritt im Bereich der semantischen

Technologien zur De-Anonymisierung von Daten führen. Problematisch ist vor allem, dass die Beständigkeit der Anonymität in vielen Fällen nicht vorhergesagt respektive garantiert werden kann. Kritisch ist hier auch die mögliche Haftung. Prof. Rolf Weber von der Universität Zürich wird ausführlich auf dieses Thema eingehen.

6.6 Gesellschaftliche Auswirkungen

Ein weiteres, politisch brisantes Governance-Problemfeld sind die **gesellschaftlichen Auswirkungen von Big Data**. Sie beginnen erst allmählich spürbar zu werden und werden noch kaum verstanden. Wir können beobachten, dass Internet-Dienstleister und Staaten zu mächtigen Datenherren werden. Das Verhältnis zwischen (für sich unbedeutenden) Datenlieferanten und (mächtigen) Datenherren ist dabei häufig ein sehr ungleiches.

So kann etwa Google mit meinen Daten Geld verdienen, weil es Dritten personalisierte Werbung an mich ermöglicht. Diesen Problemkreis wird Prof. Hafen von der ETH Zürich ausführlich behandeln und dabei auch auf verwandte Themen eingehen.

Auf einen Aspekt möchte ich aber noch selber besonders hinweisen. Eine transparente Gesellschaft hat grundsätzlich einen anderen Charakter als eine Gesellschaft mit geschützter Privatsphäre: Gewisse Formen der unfairen Machtausübung werden unmöglich, dafür tauchen neue Formen auf. Wie sich diese Veränderungen auf unser Zusammenleben, die menschliche Kommunikation und unsere demokratische Kultur auswirken werden, ist noch kaum erkennbar.

Hier ist eine politische Debatte dringend nötig!

6.7 Automatisierungsproblematik

Als letztes Problemfeld möchte ich die **Automatisierungsproblematik** ansprechen. Der Einsatz von Big Data steht sowohl bei Sachentscheiden der Verwaltung als auch bei Gerichtsverfahren vor der Tür, sofern er nicht schon Einzug gehalten hat. Schon jetzt verwenden etwa Versicherungen Stimmanalysatoren, die Schadensmeldungen per Telefon automatisch danach beurteilen, ob es sich um eine betrügerische Meldung handelt. Oder Steuerämter schätzen bei Steuerklärungen automatisiert die Wahrscheinlichkeit grösserer Inkorrektheiten ab.

Auf den ersten Blick erscheint es deshalb verlockend, beispielsweise bei Entscheidungen über den rechtmässigen Anspruch auf staatliche Unterstützungsleistungen die gesammelten Kontextdaten über den Antragsteller zur Fallbeurteilung heranzuziehen – so etwa im Sozialwesen.

Auch die Computer-Forensik wird in Zukunft ein Thema für das Gerichtswesen werden. Durch richtige Anwendung von Big Data lassen sich die Wahrscheinlichkeiten für entscheidungsrelevante Zusammenhänge sehr gut schätzen.

In welchem Ausmass wir solche Möglichkeiten auch in der Schweiz einsetzen wollen, braucht aber eine ausführliche Diskussion. Wahrscheinlichkeiten mögen zwar hilfreich sein, sie schaffen aber eben keine völlige Sicherheit. Wir dürfen den Blick für Individualgerechtigkeit nicht verstellen lassen durch Daten, die uns Zusammenhänge offen legen – oder aber im Einzelfall eben vorgaukeln.

Hier abzuwarten ist aber eine schlechte Strategie, weil wir nicht wissen, wie schnell die Einsatzmöglichkeiten von Big Data bei Gericht Einzug halten. Big Data darf sich nicht quasi naturwüchsig entwickeln. Sind entsprechende Möglichkeiten einmal gegeben, werden sie über die Rechtsanwältinnen und -anwälte ins Gerichtswesen eingebracht und die Diskussion wird sich durch die langsame Geschwindigkeit des Rechtssetzungsprozesses kaum mehr bremsen lassen. Wir sollten deshalb auch die Diskussion über Computer-Forensik möglichst bald beginnen.

7 Schluss

Ich komme zum Schluss:

Big Data hat ein sehr grosses Nutzenpotential in vielen Bereichen. Wir sind erst am Anfang, dieses Potential zu erschliessen.

Big Data stellt uns aber auch vor grosse praktische Herausforderungen, weil es multidisziplinäre Kompetenzen und neue Zusammenarbeitsformen verlangt.

Damit die Bundesverwaltung das Potential erfolgreich nutzen kann, sollte sie in die Ausbildung investieren und eine Datenstrategie entwickeln. Zusätzlich sind Governance Leitlinien für den Einsatz von Big Data zu formulieren. Ähnliches gilt auch für die Nutzer und Nutzerinnen in der Privatwirtschaft.

Aus rechtlicher Sicht sticht namentlich das De-Anonymisierungsrisiko ins Auge. Die ursprüngliche Abgrenzung zwischen Personendaten und anderen Daten löst sich immer mehr auf. Deshalb braucht es praktische Handlungsanleitungen, wie man damit umgehen soll – vielleicht auch neue gesetzliche Regelungen.

Diese Fragen werden auch im Rahmen der laufenden Überprüfung des Datenschutzgesetzes diskutiert. Natürlich sind dabei auch Anpassungen an die heutigen technischen Möglichkeiten ein Thema. Wir werden dem Bundesrat in nächster Zeit entsprechende Vorschläge unterbreiten.

Die grossen Chancen und die nicht zu unterschätzenden Gefahren machen in der Schweiz einen breiten gesellschaftlichen Dialog über Governance-Fragen von Big Data notwendig. Wir wollen diesen Dialog mit dieser Veranstaltung anstossen und es ist uns wichtig, dass die rechtliche Perspektive dabei eine zentrale Rolle spielt!

Lassen Sie uns das anpacken und den Diskurs über Big Data Governance beginnen!
Herzlichen Dank für Ihr Zuhören!