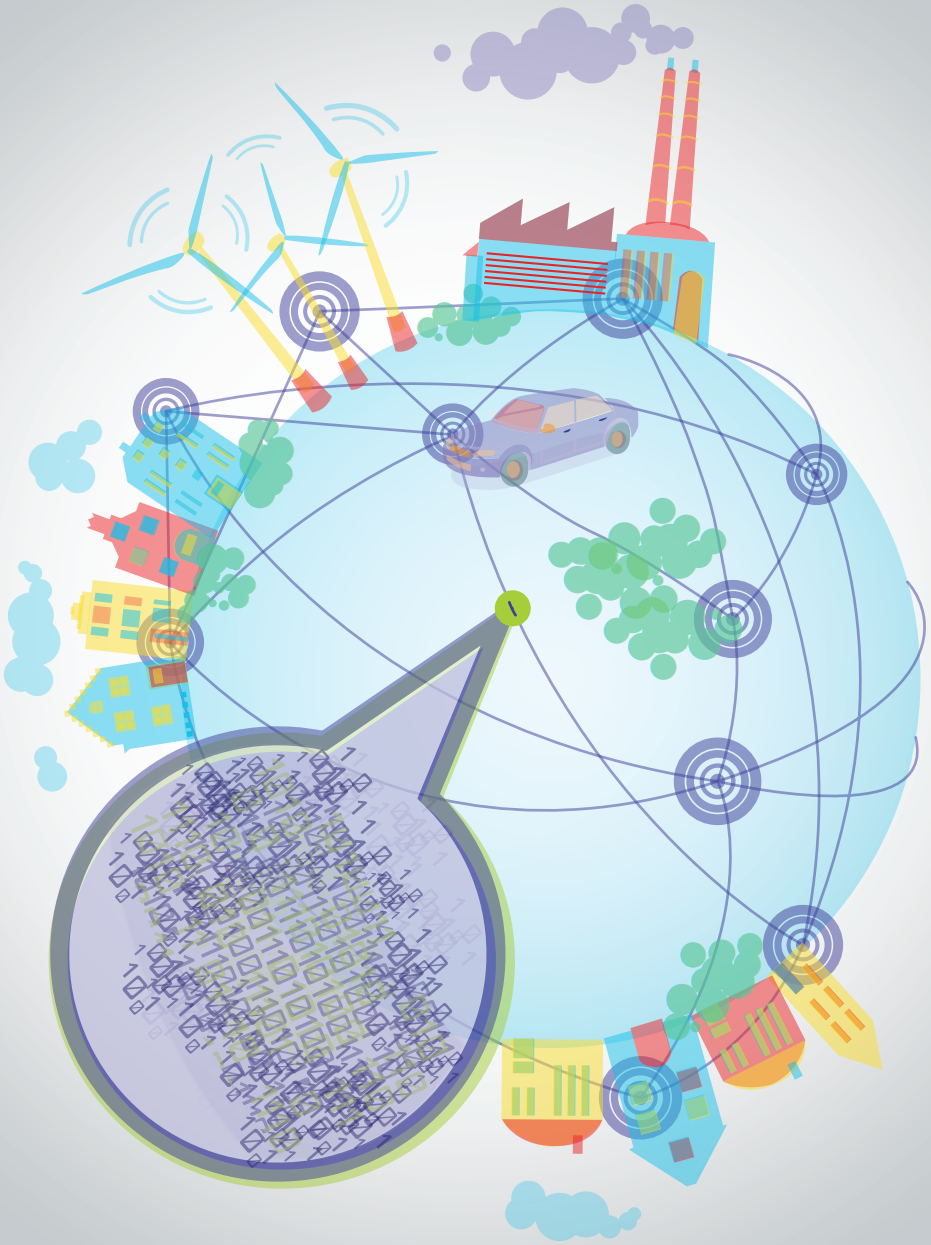
Fraunhofer
IAIS

FRAUNHOFER-INSTITUT FÜR INTELLIGENTE ANALYSE- UND INFORMATIONSSYSTEME IAIS

BIG DATA – VORSPRUNG DURCH WISSEN
INNOVATIONSPOTENZIALANALYSE



INHALTSVERZEICHNIS

EINLEITUNG ... 6
1. »Big Data – Big Opportunities« ... 6

EXECUTIVE SUMMARY ... 7
2. Die Innovationspotenzialanalyse ... 7
2.1 Ergebnisse der Recherche nach Anwendungsfällen ... 7
2.1.1 Big-Data-Anwendungen tragen zu vielen Zielen bei ... 7
2.1.2 Es liegen vielfältige Aufgabengebiete für Big-Data-Anwendungen vor ... 8
2.2 Ergebnisse der Online-Befragung ... 8
2.2.1 Potenzial von Big Data ist KMU bekannt ... 8
2.2.2 Steigerung der Wettbewerbsfähigkeit als wesentliches Ziel ... 8
2.2.3 Treiber ist das Management, nicht die IT ... 9
2.2.4 Interne Ressourcen sind knapp und werden ausgebaut ... 9
2.3 Ergebnisse aus den Branchenworkshops ... 9
2.3.1 Finanzwirtschaft: Effizienz nach innen, Individualisierung nach außen ... 9
2.3.2 Telekommunikation und Medien: Braucht es noch Marktforschung? ... 9
2.3.3 Handel: Der Traum vom IT-gestützten Tante-Emma-Laden ... 10
2.3.4 Marktforschung: Markt für Marktforschung wird größer ... 10
2.3.5 Versicherungen: Der Trend zu Versicherungen in HD ... 10
2.4 Effizienz steigern durch Big Data ... 10

RECHERCHE NACH BIG DATA ANWENDUNGSFÄLLEN ... 13
3.1 Kategorien ... 13
3.1.1 Branche und Unternehmensbereich ... 13
3.1.2 Ziele ... 15
3.1.3 Aktualität und Reaktionsweise ... 16
3.1.4 Datenquellen ... 20
3.1.5 Aufgabe ... 22
3.1.6 Analyse ... 24
3.2 Gruppen ähnlicher Anwendungsfälle ... 26
3.2.1 Personalisierte Ansprache von Kaufinteressenten ... 27
3.2.2 Personalisierte Ansprache von (potenziellen) Mitarbeitern ... 29
3.2.3 Marktmonitoring für Verkaufschancen ... 30
3.2.4 Optimierung des Betriebs ... 31
3.2.5 Finanzielle Risiken und Betrug ... 34
3.2.6 Erkennung von Attacken ... 36
3.2.7 Innovative oder verbesserte Produkte und Dienstleistungen ... 37
3.2.8 Zusammenfassung ... 40

ONLINE-UMFRAGE ZUR NUTZUNG VON BIG DATA ... 41
4.1 Methodik ... 41
4.1.1 Hypothesen ... 41
4.1.2 Zielgruppe der Befragung ... 41
4.1.3 Grundsätzliches zum Studiendesign ... 41
4.1.4 Rekrutierung ... 41
4.1.5 Repräsentativitätsproblematik ... 41
4.1.6 Ziel ... 41
4.1.7 Statistische Angaben der Teilnehmer aus der Online-Umfrage ... 42
4.1.8 Bekanntheit und Verbreitung des Begriffs Big Data ... 44
4.1.9 Bedeutung einzelner Datenarten für den Unternehmenserfolg ... 44
4.1.10 Funktionsbereiche, die Daten mit hoher Veränderlichkeit auswerten ... 45
4.1.11 Big-Data-Anwendungsfälle ... 46
4.1.12 Ziele von Big-Data-Anwendungen ... 47
4.1.13 Barrieren für Big Data ... 47
4.1.14 Technologien ... 48
4.1.15 Personal für Big Data ... 49
4.1.16 Budget ... 50
4.1.17 Erwünschte Fördermaßnahmen ... 51
4.2 Fazit der Online-Befragung ... 52

ERGEBNISSE DER BRANCHENWORKSHOPS ... 54
5.1 Methodik ... 54
5.1.1 Betrachtete Fragestellungen je Branche ... 54
5.1.2 Durchgeführte Branchenworkshops ... 54
5.2 Workshop Finanzen ... 54
5.2.1 Datenquellen ... 54
5.2.2 Prozessverbesserungen ... 54
5.2.3 Innovative Produkte ... 55
5.2.4 Fazit ... 55
5.3 Workshop Telekommunikation und Medien ... 55
5.3.1 Datenquellen ... 55
5.3.2 Prozessverbesserungen ... 56
5.3.3 Innovative Produkte ... 56
5.3.4 Fazit ... 57
5.4 Workshop Marktforschung ... 57
5.4.1 Datenquellen ... 57

5.4.2 Prozessverbesserungen ... 57
5.4.3 Innovative Produkte ... 57
5.4.4 Fazit ... 58
5.5 Workshop Handel ... 58
5.5.1 Datenquellen ... 58
5.5.2 Prozessverbesserungen ... 59
5.5.3 Innovative Produkte ... 59
5.5.4 Fazit ... 59
5.6 Workshop Versicherungen ... 59
5.6.1 Datenquellen ... 59
5.6.2 Prozessverbesserungen ... 60
5.6.3 Innovative Produkte ... 60
5.6.4 Fazit ... 60
5.7 Gemeinsame Themen in den Workshops ... 61

FAZIT ... 62
6.1 Chance »Effizientere Unternehmensführung« ... 62
6.2 Chance »Massenindividualisierung von Services« ... 62
6.3 Chance »Intelligente Produkte« ... 62

ANHANG ... 63
7.1 Steckbriefe der Big-Data-Anwendungsfälle ... 63

EINLEITUNG

1. »Big Data – Big Opportunities«

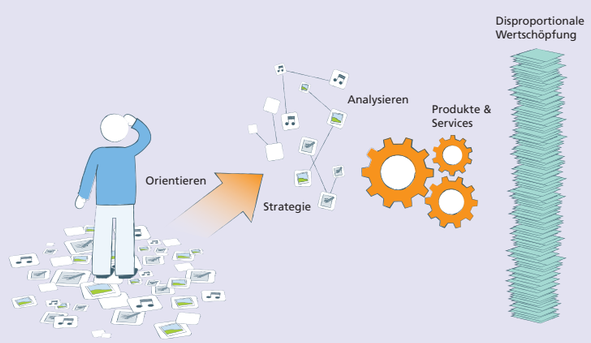
1,8 Zettabyte an Daten wurden im letzten Jahr erstmals weltweit produziert – und Prognosen zufolge verdoppelt sich das Volumen alle zwei Jahre. Das rasante Wachstum der Datenmengen, das die Digitalisierung unseres Planeten mit sich bringt, sowie ihre Analyse und Auswertung haben den Begriff »Big Data« geprägt. Große, unterschiedliche Datenmengen entstehen in hoher Geschwindigkeit. So etwas passiert zum Beispiel, wenn ein Energieerzeuger auf die jährliche Stromablesung verzichtet und dafür im Viertelstundentakt die Messdaten der Stromzähler elektronisch abliest. Oder wenn ein Flugzeugturbinenhersteller bereits nach der Landung analysieren möchte, welche Turbinen in die Inspektion müssen.

Aktuell stehen viele Unternehmen vor der Herausforderung, dass sie immer größere Datenmengen speichern, verwalten und analysieren müssen, um sie für ihre Geschäftsprozesse optimal verwerten zu können. Big Data ist nicht mehr länger nur eine Herausforderung für eine spezifische Branche – das Thema betrifft alle Wirtschaftszweige, Organisationen und Nutzer von digitalen Technologien. Weitere Beispiele sind Verbindungsdaten im Mobilfunk, Überweisungsdaten oder Echtzeitnachrichtendaten wie Twitter-Meldungen im Internet.

Neuere IT-Systeme können dabei helfen, Big Data zur Beantwortung existierender, aber auch neuer Fragestellungen auszuwerten. Das ermöglicht sowohl Prozessverbesserungen als auch Produktinnovationen. Die Vereinten Nationen etwa können mittlerweile drohende Virusinfektionen, Hungersnöte oder Unruhen über ganze Länder hinweg früher erkennen, da Millionen öffentlicher Kurznachrichten mit Fotos und Videos auswertbar sind. In Irland bietet eine Firma an, einen Wald mit Lasern zu erfassen und nach kurzer Zeit zu berechnen, welche Bäume es gibt und wie man den Wald besser bewirtschaftet. Und in naher Zukunft könnte das Smartphone bereits die attraktiven Plätze im Lieblings-Restaurant reserviert haben, wenn man gerade erst von der neuen Dienstreise erfahren hat.

Hinter all diesen Beispielen steckt Big Data – Technologien und das spezielle Know-how, aus diesem unerschöpflichen Potenzial neue Mehrwerte zu erzeugen. In der vorliegenden Dokumentation wird zusammenfassend dargestellt, welches Potenzial Big Data für Prozessverbesserungen und Produktinnovationen in deutschen Unternehmen mit sich bringt und welche offenen Fragen und Hürden aktuell bestehen.

Abbildung 1: Orientierung, Pläne und neue Produkte und Dienstleistungen durch Big Data können hohe Innovationspotenziale aufweisen



EXECUTIVE SUMMARY

2. Die Innovationspotenzialanalyse

Gefördert vom Bundesministerium für Wirtschaft und Technologie BMWi hat das Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS in Sankt Augustin eine weitreichende Untersuchung zu Nutzung und Potenzial von Big Data durchgeführt. Ziel ist es, das Thema stärker in der deutschen Wirtschaft zu etablieren.
Das Projekt umfasste drei Säulen: eine internationale Recherche zum Umgang mit Big Data und konkreten Anwendungen, eine Onlinebefragung unter Unternehmen, sowie Workshops mit Vertretern unterschiedlicher Branchen.

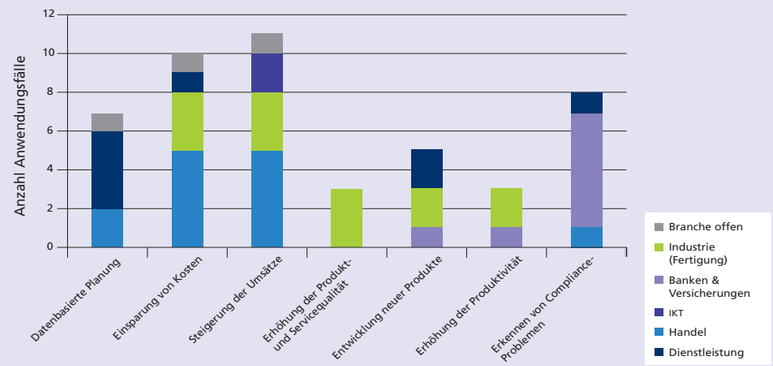
2.1 Ergebnisse der Recherche nach Anwendungsfällen

Ergebnis der Recherche sind 55 Big Data-Anwendungsfälle. Die Anwendungsfälle decken verschiedene Branchen und Unternehmensbereiche ab. Jedoch gibt es Schwerpunkte, und zwar in den Unternehmensbereichen, die für die jeweilige Branche am charakteristischsten sind: »Marketing, Vertrieb und Kundenbetreuung« im Handel, »Produktion, technische und IT-Services« in der Industrie, »Finanz- und Risiko-Controlling« in Banken und Finanzen, »Dienstleistung und Support« im Dienstleistungssektor.

2.1.1 Big-Data-Anwendungen tragen zu vielen Zielen bei

Die Steigerung der Umsätze und Einsparung von Kosten zählen zu den häufigsten Zielen, und zwar maßgeblich im Handel. Bei Banken und Versicherungen ist die Erkennung von Compliance-Problemen, vornehmlich in Zusammenhang mit Betrugsversuchen, das wichtigste Ziel. Im Dienstleistungssektor, und hier besonders durch die evidenzbasierte Medizin, steht die datenbasierte Planung im Fokus. Im Industriesektor sind die Ziele sehr vielfältig.

Abbildung 2: Ziele von Big Data-Anwendungen nach Branche



2.1.2 Es liegen vielfältige Aufgabengebiete für Big-Data-Anwendungen vor

Die Anwendungsfälle dienen verschiedenen konkreten Aufgaben. Um einen Überblick zu gewinnen und Gemeinsamkeiten herauszuarbeiten, wurden die Aufgaben so abstrahiert, dass sich Gruppen mit ähnlichen Eigenschaften ergeben:

- Marktmonitoring für Verkaufschancen
- Personalisierte Produktempfehlung
- Kündigerfrüherkennung
- Mitarbeitergewinnung
- Absatzprognose für Planung und Steuerung
- Vorausschauende Instandhaltung
- Umsichtige Steuerung
- Betrugserkennung
- Finanzielle Risikoabschätzung
- Erkennung von Attacken
- Produktverbesserung
- Innovative Produkte

Schwerpunkt im Handel sind Absatzprognose, Marktmonitoring und umsichtige Steuerung. Marktmonitoring, vorausschauende Instandhaltung sowie innovative und verbesserte Produkte sind typische Aufgaben in Industrie und Dienstleistung. Bei Banken und Versicherungen geht es um Betrugserkennung und Risikoabschätzung.

2.2 Ergebnisse der Online-Befragung

Die Online-Befragung hatte zum Ziel, das Verständnis über die Potenziale von Big Data bei kleinen und mittelständischen Unternehmen (KMU) in Deutschland abzufragen. Insgesamt nahmen Entscheider aus 82 KMU und Großunternehmen unterschiedlicher Branchen an der Onlinebefragung teil.

2.2.1 Potenzial von Big Data ist KMU bekannt

Die Online-Befragung zeigt, dass die Unternehmen mehrheitlich angeben können, welche Daten in ihrem Unternehmen vorliegen und wie sie deren Bedeutung für ihren Unternehmenserfolg einschätzen. Es wird deutlich, dass neben Daten aus Transaktions- sowie CRM-Systemen ein weites Spektrum an unstrukturierten und veränderlichen Daten in den Unternehmen existieren und ihre heutige Relevanz noch untergeordneter Natur ist. Die Teilnehmer können weiterhin benennen, welche Unternehmensbereiche Daten mit hoher Veränderlichkeit auswerten. Daraus zeichnet sich ab, dass insbesondere das Marketing und der Vertrieb sowie das Management und die Produktion zu den Big Data-affinen Unternehmensbereichen zählen.

Es ist den Teilnehmern möglich, eine Bewertung potenziell relevanter Anwendungen auf Basis von Big Data in ihrem Unternehmen vorzunehmen. So werden Anwendungen als besonders relevant eingeschätzt, die eine Prognose der Werbewirksamkeit, Abverkäufe und Mikrosegmentierung ermöglichen, gefolgt von Anwendungen zum Monitoring von Markenwahrnehmung, Wettbewerbern, Marktpreisen und Kaufinteressenten im Web sowie Besucherströmen vor Ort. Auch die automatische Preissetzung, Kündigeranalyse sowie personalisierte Kundenansprache zählen zu den favorisierten Big Data Anwendungen. Es kann davon ausgegangen werden, dass KMU die Potenziale von Big Data bekannt sind.

2.2.2 Steigerung der Wettbewerbsfähigkeit als wesentliches Ziel

Das größte Potenzial von Big Data sehen die Befragten mit deutlichem Abstand im Aufbau strategischer Wettbewerbsvorteile (69 Prozent). Es folgen die Steigerung der Umsätze (61 Prozent) und die Einsparung von Kosten (55 Prozent). Die Erhöhung der Produktivität und die datenbasierte Planung und Entscheidungsfindung sind weitere Ziele, die mit Big Data erreichbar scheinen. Diese Ergebnisverteilung macht deutlich, dass die Anwendung von Big Data nicht allein einem bestimmten Ziel dienen muss.

Je nach Anwendung können auch mehrere Ziele gleichzeitig unterstützt werden. Eine Erhöhung der Produktivität kann in einer Branche zum Aufbau strategischer Wettbewerbsvorteile beitragen, wenn in dieser Branche insbesondere über die Produktivität konkurriert wird (Automobilbau). In anderen Branchen kann die Kosteneinsparung zu den besonders relevanten Zielen gehören, um an Wettbewerbsfähigkeit zu gewinnen (Handel). Das Potenzial von Big Data Anwendungen bedient viele Ziele und steigert im Kern die Wettbewerbsfähigkeit.

2.2.3 Treiber ist das Management, nicht die IT

Mit Blick auf die Funktionen der Teilnehmer der Online-Befragung zeigt sich, dass sich der größte Teil der Teilnehmer nicht eindeutig einem Funktionsbereich zuordnen kann, die zweitgrößte Gruppe jedoch aus dem Management stammt. Häufige Titelnennungen in den Freitextantworten zur Funktion weisen auf kaufmännische Tätigkeiten hin, während vergleichsweise wenige Titel eindeutig auf IT verweisen. Das hohe Interesse an Anwendungen im Bereich Marketing und Vertrieb sowie das Ziel der Wettbewerbssteigerung deutet darauf hin, dass Big Data eher ein Management-Thema ist, als dass es durch die IT dominiert wird. Inwieweit es damit zusammen hängt, dass es datenbasierte Entscheidungen erleichtert, kann nicht eindeutig beantwortet werden.

Die datenbasierte Planung und Entscheidungsfindung ist als eines von vielen Zielen in Bezug auf den Big Data Einsatz in einem Unternehmen und einer Branche benannt.

2.2.4 Interne Ressourcen sind knapp und werden ausgebaut

Die Teilnehmer der Online-Befragung sehen sich in ihren Unternehmen aktuell nicht optimal aufgestellt, wenn es um die notwendigen Ressourcen für die Arbeit mit Big Data geht. Das beginnt mit fehlenden oder unzureichenden internen Budgets und endet mit unklaren personellen Verantwortlichkeiten und ausbaufähigen Kompetenzen. In den kommenden fünf Jahren sollte nach Angaben der Teilnehmerinnen und Teilnehmer die Personalausstattung deutlich steigen. Ähnliches gilt für das Budget: 67 Prozent sind der Meinung, dass das Budget für Big-Data-Themen steigen sollte. Die Forschungs-Hypothese wird durch die Antworten der Teilnehmer aus der Online-Befragung eher bestätigt. Die erforderlichen Ressourcen zum Arbeiten mit großen Datenmengen sind in der Mehrzahl der KMU noch nicht vorhanden, jedoch sollte sich dieses Bild in den nächsten fünf Jahren zum Positiven verändern. Um bestehenden Defizite zu ändern, wünschen sich 95 Prozent der Befragten Förderung in Form von Best Practices oder Trainings.

2.3 Ergebnisse aus den Branchenworkshops

Die Angaben aus der Online-Befragung bestätigten sich auch in den Zukunftsworkshops mit Vertretern der Branchen Telekommunikation/Medien, Handel, Finanzwirtschaft, Versicherung und Marktforschung. Die Ergebnisse der Diskussionen um Anforderungen, Ziele und Herausforderungen für den Einsatz von Big Data wurden in Roadmaps zusammengefasst. Insgesamt zeigte sich in der qualitativen Analyse, dass Big Data kein reines Technologie-, sondern vor allem ein Strategiethema ist.

2.3.1 Finanzwirtschaft: Effizienz nach innen, Individualisierung nach außen

Kurzfristig sehen die Teilnehmer des Workshops, dass die Finanzbranche stark an Effizienzgewinnung durch die Nutzung von Big Data interessiert ist. Belastbare Kosten-Nutzen-Kalkulationen stehen im Vordergrund. Allerdings wird auch gesehen, dass neue Anbieter u.a. aus fremden Branchen in der Lage sein könnten, dank Big Data potenziell interessantere Finanzprodukte anzubieten, sofern dies regulativ möglich ist. Das Innovationspotenzial von Big Data in der Finanzbranche wurde von allen Beteiligten gesehen. Allerdings ist offen, welches deutsche Unternehmen dieses Potenzial nutzen wird.

2.3.2 Telekommunikation und Medien: Braucht es noch Marktforschung?

Im Workshop Telekommunikation und Medien bildete sich die gemeinsame Auffassung, dass die Branche bereits heute über sehr viele Daten verfügt, diese jedoch erst in geringerem Maße

in die Verbesserung von Prozessen und noch weniger in die Entwicklung neuer Produkte und Dienstleistungen einfließen. Mittelfristig ist auch hier der Trend zu Prozessen und Dienstleistungen zu erkennen, die auf eine kontinuierliche Analyse von Daten aufsetzen und damit die Basis für effizienteres Unternehmensmanagement und neuartige Monitoring-Dienstleistungen für Endkunden bilden. Langfristig stellen sich die Teilnehmer vor, dass eine zunehmende Automatisierung und Individualisierung sehr komplexe Geschäftsmodelle wie das der Smart Cities und Smart Grids ermöglichen, deren konkrete Anwendungen bislang nur in groben Zügen erkennbar sind. Auch hier gilt, dass die Nutzung von Big Data immer vor dem Hintergrund der Datenschutzbestimmungen gesehen werden und die Gesellschaft angesichts der Zukunftserwartungen der Teilnehmer viele neue Antworten finden muss, um gesellschaftlich akzeptablen Fortschritt zu ermöglichen.

2.3.3 Handel: Der Traum vom IT-gestützten Tante-Emma-Laden

Der Handel sieht klare Modernisierungschancen durch Big Data. Datengetriebene Infrastrukturen werden zunehmend als Service angeboten und können dadurch auch kapitalschwächeren Handelsunternehmen oder Einzelhändlern einen schnelleren Zugang zum (Online-)Markt liefern. Mittelfristig ist es auch hier die Individualisierung (Tante-Emma-Laden als Metapher), die den Unterschied in der Kundenansprache und Kundenbindung ausmacht. Längerfristig scheinen die Daten selbst zum Produkt zu werden.

2.3.4 Marktforschung: Markt für Marktforschung wird größer

Die Marktforschung sollte nach Meinung der Experten wie kaum eine andere Branche von der Nutzung von Big Data profitieren. Neue Angebote werden entstehen, die auf Algorithmen und neuartigen Visualisierungsverfahren basieren und Kunden einen aktuelleren und detaillierteren Einblick in Märkte bieten. Sowohl Kunden als auch Anbieter arbeiten in Zukunft stärker mit computerbasierter Statistik. Eine solche Kompetenzentwicklung, auch als »Data Science« beschrieben, wird mittelfristig deutlich zunehmen. Kurzfristig sind Experten mit Data-Science-Kompetenzen wenig verfügbar, aber sehr gefragt.

2.3.5 Versicherungen: Der Trend zu Versicherungen in HD

Für die Versicherungsbranche sahen die Experten eine Reihe aktueller und künftiger Herausforderungen, die durch den Einsatz von Big Data bestehen. Im Status quo ist dank des heutigen Datenschutzes und der Einwilligungen seitens der Kunden der Umgang mit personenbezogenen Daten geregelt. In Zukunft könnten weitere Datenquellen, die z.B. durch Monitoring bzw. Prognose entstehen, die Basis für verhaltensorientierte Versicherungsprodukte bilden. Beispielhaft zu nennen sind hier Daten aus Telematik-Systemen, aus Social Media oder digitalen Patientenakten. Ob derartige Innovationspotenziale auch von den Endkunden angenommen werden, bleibt offen.

2.4 Effizienz steigern durch Big Data

Im Gesamtkontext der Innovationspotenzialanalyse fiel auf, dass in den branchenspezifischen Zukunftsworkshops ähnliche Muster in der Wahrnehmung von Nutzungsmöglichkeiten von Big Data bestehen.

So ließen sich kurzfristig in allen Branchen unmittelbare Potenziale in der Effizienzsteigerung der Unternehmensführung durch Big Data erkennen. Branchenspezifisch waren jeweils die entsprechenden Anwendungen und Datenquellen. Langfristig war in allen Branchen von einer Massenindividualisierung von Services und zunehmend intelligenten Produkten die Rede, die von Big Data und der jeweils notwendigen Analysefähigkeit profitieren. Da die Durchführung von kurz- bis mittelfristigen Projekten zur Effizienzverbesserung von belastbaren Kosten-Nutzen-Kalkulationen abhängt, sind deutlich mehr branchenspezifische Referenzbeispiele notwendig. Hier liegen speziell in Deutschland noch nicht viele öffentlich kommunizierte Referenzprojekte vor und es herrscht bei den Unternehmen eine gewisse Unsicherheit, inwieweit erfolgreiche Projekte aus den USA als Vorlage dienen können, da sich die hiesigen Regularien und Praktiken in Bezug auf Datenschutz mitunter sehr von den Vorgaben in den USA unterscheiden. Mittel- bis langfristig angelegte Produkt- und Serviceinnovationen auf Basis von Big Data erfordern entsprechend qualifiziertes Personal, welches idealerweise das strategische Wissen des Unternehmens bündelt. Hier wird es seitens der jeweiligen Unternehmensführung erforderlich sein, die notwendigen Schritte zum Aufbau der entsprechenden inhouse-Expertise und zur Investition in mittel- bis langfristig wirksame Produkte und Services auf Basis von Big Data einzuleiten. Entscheidend für beide skizzierten Stränge ist, dass die Unternehmensführung die Chancen aber auch kulturellen Herausforderungen (an-)erkennt, stärker datengetrieben zu operieren und die Fähigkeit zur intelligenten Nutzung von Daten in Produkten und Services als schwierig kopierbaren Wettbewerbsvorteil zu entwickeln.

Zusammenfassend lassen sich aus dem fachlichen Innovationspotenzial von Big Data-Anwendungen, der hohen Relevanzeinschätzung durch die KMU und den konkreten Innovationspotenzialen aus den Branchenworkshops drei Chancen für den Einsatz von Big Data in deutschen Unternehmen ableiten:

Effizientere Unternehmensführung durch Big Data

So lassen sich etwa im Einzelhandel genauere Prognosen treffen, wann welches Produkt verkauft wird und nachbestellt werden muss. Die Energiebranche kann besser vorhersagen, wie viel Strom wann benötigt wird. Und bei einfachen Prozessen wie der Postbearbeitung können lernende Systeme durch automatisierte Abläufe für mehr Effizienz sorgen.

Massenindividualisierung durch Big Data

Wenn Systeme während der Bearbeitung einer Anfrage relevante Informationen über den Kunden mitlernen, können Dienstleistungen künftig auf die einzelne Person oder das Objekt individualisiert werden. Nutzungsbezogene Versicherungsprodukte, tägliche Gesundheitsdiagnosen, individualisierte Unterhaltung oder der vorausschauende Service bei Maschinen. Bisherige standardisierte Dienstleistungen werden zunehmend individualisiert angeboten.

Intelligentere Produkte durch Big Data

Schon heute verfügen viele Maschinen und Anlagen über Sensoren, die über den Wartungszustand etc. Auskunft geben. Es gibt selbstlernende Thermostate, selbstregulierende Häuser, autonom fahrende Fahrzeuge. In Zukunft werden immer mehr Produkte und Maschinen mit eigener Big-Data-Intelligenz ausgestattet sein, um die Sensordaten direkt zu verarbeiten und zu reagieren, sich zum Beispiel auf besondere Anforderungen einzustellen und sich selbst zu reparieren.

Die deutsche Wirtschaft hat das Innovationspotenzial von Big Data trotz aller begrifflichen Unschärfen erkannt. Vielfältige Anwendungsmöglichkeiten sind sowohl im Dienstleistungsbereich als auch in der Industrie gegeben. Es sollte daher mit entsprechenden gesellschaftlichen Klärungen und Weiterentwicklungen im Bereich des Datenschutzes und der Datensicherheit möglich sein, das Innovationspotenzial durch höhere Effizienz, stärkere Individualisierung von Dienstleistungen und zunehmend intelligente Produkte zu realisieren, um auch als Volkswirtschaft an Wettbewerbsfähigkeit zu gewinnen.

INNOVATIONSPOTENZIAL-ANALYSE BIG DATA –

ERGEBNISSE IM DETAIL

RECHERCHE NACH BIG DATA ANWENDUNGSFÄLLEN

Um genauer zu verstehen, welche innovativen Anwendungsfälle für Big Data in der Wirtschaft bereits existieren, wurde zunächst eine Recherche im World Wide Web durchgeführt. Anschließend wurden die identifizierten Anwendungsfälle von Big Data aus einer fachlichen Sicht analysiert. Die wichtigsten Auswahlkriterien waren:

- Verarbeitung von umfangreichen, veränderlichen Datenmengen
- Analytische Fragestellungen
- Abdeckung möglichst verschiedener Branchen und Unternehmensbereiche
- Verschiedenartigkeit der Anwendungsfälle
- Möglichst Nennung eines Anwenders

Anwendungsfälle, wo bisherige Daten in neuartige Datenbanken transferiert werden und dadurch eine bessere Skalierbarkeit, Kosteneinsparung oder ein rein quantitativer Zeitgewinn erreicht wird, fallen damit aus der Betrachtung heraus. Durch eine beschleunigte Verarbeitung oder durch die Kombination von Datenquellen sollen in Unternehmen neue analytische Fragestellungen beantwortet werden.

Ergebnis der Recherche sind 55 Anwendungsfälle, die im Anhang als Steckbriefe gelistet werden. Um sie zu vergleichen, wurden Kategorien gebildet, die sich im Wesentlichen an der Unternehmensbefragung (siehe Abschnitt 4) orientieren.

3.1 Kategorien

Die Fälle wurden in folgende Kategorien eingeteilt:

- Unternehmensbereich
- Branche
- Ziele
- Aktualität
- Reaktion
- Datenquellen
- Aufgaben
- Analysen.

3.1.1 Branche und Unternehmensbereich

Die Kategorie »Unternehmensbereich« bestimmt sich aus den Antwortmöglichkeiten in der Befragung:

- Marketing, Vertrieb und Kundenbetreuung (im folgenden kurz »Marketing«)
- Produktion, technische und IT-Services (kurz »Produktion«)
- Dienstleistung und Support
- Distribution und Logistik
- Finanz- und Risiko-Controlling, Compliance
- Forschung und Produktentwicklung
- Management, Geschäftsführung

Die Kategorie »Branche« oder »Sektor« orientiert sich ebenfalls an der Umfrage, nimmt hier jedoch eine feinere Unterteilung des Sektors »Dienstleistung« vor (siehe Tabelle 1). So können die Anwendungsfälle für Banken und Versicherungen bzw. für IKT-Unternehmen differenziert betrachtet werden.

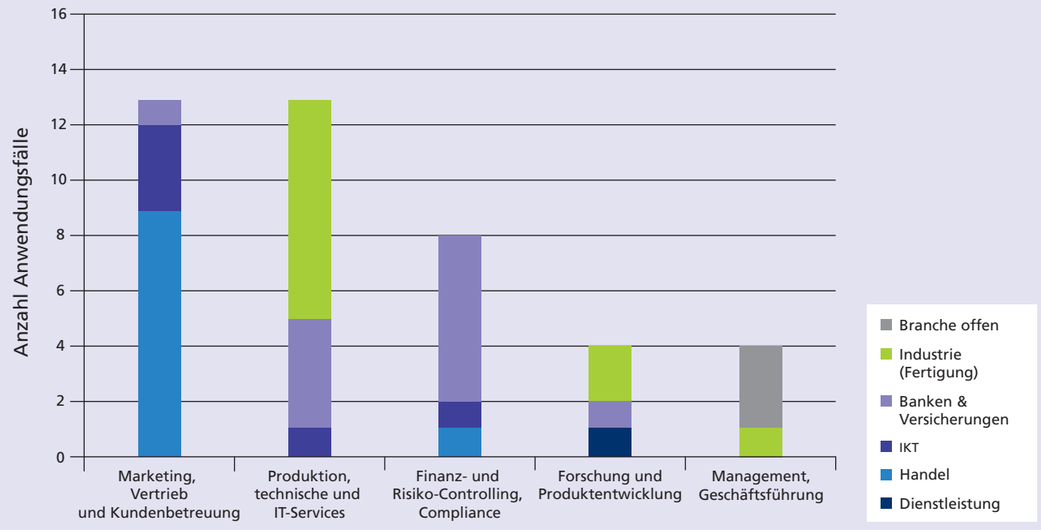
Tabelle 1: Bezug zwischen Sektor in der Umfrage und der Kategorie Sektor

Sektor in der Umfrage	Sektor in den Steckbriefen
Handel	▪ Handel, inkl. Gastronomie
Dienstleistung	▪ Banken und Versicherungen ▪ IKT ▪ Dienstleistung inklusive Medien, Medizin, öffentliche Sicherheit
Industrie	▪ Industrie, inklusive Energie

Die ausgewählten Anwendungsfälle decken die verschiedenen Branchen und Unternehmensbereiche gut ab, wie die folgende Grafik veranschaulicht. Es gibt jedoch Schwerpunkte, und zwar in den Unternehmensbereichen, die für die jeweilige Branche am charakteristischsten sind:

»Marketing, Vertrieb und Kundenbetreuung« im Handel,
»Produktion, technische und IT-Services« in der Industrie,
»Finanz- und Risiko-Controlling« in Banken und Finanzen,
»Dienstleistung und Support« im Dienstleistungssektor.

Abbildung 3: Unternehmensbereich und Branche



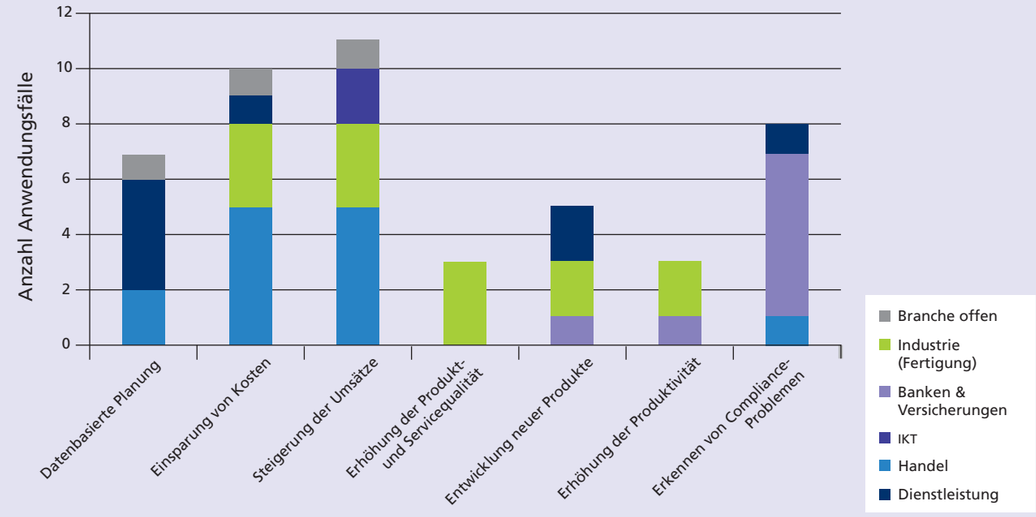
3.1.2 Ziele

Auch die Ziele der Anwendungsfälle wurden den Optionen der Befragung angelehnt und wie folgt eingeteilt:

- Steigerung der Umsätze
- Einsparung von Kosten
- Erhöhung der Produktivität
- Aufbau strategischer Wettbewerbsvorteile
- Datenbasierte Planung und Entscheidungsfindung
- Entwicklung neuer Produkte und Dienstleistungen
- Verbesserung des Kundenservices
- Erkennen von Compliance-Problemen
- Erhöhung der Produkt- und Servicequalität

Jedem Fall wurde ein möglichst spezifisches Ziel zugeordnet. Gemäß der folgenden Abbildung sind die Steigerung der Umsätze und die Einsparung von Kosten die häufigsten Ziele, und zwar maßgeblich im Handel. Bei Banken und Versicherungen ist die Erkennung von Compliance-Problemen, vornehmlich in Zusammenhang mit Betrugsversuchen, das wichtigste Ziel. Im Dienstleistungssektor, und hier besonders durch die evidenzbasierte Medizin, herrscht die datenbasierte Planung vor. Im Industriesektor sind die Ziele am vielfältigsten.

Abbildung 4: Ziele und Branche



3.1.3 Aktualität und Reaktionsweise

Die Kategorie »Aktualität« (oder Geschwindigkeit, Velocity) fragt danach, wie oft Analysen auf aktualisierten Daten durchgeführt werden. Drei Möglichkeiten werden unterschieden:

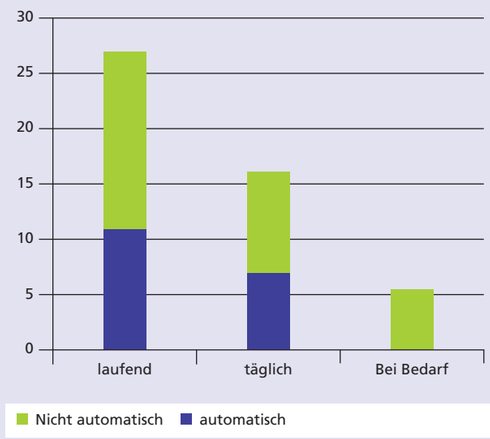
- laufend: (fast) in Echtzeit, auf Datenströmen
- täglich: ein oder mehrfach täglich
- seltener

Analysen, die nahezu in Echtzeit erfolgen, werden auf Datenströmen durchgeführt. Die Geschwindigkeit der Daten bestimmt nicht unbedingt die Häufigkeit der Analyse. Analysen werden benötigt, wenn Entscheidungen in Geschäftsprozessen anstehen. Abrechnung nach Nutzung oder Maßnahmen, um die Abwanderung von Mitarbeitern zu verhindern, sind tägliche Aktivitäten, auch wenn die Daten laufend eingehen. In einem sozialen Netzwerk können die Nutzer laufend Erfahrungen mit Medikamenten hinzufügen, die Inhalte müssen aber nicht unbedingt täglich in eine Analyseumgebung für evidenzbasierte Medizin eingebracht werden.

Die Einteilung der Anwendungsfälle nach Reaktionstyp ist nicht immer einfach, je nachdem, wo man die Grenze zwischen Analysemodul und Weiterverarbeitung im Geschäftsprozess zieht. Analyse-Module ohne automatische Reaktion können beim Kunden in Prozesse mit vollautomatischer Reaktion eingebaut werden. Ein Beispiel sind Überwachungskameras, die Besucher zählen und diese Daten in die automatisierte Personaleinsatzplanung einfließen lassen.

Die Anwendungsfälle unterscheiden sich auch darin, ob das Analyseergebnis zu einer vollautomatisierten Reaktion führt oder ob es von Menschen weiter verarbeitet wird. Abbildung 5 zeigt, dass in den meisten Fällen die Daten laufend analysiert werden, gefolgt von täglichen Analysen. Automatische und nicht automatische Reaktionen sind etwa gleich verteilt. Seltenere Analysen »bei Bedarf« führen nicht zu automatisierten Reaktionen.

Abbildung 5: Aktualität und Reaktionsweise



Nimmt man an, dass Analysen, auf die vollautomatisch reagiert wird, verlässlicher sein müssen, so lassen sich die Kategorien »Automatisierte Reaktion« und »Aktualität« gut kombinieren. Je dunkler die Farbe in Tabelle 2, umso anspruchsvoller wäre die Anwendung.

Tabelle 2: Kombination von Aktualität und Reaktionstyp

Aktualität \ Reaktion	nicht automatisiert	automatisiert
laufend	Es werden laufend anfallende Daten analysiert und die Reaktion ist nicht automatisiert	Es werden laufend anfallende Daten analysiert und die Reaktion ist automatisiert
täglich	Es werden täglich anfallende Daten analysiert und die Reaktion ist nicht automatisiert	Es werden täglich anfallende Daten analysiert und die Reaktion ist automatisiert
bei Bedarf	Die Daten werden für Analysen bei Bedarf Daten aktualisiert und die Reaktion ist nicht automatisiert	

Tabelle 2 schlüsselt Aktualität und Reaktionsweise weiter nach Branchen auf. Nun zeigt sich, dass der Schwerpunkt der Fälle in der Industrie auf laufenden Analysen liegt, auf die aber nicht automatisch reagiert wird. Zum Beispiel wird bei der präventiven Instandhaltung auf den Verdacht meist nicht vollautomatisch reagiert. Im Handel gehen viele automatisierte Reaktionen auf Echtzeitanalysen mit Produktempfehlung in Online-Systemen zusammen. Auf täglich durchgeführte Analysen wird eher manuell reagiert, zum Beispiel für die Disposition.

Bei Banken und Versicherungen verhält es sich eher umgekehrt. Auf laufende Analysen wird nicht automatisch, auf tägliche eher automatisch reagiert. Unter die täglichen Analysen fallen zum Beispiel vollautomatische Abrechnungen nach Nutzung. Nicht für alle Fälle werden Anwender genannt. Das betrifft auch die Fälle von Echtzeitanalysen für Banken und Versicherungen. In konkreten Anwendungen können die Analyseergebnisse in konkrete automatisierte Geschäftsprozesse eingebaut werden.

Das Cluster »bei Bedarf durchgeführte Analysen ohne automatische Reaktion« beherrscht den Dienstleistungssektor. Hierein fallen Analysen auf großen Datenmengen in der Medizin.

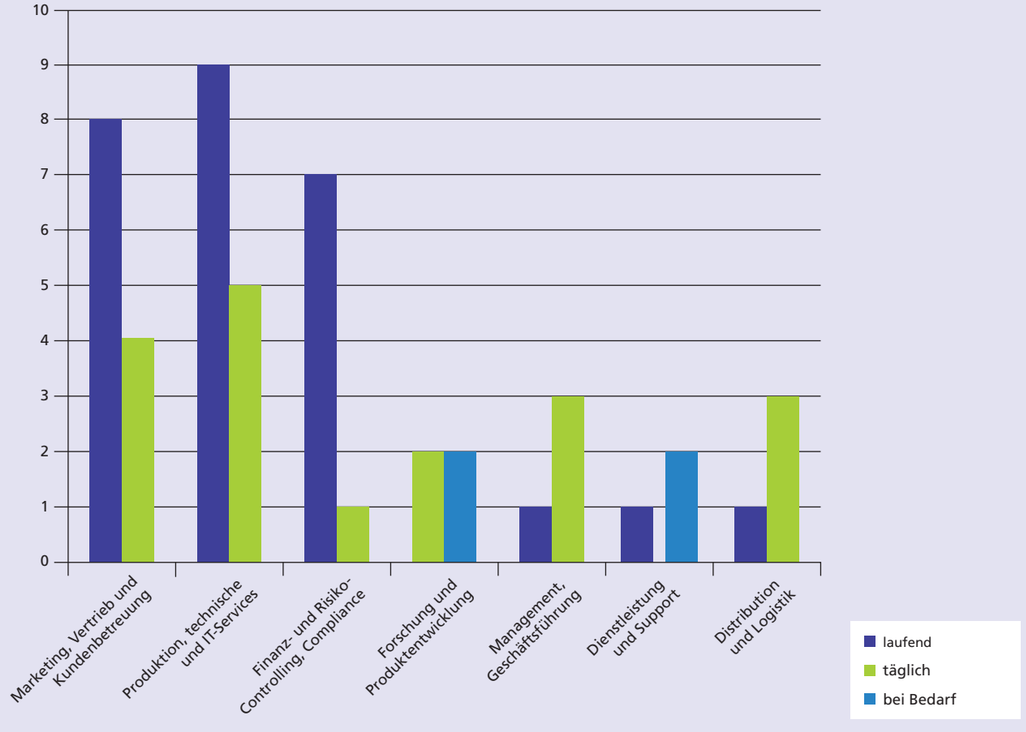
Tabelle 3: Aktualität, Reaktion und Branche

Aktualität		nicht automatisierte Reaktion	automatisierte Reaktion	
laufend				
Industrie	7			
Handel	2			5
IKT	1			1
Banken&Versicherungen	5			2
Dienstleistung	1			2
offen	1			
täglich				
Industrie	2			1
Handel	4			1
IKT				1
Banken&Versicherungen	1			3
Dienstleistung	1			
offen	2			
bei Bedarf				
Industrie	2			
Handel	1			
IKT				
Banken&Versicherungen				
Dienstleistung	3			
offen				

In allen Fällen sollten die Analysen auf aktualisierten Daten durchgeführt werden können. Das bedeutet, die Datenvorverarbeitung muss automatisiert erfolgen. Es bedeutet jedoch nicht unbedingt, dass die Analyseverfahren alle selbstlernend sind. Denn es müssen nicht immer neuartige Muster erkannt und dafür neue Regeln gelernt werden. Zum Beispiel muss eine Ähnlichkeitsfunktion für die Produktempfehlung nicht dauernd neu trainiert werden, Regeln für die Betrugserkennung aber wohl. Eine Kategorisierung der Fälle nach (selbst) lernenden Verfahren wäre sicher interessant, aber diese Information konnte aus den Fallbeschreibungen oft nicht, oder nur mit großer Unsicherheit, entnommen werden.

Abbildung 6 stellt die Aktualität der Daten den Unternehmensbereichen gegenüber. Dazu passend wurde in der Umfrage nach der Veränderlichkeit der Daten gefragt. Sich laufend verändernde Daten findet man am meisten in Anwendungsfällen aus den Unternehmensbereichen »Produktion, technische und IT-Services«, »Marketing, Vertrieb und Kundenbetreuung« sowie »Finanz- und Risiko-Controlling und Compliance«.

Abbildung 6: Aktualität der Daten und Unternehmensbereich



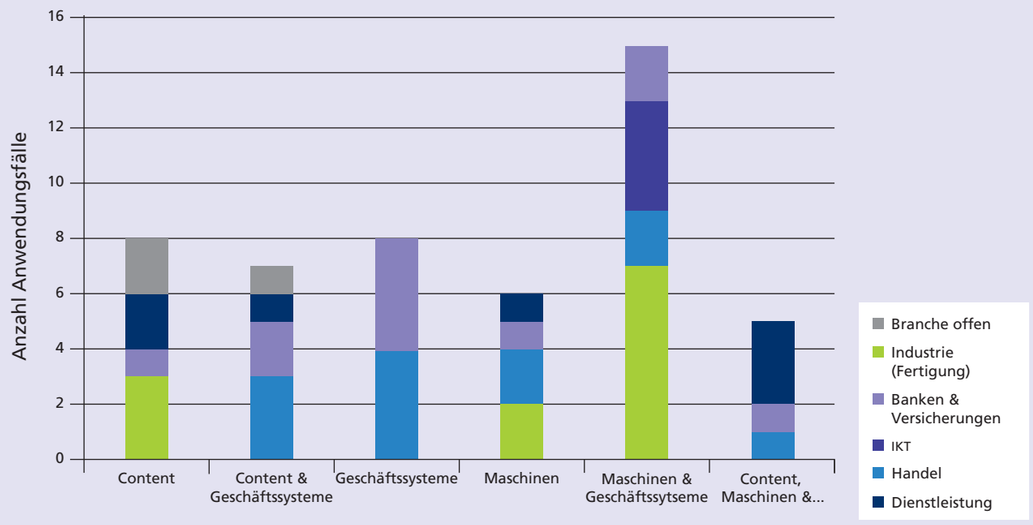
3.1.4 Datenquellen

Die Datenquellen lassen sich nach ihrem Ursprung grob in drei Gruppen unterteilen:

- Maschinengeneriert: Sensordaten, Logdaten, Sprache/Audio/Video, Datendienste
- von Menschen generierter Content: Social Media, Korrespondenz, Publikationen/Patente, Bilder, Freitext, Formulare[1], Open Data/Web Content, Sprache/Audio/Video
- Geschäftsdaten: Stammdaten/Falldaten, CRM-Daten, Transaktionsdaten

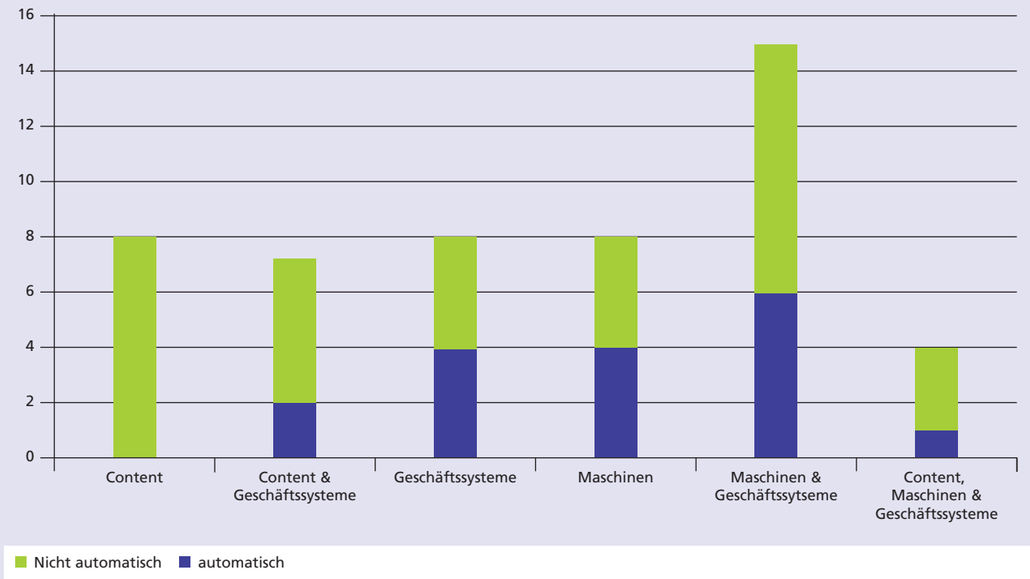
Abbildung 7 zeigt, dass die gesammelten Fälle das Spektrum der verschiedenartigen Datenquellen gut abdecken, aber ein deutlicher Schwerpunkt auf der Kombination von Maschinendaten mit Daten aus Geschäftssystemen liegt. Das ist im Wesentlichen auf den Industrie- und IKT-Sektor zurückzuführen.

Abbildung 7: Sektor und Art der Datenquellen



Laut Abbildung 8 ergibt sich zwischen der Art der Datenquellen und der Reaktionsweise kein Zusammenhang – mit der Ausnahme, dass Analysen von menschengeneriertem Content alleine keine automatischen Reaktionen erzeugen, dazu werden wohl zusätzliche Geschäftsdaten benötigt.

Abbildung 8: Art der Datenquellen und Reaktion



[1] Formulare wurde als neue Art von Content mit aufgenommen.

Ergänzt man Abbildung 8 um die Aktualität, so erhält man alle drei Dimensionen aus der Definition von »Big Data«. Das Ergebnis in Tabelle 4 zeugt von der Vielfalt der Fälle, die von Big-Data-Analysen profitieren. Dass es keine Anwendungsfälle für tägliche Analysen gibt, die alle drei Arten von Datenquellen verarbeiten, ist vermutlich Zufall.

Tabelle 4: Aktualität, Reaktionstyp und Verschiedenartigkeit der Datenquellen

Aktualität		Nicht automatisierte Reaktion	Automatisierte Reaktion	
laufend				
3 Arten von Datenquellen	2			
2	7			5
1	7			1
täglich				2
3 Arten von Datenquellen				2
2	4			
1	5			
bei Bedarf				1
3 Arten von Datenquellen	1			1
2	3			1
1	2			3

3.1.5 Aufgabe

Die Anwendungsfälle dienen verschiedenen konkreten Aufgaben. Um einen Überblick zu gewinnen und Gemeinsamkeiten herauszuarbeiten, wurden die Aufgaben so abstrahiert, dass sich Gruppen mit ähnlichen Eigenschaften ergaben. Sie werden in Abschnitt 3.2 präsentiert. Tabelle 5 gibt einen Überblick über die Aufgaben und die Themen, in die sie sich einordnen lassen.

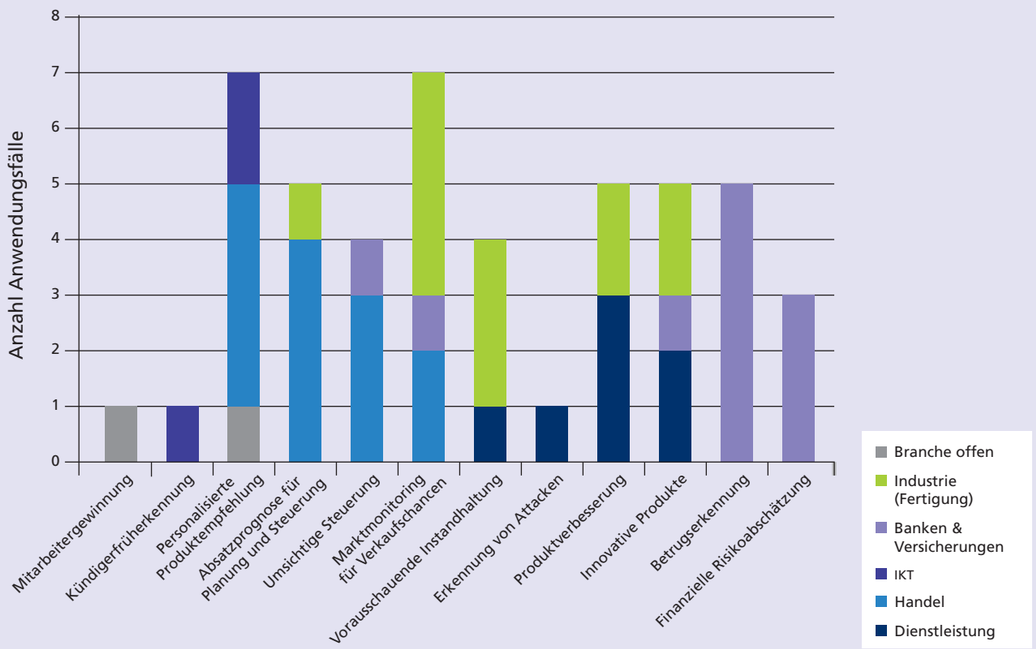
Tabelle 5: Aufgaben

Thema	Aufgaben
Verkaufschancen nutzen	▮ Marktmonitoring für Verkaufschancen
Personalisierte Ansprache von Kaufinteressenten	▮ Personalisierte Produktempfehlung ▮ Kündigerfrüherkennung
Personalisierte Ansprache von (potenziellen) Mitarbeitern	▮ Mitarbeitergewinnung
Optimierung des Betriebs	▮ Absatzprognose für Planung und Steuerung ▮ Vorausschauende Instandhaltung ▮ Umsichtige Steuerung
Finanzielle Risiken und Betrug	▮ Betrugserkennung ▮ Finanzielle Risikoabschätzung
Erkennung von Attacken	▮ Erkennung von Attacken
innovative oder verbesserte Produkte und Dienstleistungen	▮ Innovative Produkte ▮ Produktverbesserung

Abbildung 9 zeigt die Verteilung der Aufgaben auf die Branchen. Schwerpunkt im Handel sind Absatzprognose, Marktmonitoring und umsichtige Steuerung. Marktmonitoring, vorausschauende Instandhaltung sowie innovative und verbesserte Produkte sind typische Aufgaben in Industrie und Dienstleistung. Bei Banken und Versicherungen geht es vor allem um Betrugserkennung und Risikoabschätzung.

Die Grafik zeigt auch, dass die Aufgaben Mitarbeitergewinnung, Kündigerfrüherkennung und Erkennung von Attacken in wenige Anwendungsfälle aufgenommen wurden. Sie bilden trotzdem wichtige Aufgabenbereiche für Big-Data-Analysen.

Abbildung 9: Aufgabe und Branche



3.1.6 Analyse

In die Auswertung der Recherche nach Anwendungsfällen wurde des Weiteren die Kategorie des Analysetyps einbezogen:

- Prädiktiv: Ergebnis ist eine in die Zukunft gerichtete quantitative Prognose, die man zur Planung nutzen kann, die Antizipation einer Chance, die man wahrnehmen kann oder eines Risikos, das man verhindern möchte.
- Monitoring: In zeitbezogenen Daten werden Veränderungen erkannt, insbesondere mehr oder weniger komplexe Ereignisse, die man registrieren und verarbeiten möchte und auf die man ggf. reagieren möchte.
- Explorativ: Es werden verschiedene Analysen oder Analysemöglichkeiten für die aufbereiteten Daten angeboten. Hierzu zählen Anwendungen zur Entscheidungsunterstützung.
- Klassifikation: Eine Situation oder ein Fall wird bewertet, um zwischen verschiedenen Möglichkeiten der Weiterverarbeitung auszuwählen. Hierzu zählt auch das Filtern von Daten.

Abbildung 10 zeigt, wie sich Aufgaben und Analysemethode entsprechen. Wörter wie Erkennung, umsichtig oder Monitoring in der Aufgabenbezeichnung deuten indirekt auf Monitoring-Analysen; Wörter wie Früherkennung, Abschätzung, vorausschauend oder Empfehlung auf prädiktive Analysen. Explorative Analysen herrschen insbesondere bei der Produktverbesserung vor.

Abbildung 10: Aufgabe und Analysemethode

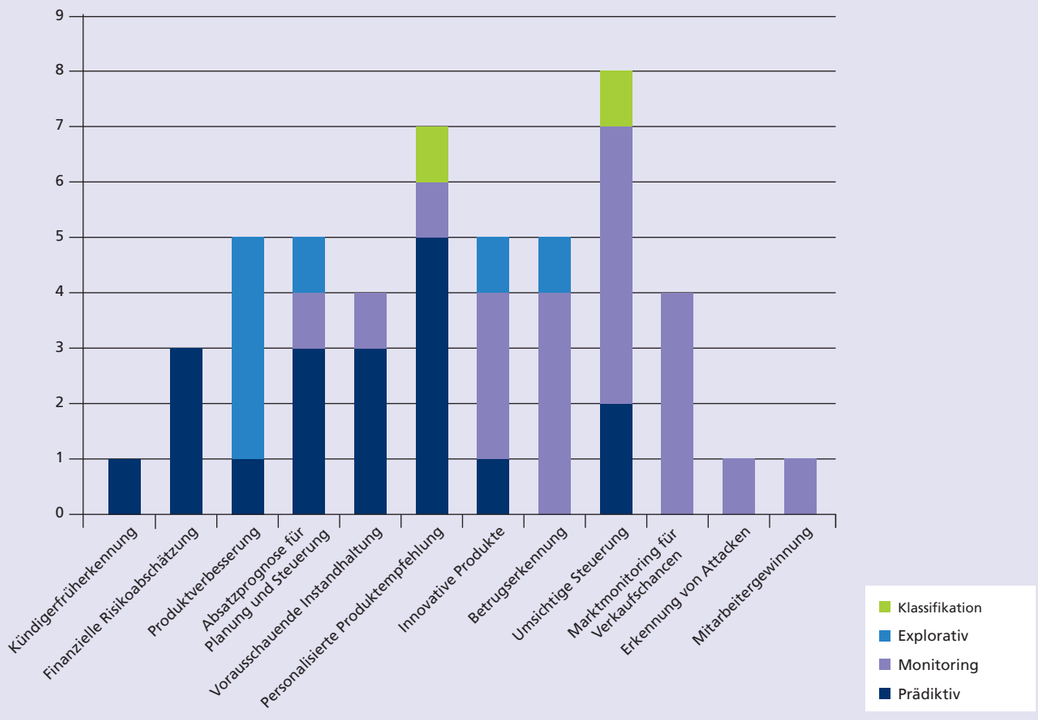


Tabelle 6 bezieht die Analysemethode auf Aktualität und Reaktionsweise. Alle Anwendungsfälle für Analysen bei Bedarf sind explorativ. Davon abgesehen gibt es alle Kombinationen von Monitoring bzw. prädiktiver Analyse mit automatischer bzw. nicht automatischer Reaktion in laufend bzw. täglich durchgeführten Anwendungen.

Tabelle 6: Aktualität, Reaktion und Analysetyp

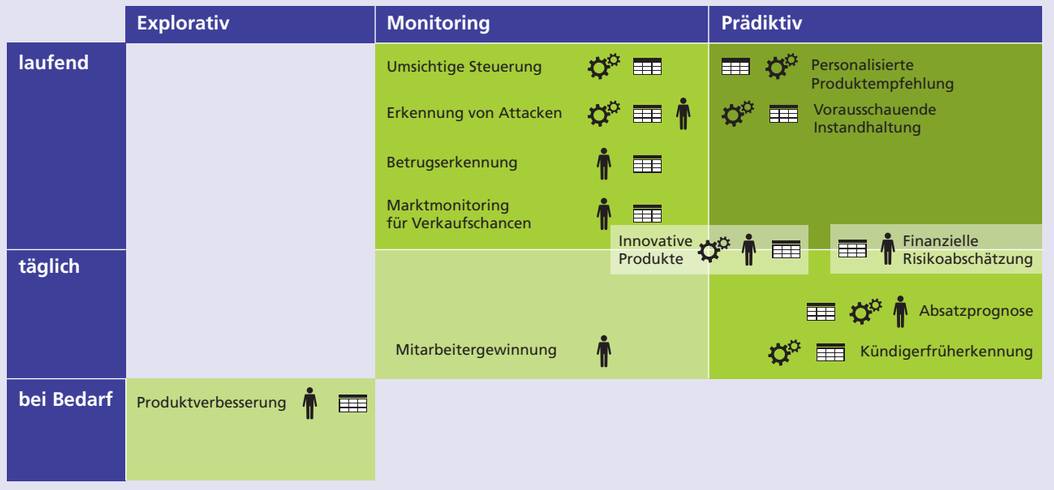
Aktualität		nicht automatisierte Reaktion	automatisierte Reaktion	
laufend				
Prädiktiv	6			4
Monitoring	10			6
Explorativ	1			
Klassifikation				
täglich				
Prädiktiv	4			4
Monitoring	4			1
Explorativ	1			
Klassifikation	1			1
bei Bedarf				
Prädiktiv				
Monitoring				
Explorativ	6			
Klassifikation				

3.2 Gruppen ähnlicher Anwendungsfälle

In diesem Abschnitt werden die Kategorien genutzt, um Gruppen ähnlicher Fälle zu bilden und zu charakterisieren. Die fachlichen Aufgabenstellungen wurden so gewählt, dass die Anwendungsfälle möglichst viele Gemeinsamkeiten aufweisen. Das betrifft insbesondere die Analysemethode und Aktualität, oft auch die Datentypen, seltener die Art der Reaktion. Wie Abbildung 11 verdeutlicht, liegen die meisten Gruppen mit ihren Schwerpunkten im laufenden Monitoring, während prädiktive Analysen laufend bis täglich stattfinden. Explorative Analysen bei Bedarf sind typisch für die Verbesserung von Produkten und Dienstleistungen. Obwohl diese Gruppe nicht die Aktualität der Daten ausnutzt, profitiert sie davon, dass verschiedenartige Daten in großen Mengen analysierbar werden. So können historische Daten für neue Fragestellungen genutzt werden, die sich erst nach der Speicherung der Daten ergeben.

Die Abbildung zeigt auch die vorherrschend verwendeten Datentypen für jede Aufgabengruppe. Bei den meisten Gruppen dominieren nicht automatische Reaktionen, aber das liegt vermutlich an Beispielen, die nur Produkte und nicht ihren Einsatz beim Anwender beschreiben.

Abbildung 11: Aufgaben, Aktualität, Analysemethode und Art der Datenquellen



3.2.1 Personalisierte Ansprache von Kaufinteressenten

Für die Umsatzsteigerung im Handel sind personalisierte Produktempfehlungen eine Art der Direktwerbung, die persönliche Daten über den Kunden nutzen. Typischerweise werden Geschäftsdaten (Kunden- und Verkaufsdaten) genutzt und nach Möglichkeit mit Daten aus anderen Quellen kombiniert: Sozialen Netzwerken, Online Spielen, Aufenthaltsdaten etc. Wichtig ist nicht nur, das passende Produkt anzubieten, sondern auch zur passenden Gelegenheit. Es handelt sich meist um laufende prädiktive Analysen mit der Kundenansprache als automatische Reaktion.

Tabelle 7. Typische Eigenschaften von Beispielen für die personalisierte Produktempfehlung

Aufgabe	Personalisierte Produktempfehlung
Unternehmensbereich	Marketing
Sektor	Handel, IKT
Ziel	Umsatzsteigerung
Automatische Reaktion	meist ja
Arten von Datenquellen	Geschäftsdaten, auch kombiniert mit Maschinendaten oder menschengeneriertem Content
Aktualität	meist laufend
Analysemethode	meist prädiktiv
Anwendungsfälle und Stichwörter	❙ Foursquare – Standortbezogenes Social Network ❙ BigPoint (GameAnalytics) – Game Analytics ❙ VMware (Lattice Engines) – sales intelligence ❙ Tesco – Next best offers ❙ Sense Networks – Geo-retargeting ❙ LinkedIn – Social Network für Berufstätige ❙ Opera Solutions – Automatische 1:1 Zielgruppenansprache

Beispiel Sense Networks: Sense Networks sammelt und analysiert Milliarden von Ortsdaten von mobilen Telefonen, Taxis, Kameras und GPS-Geräten in Echtzeit. Die eingesetzte Software gleicht nach eigenen Angaben die Koordinaten mit Geschäftsstandorten und demographischen Daten ab und ordnet den Straßenabschnitten Nutzerprofile zu. Dabei werden mehr als 2000 Attribute genutzt. Die Citisense App informiert, wo sich viele Menschen befinden. Beim Geotargeting können Personen sehr gezielt beworben werden. Beispiel: Personen, die sich drei Mal pro Woche im Umkreis von einem Kilometer von einem bestimmten Geschäft befinden, können angesprochen werden, kurz vor oder nach einem Besuch in dem Geschäft.

Kundenerhalt bedeutet Kostenersparnis in allen Branchen. Das Risiko einer Kündigung oder Abwanderung kann durch prädiktive Analysen in täglichen Geschäftsprozessen ermittelt werden, auf die automatisch oder nicht reagiert werden kann. Genutzt werden Geschäftsdaten (Kunden- und Verkaufsdaten), wenn möglich kombiniert mit Kundenverhaltensdaten aus anderen Quellen.

Tabelle 8. Typische Eigenschaften von Beispielen für die Erkennung von Kündigungsrisiken

Aufgabe	Personalisierte Produktempfehlung
Unternehmensbereich	Marketing
Sektor	alle
Ziel	Einsparung von Kosten
Automatische Reaktion	nein
Arten von Datenquellen	Maschinendaten kombiniert mit Geschäftsdaten
Aktualität	meist täglich
Analysemethode	meist prädiktiv
Anwendungsfälle und Stichwörter	❙ Telekom Italia (HP) – Verbindungsdatenanalyse

Beispiel Telecom Italia: Die Firma Telecom Italia wollte potenzielle Abwanderer rechtzeitig erkennen und durch gezielte Angebote an das Unternehmen binden. Dazu wurden nach eigenen Angaben die Verbindungsdaten der Kunden analysiert, eng vernetzte Benutzer geclustert, ihre Rolle im Kommunikationsnetz und ihre Verbindung zu Kündigern oder Nutzern aus anderen Telekommunikationssnetzen analysiert. Im Ergebnis wird ein Score für das Risiko eines Anbieterwechsels für jeden Kunden errechnet. Außerdem sollen im Kommunikationsnetz bestehender Kunden latente Neukunden identifiziert und angeworben werden.

3.2.2 Personalisierte Ansprache von (potenziellen) Mitarbeitern

Fehlendes qualifiziertes Personal ist ein Kostenfaktor. Für die Personalsuche bieten sich dem Management Monitoringdienste an, die in sozialen Medien nach unzufriedenen Fachleuten suchen. Täglich ausgewertet können die Betroffenen frühzeitig persönlich angesprochen werden (also ohne automatische Reaktion).

Tabelle 9. Typische Eigenschaften von Beispielen für die Mitarbeitergewinnung

Aufgabe	Personalisierte Produktempfehlung
Unternehmensbereich	Marketing
Sektor	viele
Ziel	Einsparung von Kosten
Automatische Reaktion	nein
Arten von Datenquellen	Content
Aktualität	täglich
Analysemethode	Monitoring
Anwendungsfälle und Stichwörter	❙ Entelo – Recruiting

Beispiel Entelo: Die Firma hat einen Algorithmus entwickelt, um Anwerber in Technologie-Unternehmen mit potenziellen neuen Mitarbeitern zusammenzubringen. Entelo erstellt nach eigener Aussage über 300 Millionen persönliche Profile aus den Aktivitäten von Personen auf Webseiten wie Github, Stackoverflow, Dribble oder Behance. Der Algorithmus priorisiert zum Beispiel Ingenieure, wenn deren Firmen schwierige Phasen durchlaufen. Das Ranking würde zum Beispiel weiter steigen, wenn der Ingenieur gerade Code gepostet hat.

3.2.3 Marktmonitoring für Verkaufschancen

Für die Umsatzsteigerung im Handel ist es wichtig, dass das Marketing Verkaufschancen frühzeitig erkennt. Dazu wird das Netz laufend überwacht: auf Veränderungen bei Konkurrenz und Zulieferern, Preisveränderungen in Online Märkten, auf Eilnachrichten und ähnliches. Wenn automatisch reagiert wird, wie bei der dynamischen Preisanpassung, dann müssen neben Daten aus dem Netz auch Geschäftsdaten verarbeitet werden.

Tabelle 10. Typische Eigenschaften von Beispielen für das Marktmonitoring für Verkaufschancen

Aufgabe	Marktmonitoring für Verkaufschancen
Unternehmensbereich	Marketing
Sektor	Handel
Ziel	Umsatzsteigerung
Automatische Reaktion	unterschiedlich
Arten von Datenquellen	Content und Geschäftsdaten
Aktualität	meist laufend
Analysemethode	meist Monitoring
Anwendungsfälle und Stichwörter	❙ Cookie's (Mercent) – Dynamic Pricing ❙ DataMinr –Signale für den Hochfrequenzhandel ❙ CTOsOnTheMove – Sale Leads ❙ Profitero – Competitive Intelligence, Live Market Intelligence

Beispiel DataMinr: DataMinr beliefert Finanzfirmen mit Eilmeldungen und versucht dabei, schneller als herkömmliche Nachrichtendienste zu sein. Die Nachrichten gewinnt die Firma nach eigenen Angaben, indem sie in Echtzeit anomale Ereignisse in Twitter erkennt und mit Themen aus deren entwickelter Wissensbasis korreliert. DataMinr reichert die Nachrichten mit analytischen und statistischen Informationen aus der eigenen Unternehmens- und Branchendatenbank an, die für die Investitionsinteressen des Kunden relevant sein könnten.

3.2.4 Optimierung des Betriebs

Absatzprognosen sind eine wichtige Information für eine datenbasierte Planung und Entscheidungsfindung und damit zusammenhängende Ziele. Speziell zu nennen sind Marketing und Disposition im Handel. Für die Prognose werden Geschäftsdaten (speziell Transaktionsdaten) sowie manchmal zusätzlich Maschinendaten (Log- und Messdaten) analysiert. Ähnlich wie bei der Absatzprognose für die Standortplanung im Handel gibt es Leistungsprognosen für die Standortwahl bei Windturbinen. Die Reaktion erfolgt meist nicht vollautomatisch.

Tabelle 11. Typische Eigenschaften von Beispielen für die Absatzprognose für Planung und Steuerung

Aufgabe	Absatzprognose für Planung und Steuerung
Unternehmensbereich	Management, Marketing
Sektor	vorwiegend Handel
Ziel	insbesondere datenbasierte Planung und Entscheidung Erhöhung der Produktivität
Automatische Reaktion	nein
Arten von Datenquellen	Geschäftsdaten, auch kombiniert mit Maschinendaten oder Content
Aktualität	meist täglich
Analysemethode	meist prädiktiv
Anwendungsfälle und Stichwörter	❙ 1010data – Trade-Area Analysis ❙ Telefonica Dynamic Insights – foot fall counting ❙ OTTO (BlueYonder) – Prognose der Abverkäufe auf Einzelproduktebene ❙ Quelle Russland (BlueYonder) – Out-of-Stock Analysis ❙ Vestas (IBM) – Leistungsprognose für Windturbinen

Beispiel OTTO: Der Versandhändler OTTO hat eine Software eingeführt, anhand derer sich der Absatz auf Artikelgrößenebene präziser und zum Teil weit vor dem Einsatz eines Artikels vorhersagen lässt. Jährlich werden nach Angaben des Herstellers 40 Millionen Einzelprognosen erstellt, die anhand einer Vielzahl von Faktoren berechnet werden. Dafür fließen wöchentlich bis zu 100 Millionen Datensätze in das System. Für die Prognose spielen Faktoren wie der Bewerbungsgrad eines Artikels online und offline sowie spezifische Artikeleigenschaften und Umfeldbedingungen eine entscheidende Rolle. Dabei lernt die Software, dass sich die Bedeutung von Einflussfaktoren wie Ansprache und Artikel-Ranking, oder im Printbereich Seitenanteil und Katalogausstoßmenge im Artikellebenszyklus ändert.

Ein weiteres Instrument der Betriebsoptimierung ist die vorausschauende Instandhaltung. Dadurch können in der Industrie Kosten eingespart oder die Produktivität erhöht werden. Das betrifft vor allem den Produktionssektor, aber zum Beispiel auch Logistiker. Um die Maschinen(teile) zu warten oder zu ersetzen bevor sie ausfallen, werden Maschinendaten (Log- und Messdaten) laufend erfasst und mit Geschäftsdaten kombiniert. Die Reaktion erfolgt meist nicht vollautomatisch.

Tabelle 12. Typische Eigenschaften von Beispielen für die vorausschauende Instandhaltung

Aufgabe	Vorausschauende Instandhaltung
Unternehmensbereich	meist Produktion
Sektor	vorwiegend Produktion
Ziel	Erhöhung der Produktivität, Einsparung von Kosten
Automatische Reaktion	meist nein
Arten von Datenquellen	Maschinendaten, meist mit Geschäftsdaten
Aktualität	laufend
Analysemethode	prädiktiv
Anwendungsfälle und Stichwörter	❙ Rolls Royce – Power by the hour ❙ McLaren – Preventive Maintenance ❙ Vimana (System Insights) – Preventive Maintenance ❙ US Xpress – Flottenoptimierung

Beispiel US Xpress: Das Logistikunternehmen verfolgt das Ziel, alle Fahrzeuge in Bewegung zu halten und insbesondere Wartungszeiten zu minimieren. Dazu werden nach Angaben des Dienstleisters 900 Datenelemente von Zehntausenden Fahrzeugen (Reifen- und Spritverbrauch, Maschinendaten, GPS) und Fahrer-Feedback aus Social Media in Echtzeit gesammelt und parallel in verschiedene operationelle und analytische Systeme übertragen.

Ganz allgemein können zur Einsparung von Kosten und zur Erhöhung der Produktivität die Prozesse laufend überwacht werden, damit man unmittelbar eingreifen kann. Anwendungsfälle sind Kameradaten für die Personaleinsatzplanung im Handel, die Überwachung von der Bohrung und Förderung von Ölfeldern, die Überwachung von Stromnetzen und die Steuerung von unproblematischen Anträgen in Fast Tracks ohne personal- und kostenintensive gründliche Prüfungen. Die dabei gewonnenen Daten können genutzt werden, um Szenarien mit Simulationen mit prädiktiven Analysen durchspielen zu lassen. Die Reaktion erfolgt meist nicht vollautomatisch.

Tabelle 13. Typische Eigenschaften von Beispielen für die umsichtige Steuerung

Aufgabe	Umsichtige Steuerung
Unternehmensbereich	Marketing, Produktion
Sektor	meist Handel, Industrie
Ziel	Erhöhung der Produktivität, Einsparung von Kosten
Automatische Reaktion	nein
Arten von Datenquellen	meist Maschinendaten, teils kombiniert mit Geschäftsdaten
Aktualität	laufend
Analysemethode	meist Monitoring, seltener prädiktiv
Anwendungsfälle und Stichwörter	❙ Darre´ (Visapix), Adidas Group (Experian) – In-store behaviour analysis ❙ Chevron – Digital oilfield ❙ GE Energy– Smart energy, Demand Response Optimization ❙ GE Energy – Lokalisierung von Stromausfällen ❙ Infinity Insurance – Fast Track Bearbeitung ❙ Autogrid – Smart Grid Analytics, Demand-Response Management ❙ SAP – Verfolgung von Lieferketten durch RFIDs

Beispiel GE Energy: Dieses Angebot ist Teil eines Systems zur optimierten Steuerung des Energieverbrauchs. Es benutzt 2-Wege-Kommunikation mit smarten Zählern, programmierbaren vernetzten Thermostaten, smarten Anwendungen, elektrischen Fahrzeugen und Ladestationen sowie Haushalts-Energieportalen. Die Daten werden laufend gesammelt und für lernende Algorithmen verwendet. Außer historischen Daten nutzen sie die aktuelle Load-Prognose, Preis- und Wetterdaten und prognostizieren den Verbrauch für verschiedene Szenarien, nach denen man durch ausgewählte Angebote an ausgewählte Verbraucher für eine bestimmte Zeit Verbrauchsspitzen verschieben kann.

3.2.5 Finanzielle Risiken und Betrug

Im Finanzsektor, bei Banken und Versicherungen lassen sich Kosten einsparen, wenn Schäden durch Betrug reduziert werden. Die Anwendungsfälle zeigen verschiedene Ansätze zur Erkennung von Kreditkartenbetrug. Dazu werden laufend Geschäftsdaten (Transaktionen), oft auch Online-Quellen überwacht. Da die Anwendungsfälle von Anbietern und nicht von Anwendern stammen, fehlt die Einbindung in die Geschäftsprozesse und damit eine automatische Reaktion.

Tabelle 14. Typische Eigenschaften von Beispielen für die Betrugserkennung

Aufgabe	Betrugserkennung
Unternehmensbereich	
Sektor	Banken und Versicherungen
Ziel	Einsparung von Kosten
Automatische Reaktion	meist nein
Arten von Datenquellen	meist Geschäftsdaten, oft auch Content
Aktualität	meist laufend
Analysemethode	Monitoring
Anwendungsfälle und Stichwörter	❙ Netcraft – AntiPhishing ❙ Billguard – Überwachung von Kreditkartenabrechnungen ❙ Paymint – Betrugserkennung bei Kreditkartentransaktionen ❙ Digital Reasoning – Compliance Monitoring ❙ Fraunhofer IGD – Kreditkartenbetrug

Beispiel Netcraft: Mit dem Dienst der Firma Netcraft können insbesondere Banken und Finanzunternehmen ihre Namen, Marken, Handelsmarken und Slogans im Internet verfolgen, um Phishing-Angriffe und Betrugsversuche durch Vorspielen einer falschen Identität zu identifizieren. Dazu nutzt Netcraft die Host- und Domänennamen von über 200 Millionen Websites und Inhalte der Startseiten von über 50 Millionen Websites zum Vergleich. Über eine im Browser installierbare Leiste kann man Betrugsversuche melden. Auf diese Weise konnten nach eigenen Angaben bis September 2012 über fünf Millionen Phishing Sites identifiziert werden.

Bei Banken und Versicherungen lässt sich der Umsatz steigern, wenn man Risiken besser prognostizieren kann. Laufend oder täglich werden dazu Geschäfts- oder Maschinendaten analysiert. Die Reaktion erfolgt häufig vollautomatisch.

Tabelle 15. Typische Eigenschaften von Beispielen für die Risikoabschätzung

Aufgabe	Finanzielle Risikoabschätzung
Unternehmensbereich	
Sektor	Banken und Versicherungen
Ziel	Steigerung des Umsatzes
Automatische Reaktion	eher ja
Arten von Datenquellen	Geschäftsdaten oder (auch) Maschinendaten
Aktualität	laufend bis täglich
Analysemethode	prädiktiv
Anwendungsfälle und Stichwörter	❙ United Overseas Bank – Risikoprognose ❙ The Climate Corporation – Micro-Predictions ❙ QuantConnect – Algorithmic Trading

Beispiel Climate Corporation: Dieser relativ neue Dienstleister versichert Risiken in der Landwirtschaft. Dazu prognostiziert die Firma Risiken und ermittelt aktuelle Schäden auf Mikroebene für 20 Millionen Farmen in den USA. Climate Corporation sammelt Temperatur-, Niederschlags-, Bodenfeuchtigkeits- und Ertragsdaten über ein großes Sensornetz. Die Auszahlungen berechnen sich automatisch danach, ob diese Faktoren pro Tag und Farm eine Schwelle überschreiten, bei der die Ernte gefährdet ist.

3.2.6 Erkennung von Attacken

In dieses Anwendungsfeld gehört die laufende Überwachung von realen Anlagen, Infrastrukturen, aber auch IKT-Anlagen und -netzen, um Attacken zeitnah zu erkennen. Analysiert werden vor allem Maschinendaten (Mess-, Log-, Kameradaten), die Reaktion erfolgt zumeist nicht automatisch.

Tabelle 16. Typische Eigenschaften von Beispielen für die Überwachung von Arealen und Infrastruktur

Aufgabe	Erkennung von Attacken
Unternehmensbereich	
Sektor	alle
Ziel	Einsparung von Kosten
Automatische Reaktion	nein
Arten von Datenquellen	Maschinendaten. Geschäftsdaten und auch Content
Aktualität	laufend
Analysemethode	Monitoring
Anwendungsfälle und Stichwörter	❙ New York (Microsoft) – Domain Awareness

Beispiel New York City: Das »Domain Awareness System« in New York dient der Verbrechensvorbeugung und Terrorismusabwehr. Es erhält nach eigenen Angaben Informationen von ca. 3000 Straßenkameras, 2600 mobilen Strahlungsmessern der Polizei und 100 Kennzeichenlesern an Straßen und in Polizeiautos. Das System kann mit menschlicher Unterstützung verdächtige Aktivitäten erkennen und Alarme in Echtzeit generieren. Außerdem bietet es Zugriff auf das Verbrechensregister, Notrufe, Kennzeichenregister und andere Datenquellen.

3.2.7 Innovative oder verbesserte Produkte und Dienstleistungen

Unternehmen in Industrie, Dienstleistungssektor, Banken und Versicherungen arbeiten in ihren Forschungs- und Entwicklungs-Abteilungen daran, innovative Produkte zu entwickeln. Zur Unterstützung bieten sich verschiedene Quellen von Big Data an: Nutzungs- und Verhaltensdaten, Trends in Social Media oder die Überwachung von Online-Portalen. Meist werden Monitoring oder prädiktive Analysen eingesetzt, die laufend oder täglich durchgeführt werden. Die Reaktion kann automatisch oder nicht automatisch erfolgen. In diese Gruppe lassen sich auch Rolls Royce (vorausschauende Überwachung) und The Climate Corporation (Prognose von Versicherungsrisiken) einordnen.

Tabelle 17. Typische Eigenschaften von Beispielen für innovative Produkte durch Nutzung neuer technischer Möglichkeiten

Aufgabe	Innovative Produkte durch Nutzung neuer technischer Möglichkeiten
Unternehmensbereich	Forschung und Entwicklung, Produktion
Sektor	Industrie, Dienstleistung, Banken und Versicherungen
Ziel	Entwicklung neuer Produkte
Automatische Reaktion	unterschiedlich
Arten von Datenquellen	Maschinendaten oder Content, oft kombiniert mit Geschäftsdaten
Aktualität	laufend bis täglich
Analysemethode	Monitoring oder prädiktiv
Stichwörter	❙ TuneSat – Audio fingerprints, music recognition ❙ Usage-based pricing ❙ Progressive – Usage-based Insurance ❙ Bild (Unbelievable Machines) – Social Media Storyteller ❙ Nestlabs – Smart Thermostats ❙ Nike – Quantified self ❙ The Climate Corporation – Micro-Predictions ❙ Rolls Royce – Power by the Hour, Preventive Maintenance

Beispiel Nike: Der Sportartikelhersteller integriert zunehmend in seine Sportprodukte Sensorik, die die Benutzung der Artikel beim Endkunden misst. Die Kunden können ihre Trainingsdaten auf der Community-Webseite Nikeplus.com hochladen, um zum Beispiel ihre Technik zu verbessern oder sich mit anderen Läufern zu messen. Nike möchte damit die Kunden besser an die Marke binden und so die Umsätze erhöhen. Längerfristig sollen die Daten auch für personalisierte Werbung beim Betreten einer Filiale oder per Mobiltelefon genutzt werden.

Die Erhöhung der Produkt- und Servicequalität ist ein weiteres Ziel der Forschungs- und Entwicklungs-Abteilungen, speziell im Dienstleistungs- und Industriesektor. Dazu können Meinungen und Feedback aus sozialen Medien gewonnen oder große Content-Datenbanken durchforstet werden. Meist werden explorative Analysen bei Bedarf durchgeführt. Die meisten Anwendungsfälle in dieser Gruppe stammen aus der evidenzbasierten Medizin. Ähnlich gibt es in der Industrie das Konzept des value-driven Designs.

Tabelle 18. Typische Eigenschaften von Beispielen für die Verbesserung von Produkten und Dienstleistungen

Aufgabe	Verbesserung der Produkte und Dienstleistungen
Unternehmensbereich	Forschung und Entwicklung, Produktion
Sektor	vorwiegend Dienstleistung (Medizin) und Industrie
Ziel	Erhöhung der Produkt- und Servicequalität
Automatische Reaktion	nein
Arten von Datenquellen	Content, teils kombiniert mit Geschäftsdaten
Aktualität	meist bei Bedarf
Analysemethode	meist explorativ
Beispiele und Stichwörter	❙ Domino`s Pizza – value-driven design, design-to-value ❙ Treato – Social Network für Patienten ❙ HealthNow (Informatica Data Services) – evidenzbasierte Medizin, comparative effectiveness research, evidence-based decision making, rapid learning healthcare ❙ Sloan Cancer Center – comparative effectiveness research, evidence-based decision making, rapid learning healthcare ❙ Linguamatics – Patent Mining

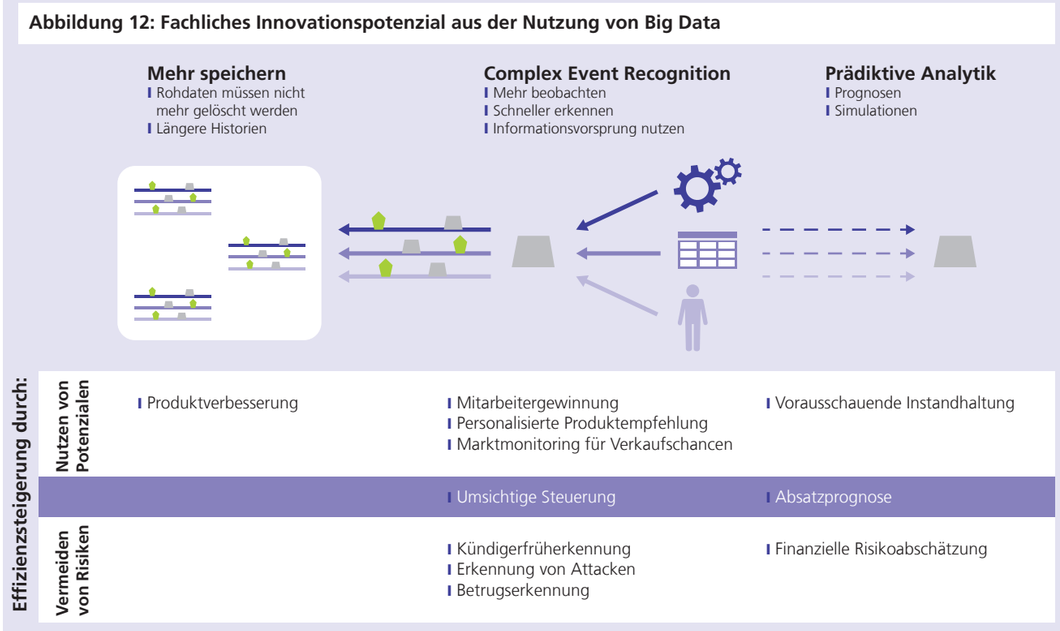
Beispiel Treato: Im Internet tauschen Patienten und Ärzte in Tausenden von Foren und Blogs ihre Erfahrungen mit Behandlungsmethoden von Krankheiten aus. Die Firma Treato speichert und analysiert nach eigenen Angaben bis zu 200 Millionen Posts pro Tag. Die Posts bestehen aus einer Mischung aus Text und Zahlen und sind sprachlich stark unterschiedlich. Sie enthalten nur wenig Fachterminologie, müssen aber in Zusammenhang mit fachlicher Terminologie gebracht werden. Die Analysen helfen zum Beispiel dabei herauszufinden, welche Nebenwirkungen eines Medikaments Patienten am häufigsten beschreiben oder mit welcher Behandlung sie bei vergleichbaren Erkrankungen zufrieden sind. Bisher dauerte es oft mehrere Jahre, bis Behörden ein Medikament mit starken Nebenwirkungen vom Markt genommen haben oder bis ein erfolgreiches Medikament für bis dahin nicht bekannte Krankheitsbilder zum Einsatz kam.

3.2.8 Zusammenfassung

Wie kann die Nutzung von Big Data zu innovativen Anwendungen führen?

Die folgende Grafik versucht die wesentlichen Erkenntnisse aus der Fallanalyse zusammenzufassen. In der oberen Hälfte sind die neuen Möglichkeiten skizziert, die sich aus der Technologie ergeben: Daten aus unterschiedlichen Quellen lassen sich schnell analysieren, um komplexe Ereignisse schnell zu erkennen. Die erfassten Daten werden länger gespeichert als früher. Damit wächst die Datenmenge, aus der man bei Bedarf immer neues Wissen generieren kann, sei es durch manuelle Exploration oder automatische Mustererkennung. Die grünen Formen in der Datenbank links deuten Anwendungsfälle an, die bei der Speicherung noch gar nicht als solche erkannt wurden. Erst durch das Auftreten eines neuen zu behandelnden Phänomens würden diese Anwendungsfälle wichtig. Insbesondere werden aus den gespeicherten Daten Regeln zur Definition der komplexen Ereignisse extrahiert, die beim Monitoring erkannt werden sollen, oder prädiktive Modelle zur Prognose und Simulation von Szenarien. Diese erlauben, die Zukunft besser einzuschätzen.

Abbildung 12: Fachliches Innovationspotenzial aus der Nutzung von Big Data



Die untere Hälfte der Abbildung zeigt, wie die verschiedenen Gruppen von Anwendungsfällen die neuen Möglichkeiten des Monitorings, der Vorhersage oder Exploration nutzen, um Chancen zu ergreifen, Risiken zu vermeiden oder einfach nur datenbasiert zu entscheiden.

ONLINE-UMFRAGE ZUR NUTZUNG VON BIG DATA

4.1 Methodik

Im Rahmen der Innovationspotenzialanalyse sollte das Konzept »Big Data« Entscheidern aus kleinen und mittleren Unternehmen (KMU) vorgestellt werden, um den Stand der Nutzung sowie das Interesse an entsprechenden Technologien und Anwendungen abzufragen.

4.1.1 Hypothesen

Die Erhebung basierte konzeptionell auf folgenden Forschungs-Hypothesen:

- Den Unternehmen sind die Potenziale von Big Data oft noch nicht bekannt
- Die Bekanntheit der vorhandenen Fachbegriffe, Analysemethoden und der damit verbundenen Potentiale ist niedrig
- Ressourcen zum Arbeiten mit großen Datenmengen sind in der Mehrzahl der KMU noch nicht vorhanden
- Big Data ist primär ein Management-Thema, da es datenbasierte Entscheidungen erleichtert

4.1.2 Zielgruppe der Befragung

Aufgrund dieser Annahmen war es das Untersuchungsziel, Entscheider verschiedenster Unternehmensbereiche zum Thema Big Data zu befragen – und nicht wie bisher vornehmlich IT-Fachkräfte bzw. -Entscheider. Auch fokussierte die Befragung auf KMU, die eventuell mit Hürden bei der Einführung oder Umsetzung von Big Data zu kämpfen haben. Entsprechend sah das Studiendesign vor, dass die Beantwortung der gestellten Fragen selbst Entscheidern möglich sein musste, die bisher wenig oder nichts mit Big Data zu tun hatten, sich jedoch in Branchen aufhalten, die gemäß unserer Einschätzung von der Datenlage her prädestiniert sind. Ausgeschlossen waren somit Betriebe, in denen per se wenig Daten anfallen, wie etwa Handwerksbetriebe. Basierend auf dem Vorangegangenen erfolgte im Sinne der Gewährleistung von Praktikabilität und Expertenvalidität eine Befragung von verschiedensten Funktionsbereichs-Experten und nicht »nur« die Befragung von IT -Experten.

4.1.3 Grundsätzliches zum Studiendesign

Es wurde in Bezug auf die Inhalte und die Fragenformulierung ein Fragebogen konzipiert, der im Studiendesign die o.g. Anforderungskriterien (z.B. Verständlichkeit, Erläuterung des Big-Data-Begriffs vorab) sowie die gängigen psychologischen und marktforscherischen Anforderungen erfüllte. Dennoch war das Niveau des Fragebogens als anspruchsvoll einzustufen. Die Bearbeitungszeit wurde daher auf maximal zehn Minuten begrenzt. Vorlaufstudien ergaben auch niedrigere Bearbeitungszeiten, je nach Vorbildung der Probanden. Studienform war eine quantitative Online-Computerbefragung, da sich diese Form der Befragung für datenaffine Entscheider bewährt hat.

4.1.4 Rekrutierung

Über verschiedenste Medien erfolgte die Einladung an Entscheider zur Teilnahme, z.B. mittels postalischem Anschreiben, PR-Maßnahmen (Pressemitteilungen an Online-Fachportale und Entscheider-Medien) oder als Meldung auf Portalen von Kooperationspartnern des Instituts. Die breite Auswahl der Medien hatte zum Ziel, dass möglichst viele Unternehmen der Big-Data-Potenzialbranchen von der Studie erfuhren – d.h. die PR-Arbeit fokussierte auf die von Fraunhofer vorab identifizierten Potenzialbranchen.

4.1.5 Repräsentativitätsproblematik

Aufgrund der finalen Fallzahl n < 100 kann eine Repräsentativität jedoch nicht – und in keinem Fall innerhalb einzelner Branchen oder Antwortkategorien – erwartet werden. Die quantitative Befragung liefert eher Hinweise grundsätzlicher und qualitativer Art zur weiteren Exploration des Innovationspotenzials von Big Data. Um den qualitativen Aspekt der Online-Befragung zu unterstützen, inhaltlich zu füllen und zu validieren, fanden zusätzlich qualitative Expertenworkshops statt.

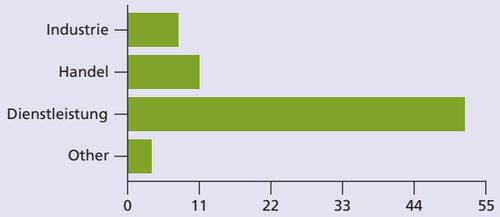
4.1.6 Ziel

Mit Hinblick auf den eingangs beschriebenen Forschungsbedarf sowie die Hypothesen und die Zielgruppe der Studie, wurden zunächst die Big Data Potenziale aus den recherchier-

ten und systematisierten Anwendungsfällen sowie die zur Hebung des Potenzials notwendigen und von Entscheiderseite gewünschten Maßnahmen, die zur Ableitung von Handlungsempfehlungen dienten, identifiziert.

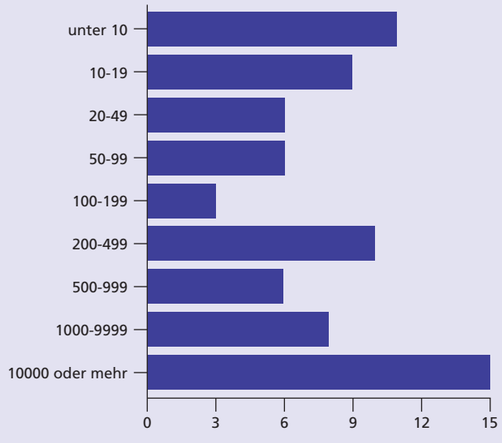
4.1.7 Statistische Angaben der Teilnehmer aus der Online-Umfrage

Abbildung 13: Herkunft der Teilnehmer an der Online-Umfrage nach Wirtschaftssektor



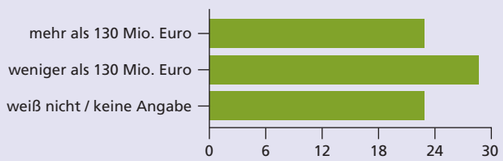
Die Unternehmensgrößen der Teilnehmer waren gut verteilt und reichten von sehr kleinen Unternehmen bis hin zu Großunternehmen. Im Vergleich mit den recherchierten Anwendungsfällen in Abbildung 3 fiel auf, dass nur wenige Anwendungsfälle aus der Dienstleistungsbranche identifiziert wurden, dagegen viele Fälle aus dem Handel, der Industrie und Banken und Versicherungen bekannt sind. Für eine stärkere Verbreitung von Big Data wäre es also künftig wichtig, mehr Beispiele für Big Data Anwendungen speziell für den Dienstleistungssektor bereit zu stellen.

Abbildung 14: Verteilung der Teilnehmer nach Unternehmensgröße



Die Umsätze der Teilnehmer aus den Unternehmen verteilten sich auf drei Klassen, wovon die Klasse »weniger als 130 Mio. Euro p.a.« schwach dominierte.

Abbildung 15: Verteilung der Teilnehmer nach Umsätzen

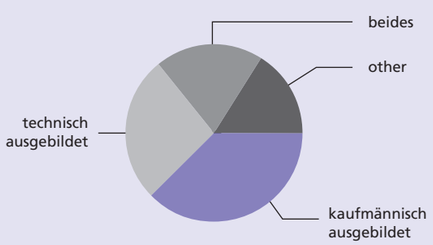


Der größte Teil der Teilnehmer konnte sich nicht eindeutig einem Funktionsbereich zuordnen. Die zweitgrößte Gruppe stammte aus dem Management, gefolgt von Produktentwicklung sowie Marketing und Vertrieb.

Abbildung 16: Verteilung der Teilnehmer nach Funktionsbereich

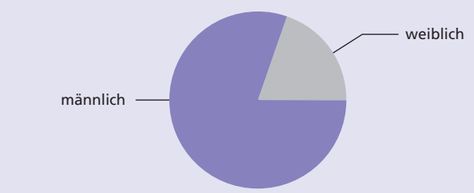
Hier fällt im Vergleich mit den recherchierten Anwendungsfällen auf, dass es viele Anwendungsfälle zur Produktion gab, aber kaum Personen mit Bezug zur Produktion an der Umfrage teilnahmen. Umgekehrt gab es viele Teilnehmer aus dem Management, aber relative wenige dedizierte Anwendungsfälle für das Management. Es ist jedoch anzunehmen, dass Manager an Fällen für alle Unternehmensbereiche interessiert sind. Beispiele und Teilnehmer im Bereich Marketing

Abbildung 17: Ausbildung



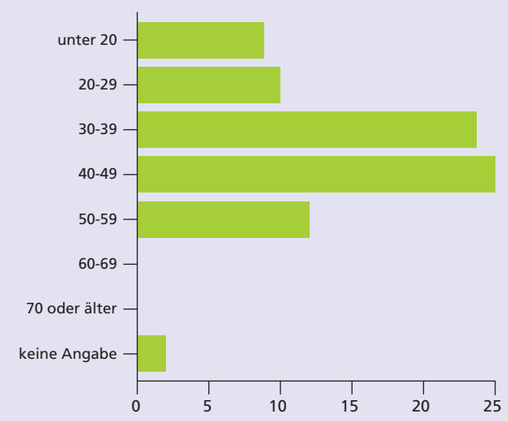
und Vertrieb sind gut vertreten. Es herrschte insgesamt eine annähernd ausgeglichene Verteilung zwischen technischer und kaufmännischer Ausbildung vor. Der mit Abstand größte Teil der Teilnehmer war männlich. Frauen waren mit ca. 1/4 vertreten.

Abbildung 18: Geschlecht



Die Altersgruppe 30 bis 49 war zahlenmäßig am häufigsten vertreten.

Abbildung 19: Alter



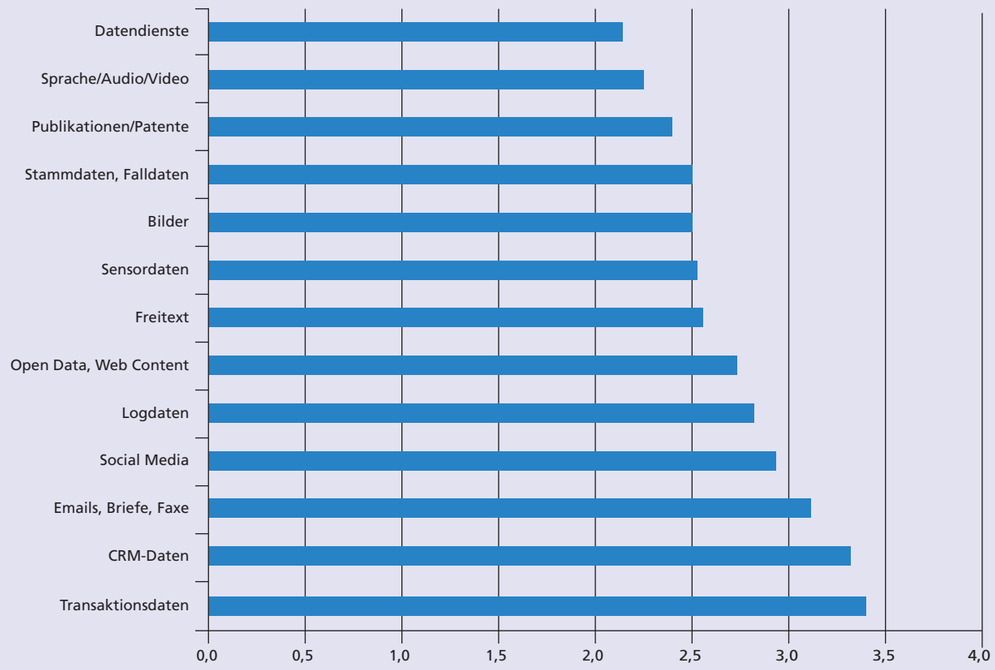
4.1.8 Bekanntheit und Verbreitung des Begriffs Big Data

Die befragten Unternehmen erfüllen zu 70 Prozent mindestens eines der drei Big Data-Kriterien (Umfang, Unterschiedlichkeit, Veränderung). Umgekehrt heißt das, dass 30 Prozent der befragten Unternehmen noch keine Big Data-Kandidaten sind. 20 Prozent der befragten Unternehmen erfüllen alle drei Big-Data-Kriterien gleichzeitig - und unterliegen damit hinsichtlich ihrer Datenlage einer hohen Dynamik. Jeder dritte Befragte kennt den Begriff Big Data bereits und bestätigte gleichzeitig, dass im Unternehmen entsprechende Projekte durchgeführt werden. Jeder fünfte Befragte kannte zum Beginn der Befragung den Begriff Big Data noch nicht.

4.1.9 Bedeutung einzelner Datenarten für den Unternehmenserfolg

Die Relevanz von Daten wurde hier wie folgt gewichtet: sehr wichtig = 4, wichtig = 3, eher unwichtig = 2, unwichtig = 1. Für das Ranking wurde im Sinne einer besseren Veranschaulichung der Durchschnitt berechnet.

Abbildung 20: Einschätzung: Wichtigkeit der Daten für den Unternehmenserfolg



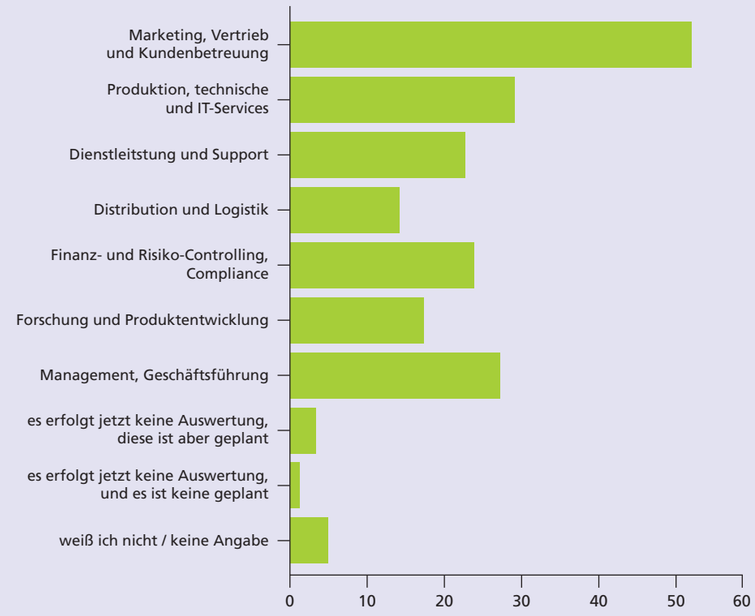
Auffällig: Bis zu zehn Prozent der Befragten stiegen bei der Bewertung nach der Wichtigkeit aus, antworteten an späterer Stelle aber weiter. Es handelte sich also nicht um Abbrüche, sondern die Frage wurde offenbar bewusst nicht beantwortet. Es drängt sich der Verdacht auf, dass das Wissen fehlt, um die Bedeutung der Daten abschätzen zu können. Hierbei wird auf die Ergebnisse der Zukunftsworkshops mit BranchenvertreterInnen verwiesen.

Aus Abbildung 20 geht hervor, dass in den meisten Anwendungsfällen Geschäftsdaten verarbeitet werden, zu denen die in den Antworten hocheingeschätzten Transaktionsdaten und CRM-Daten sowie die Stamm- und Falldaten zählen. Hoch eingeschätzt wurde ebenfalls Content, also E-Mails, Briefe, Fax, Social Media, Open Data Web Content, Bilder, Publikationen und Patente sowie Sprache/Audio/Video. Hier könnten die Befragungsteilnehmer Interesse an weiteren Anwendungsfällen haben.

4.1.10 Funktionsbereiche, die Daten mit hoher Veränderlichkeit auswerten

Speziell Marketing und Vertrieb werten Daten mit hoher Veränderlichkeit aus, gefolgt von Produktion und Management/Geschäftsführung.

Abbildung 21: Funktionsbereiche, die Daten mit hoher Veränderlichkeit auswerten

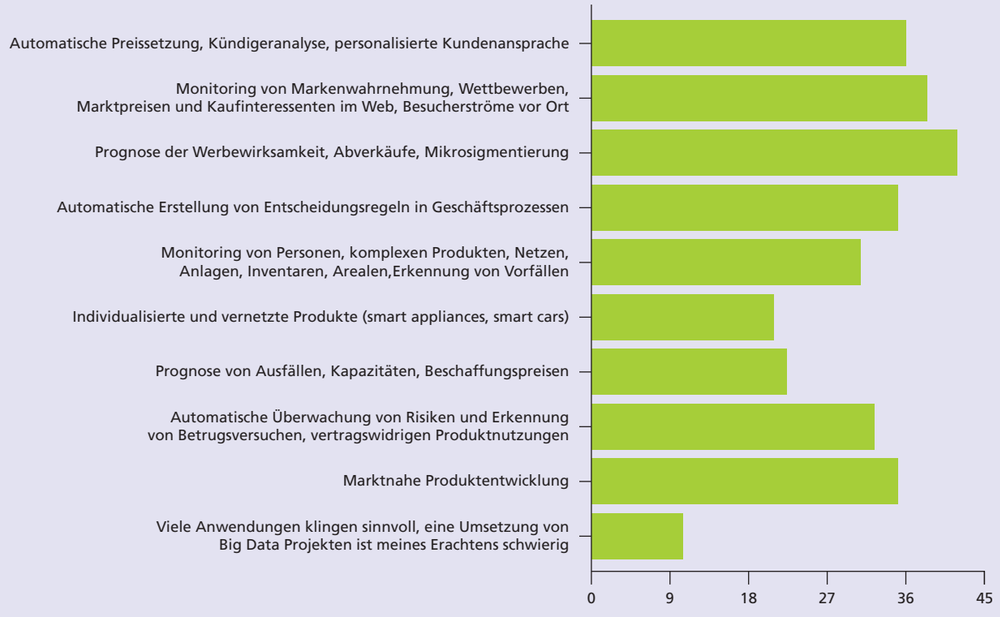


Dem Vergleich mit dieser Frage dient die Abbildung 6. Demnach finden sich mit Abstand die meisten Anwendungsfälle, in denen laufende Daten analysiert werden, in den Unternehmensbereichen Produktion, technische und IT-Services, in Marketing, Vertrieb und Kundenbetreuung sowie in Finanz-, und Risiko-Controlling und Compliance. Das passt zu den hier gegebenen Antworten. Diskrepanzen zu den Antworten in der Umfrage gibt es für Forschung und Produktentwicklung, sowie für Management und Geschäftsführung. Auch hier könnte es sich also lohnen, weitere Anwendungsfällen zu eruieren.

4.1.11 Big-Data-Anwendungsfälle

Insbesondere die Prognose der Werbewirksamkeit, Abverkäufe und Mikrosegmentierung, gefolgt von Monitoring von Markenwahrnehmung, Wettbewerbern, Marktpreisen, Kaufinteressenten im Web und Besucherströme vor Ort werden als interessante Big-Data-Anwendungsfälle eingestuft. Geringer ausgeprägt ist das gegenwärtige Interesse an der Prognose von Ausfällen, Kapazitäten, Beschaffungspreisen sowie individualisierten und intelligenten Produkten.

Abbildung 22: Interesse an Big-Data-Anwendungsfällen



4.1.12 Ziele von Big-Data-Anwendungen

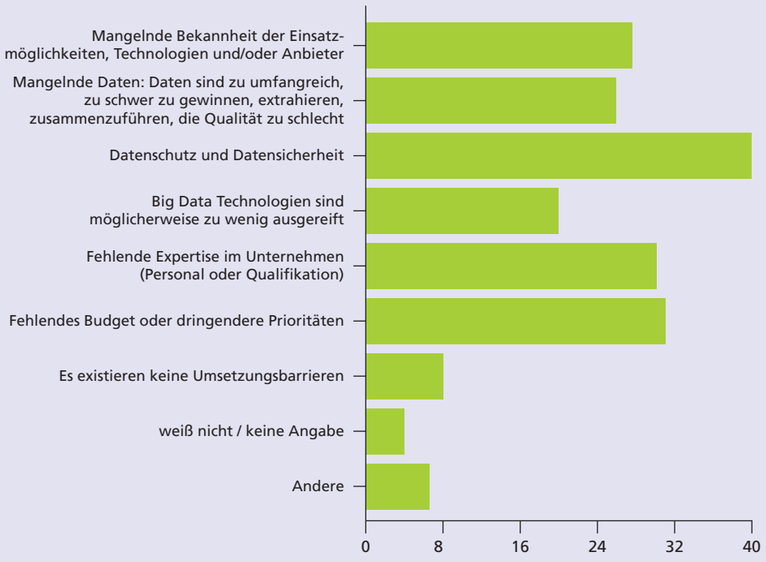
Am häufigsten zählt der Aufbau strategischer Wettbewerbsvorteile zu den Zielen, die mit Big Data erreicht werden sollen. Erst dann folgen weitere Ziele wie die Steigerung der Umsätze und die Einsparung von Kosten.

Abbildung 23: Ziele von Big-Data-Anwendungen

4.1.13 Barrieren für Big Data

Als wesentliche Barrieren für Big Data stechen der Datenschutz und die Datensicherheit heraus. Erst mit etwas Abstand folgen die mangelnde Bekanntheit der Einsatzmöglichkeiten, Technologien und/oder Anbieter sowie Expertise und Budget.

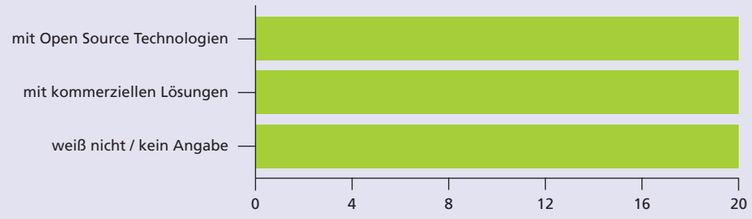
Abbildung 24: Barrieren für Big Data



4.1.14 Technologien

Zu einem Drittel setzen die Teilnehmer der Online Befragung Open Source und zu einem Drittel kommerzielle Entwicklungen ein. Die restlichen Teilnehmer sind sich unsicher, was eingesetzt wird oder geben in einem Freitext-Feld an, Eigenentwicklungen vorzunehmen.

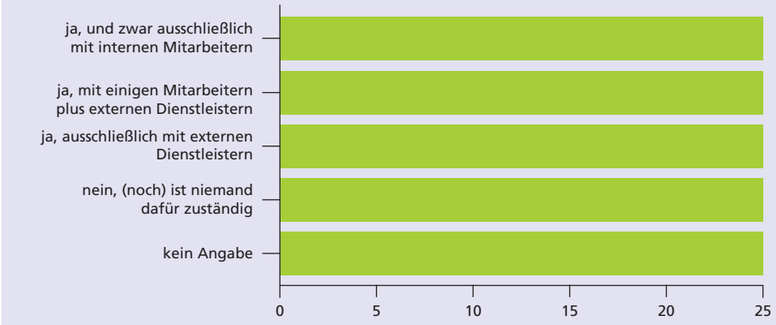
Abbildung 25: Technologien

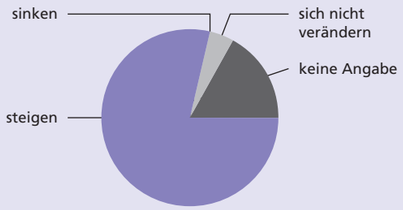


4.1.15 Personal für Big Data

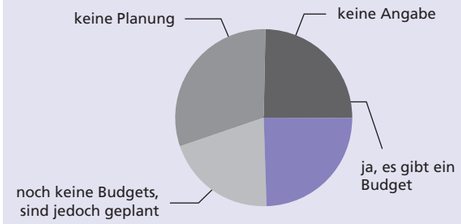
Der Umgang mit Big Data ist zumeist personell noch nicht organisiert. Falls dieser bereits organisiert ist, dann mit eigenen Mitarbeitern sowie mit externen Dienstleistern. Das deutet darauf hin, dass Expertise für Big Data zu wichtig ist, um sie ausschließlich externen Dienstleistern zu überlassen, jedoch Gründe bestehen müssen, externe Dienstleister einzubinden.

Abbildung 26: Personalressourcen für Big-Data-Aufgaben



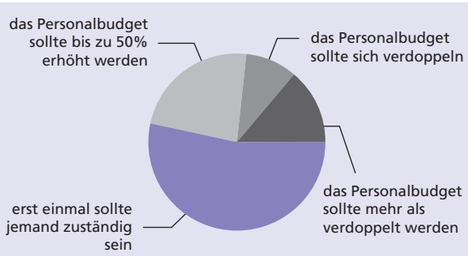


In den kommenden fünf Jahren sollte die Personalausstattung für das Aufgabengebiet Big Data in den Unternehmen deutlich steigen. Einige befragte Personen konnten keine Angabe machen. Niemand spricht sich für ein Sinken der Personalressourcen aus.

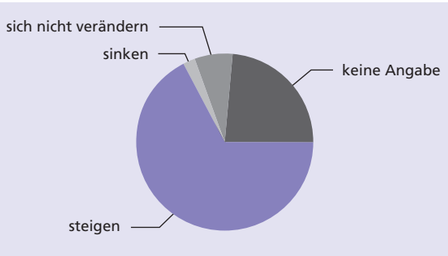


4.1.16 Budget

Etwa ein Viertel aller Unternehmen verfügt bereits über ein Budget für Big Data-Themen, d.h. Technologien, Analysen und Datenquellen. Ein knappes Viertel plant ein solches Budget, während die Mehrheit entweder keines plant oder keine Angabe macht.

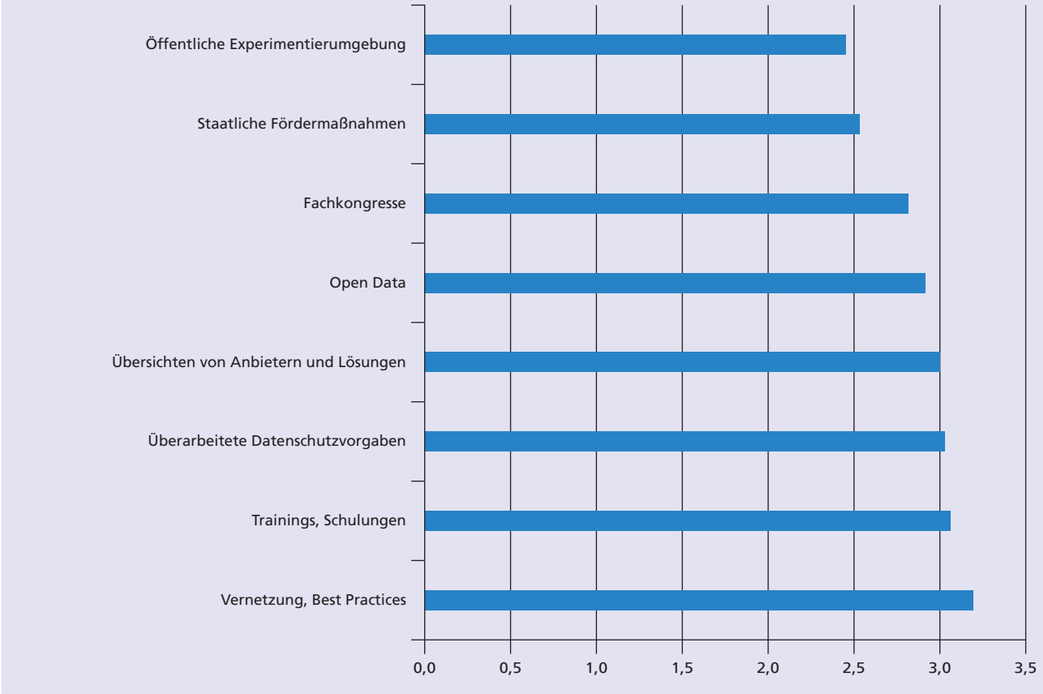


Wie es für ein neues Thema zu erwarten ist, scheint die personelle Zuordnung in vielen Unternehmen noch nicht klar geregelt zu sein. Falls dies bereits der Fall ist, würden die meisten ihr zuständiges Personal um mehr als 50 Prozent erhöhen.



In den kommenden fünf Jahren sollte sich für die absolute Mehrheit das Budget (außer Personal) für das Aufgabengebiet Big Data in Ihrem Unternehmen steigern.

Abbildung 27: Wichtigkeit der Förderungsmaßnahme:



4.1.17 Erwünschte Fördermaßnahmen

Die Wichtigkeit wurde hier gewichtet wie folgt: sehr wichtig = 4, wichtig = 3, eher unwichtig = 2, unwichtig = 1. Für das Ranking wurde im Sinne einer besseren Veranschaulichung der Durchschnitt berechnet.

Alle angebotenen Maßnahmen wurden als wichtig erachtet. Der Wunsch nach Förderung ist deutlich erkennbar. Insbesondere Best Practices und Trainings sind erwünscht.

4.2 Fazit der Online-Befragung

Die Online-Befragung hatte zum Ziel, das Verständnis über die Potenziale von Big Data bei KMU in Deutschland abzufragen. Anhand von Forschungs-Hypothesen wurde die Online-Befragung konzipiert. Nun sollen die Hypothesen anhand der vorliegenden Antworten diskutiert werden.

Hypothese 1: Den Unternehmen sind die Potenziale von Big Data oft noch nicht bekannt.

Die Online-Befragung zeigt, dass die Unternehmen mehrheitlich angeben können, welche Daten in ihrem Unternehmen vorliegen und wie sie deren Bedeutung für ihren Unternehmenserfolg einschätzen. Es wird deutlich, dass neben Daten aus Transaktions- sowie CRM-Systemen ein weites Spektrum an unstrukturierten und veränderlichen Daten in den Unternehmen existieren und ihre heutige Relevanz noch untergeordneter Natur ist. Die Teilnehmer können zudem benennen, welche Unternehmensbereiche Daten mit hoher Veränderlichkeit auswerten. Daraus zeichnet sich ab, dass insbesondere das Marketing und der Vertrieb sowie das Management und die Produktion zu den Big-Data-affinen Unternehmensbereichen zählen. Es ist den Teilnehmern weiterhin möglich, eine Bewertung potenziell relevanter Anwendungen auf Basis von Big Data in ihrem Unternehmen vorzunehmen. So werden Anwendungen als besonders relevant eingeschätzt, die eine Prognose der Werbewirksamkeit, Abverkäufe und Mikrosegmentierung ermöglichen, gefolgt von Anwendungen zum Monitoring von Markenwahrnehmung, Wettbewerbern, Marktpreisen und Kaufinteressenten im Web sowie Besucherströmen vor Ort. Auch die automatische Preissetzung, Kündigeranalyse sowie personalisierte Kundenansprache zählen zu den favorisierten Big-Data-Anwendungen.

Auf Basis dieser Antworten kann nicht davon ausgegangen werden, dass KMU die Potenziale von Big Data nicht bekannt sind. Es ist vielmehr davon auszugehen, dass eine Klärung des Begriffs Big Data zu Beginn hilfreich ist und die Nennung beispielhafter Datenquellen und Anwendungen den Transfer auf das eigene Unternehmen erleichtert.

Hypothese 2: Die Bekanntheit der vorhandenen Fachbegriffe, Analysemethoden und der damit verbundenen Potenziale ist niedrig.

Die Teilnehmer haben nach der Nennung relevanter Datenquellen und Anwendungen auf Basis von Big Data ihre Einschätzung abgegeben, welche Ziele mit Big Data in ihren Unternehmen bzw. ihrer Branche erreichbar sind. Mit deutlichem Abstand sehen sie das größte Potenzial von Big Data im Aufbau strategischer Wettbewerbsvorteile. Dann folgen Ziele der Steigerung der Umsätze und der Einsparung von Kosten. Die Erhöhung der Produktivität und die datenbasierte Planung und Entscheidungsfindung sind weitere Ziele, die mit Big Data erreichbar scheinen.

Diese Ergebnisverteilung macht deutlich, dass sich die Anwendung von Big Data nicht auf ein bestimmtes unternehmerisches Ziel kapriziert. Je nach Anwendung können auch mehrere Ziele gleichzeitig unterstützt werden. Eine Erhöhung der Produktivität kann in einer Branche zum Aufbau strategischer Wettbewerbsvorteile beitragen, wenn in dieser Branche insbesondere über die Produktivität konkurriert wird (z.B. Automobilbau). In anderen Branchen kann die Kosteneinsparung zu den besonders relevanten Zielen gehören, um an Wettbewerbsfähigkeit zu gewinnen (z.B. Handel). Aus diesen Antworten lässt sich für die Mehrheit der teilnehmenden Unternehmen erkennen, dass die Potenziale von Big Data nicht unbekannt sind. Vielmehr scheint es den Teilnehmern möglich zu sein, Big-Data-affine Unternehmensbereiche zu benennen, relevante Big-Data-Anwendungen zu bewerten sowie den daraus resultierenden Beitrag zu verschiedenen Zielsetzungen ableiten zu können. Das Potenzial von Big-Data-Anwendungen scheint vielfältiger Natur zu sein und zielt dennoch primär auf die Steigerung der Wettbewerbsfähigkeit ab.

Hypothese 3: Big Data ist primär ein Management-Thema, da es datenbasierte Entscheidungen erleichtert.

Mit Blick auf die Funktionen der Teilnehmer der Online-Befragung zeigt sich, dass sich der größte Teil der Teilnehmer nicht eindeutig einem Funktionsbereich zuordnen kann, die zweitgrößte Gruppe jedoch aus dem Management stammt. Häufige Titelnennungen in den Freitextantworten zur Funktion weisen auf kaufmännische Tätigkeiten hin, während vergleichsweise wenige Titel eindeutig auf IT verweisen. Das hohe Interesse an Anwendungen im Bereich Marketing und Vertrieb sowie das Ziel der Wettbewerbssteigerung deutet darauf hin, dass Big Data eher ein Management-Thema ist, als dass es durch die IT dominiert wird. Inwieweit es damit zusammen hängt, dass es datenbasierte Entscheidungen erleichtert, kann so nicht eindeutig beantwortet werden. Die datenbasierte Planung und Entscheidungsfindung ist als eines von vielen Zielen in Bezug auf den Big-Data-Einsatz in einem Unternehmen und einer Branche benannt.

Hypothese 4: Ressourcen zum Arbeiten mit großen Datenmengen sind in der Mehrzahl der KMU noch nicht vorhanden.

Die Teilnehmer der Online-Befragung sehen sich in ihren Unternehmen aktuell nicht optimal aufgestellt, wenn es um die notwendigen Ressourcen für die Arbeit mit Big Data geht. Das beginnt mit fehlenden oder unzureichenden internen Budgets und endet mit unklaren personellen Verantwortlichkeiten und ausbaufähigen Kompetenzen. In den kommenden fünf Jahren, so schätzen die Teilnehmer, sollte die Personalausstattung jedoch deutlich steigen. Ähnliches gilt für das Budget. Niemand spricht sich für ein Sinken eines entsprechenden Budgets aus.

Die Forschungs-Hypothese wird durch die Antworten der Teilnehmer aus der Online-Befragung eher bestätigt. Die erforderlichen Ressourcen zum Arbeiten mit großen Datenmengen sind in der Mehrzahl der KMU noch nicht vorhanden, jedoch sollte sich dieses Bild in den nächsten fünf Jahren zum Positiven verändern.

Um die bestehenden Defizite zu beheben, wünschen sich 95 Prozent der Befragten, und damit auch Befragte aus nicht Big-Data-Unternehmen, Förderungen in Form von Best Practices oder Trainings.

ERGEBNISSE DER BRANCHENWORKSHOPS

5.1 Methodik

Die Branchenworkshops hatten die Zielsetzung, zukünftige Entwicklungen zur Nutzung von Big Data in ausgewählten deutschen Branchen in Form einer Roadmap gemeinsam zu erarbeiten. Moderiert wurden die Workshops von Herrn Dr. Koch aus dem Bereich Management von Innovationen und Technologien des Fraunhofer ISI.

5.1.1 Betrachtete Fragestellungen je Branche

- Welche zukunftsorientierten Potenziale ergeben sich durch das Thema Big Data?
- Wie werden bestehende Produkte, Prozesse und Marktlösungen durch Big Data beeinflusst?
- Welche zukünftigen veränderten oder vollkommen neuen Produkte, Prozesse und Marktlösungen sind durch Big Data denkbar?

5.1.2 Durchgeführte Branchenworkshops

- 12.11.2012: Finanzen
- 15.11.2012: Telekommunikation und Medien
- 19.11.2012: Marktforschung
- 26.11.2012: Handel
- 30.11.2012: Versicherungen

5.2 Workshop Finanzen

Der halbtägige Workshop fand am 12.11.2012 in Frankfurt am Main statt. Nach einer inhaltlichen Einführung in das Thema Big Data durch das Fraunhofer IAIS wurde mit allen Teilnehmern eine Roadmap erarbeitet, welches Innovationspotenzial Big Data in der Finanzbranche aufweist. Zu den teilnehmenden Firmen des Workshops gehörten Direktbanken, Wirtschaftsinformationsanbieter, Transaktionsdienstleister und Strategie-Berater mit Schwerpunkt Finanz-IT.

5.2.1 Datenquellen

Zunächst wurden gemeinsam Datenquellen gesammelt, die potenzialreich erscheinen, kurzfristig verfügbar sind und die Basis für kurzfristig neue bzw. optimierte Produkte und Dienstleistungen bilden können. Dazu zählen für die Teilnehmer im Bereich öffentlich verfügbarer Datenquellen insbesondere regionale Nachrichten, Geodaten, Finanzmarktdaten, Wetterdaten, Meinungsdaten, Reputationsdaten aus Social Networks, Marktdaten und Forenbeiträge. Innerhalb der Unternehmen werden Stammdaten (Personen, Waren, Lieferanten), Zahlungsverkehrsdaten, Vertragstexte, Firmennachrichten, Kundenstammdaten, Mitarbeiterdaten, Daten zu Branchen und Wettbewerbern, Reklamationen und Beschwerden sowie längerfristig auch bestimmte Daten von Mitbewerbern im Rahmen von Data Pools für Marketing, Support und Fraud-Bekämpfung als relevant erachtet. Auf der Ebene der Kunden sind Bewegungsdaten, gepostete persönliche Events, technische Einstellungen auf den Endgeräten und Online-Aktivitäten von Relevanz.

5.2.2 Prozessverbesserungen

In einem zweiten Schritt wurde diskutiert, welche Verbesserungen auf Seiten der Prozesse in der Finanzbranche möglich sind, wenn Big Data genutzt wird. Kurzfristig umsetzbare Beispiele sind ein gezielteres Cross- und Upselling, nutzerorientierte Werbung mit regionalem Fokus, Erkennen neuer in Foren kommunizierter Tricks von Kreditkartenbetrügern, sowie die Vereinfachung von Kundendialogen. Auf mittelfristige Sicht könnte die Finanzbranche zu einem »Now-Casting« gelangen, in dem der aktuelle Stand des eigenen Geschäfts jederzeit abrufbar ist und es damit möglich wird, das Geschäft an die aktuellen Bedingungen besser anzupassen (Regelsteuerung). Durch Kommunikationsanalysen zwischen Unternehmen und Kunden wären Kundendialoge möglich, die das bevorzugte Kommunikationsverhalten des Kunden berücksichtigen. Aus Sicht des Marketings wird das Multiplikatorenmarketing wichtiger werden, wenn erkannt wird, welche Kunden vom Service begeistert sind und dies öffentlich kommunizieren. Ebenfalls denkbar sind bessere Angebote für die Betrugserkennung, wenn Daten von Banken und Kunden integriert ausgewertet werden. Längerfristige Prozessverbesserungen sind durch den Einzug von A/B Tests möglich, die darauf abzielen, Varianten von Prozessen im Echtbetrieb miteinander zu vergleichen und die beste Variante schlussendlich umzusetzen. Auch der Produktentwicklungsprozess kann durch Big Data optimiert werden. Durch eine kontinuierliche Auswertung von Suchanfragen großer Suchmaschinen könnten auch kurzfristige Bedarfe an neuen oder veränderten Finanzprodukten erkannt werden. Durch den Einbezug komplexerer Daten wäre es längerfristig ebenfalls denkbar, auf Authentifizierungen nach heutigem Stand zu verzichten. Bis 2020 wird erwartet, dass die Produktion von Standardprodukten in der Finanzbranche abnimmt und massenindividualisierte Produkte Einzug halten. Welche Produkte das sein könnten, wurde im letzten Schritt des Workshops untersucht.

5.2.3 Innovative Produkte

In einem dritten Schritt wurden optimierte oder neue Produkte identifiziert, die auf Basis der Nutzung von Big Data möglich sind. Auch hier wurde zunächst damit begonnen, die kurzfristig realisierbaren Produkte zu nennen. Beispiele dafür sind persönliche Finanz-Cockpits, die ausführlichere Informationen zur eigenen Finanzsituation und zu Chancen der Geldanlage liefern. Durch die Nutzung besserer Profilierungs- und Anti-Fraud Verfahren sind kooperative Geldanlagen bzw. Geldleihen möglich, die bislang ein zu großes Sicherheitsrisiko darstellen. Mittelfristig denkbar sind maßgeschneiderte Informationsprodukte zur eigenen Finanzsituation, die auf automatisierter Inhalteerstellung (Content Creation) basieren und dadurch finanziell darstellbar werden. Längerfristig sind Kundenakten denkbar, die es einem Kunden ermöglichen, Anbieter zu wechseln und dabei ihre Datenhistorie mit den eigenen Hintergründen und Präferenzen einzubringen. 2020 könnte die Finanzbranche mit Hilfe von massenindividualisierten Dienstleistungsprozessen in der Lage sein, etwa durch Assisted Finance, jedem Kunden zu jeder Zeit dabei zu helfen, seinen finanziellen Status zu verstehen und zu optimieren.

5.2.4 Fazit

Kurzfristig waren sich die Teilnehmer einig, dass die Finanzbranche stark an Effizienzgewinnung durch die Nutzung von Big Data interessiert ist. Belastbare Kosten-Nutzen-Kalkulationen stehen im Vordergrund. Allerdings wird auch gesehen, dass neue Anbieter, u.a. aus fremden Branchen, in der Lage sein könnten, dank Big Data potenziell interessantere Finanzprodukte anzubieten, sofern dies regulativ möglich ist. Das Innovationspotenzial von Big Data in der Finanzbranche wurde von allen Beteiligten gesehen. Allerdings ist offen, welches deutsche Unternehmen dieses Potenzial nutzen wird.

5.3 Workshop Telekommunikation und Medien

Der halbtägige Workshop fand am 15.11.2012 in Sankt Augustin statt. Nach einer inhaltlichen Einführung in das Thema Big Data durch das Fraunhofer IAIS wurde mit allen Teilnehmern eine Roadmap erarbeitet, welche das Innovationspotenzial von Big Data in der Branche der Telekommunikation darstellt. Zu den teilnehmenden Firmen des Workshops gehörten Telekommunikationsanbieter aus Europa, Verbandsvertreter, Software-Anbieter im Bereich Big Data und Strategie-Berater mit Schwerpunkt Telekommunikation.

5.3.1 Datenquellen

In einem ersten Schritt wurden Datenquellen gesammelt, die kurzfristig verfügbar sind und die Basis für neue bzw. optimierte Dienstleistungen in der Telekommunikation bilden können. Im Bereich öffentlich verfügbarer Datenquellen sind für die Teilnehmer das Web inkl. Patenten, Public Open Data zu Orten, Personen, Firmen, Textdaten, Bilddaten, Wetterdaten, Mediendaten und Wifi-Daten von Bedeutung. Mittelfristig erscheinen Sensordaten aus dem Bereich M2M als besonders relevant. Längerfristig werden Open Data Pools anvisiert, die Daten verschiedener Anbieter bündeln und bereit stellen. Auf Ebene der Unternehmen werden kurzfristig Kundenstammdaten, Logdaten der Infrastruktur und speziell Call-Data-Records und Voice genannt, die relevant sind.

Beispielhaft zählen dazu Logdaten aus Switches und Proxies des Mobil- und Festnetzes. Auf einer mittelfristigen Zeitachse sehen die Teilnehmer des Workshops Daten aus IP-TV-Nutzung, M2M, Car2Car sowie Smart Metering als relevant an. Längerfristig erscheinen Daten aus Internet-enabled Devices als relevant, wobei diese Daten sowohl von Unternehmen (B2B) auch aus internetfähigen Geräten von Konsumenten stammen können (B2C).
Auf Kundenseite werden bereits heute Daten erzeugt, die für die Unternehmen aus der Telekommunikation relevant erscheinen. Dazu zählen Daten aus der internen Sensorik der Geräte, etwa Ort, Ausrichtung, Licht, Lautstärke oder Vibration. Auch das Konsumverhalten, welches sich durch Bezahl- und Abrechnungsdaten, getätigte Anrufe, App-Nutzung, Messaging- und Medianutzung in Form von Daten ausdrückt, wird als wichtig eingeschätzt. Social Networks werden ebenfalls als relevante Datenquelle betrachtet. Mittelfristig von Relevanz sind analytisch gewonnene Daten zur Bewegung im Raum, Bezahlverhalten, Konsumverhalten, Gewohnheiten, Interessen, Hobbies, Netzwerken, Freunden, Gesundheit, Kenntnissen. Im B2B-Bereich sind es Energiedaten sowie Daten aus NFC. Langfristig werden Daten aus Voice-to-Text und Stimmungsanalysen auf Textdaten als relevant eingeschätzt. Als eine künftige Quelle diverser konsumentenbezogener Daten werden digitale Hüllen, d.h. vom Konsumenten steuerbare Datenpools, gesehen.

5.3.2 Prozessverbesserungen

In einem zweiten Schritt wurde diskutiert, welche unternehmensinternen Prozess-Verbesserungen möglich sind, wenn die genannten Datenquellen genutzt werden. Kurzfristig umsetzbare Beispiele sind bessere Fraud-Analysen, Churn-Analysen, Ressourcenplanungen, Netzwerkoptimierungen, Availability-Management und detailliertere Betriebsunterstützungen. Kurz- bis mittelfristig sehen die Teilnehmer eine Verbesserung der internen Planung von Ressourcen und Prognosen der Finanzlage sowie kurzfristigere Produktentwicklungen und gezielte Serviceverbesserungen durch die Auswertung aktueller Suchmaschinenanfragen. Zum Kunden entstehen Personalisierungen von Produktumgebungen, die deutlich über bisherige Angebote hinausgehen und bis zur 1:1 Individualisierung reichen (auch als Mass Customization bekannt). Mittelfristig entstehen neue Prozesse, die ein Data Provisioning an Externe ermöglichen. Langfristig wird eine stärkere Automation von Prozessen erwartet, die zu einer höheren Gesamteffizienz des Unternehmens beiträgt.

5.3.3 Innovative Produkte

In einem dritten Schritt wurden im Workshop optimierte oder neue Produkte und Dienstleistungen identifiziert, die durch die Nutzung von Big Data möglich werden. Auch hier wurde zunächst damit begonnen, die kurzfristig realisierbaren Produkte und Dienstleistungen zu nennen. Kurzfristig sehen die Teilnehmer des Workshops bessere Angebote im Bereich Verkehrsmanagement (in Analogie zu TomTom HD oder Google Traffic) sowie Maintenance Monitoring mit Ereigniserkennung bei Maschinen und Anlagen. Auf Kundenseite sind Smart Metering mit Energiesteuerungs-Apps und Energieoptimierungen möglich. Location-based services, die eine individualisierte Bewerbung, Empfehlung von Produkten und sicherere Bezahlung ermöglichen, werden ebenfalls genannt. IP-TV ermöglicht individualisierte Programme und auf Nutzer abgestimmte Werbung. Daran schließen sich neue datenbasierte Produkte über Orte, Konsumenten und Ereignisse an, die im B2B-Bereich z.B. für Mikro-Wetterdienste, Tourismusangebote, Werbewirtschaft, sichere Logistik oder Städteplanung, Eventorganisation und Katastrophenschutz vermarktbar sind.
Langfristig wird erwartet, dass zunehmend intelligente Produkte vermarktet werden, die auf einer starken Automatisierung und Individualisierung basieren. Beispielhaft genannt wurden Produkte zur vorausschauenden Gesundheitsüberwachung (Automated Health), Life Assist (persönliche Assistenten) sowie vorausschauende Wartung von Maschinen und Anlagen (Predictive Maintenance). Auf Unternehmensebene werden neue Geschäftsmodelle erwartet, die auf einer Kombination vormalig getrennter Industrien beruhen. Hier wurden Konsortien mit Energie-Providern, IT-Providern und weiteren als Beispiele angeführt. Diese werden speziell im Kontext von Smart Cities und der Energiewende erwartet.

5.3.4 Fazit

Im Workshop Telekommunikation bildete sich die gemeinsame Auffassung heraus, dass die Branche bereits heute über sehr viele Daten verfügt, diese jedoch erst in geringerem Maße in die Verbesserung von Prozessen und noch weniger in die Entwicklung neuer Produkte und Dienstleistungen einfließen. Mittelfristig ist auch hier der Trend zu Prozessen und Dienstleistungen zu erkennen, die auf eine kontinuierliche Analyse von Daten aufsetzen und damit die Basis für effizienteres Unternehmensmanagement sowie neuartige Monitoring-Dienstleistungen für Endkunden bilden. Langfristig stellen sich die Teilnehmer vor, dass eine zunehmende Automatisierung und Individualisierung möglich wird, die sehr komplexe Geschäftsmodelle wie das der Smarter Cities und Smart Grids ermöglicht, deren konkrete Anwendungen bislang nur in groben Zügen erkennbar sind. Auch hier gilt, dass die Nutzung von Big Data immer vor dem Hintergrund der Datenschutzbestimmungen gesehen werden muss und die Gesellschaft angesichts der Zukunftserwartungen der Teilnehmer viele neue Antworten finden muss, um gesellschaftlich akzeptablen Fortschritt zu ermöglichen.

5.4 Workshop Marktforschung

Der halbtägige Workshop fand am 19.11.2012 in München statt. Nach einer inhaltlichen Einführung in das Thema Big Data durch das Fraunhofer IAIS wurde mit allen Teilnehmern eine Roadmap erarbeitet, welche das Innovationspotenzial von Big Data in der Branche der Marktforschung darstellt. Zu den teilnehmenden Firmen des Workshops gehörten Vertreter internationaler und nationaler Marktforschungsunternehmen verschiedener Größenklassen sowie überregionaler Zeitschriften für Marketing und Marktforschung.

5.4.1 Datenquellen

Bereits heute lebt die Marktforschung davon, Daten im Kundenauftrag zu erheben und auszuwerten. Daher ist es spannender zu fragen, wie sich die Branche der Marktforschung kurz-, mittel- und langfristig verändern wird.

5.4.2 Prozessverbesserungen

Kurzfristig ermöglicht der Zugang zu detaillierteren, aktuelleren und komplexeren Daten über Konsumenten und Organisationen aus sozialen Netzwerken und Webseiten sowie Blogs eine Steigerung der Effizienz. Viele Daten müssen nicht mehr selbst durch Umfragen und ähnliches erzeugt werden, sondern liegen bereits online vor. Intern müssen dazu bessere Verfahren zur Datenbeschaffung umgesetzt werden, um auch Rohdaten aus unterschiedlichen Quellen in kurzer Zeit zur Analyse verfügbar zu machen. Eng daran geknüpft sind Verbesserungen in der Modellentwicklung. Der alleinige Zugriff auf größere, aktuellere und heterogene Daten stellt neue Anforderungen an die Mitarbeiter, geeignete Modelle zu entwickeln, welche die marktforscherischen Fragen mit hinreichender Qualität beantworten können. In diesem Zusammenhang werden auch neuartige Visualisierungsmöglichkeiten Einzug halten, die es Marktforschern wie auch Kunden erlauben, die Ergebnisse leichter zu interpretieren und schneller zu Handlungsempfehlungen zu kommen. Mittel- bis längerfristig werden Prozesse in der Marktforschung stärker als bislang auf passive Sensorik aufsetzen. Das können Sensordaten aus Geräten sein, die genauere Einsichten in das Nutzungsverhalten ermöglichen. Längerfristig erwarten daher die Teilnehmer, dass ein kontinuierliches und agiles Messen und Auswerten üblich wird. Dadurch lassen sich Produktentwicklungszeiten verkürzen und neue Trends schneller entdecken.

5.4.3 Innovative Produkte

Auf der Ebene der Produkte und Dienstleistungen wird die Marktforschung stärker als bislang die Entwicklung von Algorithmen im Kundenauftrag realisieren. »Computational Statistics« wird die Basis für neue Angebote bilden, die den

Kunden das zentrale Problem lösen, relevante Informationen aus den eigenen und/oder öffentlichen Daten zu ziehen. Künftige Angebote werden stärker als bislang mit Dashboards und Visualisierungen arbeiten, um beim Transfer der Informationen in die betriebliche Praxis zu helfen. Längerfristig vorstellbar sind Analyselösungen, die beim Kunden installiert werden und dem Kunden bei der kontinuierlichen Messung und Analyse eigener Daten aus der Produktnutzung helfen. Dies wird begleitet durch Schulungen für die Weiterentwicklung der dafür benötigten Algorithmen und der Interpretation der Ergebnisse. Die Marktforschung wird somit stärker als bislang in viele Bereich diffundieren und nicht zuletzt durch neue, datengetriebene Wettbewerber auch dazu führen, dass der Markt für kostengünstigere Marktforschungslösungen größer wird.

5.4.4 Fazit

Die Teilnehmer des Workshops waren sich einig, dass die Branche der Marktforschung wie kaum eine andere von der Nutzung von Big Data profitieren wird. Neue Angebote werden entstehen, die auf Algorithmen und neuartigen Visualisierungsverfahren basieren und Kunden einen sehr viel aktuelleren und detaillierteren Einblick in Märkte bieten. Sowohl Kunden als auch Anbieter werden in Zukunft stärker mittels computerbasierter Statistik arbeiten. Eine derartige Kompetenzentwicklung, zuweilen auch als Data Science beschrieben, wird mittelfristig deutlich zunehmen. Kurzfristig sind Personen mit derartigen Kompetenzen wenig verfügbar und hoch gefragt.

5.5 Workshop Handel

Der halbtägige Workshop fand am 26.11.2012 in Sankt Augustin bei Bonn statt. Nach einer inhaltlichen Einführung in das Thema Big Data durch das Fraunhofer IAIS wurde mit allen Teilnehmern eine Roadmap erarbeitet, welches Innovationspotenzial Big Data in E-Commerce und Versandhandel aufweist. Zu den teilnehmenden Firmen des Workshops gehörten Versender, E-Commerce-Händler und Strategie-Berater mit Schwerpunkt Versandhandel.

5.5.1 Datenquellen

In einem ersten Schritt wurden gemeinsam Datenquellen gesammelt, die potenzialreich erscheinen, kurzfristig verfügbar sind und die Basis für kurzfristig neue bzw. optimierte Dienstleistungen im Handel bilden können. Dazu zählen für die Teilnehmer im Bereich öffentlich verfügbarer Datenquellen insbesondere Preise des Wettbewerbs, Social Media Beiträge zu Produkten und Firmen, Forenbeiträge, aber auch Daten über Kaufabsichten z.B. durch Online-Tracking. Weitere relevante Daten aus öffentlichen Quellen stammen von Adressregistern und Meldeämtern, Haushaltspanel-Anbietern, von Wetterdiensten und Geodatenanbietern sowie Herstellern mit Rückrufaktionen. Längerfristig von Interesse sind Daten statistischer Ämter, die Auskunft über demographische Veränderungen bieten. Der Wunsch nach einem Zugriff auf a priori nicht selektive Daten aus großen Suchmaschinen ist ebenfalls da, wird jedoch als weder kurz- bis mittelfristig verfügbar angesehen.
Auf Ebene der Unternehmen sind aktuelle und vollständige Meta- und Masterdaten zu Produkten essentiell. Kundenstammdaten sind ebenso wie Transaktionsdaten aus Web-Shops, Kassen, Call-Centern bis hin zu Briefen und Faxen wichtig und kurzfristig verfügbar. Kurzfristig sind für Web-Shop-Betreiber auch Daten über die Kunden-Reise verfügbar, die dank Tracking und Retargeting zunehmend aktuell und hinreichend detailliert sind. Langfristig werden Datenpools mit konkurrierenden Anbietern als relevante Datenquelle angesehen.
Auf Kundenseite werden ebenfalls Daten erzeugt, die für die Unternehmen im Bereich Handel von Relevanz sind. Dazu zählen insbesondere Daten zum Ort und Gerät des Nutzers. Mittelfristig sind die genutzten Medien und die Entscheidungswege im Kaufprozess von Interesse. Daten über Entscheidungswege können sowohl anhand von Klickpfaden entstehen oder auch durch Bewegungen im realen Raum, die durch Sensoren aufgezeichnet und in Form von Bewegungsprofilen maschinell bearbeitbar werden. Langfristiges Interesse wird an Daten über die emotionale Situation der Konsumenten geäußert. Hier stellt sich im Detail die Schwierigkeit, die verschiedenen Daten aus dem Multichannel-Ansatz zu integrieren. Das gilt insbesondere bei Daten, die aus dem Support bzw. der Bewerbung kommen (Werbereaktionsdaten).

5.5.2 Prozessverbesserungen

In einem zweiten Schritt wurde diskutiert, welche Verbesserungen auf Seiten der Prozesse möglich sind, wenn Big Data genutzt wird. Kurzfristig umsetzbare Beispiele sind individualisierte Markenbotschaften, ereignisgesteuerte automatisierte Kommunikation, die optimierte Warenlogistik für Sonderaktionen mittels Bedarfsvorhersage und die Serviceverbesserung via Social Media, die auf Verfahren des Opinion Minings setzt. Mittelfristig stehen umfassende Automatisierungen von Prozessen und Entscheidungen im Fokus der Prozessverbesserungen. Daran anschließend wird mittel- bis langfristig die Möglichkeit gesehen, neue Nischenmärkte mittels Analytik besser identifizieren zu können, als dies bisher mangels Daten möglich ist. Die Prozesse im Handel werden langfristig dazu in der Lage sein, den Ort, die Historie des Konsumenten und die aktuelle Situation zu berücksichtigen, um das richtige Produkt einer konkreten Person anzubieten.

5.5.3 Innovative Produkte

In einem dritten Schritt wurden im Workshop optimierte oder neue Produkte und Dienstleistungen identifiziert, die durch die Nutzung von Big Data möglich werden. Auch hier wurde zunächst damit begonnen, die kurzfristig realisierbaren Produkte und Dienstleistungen zu nennen. Kurzfristig bis mittelfristig sahen die Teilnehmer des Workshops insbesondere das Aufkommen neuer Software-as-a-Service Lösungen als Infrastruktur für den Handel. Mittelfristig wurden Analytics-as-a-Service angedacht, die den Handel optimieren können. Dazu zählen Dienste zur produktbezogenen Analyse von Meinungen und Emotionen sowie weitere Geschäftsmodelle im Bereich der Handels-Infrastruktur wie Risk-Management, Financial Services, Shop-Betrieb oder Micro-Payment-Dienste. Längerfristig werden zunehmend Daten als solches zum Produkt, was am Beispiel von Telefonica mit dem Verkauf von Bewegungsprofilen illustriert wurde. Allerdings wurde auch hier darauf hingewiesen, dass das datenschutzrechtlich konforme Handling von Daten selbst eine wesentliche Dienstleistung darstellen wird.

5.5.4 Fazit

Kurzfristig waren sich die Teilnehmer einig, dass speziell datengetriebene Infrastrukturleistungen den Handel modernisieren werden. Diese werden zunehmend als Service angeboten und dadurch auch kapitalschwächeren Handelsunternehmen oder Einzelhändlern einen schnelleren Zugang zum (Online-)Markt liefern. Mittelfristig sind es Analyse-Dienstleistungen, die den Unterschied in der Kundenansprache und Kundenbindung ausmachen. Längerfristig scheinen Daten selbst zum Produkt zu werden – das alles vor dem Hintergrund zunehmend restriktiverer Datenschutzbestimmungen.

5.6 Workshop Versicherungen

Der Workshop fand am 30.11.2012 in Sankt Augustin statt. Nach einer inhaltlichen Einführung in das Thema Big Data durch das Fraunhofer IAIS wurde mit allen Teilnehmern eine Roadmap erarbeitet, welche das Innovationspotenzial von Big Data in der Branche der Versicherungen darstellt. Zu den teilnehmenden Firmen des Workshops gehörten Vertreter von großen Versicherungsunternehmen aus Deutschland.

5.6.1 Datenquellen

In einem ersten Schritt wurden Datenquellen gesammelt, die kurzfristig verfügbar sind und die Basis für kurzfristig neue bzw. optimierte Dienstleistungen bilden können. Im Bereich öffentlich verfügbarer Datenquellen zählen für die Teilnehmer sozio-demographische Daten auf PLZ-Ebene, Geodaten, die Auskunft über Kaufkraft, Bebauung, Alter etc. geben, kommerzieller Web Content, News, Blogs und Social Media, natürliche Katastrophen, Klima-Daten und Ereignisse sowie User Generated Content wie etwa Open Street Maps und andere. Mittelfristig relevant erscheinen Verbrauchsdaten im Bereich Energie.

Auf Ebene der Unternehmen werden kurzfristig verfügbare Daten im Bereich Policen, Kunden, Gesundheit, Schäden sowie aus der Abwicklung und dem Kundenkontakt als relevant erachtet. Im unmittelbaren Fokus stehen weiterhin genauere Daten zu den Kunden, etwa zum Gehalt, zur familiären Situation, zum Beruf etc. Mittelfristig verfügbar und relevant sind Daten zu eigenen und fremden Marken sowie über das Verhalten in der Kundenkommunikation.
Auf Kundenseite werden bereits heute Daten erzeugt, die für die Unternehmen aus der Versicherungsbranche relevant erscheinen. Dazu zählen Daten aus Social Media, die Auskunft über Interessen, Vorlieben und Produkt-Priorisierungen geben. Mittelfristig erscheinen Verhaltensdaten aus der Nutzung des Internets einschließlich Social Media sowie aus Telematik-Systemen, die Auskunft über das Fahrerverhalten und Wegenutzung geben, sowie location based social networks die Auskunft über das räumliche Bewegungs- und Kommunikationsverhalten geben. Diese Daten könnten zu dynamischen Lebensmodellen aggregiert werden. Längerfristig relevant erscheinen Gesundheitsdaten aus Patientenakten.

5.6.2 Prozessverbesserungen

In einem zweiten Schritt wurde diskutiert, welche unternehmensinternen Prozess-Verbesserungen möglich sind, wenn die genannten Datenquellen genutzt werden. Kurzfristig umsetzbare Beispiele sind ein besseres Schaden- und Betrugsmanagement, ein besseres Risikomodelling, Echtzeit-Betrugserkennung und Regelbildung aus Schadensdaten. Das alles könnte direkt in das Portfoliomanagement bzw. die Zeichnungspolitik einfließen, die sich dadurch optimieren ließen.
Mittelfristig taucht die dynamische Preisfindung sowie das echtzeitfähige Cross- und Upselling auf. Eine Online-Beantragung mit direkter Policierung wie auch eine nachgeschaltete Storno-Prävention werden als relevante Prozessverbesserungen betrachtet. In Richtung Kunden soll es künftig möglich sein, transparenter und schneller als bislang Auskunft darüber geben zu können, welche Daten das Unternehmen gespeichert hat. Längerfristig zieht die Automatisierung ein, die über Produkte und Prozesse hinweg eine automatische Leistungsbearbeitung ermöglicht.

5.6.3 Innovative Produkte

In einem dritten Schritt wurden im Workshop optimierte oder neue Produkte und Dienstleistungen identifiziert, die durch die Nutzung von Big Data möglich werden. Kurzfristig erwarten die Teilnehmer des Workshops individuelle Produktbündel, die auf bestimmte Konsumentengruppen (»20somethings, Golden Ager«) zugeschnitten sind. Darin sind Bedarfe und Risiken enthalten, die bislang in statischen Modellen nicht abgebildet werden.
Mittelfristig werden Tarife erwartet, die vom gemessenen Verhalten abhängen. Beispielhaft sind hier Autoversicherungen zu nennen, die das Fahrverhalten in die individuelle Tarifierung einbeziehen. Dafür sind vom Kunden Daten einzubeziehen, die per Einwilligung erhoben und verarbeitet werden und dafür im Gegenzug Rabatte versprechen. Zunehmen wird die Individualisierung und Dynamisierung von Versicherungsprodukten. So sind mittelfristig Mikroversicherungen denkbar, die eine ganz bestimmte Situation kurz vor dem Eintreten der Situation versichern. Bereits heute finden sich erste Beispiele dieser Art (z.B. Tages-Versicherung für Skiabfahrten). In Zukunft wird dies genauer auf den Ort, die Zeit und die Person bezogen geschehen können.
Neue Dienstleistungen auf Basis von Big-Data-Analysen sind vorstellbar, um Scoring-, Klima- und Rentabilitätsrechnungen zu optimieren. Längerfristige Produktverbesserungen in der Branche der Versicherungen zielen auf 24/7, also eine Jederzeit-Rundumbetreuung, die proaktiv Kundenbedürfnisse erkennt und Vorschläge unterbreitet (»Rundum sorglos Paket«). Sie könnte am Ende in einer Art Lebensassistenz als intelligentes Produkt münden, welches die Taxierung für die Versicherung in Echtzeit vornimmt.

5.6.4 Fazit

Die Teilnehmer des Workshops sehen eine Reihe aktueller und künftiger Herausforderungen, die durch den Einsatz von Big Data in der Branche der Versicherungen bestehen. Im Status quo ist dank des heutigen Datenschutzes und der Einwilligungen seitens der Kunden der Umgang mit personenbezogenen Daten geregelt. In Zukunft erscheinen weitere Datenquellen, die z.B. durch Monitoring bzw. Prognose entstehen und die Basis für verhaltensorientierte Versicherungsprodukte bilden. Beispielhaft zu nennen sind hier Daten aus Telematik-Systemen, aus Social Media oder digitalen Patientenakten. Wie werden Arrangements aussehen, um diese als relevant eingeschätzten Daten nutzen zu können? Jüngere Generationen haben bereits heute ein anderes Verhalten bezüglich ihrer Daten. Wie werden Personen aus älteren Generationen ihren Versicherungsschutz erhalten? Auch ist die Frage zu klären, was mit Personen geschieht, die aufgrund ihres als riskant eingeschätzten Lebensverhaltens nicht mehr in den Genuss einer Versicherung kommen. Längerfristig scheint sich auch die Rolle der Versicherungen zu erweitern. Durch die Analyse von Verhaltensdaten und die Abschätzung des individuellen Risikos wird es bereits mittelfristig möglich sein, Hinweise zur Lebensanpassung zu liefern, die zumindest auf Basis der verwendeten Daten und Modelle ein längeres und sichereres Leben ermöglichen könnten. Ob derartige Innovationspotenziale auch von den Endkunden angenommen werden, bleibt offen.

5.7 Gemeinsame Themen in den Workshops

Auffällig war, dass in den branchenspezifischen Zukunftsworkshops ähnliche Muster in der Wahrnehmung von Nutzungsmöglichkeiten von Big Data bestehen. So ließen sich kurzfristig in allen Branchen unmittelbare Potenziale in der Effizienzsteigerung der Unternehmensführung durch Big Data erkennen. Branchenspezifisch waren jeweils die entsprechenden Anwendungen und Datenquellen. Langfristig war in allen Branchen von einer Massenindividualisierung von Services und zunehmend intelligenten Produkten die Rede, die von Big Data und der jeweils notwendigen Analysefähigkeit profitieren. Da die Durchführung von kurz- bis mittelfristigen Projekten zur Effizienzverbesserung von belastbaren Kosten-Nutzen-Kalkulationen abhängt, sind deutlich mehr branchenspezifische Referenzbeispiele notwendig. Hier liegen speziell in Deutschland noch nicht viele öffentlich kommunizierte Referenzprojekte vor und es herrscht bei den Unternehmen eine gewisse Unsicherheit, inwieweit erfolgreiche Projekte aus den USA als Vorlage dienen können, da sich die hiesigen Regularien als auch Praktiken in Bezug auf Datenschutz mitunter sehr von den Vorgaben in den USA unterscheiden. Mittel- bis langfristig angelegte Produkt- und Serviceinnovationen auf Basis von Big Data erfordern entsprechend qualifiziertes Personal, welches idealerweise Teil des strategischen Wissens des Unternehmens ist. Hier wird es seitens der jeweiligen Unternehmensführung erforderlich sein, die notwendigen Schritte zum Aufbau der entsprechenden inhouse Expertise und der Investition in mittel- bis langfristig wirksame Produkte und Services auf Basis von Big Data zu investieren. Entscheidend für beide skizzierten Stränge ist, dass die Unternehmensführung die Chancen, aber auch kulturellen Herausforderungen (an-)erkennt, stärker datengetrieben zu operieren und die Fähigkeit zur intelligenten Nutzung von Daten in Produkten und Services als schwierig kopierbaren Wettbewerbsvorteil zu entwickeln.

FAZIT

Die Innovationspotenzialanalyse gibt Antworten auf bislang unbekannte Aspekte zum Potenzial von Big Data in deutschen Unternehmen. So ist die bisherige Zurückhaltung vieler Unternehmen einer Unwissenheit über mögliche Anwendungsfälle von Big Data geschuldet. Zudem spielen Datenschutzbedenken eine große Rolle. Es bestehen Hemmungen, Anwendungen zu testen, die zwar juristisch möglich wären, jedoch in der öffentlichen Wahrnehmung zu Missverständnissen führen könnten. Außerdem wird mangelnde Kreativität als Schlüsselfaktor genannt, um aus vorliegenden Daten mehr geschäftlichen Nutzen zu ziehen. Fehlende interne Kompetenzen sind ebenfalls eine durchgängig genannte Barriere zur Adoption von Big Data. Überraschend war die Erkenntnis über die hohe strategische Relevanz des Themas sowie die explizite Einstufung als nicht-technisches Thema, welches früher oder später in jeder Branche einziehen wird.

Aus den Ergebnissen ließen sich drei branchenübergreifende Chancen als Innovationspotenziale durch Big Data ableiten:

6.1 Chance »Effizientere Unternehmensführung«

Kurzfristig wird ein hohes Potenzial in der betrieblichen Effizienzsteigerung gesehen. In einem ersten Schritt geschieht dies durch den Zugriff auf Echtzeit-Informationen zu beliebigen betrieblichen Prozessen, die dadurch die unternehmerischen Einzelentscheidungen prägen. In einem zweiten Schritt wird eine höhere Effizienz durch taktische Prognoseverfahren erwartet, die zu günstigeren Beschaffungs-, Produktions- und Lagerkosten oder Personaleinsätzen führen. In einem dritten Schritt folgen Automatisierungen von Entscheidungen, die bislang manuell erfolgten. Beispiele hierfür finden sich in der Posteingangsbearbeitung bei Versicherern, die fallbasierte Entscheidungen für wenig komplizierte Fälle mittlerweile fast vollständig automatisieren. Ähnliches wird Einzug in viele weitere Branchen finden.

6.2 Chance »Massenindividualisierung von Services«

Mittel- bis langfristig war in allen Branchen von einer Massenindividualisierung von Services die Rede, die durch Big Data möglich werden. Diese Services sind auf den einzelnen Konsumenten zugeschnitten, wenn historische Daten z.T. im Sekundenbereich ausgewertet werden, um das richtige Angebot zur richtigen Zeit in der richtigen Form zu präsentieren. Hier ist speziell die Online-Werbeindustrie der technologische Vorreiter in der Anwendung von Big Data, die im Sub-Sekundenbereich eine massenindividualisierte Werbeeinblendung realisiert. Das ist für weitere, datengestützte, Dienstleistungen ebenfalls denkbar. Der Handel etwa hat die Leitidee eines »IT-gestützten Tante Emma-Ladens«, der auf diesem Prinzip basiert und nicht nur diese Branche verändern wird.

6.3 Chance »Intelligente Produkte«

Langfristig werden intelligente Produkte und damit maschinelles Lernen auf Big Data Einzug in diverse Produkte halten. Dazu gehören komplexe Anlagen (Ölplattformen), Maschinen bis hin zu profanen Alltagsgegenständen (Thermostate). Produkte erfassen dann lokal viele unterschiedliche Sensor-Daten, werten diese auf Basis des bereits Gelernten aus und stimmen sich punktuell mit einer Zentrale zu bestimmten Situationen ab – etwa wann eine Wartung sinnvoll erscheint oder was getan werden könnte, um noch effizienter zu arbeiten (z.B. intelligente Baumaschinen, Assisted Finance, Energy Control uvm.). Hier erscheint mit Blick auf den ausgeprägten Maschinen- und Anlagenbau sowie die Automobilwirtschaft in Deutschland ein sehr großes Potenzial vorzuliegen, mit intelligenteren Produkten dank Big Data die Wettbewerbsfähigkeit ganzer Schlüsselbranchen zu stärken.

ANHANG

7.1 Steckbriefe der Big-Data-Anwendungsfälle

Dieser Anhang listet steckbriefartig die Anwendungsfälle auf, die bei der Recherche ausgewählt wurden.

ID	Titel	Beschreibung	URL
Adidas Group	Analysieren der Besucherströme in Filialen und Standorten	Um Geschäftsperformanz-Indikatoren für die Läden im asiatisch-pazifischen Raum zu erheben, hat Adidas eine Footfall-Lösung von Experian in derzeit 200 Geschäften installiert. Sie besteht aus Wärmebildkameras zum Zählen der Besucher in den verschiedenen Bereichen und einer Reporting-Software zur Analyse der Daten auf Stundenbasis. Die Zahlen werden insbesondere zur Personaleinsatzplanung genutzt und haben geholfen, die Konversationsrate zu steigern.	www.footfall.com
AutoGrid	Simulation von Energienetzen zur Prognose	Der Kern von AutoGrids Energy Data Platform ist eine Optimierungsmaschine, welche die komplexen Beziehungen zwischen Energieerzeugung, -übertragung, -verteilung und Verbrauch in einem Gebiet oder einer Anlage modelliert. Strukturierte und unstrukturierte Daten aus dem Netz, von Erzeugern und Verbrauchern werden genutzt, um aus Millionen von Variablen den zukünftigen Verbrauch und Stresszustände im Netz zu prognostizieren. Es können Szenarien berechnet werden, um Einflussfaktoren wie Preissignale zu verstehen.	www.auto-grid.com
BigML	Lernen für alle	BigML bietet einen cloud-basierten Dienst für Data Mining an, der sich nach der Größe der Daten, des Modells und der Anzahl der Modellnutzungen richtet. Die Nutzer laden ihre Daten hoch, können Felder löschen und die Daten visuell betrachten. Sie können dann ein pädiktives Modell erstellen und visuell überprüfen. Das Modell kann über ein Webformular, ein Widget oder per API des Kunden von einem Smart Phone benutzt werden.	bigml.com

BillGuard	Überwachung von Kreditkartenabrechnungen	Billguard ist ein persönlicher Finanzdienst, der seine Nutzer auf Buchungsfehler sowie falsche oder betrügerische Buchungen hin weist. Er benötigt Lesezugriff auf die Webseite mit den Kreditkartenaktivitäten des Benutzers. Er prüft sie täglich anhand von hunderten automatisierten Sicherheitstests, darunter Beschwerden im Web und bei Banken. Die Nutzer können Hinweise, Fragen und Kommentare zu Transaktionen und Händlern teilen, um sich gemeinsam zu schützen.	www.billguard.com
Chevron	Echtzeit-Überwachung von komplexen Anlagen	Die Öl-Industrie begann schon vor 10 Jahren verteilte Sensoren, Hochgeschwindigkeitsnetze und Data Mining einzusetzen, um Bohrungen zu überwachen, zu steuern und Störungen zu vermeiden. Diese Maßnahmen zur Produktivitätssteigerung bezeichnet man in der Branche als »Digitales Ölfeld«, bei Chevron, »i-field«, bei BP »Field of the Future« und bei Royal Dutch Shell »Smart Fields«. Damit sollen unnötige Inspektionen vermieden, Störungen schneller beseitigt und Gefahren frühzeitig erkannt werden.	www.chevron.com
Cookie's	Dynamische Preisfindung	Cookie's, ein amerikanisches Kinderbekleidungsgeschäft ändert alle 15 Minuten die Preise seiner 2 Millionen Produkte, um Amazon zu unterbieten. Ähnlich schnell werden Preise für Verbraucherelektronik, Kleidung, Schuhe, Schmuck und Haushaltswaren geändert. Die Software kalkuliert die Preise abhängig von Preis und Lieferkosten der Konkurrenz, Preiseinschränkungen der Lieferanten und Ausverkaufszeiten.	www.mercent.com
CTOsOnTheMove	Echtzeit-Erkennen von Karrierewechseln	CTOsOnTheMove.com bietet einen Dienst, der zeitnah Verkaufshinweise generiert, mit denen seine Kunden ihren Umsatz steigern können. Dazu werden in Realzeit Tausende von Medienquellen, Unternehmenswebsites, Jobbörsen, Regierungspublikationen und Social Media überwacht, um Änderungen in der Geschäftsführung und andere Verkaufs-Trigger zu erkennen.	www.ctosonthemove.com

Dataminr	Signale für den Hochfrequenzhandel	DataMinr beliefert Finanzfirmen mit Eilmeldungen, und versucht dabei, schneller als Nachrichtendienste zu sein. Die Nachrichten gewinnt es, indem es in Echtzeit anomale Ereignisse auf Twitter erkennt und mit Themen aus seiner Wissensbasis korreliert. Es reichert die Nachrichten mit analytischen und statistischen Informationen aus seiner eigenen Unternehmens- und Branchendatenbank an, die für die Investitionsinteressen des Kunden relevant sind.	www.dataminr.com
DigitalReasoning	Compliance Monitoring	Mit seiner Software möchte Digital Reasoning Banken helfen, interne Betrugsversuche zu entdecken und zum Beispiel Warnungen zu erzeugen. Dazu dienen maschinelle Lernverfahren, die unstrukturierte Texte wie E-mails, Tweets und Dokumente zerlegt, verknüpft und analysiert.	www.digitalreasoning.com
Enercast	EE-Vorhersagen	Der Onlineservice erstellt Leistungsprognosen und Einspeisehochrechnungen für Solaranlagen- und Netzbetreiber. enercast.de basiert auf Algorithmen, die vom Fraunhofer IWES entwickelt wurden.	www.enercast.de
Entelo	Geeignete Mitarbeiter finden	Entelo hat einen Algorithmus entwickelt, um Anwerber in Technologie-Unternehmen mit potenziellen neuen Mitarbeitern zusammenzubringen. Entelo hat über 300 Millionen persönlicher Profile aus den Aktivitäten von Personen auf Webseiten wie Github, Stackoverflow, Dribble oder Behance erstellt. Der Algorithmus priorisiert zum Beispiel Ingenieure, wenn deren Firmen schwierige Phasen durchlaufen.	www.entelo.com
Facebook	FB	Das Social-Network Facebook verarbeitet täglich 2,5 Milliarden Inhalte, 2,7 Milliarden Likes und 300 Millionen Fotos. Zusammengenommen ergibt das ein tägliches Datenvolumen von mehr als 500 Terabyte. Jede halbe Stunde scannt das System etwa 105 Terabyte.	www.facebook.com

Farmeron	Dem Puls der Farm näher sein	Farmeron möchte die Verwaltung von Viehfarmen durch intelligentes Datenmanagement vereinfachen und durch entsprechende Analysen ein genaueres Bild zur datenbasierten Entscheidungsfindung liefern.	www.farmeron.com
Food Genius	Big Data fürs Essen	Food Genius unterstützt die Entwicklung neuer Nahrungsmittelprodukte und Rezepte. Die Nutzer können eine Datenbasis mit 230.000 Restaurant-Menüs und 16 Millionen Menüeinträgen nach Zutaten und Zubereitungsmethoden explorieren, um nationale und regionale Trends zu identifizieren.	getfoodgenius.com
Foursquare	Standortbezogenes Social Network	Foursquare ist eine Web- und Mobiltelefon-App, die es registrierten Benutzern ermöglicht, sich mit Freunden zu verbinden und ihren aktuellen Standort bekanntzugeben. Die Nutzer können Tipps zu den Standorten abgeben und austauschen. Die Tipps werden für personalisierte Empfehlungen genutzt. Für jeden Check-in und Tipp werden Punkte verliehen, die zu Vergünstigungen führen.	de.foursquare.com
Implisense	Finden und Binden von Geschäftskunden durch Web-Monitoring	Implisense extrahiert Daten zu Firmen aus öffentlichen Online-Quellen in Echtzeit. Dazu gehören Stammdaten, Aktivitätsdaten sowie Branchendaten aus Deutschland, Österreich und der Schweiz. Anwender können mehrere hundert Firmen gleichzeitig beobachten und aufgrund von aktuellen Ereignissen beim Geschäftskunden eine intelligentere Akquise bzw. Kundenbindung realisieren.	www.implisense.com

Fraunhofer IGD	Exploration zur Erkennung von Kreditkartenbetrug	Der Zeitdruck, um nach der ersten Reklamation eines Kunden effektive Alarmierungsregeln zur Abstellung etwaiger Betrugsstrategien zu finden, ist bei europäischen Spezialisten für Kreditkartensicherheit hoch. Die Lösung sucht dazu in über 25 Attributen (Höhe der Transaktionen, Ort, Währung etc.) der Millionen Transaktionen pro Tag neuartige Muster. Durch die Kombination von automatischer und interaktiver Datenanalyse können gerade sich dynamisch verändernde Muster effektiv untersucht werden.	www.igd.fraunhofer.de
Game Analytics	Platzierung von Angeboten in Online-Spielen	GamesAnalytics bieten Lösungen an, um Online-Spieler zu segmentieren, ihr Verhalten zu prognostizieren und dadurch das Spielerlebnis individuell zu verbessern. Indem den Spielern im passenden Augenblick eine individuelle Nachricht geschickt wird, kann der Umsatz gesteigert werden.	www.gamesanalytics.com
GE Energy	Energieverbrauchssteuerung	GEs Demand Response Management System ist Teil eines Systems zur optimierten Steuerung des Energieverbrauchs. Es benutzt 2-Wege-Kommunikation mit smarten Zählern, programmierbaren vernetzten Thermostaten, smarten Anwendungen, elektrischen Fahrzeugen und Ladestationen sowie Haushalts-Energieportalen. Die Daten werden laufend gesammelt und für lernende Algorithmen verwendet.	www.gedigitalenergy.com
Ginger.io	Mobile Lebens-Monitore	Ginger.io ist eine Ausgründung vom MIT Media Lab mit Bezügen zum Forschungsprogramm Reality Mining von Alex Pentland. Die Ausgründung möchte das Kommunikationsverhalten und die Standortdaten von Smartphone-Nutzern auswerten, um frühzeitig Rückschlüsse auf einen ungesunden Lebenswandel zu ziehen. Davon profitieren sollen Patienten aber auch Krankenversicherungen.	www.ginger.io

Google.org	GoogleFlu Tracker	Google hat festgestellt, dass die Häufigkeit bestimmter Suchbegriffe Anhaltspunkt für die Häufigkeit von Grippeerkrankungen sein kann. Für Google Grippe-Trends werden Daten der Google-Suche gesammelt und ausgewertet. Auf Grundlage der Ergebnisse wird anschließend nahezu in Echtzeit die Häufigkeit von Grippeerkrankungen geschätzt.	www.google.orgflutrends/
HealthNow	Entscheidungs-unterstützung in der Beratung oder Behand-lung	HealthNow, Krankenversicherer und Anbieter medizinischer Leistungen, hat 16 Unternehmens-datenquellen integriert. Zusammen mit Daten aus sozialen Netzwerken, unstrukturierten Daten aus Forderungen und medizinischen Aufzeichnungen werden analytische Vorhersagen zu Problemen in der Gesundheitspflege ermöglicht.	www.informatica.com
IBM	Entscheidungs-unterstützung in der Diagnose und Behand-lung	Genetik und Molekularbiologie machen so schnelle Fortschritte, dass die Diagnose von Krebs und die Wahl der Therapie immer schwieriger wird. Um den evidenzbasierten medizinischen Ansatz im Memorial Sloan-Kettering Cancer Center zu unterstützen, wird IBMs Watson auf die große klinische Datenbasis angesetzt und mit den Analyse- und Entscheidungsunterstützungssystemen des Zentrums kombiniert.	www.mskcc.org
Vestas	Leistungs-prognose für alternative Energien	Bevor eine Windtturbine installiert wird, verlangen Vestas Kunden eine Prognose der erzeugten Energie und des Return of Investment. Im Betrieb soll die Software genutzt werden, um die Performanz vorherzusagen, die Reaktion der Rotorblätter auf Wetteränderungen zu analysieren und die beste Zeit für die Wartung zu bestimmen.	www.vestas.com

IntroHive	Wer kennt wen	IntroHive möchte durch eine Analyse der betriebsinternen Kommunikationsdaten aus E-Mail, LinkedIn, Twitter u.ä. das firmenspezifische Netzwerk erfassen, um Kaltanrufe (Cold calls) bei neuen Kunden zu erleichtern. Die Idee dabei ist, dass viele Beziehungen zu neuen Kunden bereits im Unternehmen bestehen, aber nicht bekannt sind.	www.introhive.com
Lattice Engines	Vertriebsunter-stützung durch Big Data	Um die Umsätze zu steigern, sollen sich die Vertreter auf profitable Segmente konzentrieren. Mit der Software SalesPrism lassen sich Kunden profilieren, Trigger auf Ereignisse setzen und Voraussagen generieren, um Angebote zur rechten Zeit zu unterbreiten. Die Software integriert und analysiert laufend Daten aus Abrechnungssystem, CRM-System, Marketingsystem und von externen Datenverkäufern.	www.lattice-engines.com
Linguamatics	White-Spot Analyse in Patenten	Der Patentdienst von Linguamatics ermöglicht die Analyse von über 10 Millionen Dokumenten aus den größten Patentdatenbanken durch Textmining-Software. In Echtzeit können die Benutzer suchen und zum Beispiel Verknüpfungen, Kategorien und Cluster bilden.	www.linguamatics.com
McLaren	Echtzeit-Überwachung von komplexen Produkten zur Prävention	McLarens Formel-1-Rennwagen produzieren einen Datenstrom, der in Echtzeit mit SAPs HANA In-Memory-Technologie analysiert wird. Prädiktive Modelle vergleichen aktuelle mit historischen Daten, so dass das McLaren Team sofortige Vorbeugemaß-nahmen ergreifen kann.	www.mclaren.com

Mint	Assisted Finance	Mint.com versteht sich als Realzeit-Wirtschaftsindex für Verbraucher. Das Angebot basiert auf anonymen Verkaufsdaten, die das Unternehmen sammelt. Die Ausgaben werden nach Händler, Kategorie, Region ausgewiesen, und zwar als Durchschnittspreise der Transaktionen pro Monat oder nach Popularität (auf Basis der Anzahl der Transaktionen pro Monat). Der Dienst finanziert sich durch personalisierte Werbung.	www.mint.com
Nest Labs	Vernetzte intelligente Produkte	Das smarte Thermostat von NestLabs hilft beim Energiesparen, indem es das Verhalten der Bewohner studiert und prognostiziert, wann Heizung oder Kühlung gebraucht wird. Thermostate in verschiedenen Räumen kommunizieren miteinander und mit der Cloud, wo alle Daten gesammelt, mit weiteren Daten, wie Wetterprognosen kombiniert und analysiert werden. Versprochen wird eine Reduktion der Heiz- und Kühlkosten für Endnutzer.	www.nest.com
Netcraft	AntiPhishing	Mit dem Dienst von Netcraft können Banken und Finanzunternehmen ihre Namen, Marken, Handelsmarken und Slogans im Internet verfolgen, um Phishing Angriffe und Betrugsversuche durch Vorspielen einer falschen Identität zu erkennen. Dazu nutzt der Dienst Host- und Domänennamen von über 200 Millionen Websites und Inhalte der Startseiten von über 50 Millionen Websites zum Vergleich.	audited.netcraft.com
New York	Feinräumige Überwachung großer komplexer Areale	Das »Domain Awareness System« in New York dient der Verbrechensvorbeugung und Terrorismusabwehr. Es erhält Information von ca. 3.000 Straßenkameras, 2.600 mobilen Strahlungsmessern der Polizei und 100 Kennzeichenlesern an Straßen und in Polizeiautos. Es kann verdächtige Aktivitäten erkennen und Alarme in Echtzeit generieren. Außerdem bietet es Zugriff auf das Verbrechensregister, Notrufe, Kennzeichenregister und andere Datenquellen.	www.nyc.gov

Nike	Smarte Sportprodukte	Nike integriert in seine Sportprodukte (Golfschläger, Schuhe, Armband, ...) Sensorik, die die Benutzung misst. Die Kunden können ihre Trainingsdaten auf die Community-Seite Nikeplus.com laden, um zum Beispiel ihre Technik zu verbessern oder sich mit anderen zu messen. Nike möchte damit die Kunden besser an die Marke binden und so die Umsätze erhöhen. Längerfristig sollen die Daten auch für personalisierte Werbung beim Betreten einer Filiale oder per Mobiltelefon genutzt werden.	nikeplus.nike.com/plus
Opera Solutions	Automatische 1:1 Zielgruppenansprache	Ein Lebensmittellieferant hatte sei 5 Jahren schwindende Umsätze. Eine dynamische Langzeitanalyse des Einkaufsverhaltens jedes Haushalts wurde durchgeführt und die Haushalte wurden so geclustert, dass die Verkäufer jetzt individuelle Empfehlungen geben. Sie berücksichtigen das Verhalten von vergleichbaren Haushalten und die Zeit, die seit dem letzten Kauf ein Produkts vergangen ist. Täglich werden 1 Million Empfehlungen generiert.	www.operasolutions.com
Otto	Tägliche Prognose der Abverkäufe	OTTO versucht mit Hilfe von Big Data, den Absatz auf Artikelgrößenebene präziser vorherzusagen. Jährlich werden 40 Millionen Einzelprognosen erstellt, die anhand einer Vielzahl von Faktoren berechnet werden. Dafür fließen täglich bzw. wöchentlich bis zu 100 Millionen Datensätze ins System. Für die Prognose spielen Faktoren, wie der Bewerbungsgrad eines Artikels online und offline, sowie spezifische Artikeleigenschaften und Umfeldbedingungen eine entscheidende Rolle.	www.blue-yonder.com

Paymint	Betrugser-kennung bei Kreditkarten-transaktionen	Mit seiner Fraud Management Lösung MINTify unterstützt PAYMINT seine Kunden dabei, durch die Erkennung von Betrugsversuchen in Echtzeit erhebliche Einsparungen zu realisieren. Das Modul »MINTify rule« erzeugt automatisch Regeln zur Erkennung neuer Betrugsmuster. Je nach Risikobereitschaft und Kundenorientierung können Nutzer flexibel entscheiden, ob die Regel im Echtzeit-Präventionssystem verdächtige Transaktionen abblocken oder lediglich Warnmeldungen generieren soll.	www.paymint.de
Profitero	Echtzeit-Über-wachung der veröffentlichten Preise und Leistungen von Wettbewerbern	Profitero ist ein Dienst im Web, der den Verant-wortlichen für die Preisbildung täglich genaue Informationen über ihre Wettbewerber und Tausende von Online-Produkten gibt. Man kann zusätzlich den Lagerbestand der Konkurrenz und ihre Lieferkosten beobachten und mit den eigenen vergleichen. Die Genauigkeit wird erreicht, indem ein algorithmisches Matching mit einer manuellen Bestätigung kombiniert wird. Der Kundennutzen besteht in der Zeit- und Kostenersparnis.	www.profitero.com
Progressive	Bepreisung nach Nutzung	Der Autoversicherer Progressive bietet sein Snapshot Programm in 40 US-Staaten an. Ein Telematikgerät, das mit dem OnBoard Diagnostic (OBD-II) Port im Auto verbunden wird, überträgt Uhrzeit, Distanzen und Bremsverhalten an die Versicherung. Die Versicherung ist günstiger, wenn das Auto weniger, zu weniger gefährlichen Zeiten und weniger riskant gefahren wird. Die Kunden können damit bis zu 30% sparen.	www.progressive.com

QuantConnect	Algorithmisches Handeln	QuantConnect bietet eine cloudbasierte Simulations- und Testumgebung mit hochaufge-lösten Daten für Entwickler von automatischen Finanztransaktionsalgorithmen und Investoren, die solche Algorithmen suchen. Die Algorithmen sollen auf minimale Standards geprüft werden. Über Partnerschaften sollen Investoren Algorithmen lizensieren können.	quantconnect.com
Quelle Russland	Ausverkäufe vermeiden	Mit einer Predictive-Analytics-Software kann Quelle Russland seine Daten in Echtzeit auswerten und die Kundennachfrage frühzeitig ermitteln. Zu bestimm-ten Terminen wird die Absatzmenge pro Artikel und die Retourenquote sowohl für den Printkatalog als auch den Online-Shop vorhergesagt.	www.blue-yonder.com
Recorded Future	Die Zukunft vorhersagen	Recorded Future analysiert nach eigenen Angaben mehrere hundert Online-Quellen wie z.B. Blogs, Webseiten, Patentanmeldungen, um darin Personen-, Organisations- und Datumsbezeich-nungen mittels Text Mining zu extrahieren. Aus diesen Einzeldaten werden strukturierte Abfragen über künftige Ereignisse möglich, die für die Wettbewerbsanalyse, das Trendmonitoring oder Sicherheitsüberwachungen einsetzbar sind.	www.recordedfuture.com
Rolls Royce	Echtzeit-Überwachung von Maschinen zur Prävention	Das Konzept »power by the hour« geht auf Rolls Royce zurück. Wartungsverträge werden für einen festen Betrag pro Flugzeit angeboten. Das erlaubt dem Unternehmen, die Wartungskosten gut abzuschätzen. Über Sensoren in Maschinen werden bei Rolls Royce Leistungsdaten laufend zentral gesammelt. Dort werden sie analysiert und mit Mustern aus historischen Daten verglichen, um Ausfälle zu prognostizieren und entsprechende Maßnahmen einzuleiten. Außerdem wird darüber leistungsorientiert abgerechnet.	www.logica.dk

SAP	Betrugser-kennung in Lieferketten	Mit RFIDs mit eindeutigen Objektbezeichnungen (Electronic Product Codes, EPCs) können Produkte, zum Beispiel Medikamente, nahtlos vom Produzenten zum Verbraucher in Realzeit verfolgt werden. SAPs GoRFID gestattet, Objektdaten zwischen unabhängigen Stellen in der Lieferkette auszutauschen. Die resultierenden großen Datenmengen können in Sekunden durchsucht und aggregiert werden. So könnte beim Verkauf in der Apotheke die Echtheit eines Medikaments geprüft werden.	epic.hpi.uni-potsdam.de/Home/GoRFID
Sense Networks	Geo-retargeting	Sense Networks sammelt und analysiert Milliarden von Ortsdaten von mobilen Telefonen, Taxis, Kameras und GPS-Geräten in Echtzeit. Dabei werden die eingesetzten maschinellen Lernverfahren ständig verbessert. Die Software gleicht die Koordinaten mit Geschäftsstandorten und demografischen Daten ab und ordnet den Straßenabschnitten Nutzerprofile zu. Dabei werden mehr als 2000 Attribute genutzt. Die Citisense App informiert, wo sich viele Leute befinden. Beim Georetargeting können Personen sehr gezielt beworben werden. Beispiel: Personen, die sich drei Mal pro Woche in einem Kilometer Umkreis von einem bestimmten Geschäft befinden, können kurz vor oder nach einem Besuch in dem Geschäft angesprochen werden. Wertbebotschaften werden innerhalb von Millisekunden erzeugt.	www.sensenetworks.com

Telecom Italia	Churn Analyse	Telecom Italia möchte potenzielle Abwanderer rechtzeitig erkennen und durch gezielte Angebote weiterhin an das Unternehmen binden. Dazu werden die Verbindungsdaten der Kunden analysiert, eng vernetzte Benutzer geclustert, ihre Rolle im Kommunikationsnetz und ihre Verbindung zu Kündigern oder Nutzern aus anderen Telekommuniktaitonsnetzen analysiert . Im Ergebnis wird ein Score für das Risiko eines Anbieterwechsels für jeden Kunden errechnet. Außerdem sollen im Kommunikationsnetz bestehender Kunden latente Neukunden identifiziert und angeworben werden.	www.telecomitalia.com
Telefonica Dynamic Insights	Soziogramme im stationären Handel	Smart Steps zeigt, wie viele Personen ein Gebiet pro Stunde besuchen, differenziert nach Alter und Geschlecht. Es entsteht eine zeitbezogene »Heatmap« mit Rasterfeldern von wenigen 100 qm. Genutzt werden anonymisierte aggregierte Daten aus dem Mobilfunknetz. Mit den Informationen können Geschäfte ihre Werbung spezialisieren sowie Standorte und Filialformate besser planen.	dynamicinsights.telefonica.com
Tesco	Aus einem breiten Spektrum an bestehenden Angeboten wird das für den jeweiligen Kunden am besten passende ausgewählt	Sogenannte «Next Best Offers" sind spezifische Angebote, die aufgrund von demographischen Daten des Verbrauchers und seinem Einkaufsverhalten automatisch erzeugt werden. Mit seiner NBO-Strategie möchte Tesco die Umsätze mit regulären Kunden vergrößern und sie mit spezifischen Coupons im Clubcard-Programm binden.	www.tesco.uk

The Climate Corporation	Preisgestaltung durch bessere Risikoabschätzung lokaler Bedingungen	Die Climate Corporation versichert Risiken in der Landwirtschaft. Dazu prognostiziert sie Risiken und ermittelt aktuelle Schäden auf Mikroebene für 20 Millionen Farmen in den USA. Sie sammelt Temperatur-, Niederschlags-, Bodenfeuchtigkeits- und Ertragsdaten über ein großes Sensornetz. Die Auszahlungen berechnen sich automatisch danach, ob diese Faktoren eine Schwelle überschreiten, bei der die Ernte gefährdet ist, und zwar pro Tag und Farm.	www.climate.com
Treato	Social Network für Patienten	Im Internet tauschen Patienten und Ärzte in Tausenden von Foren und Blogs ihre Erfahrungen mit Behandlungsmethoden von Krankheitsbildern aus. Die Analysen helfen zum Beispiel herauszufinden, welche Nebenwirkungen eines Medikaments Patienten am häufigsten beschreiben oder mit welcher Behandlung sie bei vergleichbaren Erkrankungen zufrieden sind. Bisher dauerte es oft mehrere Jahre, bis Behörden ein Medikament mit starken Nebenwirkungen vom Markt genommen haben oder bis ein erfolgreiches Medikament für bis dahin nicht bekannte Krankheitsbilder zum Einsatz kam.	www.treato.com
TreeMetrics	Den Wald vor lauter Bäumen sehen	TreeMetrics has developed a unique set of solutions to better measure and improve the management of the forest resource. Treemetrics is the first company in the world to offer both Aerial and Terrestrial 3d laser scanning systems. Both systems combined offer incredible new information on the structure of a forest.	www.treemetrics.com
TuneSat	Abrechnung der tatsächlichen Nutzung von digitalen Gütern	TuneSat erzeugt eindeutige Audio-Fingerabdrücke und überwacht damit Hunderte von Fernsehkanälen und Millionen von Webseiten, so dass die Urheber die ihnen zustehenden Gebühren kassieren können.	www.tunesat.com

United Overseas Bank	Überwachung von Zahlungsströmen auf mögliche Betrugsversuche	Bei der Unitied Overseas Bank in Singapur sind die Risiken auf etwa 45.000 verschiedene Finanzinstrumente verteilt und werden bestimmt über etwa 100.000 Marktparameter (Preise, Fristen, Fälligkeiten, etc.). Mit In-Memory-Technologie und »Complex Event Processing« von SAS konnte die Berechnung des Gesamtrisikos von 18 Stunden auf wenige Minuten gesenkt werden. Schnell sich ändernde Parameter können quasi in Echtzeit einbezogen werden. Im operativen Geschäft können damit Handelsstrategien im Voraus geprüft und neue Marktereignisse in ihren Wirkungen schneller eingeschätzt werden.	www.uob.com.sg
US Xpress	Flottenoptimierung	Das Ziel von US Xpress ist, alle Fahrzeuge in Bewegung zu halten und insbesondere Wartungszeiten zu minimieren. Dazu werden in Echtzeit 900 Datenelemente von zehntausenden Fahrzeugen (Reifen- und Spritverbrauch, Maschinendaten, GPS) und Fahrer-Feedback aus Social Media gesammelt und parallel in verschiedene operationelle und analytische Systeme übertragen. Ein Ergebnis daraus ist die Vorgabe, dass bei Kreuzungen in den USA stets rechts abgebogen werden soll, um jährlich 10 Mio. Dollar durch vermiedene Unfälle und Wartezeiten zu sparen.	www.usexpress.com
Visapix	Analysieren der Besucherströme in Filialen und Standorten	Um den Kundenservice nicht nur an zahlenden Kunden auszurichten, sondern auch Besucher in die Bedarfsplanung einzubeziehen, hat sich das Schuhhaus Darre´ für die Besucherfrequenzmessung per Videoanalyse-Software entschieden. Durch die Auswertung werden jetzt alle Besucher, die das Geschäft betreten, und ihre Bewegung zwischen den einzelnen Abteilungen erfasst. Die Ergebnisse werden für die Berechnung des Personalbedarfs im Tages-, Wochen- und Saisonverlauf an die Personalplanungs-Software weitergeleitet.	www.visapix.de

IMPRESSUM

HERAUSGEBER

Fraunhofer-Institut für Intelligente Analyse-
und Informationssysteme IAIS
Schloss Birlinghoven
53757 Sankt Augustin
Telefon 02241 14-2252
pr@iais.fraunhofer.de
www.iais.fraunhofer.de/bigdata

Die Innovationspotenzialanalyse Big Data wurde gefördert mit Mitteln des
Bundesministeriums für Wirtschaft und Technologie im Kontext des THESEUS-Programms.

KOORDINATION DER STUDIE

Dr. Michael May
Abteilung Knowledge Discovery

AUTOREN

Dr. Andreas Schäfer, Dr. Melanie Knapp, Dr. Michael May, Dr. Angelika Voß
Abteilung Knowledge Discovery

REDAKTION

Katrin Berkler M.A., Robert Müller B.A.
Presse und Öffentlichkeitsarbeit

GESTALTUNG & GRAFIKEN

Kristina Reusch, Mediengestalterin Digital und Print

© Fraunhofer-Institut für Intelligente Analyse- und
Informationssysteme IAIS, Sankt Augustin, 2012