



Bundesministerium  
für Wirtschaft  
und Energie



Smart Data

---

# Smart Data – Innovationen aus Daten

---

*Ergebnisbroschüre*



## Impressum

### Herausgeber

Bundesministerium für Wirtschaft und Energie (BMWi)  
Scharnhorststraße 34-37  
10115 Berlin  
Tel.: 030/186 15 – 6141  
Fax: 030/186 15 – 5208  
E-Mail: [info@bmwi.bund.de](mailto:info@bmwi.bund.de)  
[www.bmwi.de](http://www.bmwi.de)

### Text und Redaktion

LoeschHundLiepold Kommunikation GmbH, Berlin

### Gestaltung und Produktion

PRpetuum GmbH, München

### Stand

November 2017

### Druck

Druck- und Verlagshaus Zarbock GmbH & Co. KG,  
Frankfurt

### Bildnachweis

Victoria – Fotolia (Titel),  
Elnur – Fotolia (S. 7),  
KARL STORZ GmbH & Co. KG (S. 10),  
Projekt SAHRA/ SAS Institute (S. 13),  
Ravil Sayfullin – Fotolia (S. 16),  
Vege – Fotolia (S. 20),  
SD4M – Smart Data for Mobility (S. 23),  
StEB Köln (S. 25),  
Kara – Fotolia (S. 27),  
Fraunhofer FOKUS (S. 30),  
NicoElNino – Fotolia (S. 32),  
industrieblick – Fotolia (S. 34),  
Patrick P. Palej – Fotolia (S. 36),  
MAGNIFIER – Fotolia (S. 40)  
DFKI (S. 43/48),  
Flughafen Stuttgart GmbH (S. 51),  
peshkova – Fotolia (S. 54)

Diese Broschüre ist Teil der Öffentlichkeitsarbeit des Bundesministeriums für Wirtschaft und Energie. Sie wird kostenlos abgegeben und ist nicht zum Verkauf bestimmt. Nicht zulässig ist die Verteilung auf Wahlveranstaltungen und an Informationsständen der Parteien sowie das Einlegen, Aufdrucken oder Aufkleben von Informationen oder Werbemitteln.



Das Bundesministerium für Wirtschaft und Energie ist mit dem audit berufundfamilie® für seine familienfreundliche Personalpolitik ausgezeichnet worden. Das Zertifikat wird von der berufundfamilie gGmbH, einer Initiative der Gemeinnützigen Hertie-Stiftung, verliehen.



Diese und weitere Broschüren erhalten Sie bei:  
Bundesministerium für Wirtschaft und Energie  
Referat Öffentlichkeitsarbeit  
E-Mail: [publikationen@bundesregierung.de](mailto:publikationen@bundesregierung.de)  
[www.bmwi.de](http://www.bmwi.de)

Zentraler Bestellservice:  
Telefon: 030 182722721  
Bestellfax: 030 18102722721

# Inhalt

Smart Data – Das Programm .....	2
Smart Data – Innovationen aus Daten .....	3
KDI – Künstliche Datenintelligenz .....	6
InnOPlan- Innovative, datengetriebene Effizienz OP-übergreifender Prozesslandschaften .....	9
SAHRA – Smart Analysis – Health Research Access .....	12
FastGenomics – Fast Analysis of Single Cell Transcriptomics .....	15
Fachgruppe Rechtsrahmen .....	17
iTesa – Intelligent Traveller Early Situation Awareness .....	19
SD4M – Smart Data for Mobility .....	22
sd-kama – Smart-Data-Katastrophenmanagement .....	24
ExCELL – Echtzeitanalyse und Crowdsourcing für eine selbstorganisierte City-Logistik .....	26
VIRTUOSE-DE – Service-Plattform für echtzeitfähige Big-Data-Videoanalyse und -verarbeitung in der Cloud .....	29
Fachgruppe Sicherheit .....	32
SAKE – Semantische Analyse Komplexer Ereignisse .....	34
PRO-OPT – Big-Data-Produktionsoptimierung in Smart Ecosystems .....	36
SIDAP – Skalierbares Integrationskonzept zur Datenaggregation, -analyse, -aufbereitung von großen Datenmengen in der Prozessindustrie .....	39
Smart Data Web – Datenwertschöpfungsketten für industrielle Anwendungen .....	42
Fachgruppe Wirtschaftliche Potenziale und gesellschaftliche Akzeptanz .....	45
SmartRegio – Trend-Analysen auf Basis heterogener Massendaten .....	47
SmartEnergyHub – Datendrehscheibe für intelligente Energienutzung .....	50
EDV – Einfaches Digitales Vergessen .....	53

# Smart Data – Das Programm

Die Entwicklung innovativer Datenanalyseverfahren sowie beeindruckender Verfahren zur Speicherung und Suche großer, hochdimensionierter Datenmengen hat unsere Welt verändert. Unser Leben wird immer digitaler, damit aber auch viel abhängiger von der Verfügbarkeit digitaler Kompetenz und vom vernünftigen Umgang mit diesen Daten. Das Programm „Smart Data - Innovationen aus Daten“ des Bundesministeriums für Wirtschaft und Energie (BMWi) greift diese Thematik nicht nur auf, sondern geht noch einen Schritt weiter.

Denn neben der Fokussierung auf die Verarbeitung großer Datenmengen, wird ebenfalls deren Verknüpfung und damit eine noch bessere Interpretation und Verwertung der Daten als nutzbare Information in den Blick genommen. Daher der Name „smart“, der darauf hinweisen soll, dass die Nutzung der gewonnenen Information aus den großen Datenmengen ein Schlüssel für digitale Innovation ist und auf diese Weise letztlich auch ein Schlüssel zum Erhalt der Wettbewerbsfähigkeit unserer Industrie ist.

Das Programm „Smart Data“ hat sich auf die vier speziellen Anwendungsfelder „Industrie“, „Mobilität“, „Energie“ und „Gesundheit“ fokussiert, die zum einen wichtige Pfeiler unserer Wirtschaft sind und von denen zum anderen gleichermaßen eine erhebliche gesellschaftliche Relevanz aus-

geht. Nicht zuletzt sollen sie als Muster für andere Anwendungsfelder dienen.

Neben den technischen Ergebnissen der Einzelprojekte finden Sie im vorliegenden Bericht übergeordnete Themen, die im Laufe des Programms eine zentrale Rolle eingenommen haben. Diese Themen dürften ebenso wie die Anwendungsfelder eine entscheidende Funktion für weitere Entwicklungen unserer Wirtschaft und Gesellschaft einnehmen.

Zwischen all diesen Themen ist besonders ein Kernthema der Digitalisierung, das als Identity Management zu benennen ist, herausgestochen. Denn digitale Identitäten nehmen in der heutigen Zeit einen wichtigen Stellenwert ein. Dennoch hinterfragen wir viel zu selten die Herkunft und Vertrauenswürdigkeit dieser Daten. Mehr zu dieser Problematik und zu weiteren, spannenden Themen können Sie im Folgenden nachlesen.

Ich wünsche Ihnen viel Spaß bei der Lektüre und hoffe, dass Sie dabei neue Impulse, Ideen und Einsichten zum Thema „Smart Data“ erhalten.

Zuletzt wünsche ich Ihnen wie uns viel Erfolg auf dem Weg zu einer „smarten Gesellschaft“.

*Ihr Prof. Dr. Dr. Stefan Jähnichen  
Leiter der Begleitforschung „Smart Data – Innovationen aus Daten“*

# Smart Data – Innovationen aus Daten

Innovationen aus Daten – unter diesem Motto ist das Technologieprogramm Smart Data vor rund vier Jahren gestartet. Daten sind ein immer wertvolleres Wirtschaftsgut, aus dem gewinnbringende Resultate erzeugt werden können. Unzählige Studien prognostizieren ein rasantes Wachstum, so dass sich die verfügbare Menge an Daten und damit auch deren wirtschaftliche Nutzung um ein Vielfaches steigern werden – und das in kürzester Zeit. So wird sich laut dem „Cisco Visual Networking Index™ (VNI)“ der weltweite Datenverkehr bis zum Jahr 2021 verdreifachen und erreicht dann ein Volumen von 3,3 Zettabyte pro Jahr.<sup>1</sup>

Bereits heute nutzt ein Fünftel der deutschen Unternehmen Big Data.<sup>2</sup> Die riesigen Datenmengen dienen als Rohstoff, den es automatisiert aufzubereiten gilt, um sein gesamtes wirtschaftliches Potenzial zu nutzen. Durch die Strukturierung und Analyse von Massendaten und ihre anschließende Anreicherung bspw. mit semantischen Modellen, kann gewinnbringendes Wissen erzeugt werden – so wird Big Data zu Smart Data. In den vergangenen Jahren ist ein regelrechter „Trend“ im Bereich der Smart-Data-Lösungen zu verzeichnen, der sich in einem rasanten Wachstum der wirtschaftlichen Nutzung widerspiegelt. Der Umsatz von Smart-Data-Lösungen wird zum Jahr 2025 auf 85 Milliarden Euro prognostiziert, wie die Kurzstudie „Smart-Data-Business – 10 Thesen zur Nutzung von Big-Data-Lösungen in der Wirtschaft“ aus dem Jahr 2015 berichtet.<sup>3</sup>

Das Programm „Smart Data – Innovationen aus Daten“ des Bundesministeriums für Wirtschaft und Energie (BMWi) neigt sich 2018 dem Ende zu. Das Förderprogramm startete im Jahr 2014 und hat einen wichtigen Beitrag geleistet, den Markt für Smart-Data-Technologien in Deutschland breitenwirksam zu erschließen. Sechzehn Leuchtturmprojekte mit Vorbildcharakter insbesondere für die Anwendung bei kleinen und mittelständischen Unternehmen werden aktuell mit rund 36 Millionen Euro gefördert. Vier Anwendungsbereiche wurden dazu in den Fokus gerückt: Industrie, Mobilität, Energie und Gesundheit.

In der Industrie kann die intelligente Analyse von Daten Qualitäts-, Optimierungs- und Effizienzsteigerungen herbeiführen, aber auch die individuelle Anpassung von Pro-

dukten an Kundenwünsche gewährleisten. Nur durch die Auswertung von Produktionsdaten können autonome Systeme vorangetrieben und die effiziente Fertigung von individuellen Produkten in „Losgröße 1“ realisiert werden. Das Zukunftsthema „Industrie 4.0“ ist somit in großem Maße von Smart-Data-Technologien abhängig.

Im Anwendungsbereich der Mobilität handelt es sich bei den Daten überwiegend um Informationen aus Fahrzeugen und Infrastrukturkomponenten. Sie können beispielsweise bei der Vermeidung von Staus und Kollisionen helfen, aber auch neue zielgruppenspezifische Mobilitätskonzepte entwickeln und unterstützen.

Im Energiesektor gibt es durch den vermehrten Einsatz intelligenter Netzkomponenten, Stromzählern und vernetzten Verbrauchsgeräten ein erhöhtes Datenaufkommen. Mit deren Analyse können wichtige Handlungsfelder wie die Integration des Stroms aus erneuerbaren Energiequellen in die vorhandenen Netzstrukturen, die Steuerung von Energieströmen und das Umgehen von Spannungsschwankungen und Stromausfällen erschlossen werden.

Auch in der Medizin sorgen Smart-Data-Technologien in Anwendungsbereichen wie Medizintechnik, Gesundheitsmanagement, klinische Forschung oder Biotechnologie für neue Wege. Ebenso profitieren die Diagnostik und Therapie von der intelligenten Analyse der Daten. Smart-Data-Lösungen bilden auch die Grundlage für viele individualisierte und fallspezifische Behandlungen, die den Diagnoseprozess und die anschließenden Behandlungen enorm verbessern.

## Vier Jahre „Smart Data“ – Ergebnisse, Erfolge und Herausforderungen

Das Technologieprogramm „Smart Data – Innovationen aus Daten“ hat anhand seiner sechzehn Leuchtturmprojekte gezeigt, wie große Datenmengen – Big Data – zu nutzbringender Information aufbereitet werden können und damit maßgeblich zur Entwicklung neuer intelligenter Dienste sowie der Verbesserung bestehender Anwendungen beitragen.

1 Cisco Systems, Inc., Cisco Visual Networking Index™ (VNI)

2 Bundesministerium für Wirtschaft und Energie (BMWi); Monitoring-Report| Kompakt – Wirtschaft DIGITAL 2017

3 Smart-Data-Begleitforschung in Zusammenarbeit mit dem Business Application Research Center (BARC) und dem Bundesverband der IT-Anwender e.V. (VOICE); Smart-Data-Business – 10 Thesen zur Nutzung von Big-Data-Lösungen in der Wirtschaft

Um den langfristigen Nutzen der entwickelten Technologien sicherzustellen, legte die Begleitforschung großen Wert darauf, den Markteintritt optimal vorzubereiten. Zentrales Anliegen ist in diesem Zusammenhang die nachhaltige Nutzung der Technologie sowie deren Übertragbarkeit auf verwandte Anwendungsbereiche. Vor diesem Hintergrund evaluiert die Begleitforschung gemeinsam mit den Projektkonsortien die wirtschaftlichen Potenziale, die gesellschaftliche Akzeptanz sowie die rechtlichen Rahmenbedingungen und Möglichkeiten einer angemessenen IT-Sicherheit – jeweils in einem individuellen, auf die konkrete Projektarbeit bezogenen Ansatz.

Trotz der individuellen Herausforderungen einzelner Projekte, die sich auch im Evaluationsprozess herauskristallisiert haben, lassen sich durchaus einige Parallelen feststellen. So wurde im Laufe der Projektarbeit deutlich, dass insbesondere das Inkrafttreten der Datenschutz-Grundverordnung und ihr absehbarer Anwendungsbeginn Einfluss auf den Entwicklungsprozess einer Technologie haben kann. Mit Blick auf die Sensibilität von Daten hat die neue Norm zu einem Umdenken geführt und dazu beigetragen, dass Datenschutz nun am Anfang des Entwicklungsprozesses steht. Auch in Sicherheitsfragen konnte festgestellt werden, dass sich die Projektkonsortien entsprechenden Herausforderungen immer früher stellen und nicht erst bei bereits fortgeschrittenem Projektverlauf. Dies generiert Vertrauen und Akzeptanz in einer Gesellschaft, in der Sicherheit nicht mehr nur eine technische Spielerei, sondern obligatorischer Bestandteil einer Technologie sein soll.

Deutlich wurde im Programm auch, welchen hohen Stellenwert insbesondere die Bearbeitung projektübergreifender Themen innehat. Im Fokus des Technologiestandards standen nicht nur die einzelnen Förderprojekte, sondern insbesondere die projektübergreifende Zusammenarbeit bei erfolgskritischen Themen, die es bei Forschung und Entwicklung zu berücksichtigen gilt. In den Arbeitsgruppen „Wirtschaftliche Potenziale“, „Rechtsrahmen“ und „Sicherheit“ wurden solche projektübergreifenden Themenstellungen analysiert und diskutiert, Lösungsansätze erarbeitet sowie Synergien zwischen den Projekten aufgezeigt. Analysen und Forderungen zur Nutzung von Open Data, Diskussionspapiere zum Datenschutz oder den künftigen Anforderungen an Smart-Data-Fachkräfte stehen beispielhaft für den breiten Themenkanon, der hier bearbeitet wurde und durch entsprechende Publikationen einen nachhaltigen und öffentlichen Mehrwert hervorbrachte.

Dabei wurden die Projekte nicht nur in ihrer Entwicklung begleitet und gemeinsame Herausforderungen herausgearbeitet, sondern auch weiterhin offene Fragestellungen identifiziert. Sehr früh kristallisierte sich hier die Verwaltung und Prüfung von Identitäten, das sogenannte Identity Management, als eine zentrale Herausforderung für die Entwicklung der Smart-Data-Projekte heraus. Woher stammen Daten? Sind sie vertrauenswürdig? Wer ist der Absender? Ist der Empfänger für den Erhalt bestimmter Daten legitimiert und darf er sie weitergeben? Die Bearbeitung dieser Fragen zum Identitätsmanagement steht sinnbildlich für die Verschränkung der Inhalte der drei Arbeitsgruppen – und die Herausforderung, dass hier an vielen Stellen noch Pionierarbeit geleistet werden muss, zum Beispiel bei der Erarbeitung von technischen Schutzmechanismen.

Die Analyse des ISÆN-Konzepts (Individual perSonal data Auditable addrEss) durch die Smart-Data-Begleitforschung hat dieses Konzept als einen vielversprechenden Ansatz für die Kontrolle über personenbezogenen Daten im Internet bewertet und steht damit als ein Ergebnis der Arbeitsgruppen, das zeigt, welche Ansätze und Konzepte es bereits für viele der Herausforderungen gibt. Es bleibt jedoch ein Prozess mit noch vielen offenen Fragestellungen, die die Erarbeitung zukünftiger Smart-Data-Technologien beeinflussen.

Lösungen dürfen aber nicht nur auf nationaler Ebene gesucht werden, sondern müssen auf europäischer oder vielmehr noch internationaler Ebene gefunden werden. Die neue Datenschutzgrundverordnung (DSGVO) der Europäischen Union, die ab Mai 2018 gilt, und Datenschutzregeln für Unternehmen und Behörden deutlich verschärft, ist mehr als nur ein Indiz für die Dringlichkeit dieses Themas.

Um Smart-Data-Lösungen zu entwickeln, prüfen und umzusetzen, reicht es aber nicht aus, rechtlich auf internationaler Ebene zu agieren. Vielmehr müssen internationale Standards institutionalisiert werden, um eine globale Ausstrahlung und vor allem die Erhaltung der Wettbewerbsfähigkeit zu gewährleisten. Die Begleitforschung konnte gemeinsam mit Projekten des Programms unter anderem eine DIN Spec (DIN Spec 91349) erarbeiten, die einen Rahmen für die Beschreibung und den Einsatz von Normen, Standards aber auch für Vorgehensweisen und Methoden im Umgang mit Massendaten schafft. Eine Legitimation der Norm auf europäischer Ebene soll im nächsten Schritt erfolgen. Es ist zu wünschen, dass dieser Erfolg dazu beiträgt,

auch die Zukunftsfähigkeit der Förderprojekte, sowohl national als auch international, voranzutreiben.

Wie die zahlreichen, auch internationalen Auftritte der Smart Data-Projekte belegen, besteht nicht nur in Deutschland ein großes Interesse an der Verwertung der gewonnenen Ergebnisse. Das Technologieprogramm hat an zahlreichen Stellen konkrete Lösungen hervorgebracht, die vielerorts schon aktiv in Unternehmen oder Forschungseinrichtungen eingesetzt werden oder weiterentwickelt werden sollen.

Zur vollen Ausschöpfung des Potenzials einer Technologie, ist eine entsprechende gesellschaftliche Akzeptanz notwendig. Während die Auswirkungen und Implikationen der DSGVO noch nicht in vollem Umfang absehbar sind, gewinnen die Themen maschinelles Lernen und künstliche Intelligenz zunehmend an Stellenwert. Die Ergebnisse des Technologieprogramms „Smart Data – Innovationen aus Daten“ zeigen die Korrelation rechtlicher Rahmenbedingungen und Sicherheit sowie gesellschaftlicher Akzeptanz auf. Eine ganzheitliche und weitsichtige Vorgehensweise wurde in diesem Kontext als Chance erkannt: Die Projekte legen ihren Fokus bereits auf diese Aspekte und sind vor diesem Hintergrund als Vorreiter auf diesen Gebieten zu bewerten.

# KDI – Künstliche Datenintelligenz

## Patientendaten verknüpfen, Behandlung optimieren

### Das Projekt in Kürze

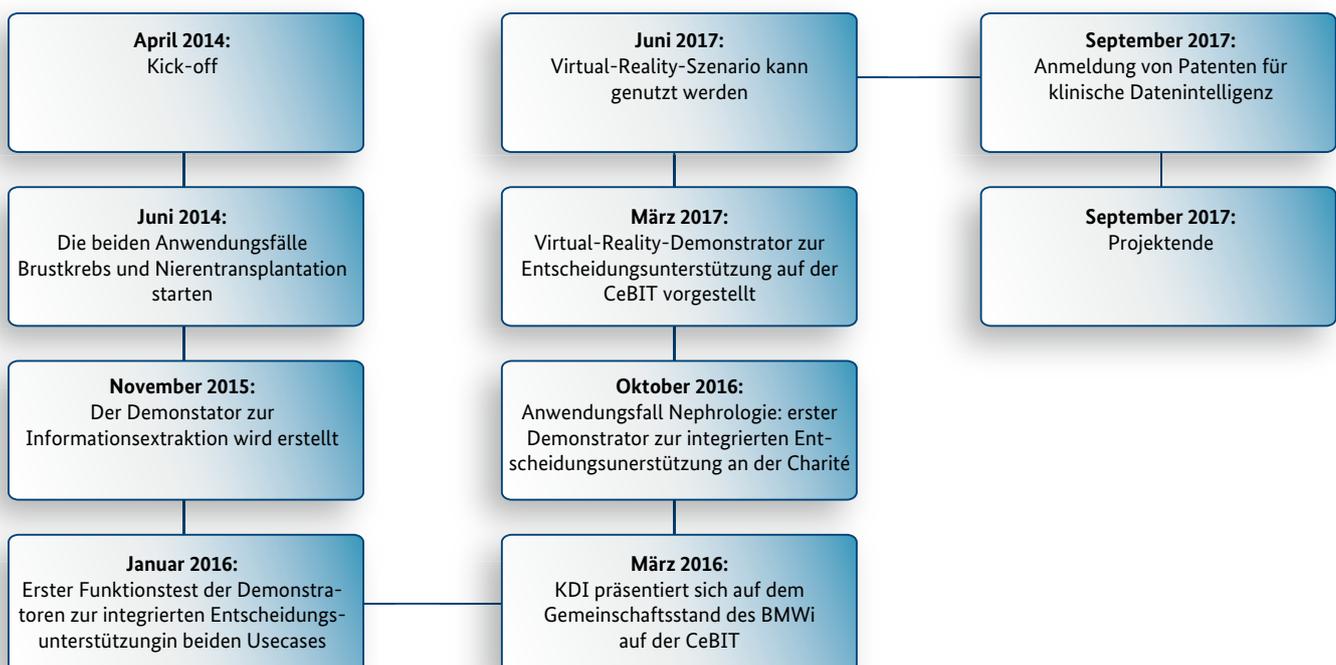
Ob Informationen über Diagnosen, Krankheitsverläufe, Medikationen oder Therapien – Patientendaten im klinischen Umfeld stammen heutzutage aus zahlreichen Quellen. Die Daten liegen in unterschiedlichen Formaten und Strukturen vor, etwa als Arztbrief, Röntgenbild oder als Diagnosedaten aus EKG oder CT, in Schriftform, als Bild oder auch als Audiodatei. Datensätze sind in der Regel gar nicht oder nur unzureichend untereinander vernetzt und behandelnde Ärztinnen und Ärzte besitzen eingeschränkte fallangepasste Datensichten. Dabei ist eine gesamtheitliche Betrachtung der Patientinnen und Patienten mit einer Verknüpfung aller vorhandenen Informationen wichtig, um Krankheitsrisiken besser zu prognostizieren und bei Erkrankungen, zum Beispiel Krebs, eine optimale Therapie zusammenstellen zu können.

Im Projekt KDI wurden umfangreiche und komplexe medizinischen Datensätze zusammengetragen und automatisiert ausgewertet, um klinische Entscheidungsprozesse zu unterstützen. Die Daten aus unterschiedlichen Quellen werden dafür in einem Patientendaten-Modell zusammengeführt. Mit Methoden des Maschinellen Lernens wurden

Module zur Entscheidungsunterstützung entwickelt, die die Gesamtheit der zusammengeführten Daten auswerten, darauf basierend Entscheidungen vorschlagen und deren Konsequenzen bewerten. So sollen bessere und beschleunigte Behandlungserfolge erreicht werden. Zudem können Prozesse über verschiedene Kliniken und Zeiträume hinweg analysiert werden, um neue innovative Dienste für die Versorgung und Forschung zu entwickeln. Durch eine intelligente Suche (eine sogenannte „facettierte Suche“) können gezielter bestimmte Patientengruppen anhand eines automatisch vorgeschlagenen Kriteriums ausgewählt werden, wodurch eine verbesserte Analyse und Erkennung bestimmter Muster von möglichen Krankheitsbildern oder Ähnlichem möglich wird.

Die erarbeitete Lösung lässt sich prinzipiell auf jedes Krankheitsbild anwenden. Im Projekt wurde exemplarisch ein Brustkrebs-Demonstrator entwickelt, der Ärztinnen und Ärzten bei Behandlungen hilft, sowie eine Anwendung, die das Informationsmanagement und die Therapie bei Nierentransplantationen begleitet.

### Meilensteine im Projektverlauf



## Anwendung in der Praxis und Ausblick

Im ersten Anwendungsszenario zum Thema Brustkrebs konnte eine große Zahl von Fällen aus der „Bavarian Breast Cancer Cases and Controls Studie“ genutzt werden. Für alle Patientinnen lagen Mammografie-Daten aus dem radiologischen Archiv des Universitätsklinikum Erlangen vor, die mit weiteren genetischen Daten, wie z. B. Genexpressionsprofilen des Tumormaterials und Informationen aus der elektronischen Patientenakte, verknüpft werden konnten. Ziel des ersten Anwendungsszenarios war es, die unterschiedlichen Datentypen datenschutzkonform in eine Plattform zu integrieren und allen behandelnden Ärztinnen und Ärzten sowie den Patientinnen über eine App zur Verfügung zu stellen. Hier ist es gelungen, eine Arbeitsplattform für Ärzte bereitzustellen, die integrierte Entscheidungsunterstützung bei den Behandlungsschritten und der Auswahl der Medikation leistet. Prädiktive Modelle, die mit Methoden des Maschinellen Lernens trainiert wurden, schlagen Therapieentscheidungen vor, bewerten die Konsequenzen der Entscheidungen und können auf ungewöhnliche Entscheidungen der Ärztin oder des Arztes hinweisen.

Bei dem Anwendungsfall zur Nierentransplantation konnte auf Daten der Charité zurückgegriffen werden: Informatio-

nen zu Grund- und Begleiterkrankungen von Patientinnen und Patienten, aber auch Medikations- und Therapiepläne, Transplantations- und Spenderdaten sowie Untersuchungsbefunde konnten genutzt werden. Die Kombination dieser Daten war besonders erfolgsversprechend, weil in der Transplantationsdatenbank Ergebnisdaten zur Verfügung standen, die den Erfolg bestimmter Therapien bestätigten. Therapeutische Entscheidungen im Bereich der Transplantationsmedizin werden häufig auf der Basis von komplexen Informationen aus verschiedenen Datenquellen getroffen: So werden bei der Zuweisung von Organen eine Vielzahl von Parametern aus unterschiedlichen Datenquellen berücksichtigt, die schließlich in Echtzeit zu Entscheidungen mit hoher Tragweite für den individuellen Patienten führen. Gleichzeitig sind die Entscheidungsprozesse auch modellhaft für andere Disziplinen der klinischen Medizin. Auch in diesem Anwendungsfall, der Nierentransplantation, konnte ein Echtzeitsystem für die automatische Entscheidungsunterstützung entwickelt werden. Zudem wurde eine facetiierte Suche an der Charité installiert. Diese kommt unter anderem bei der Auswahl von Patientinnen und Patienten zur Teilnahme an Studien (Kohortenselektion) zum Einsatz.



## Die Projekt-Ergebnisse auf einen Blick

Früher	Mit KDI
In der klinischen Praxis werden der Ärztin/dem Arzt medizinischen Bilder ohne explizite Verlinkung zu dazugehörigen Textbefunden angezeigt. Er benötigt viel Zeit, um die krankheitsbezogenen Informationen der Bilder mit dem Text zu kombinieren.	KDI ermöglicht es, Informationen aus medizinischen Bildern (z.B. Röntgenaufnahmen) textuell zu annotieren, zu klassifizieren und zeitlich darzustellen. Entscheidungen können so schneller getroffen werden.
Zur Qualitätskontrolle werden im klinischen Alltag sämtliche Informationen zu Befunden und Therapieentscheidungen zu spezifischen Patientengruppen händisch dokumentiert und ausgewertet.	In KDI wurde eine Möglichkeit zur Informationsextraktion aufgebaut, mit der die bis dato händisch erfassten Befund- und Therapiedaten automatisiert analysiert werden können.
Patientendaten müssen aus unterschiedlichen Quellen zusammengesucht werden. Heterogene Daten werden nicht in eine gemeinsame Struktur überführt.	Datenquellen werden durch automatische Digitalisierung und Informationsextraktion von medizinischen Texten in einem Patientendatenmodell automatisch miteinander verknüpft und für eine Analyse und mögliche Entscheidungshilfe (Deep Learning) genutzt.
Es konnte keine interaktive Suche in Patientendaten stattfinden. Ergebnisse konnten nur einzeln, aber nicht für eine spezielle Patientenauswahl (Kohortenselektion) ausgegeben werden.	Die facettrierte Suche funktioniert in Echtzeit und ermöglicht Ärztinnen und Ärzten das gezielte Auswählen von Patientengruppen anhand bestimmter Merkmale. Sowohl strukturierte medizinische Daten als auch unstrukturierte Texte können auf bestimmte Fakten durchsucht werden, die Nutzung international etablierter Datenbankmodelle ermöglicht eine weltweite Zusammenarbeit.
Ärzten fehlt die Entscheidungsunterstützung bei telemedizinischen Anwendungen.	Ein Demonstrator zur Entscheidungsunterstützung in Echtzeit mittels Virtual Reality unterstützt Ärzte bei der Behandlung.
MRT-Bilder von Tumoren sind für sich genommen nicht sehr zuverlässig, um weitere Vorhersagen zu möglichen Mutationen zu treffen.	Durch die Kombination von MRT-Aufnahmen mit zusätzlichen Patienteninformationen, wie z. B. zur Familiengeschichte, können Vorhersagemethoden künftig als unterstützendes Kriterium zur gezielteren Auswahl für spezielle genetische Test und Behandlungen dienen.
Es existiert kein Beispiel eines „Clinical Data Science Centers“ in Deutschland, in dem Technologien wie beispielsweise Algorithmen und maschinelles Lernen genutzt werden, um Krankheiten zu diagnostizieren und zu behandeln.	Erste Tests mit KDI-Daten finden auf dem DFKI-Supercomputer statt.

# InnOPlan – Innovative, datengetriebene Effizienz OP-übergreifender Prozesslandschaften

## Prozesse im OP-Umfeld verbessern

### Das Projekt in Kürze

Der Operationssaal ist das Herz eines Krankenhauses, da dort wichtige klinische, patientenbezogene und administrative Prozesse zusammenlaufen. Besonders hier ist ein reibungsloser Betrieb von zentraler Bedeutung. Häufig sind die vielfältigen Prozesse jedoch nicht ideal aufeinander abgestimmt und die eingesetzten medizinischen Geräte nur zu einem geringen Grad miteinander verknüpft. Dabei bietet gerade eine stärkere Vernetzung für Krankenhäuser, Klinikpersonal und Patientinnen und Patienten großes Potenzial.

Das Projekt InnOPlan setzt genau hier an: Abläufe im OP sowie OP-übergreifende Prozesse in Krankenhäusern sollen durch die Nutzung von Smart Data optimiert werden. Der Ansatz von InnOPlan ist es, medizinische Geräte so zu erweitern, dass sie nicht mehr nur ihre unmittelbare technische Funktion (z. B. Saugen und Spülen) erfüllen, sondern zu intelligenten Datenlieferanten werden.

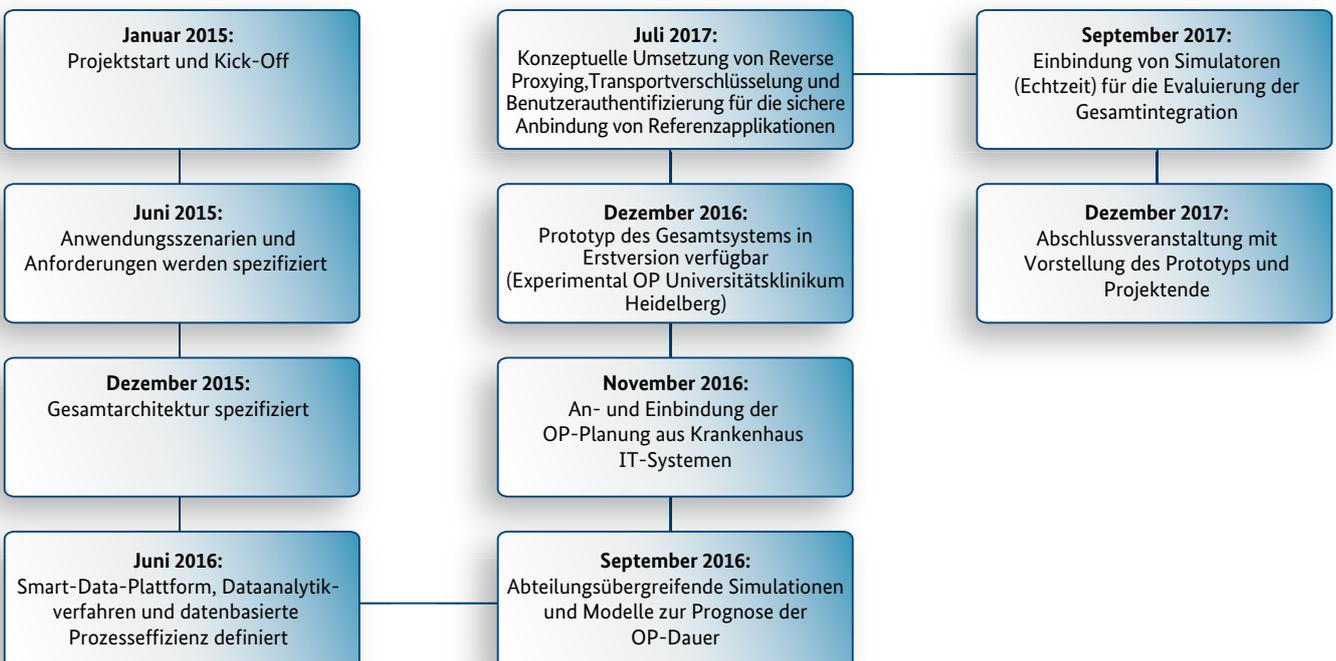
Hierfür wurde im ersten Schritt erforscht, wie Medizingeräte zu gestalten sind, damit sie zum intelligenten Datenlieferanten werden können. Zudem wurden innerhalb des Projekts

passende Methoden und Werkzeuge entworfen, die es ermöglichen, die meist heterogenen Datenmengen zu verarbeiten und wertschöpfend nutzbar zu machen.

In einem zweiten Schritt wurden diese Daten mit klinischen Prozessen innerhalb und außerhalb des Operationssaals verknüpft, um so das OP-Management, Arbeitsabläufe und Prozesse im OP-Umfeld zu verbessern. Dazu wurde eine Smart-Data-Plattform entwickelt, die bestehende Vernetzungskonzepte nutzt, um Daten von Medizingeräten und IT-Systemen zu erfassen. Auch können die auf dieser Smart-Data-Plattform aggregierten, analysierten und angeereicherten Daten – unter Berücksichtigung der Datensicherheit – Unternehmen und Dienstleistern zur Entwicklung von Applikationen und Services zur Verfügung gestellt werden.

Langfristiges Ziel ist es, der deutschen Schlüsselbranche Medizintechnik Chancen für neue, wirtschaftlich nachhaltige Dienstleistungen und Geschäftsmodelle bis hin zur Entwicklung neuer Märkte zu schaffen.

### Meilensteine im Projektverlauf



## Weg in die Praxis

Im Projekt InnOPlan durchgeführte Datenanalysen und Simulationen zeigen, dass die intelligente Einbindung von Gerätedaten die Prozesse im OP-Umfeld deutlich verbessern können, z.B. durch Schaffung von Transparenz über den OP-Fortschritt und die Bereitstellung der Informationen über OP-Grenzen hinweg. Entsprechend entwickelt das Projekt zwei Referenzapplikationen. Die Referenzapplikation I soll mithilfe der analysierten Gerätedaten die OP-Koordination im OP-Trakt verbessern. In der Referenzapplikation II werden die im Rahmen eines operativen Eingriffs erfassten Gerätedaten für die teilautomatisierte Erstellung von OP-Berichten genutzt.

Die Ergebnisse des Projekts sollen unter anderem am Universitätsklinikum Heidelberg zum Einsatz kommen, wo eine Smart-Data-Plattform zur Optimierung von Prozessabläufen eingesetzt werden soll. Zudem beabsichtigt die Firma KARL STORZ, die ebenfalls Teil des Konsortiums ist, Ergebnisse aus dem Projekt InnOPlan in neuen Projekten bzw. Produkten einzusetzen.

Über das Anwendungsgebiet der Medizin hinaus entwickelt das Projekt allgemeine Konzepte für Smart-Data-Plattformen, die es ermöglichen sollen, unterschiedliche Geräte

miteinander zu vernetzen. In diesem Zusammenhang arbeitet InnOPlan an einem wirtschaftlich und rechtlich tragfähigen Konzept, das kleinen und mittleren Unternehmen den Zugriff auf bislang nicht zugängliche Daten und darauf aufbauend die Entwicklung innovativer Dienste ermöglichen soll. Des Weiteren sollen innerhalb des Projekts erarbeitete Ergebnisse (Datensätze, Smart-Data-Plattform, etc.) als Ausgangsbasis für weitere Forschungsprojekte dienen.

Für die Weiterentwicklung der im Projekt entwickelten Ansätze ist vor allem die Verbesserung und der Ausbau einer übergreifenden Vernetzung von Geräten zentral, welche bisher durch heterogene Schnittstellen der einzelnen Hersteller erschwert wird. Zudem beschränkte sich das Projekt InnOPlan vorrangig auf nicht patientenbezogene Daten. Wenn die Ergebnisse für die Analyse personenbezogener Daten genutzt werden soll, sind – insbesondere aus rechtlichen Gründen – weitere Maßnahmen zum Schutz und der Nachverfolgbarkeit der Daten nötig.



## Die Projekt-Ergebnisse auf einen Blick

Früher	Mit InnOPlan
Informationen zum laufenden OP-Eingriff stehen nicht, nur teilweise oder eventuell fehlerhaft zur Verfügung.	Durch die Smarte OP-Geräteplattform stehen Informationen zum laufenden OP-Eingriff zur Verfügung und können für Übersichten und Notifizierungen genutzt werden.
Informationen müssen mühselig zusammengesucht werden.	Die smarte OP-Geräteplattform soll in Zukunft Geräte und Systeme miteinander vernetzen und somit einen strukturierten Zugang zu Informationen ermöglichen.
Geräte erfüllen ihre medizinische Aufgabe, stellen aber keine Informationen zu Verfügung, die ausgewertet werden können.	Die smarte OP-Geräteplattform vernetzt Geräte und Systeme und ermöglicht die automatisierte Erfassung und Bereitstellung von Informationen.
Kleine und mittlere Unternehmen haben keine Möglichkeit, auf Daten zuzugreifen und darauf aufbauend Dienste zu entwickeln.	Bereitstellung eines wirtschaftlich und rechtlich tragfähigen Konzepts, um perspektivisch zertifizierten kleinen und mittleren Unternehmen aus dem Bereich der Software-Entwicklung Möglichkeiten zu bieten, auf bislang nicht zugängliche Daten zuzugreifen und darauf aufbauend neuartige Dienste zu entwickeln.
Die Auswahl von Modellierungssprachen (künstliche Sprachen zur Erstellung von Prozessmodellen) erfolgt ohne einheitliche Kriterien.	Die Auswahl von passenden Modellierungssprachen erfolgt anhand einer Reihe von Kriterien, wie z.B. Verständlichkeit. Dies ermöglicht ein präziseres und einheitlicheres Arbeiten.
Die tatsächlichen Prozessabläufe und somit Abweichungen von Leitlinien im OP-Trakt sowie die benötigte Zeitdauer an den Schnittstellen zu anderen Abteilungen sind nur wenig bekannt.	Die Prozesse und Abläufe im OP-Trakt und die Verzahnung mit vor- und nachgelagerten Prozessen werden datenbasiert erhoben und modelliert, so dass eine bessere Grundlage für Optimierungen vorliegt.
Die Planung der OP-Belegung erfolgt auf Basis pauschaler Erfahrungswerte je OP-Typ ohne Bezug zu individuellen Faktoren.	Die smarte OP-Geräteplattform ermöglicht mit Hilfe von Machine-Learning-Verfahren eine deutlich genauere Vorhersage der OP-Dauer und bietet somit die Basis für eine bessere Planung der OPs.
Prozessänderungen rund um den OP-Trakt und deren Auswirkung auf die Optimierung sind nur ansatzweise bekannt.	Simulationen zeigen Auswirkungen von datenbasierten Optimierungen der OP-Abläufe im Zusammenspiel mit beteiligten Abteilungen auf Ressourcen (wie Personal und OP-Säle) und ermöglichen es, diese zu bewerten.

# SAHRA – Smart Analysis – Health Research Access

## Gesundheitsdaten wissenschaftlich nutzbar machen

### Das Projekt

Ob im Krankenhaus, im Rahmen der Abrechnung verschiedener Leistungen bei den Krankenkassen, im Labor oder in der Arztpraxis: Im Gesundheitswesen werden sehr viele Daten erfasst. Diese Daten werden jedoch nur selten zur Erforschung und Verbesserung der Versorgung der Patientinnen und Patienten verwendet und sind auch bei Zustimmung der Patientin oder des Patienten kaum kombinierbar, wenn sie beispielsweise in unterschiedlichen Formaten vorliegen.

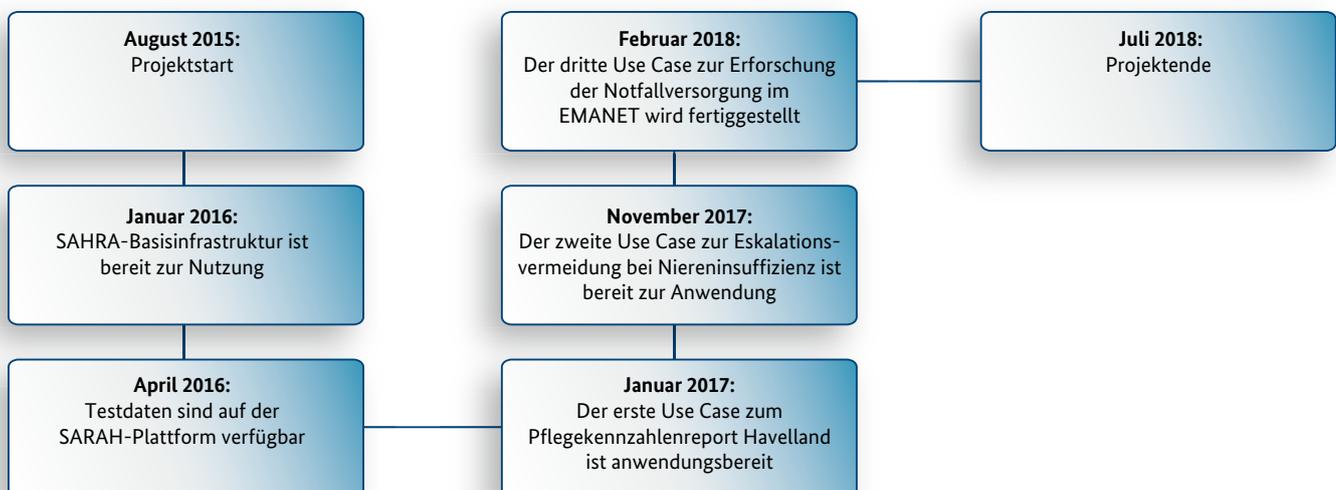
Das Projekt SAHRA hat deshalb eine Plattform entwickelt, die Abrechnungs-, Behandlungs- sowie Studien- und Referenzdaten miteinander kombinierbar und für die Versorgungsforschung zugänglich macht. Da Gesundheitsdaten hochsensibel sind, kommt dem Datenschutz im Projekt höchste Bedeutung zu. Im Projekt entstand so eine sichere, webbasierte Analyseplattform, auf der Versorgungsdaten mit weiteren Datenquellen zusammengeführt und mit wissenschaftlichen Methoden ausgewertet werden können. Die Plattform stellt für Forscherinnen und Forscher eine große Erleichterung dar, da sie durch die hohe Nutzerfreundlichkeit ihre Forschungen wesentlich unkomplizierter durchführen können. Erprobt wird SAHRA unter anderem im kommunalen Pflegebereich, zur Erforschung der Notfallversorgung und zur Eskalationsvermeidung bei Niereninsuffizienz.

### Weg in die Praxis

Das Plattform von SAHRA ermöglicht die rechtssichere Kombination von Gesundheitsdaten für die Versorgungsforschung. Sie wurde so konzipiert, dass auch medizinische Laien und Praktikerinnen und Praktiker Forschungsergebnisse und Kennzahlen nutzen können. Der kontrollierte Einblick in die regionale Versorgungslage kann beispielsweise Sozialreferentinnen und Sozialreferenten von Kommunen und Landkreisen bei ihrer Arbeit unterstützen. Zudem können Start-ups aus dem digitalen Gesundheitswesen wichtige Impulse für die Anpassung ihrer Produkte an den tatsächlichen Bedarf gewinnen.

Im ersten Anwendungsfall des Projekts konnte die Praxis-tauglichkeit bereits bewiesen werden: Auf der Basis von anonymisierten Abrechnungsdaten der AOK Nordost, die mit entsprechenden regionalen Referenzdaten kombiniert wurden, konnte ein Pflegekennzahlenreport für den Landkreis Havelland, Pilotpartner des Projekts, erstellt werden. Auf dieser Grundlage wird nun erstmals eine auf den tatsächlichen Bedarf zugeschnittene Sozialplanung, beispielsweise zur Planung der Seniorenhilfe, für die kommunale Ebene ermöglicht. Diese Kennzahlen werden noch in 2017 von fünf weiteren Landkreisen für die Pflegestrukturplanung der nächsten Jahre verwendet, stetig weiterentwickelt und auf weitere Landkreise in der Nordost-Region (bestehend aus Mecklenburg-Vorpommern, Brandenburg,

### Meilensteine im Projektverlauf



Berlin) ausgeweitet. Auch eine bundesweite Umsetzung ist denkbar.

Weitere konkrete Anwendungsfälle von SAHRA dienen der effizienten und praxiswirksamen Umsetzung öffentlich geförderter Versorgungsforschung. Aktuell werden z. B. Anwendungen mit der Charité – Universitätsmedizin Berlin, der Technischen Universität Berlin sowie mit dem Fraunhofer IZI Rostock implementiert. Bei diesen Anwendungsfällen liegt der Fokus auf der Analyse von rechtssicher verknüpften Behandlungs-, Studien- und Abrechnungsdaten zu Forschungszwecken. Das Projektteam holte dafür im August 2017 die entsprechenden Genehmigungen ein und arbeitet nun an der Zusammenführung sowie Anonymisierung der forschungsnotwendigen Informationen.

Diese Forschungsdaten werden im Anschluss auf der Plattform den jeweils berechtigten Versorgungsforscherinnen und Versorgungsforschern zur Verfügung gestellt. Die Versorgungsforscherinnen und Versorgungsforscher können dann dort die jeweils für ihre Forschungsvorhaben notwendigen Versorgungsdaten mithilfe der In-Memory-Technologie von SAHRA verwenden. Diese Technologie nutzt den Arbeitsspeicher eines Computers als Datenspeicher und hat deshalb eine hohe Zugriffsgeschwindigkeit. Durch die verfügbaren statistischen Analysewerkzeuge eignet sich die Plattform für die praxisnahe Versorgungsforschung, aber auch für den Praxistest der Forschungsergebnisse durch beteiligte Ärzte.



### Die Projekt-Ergebnisse auf einen Blick

Früher	Mit SAHRA
Die Datenübertragung für Forschungsprojekte erfolgt per Datenabzug und Datenträgeraustausch mit Löschverpflichtung. Die Forschungsergebnisse liegen nur als Publikation vor.	Der Datenzugriff auf die notwendigen Forschungsdaten erfolgt kontrolliert über die webbasierte SAHRA-Plattform, und kann gesetzeskonform entzogen und erteilt werden. Forschungsergebnisse werden somit wiederholt nutzbar.
Analysen sind nur an bestimmten Standorten durch einzelne Forscher möglich, etwa an den Computern in den jeweiligen Büros.	Versorgungsforschung wird durch einen VPN-gesicherten Webzugang auch in standortübergreifenden Forschungsteams möglich.
Ein Erkenntnisgewinn aus der Versorgungsforschung ist überwiegend Experten vorbehalten.	Auch medizinische Laien und berechtigte Praktikerinnen und Praktiker können die anonymisierten Daten auf der Plattform für genehmigte Zwecke praktisch nutzen.
Die Infrastruktur für die Datenanalyse muss mit erheblichem Finanz- und Zeitaufwand immer wieder neu errichtet werden.	Die Infrastruktur zur Datenanalyse ist durch die SAHRA-Plattform verfügbar, sodass keine bestimmte Hard- oder Software vorinstalliert werden muss.

# FASTGenomics – Fast Analysis of Single Cell Transcriptomics

## Schnellere Genomanalyse dank digitaler Technologien

### Das Projekt in Kürze

Die Genomanalyse besitzt das Potenzial, Forscherinnen und Forschern wichtige Erkenntnisse zum Beispiel über die Entstehung von Erbkrankheiten oder immunologischen Erkrankungen zu ermöglichen. Zudem bietet sie enorme Chancen für die Krebs- und Stammzellforschung. Ein relativ neues Verfahren zur Untersuchung des Genoms ist die Sequenzierung auf Einzelzell-Ebene: Während bei dem herkömmlichen Verfahren zur Genomanalyse ein Mischsignal aus dem Erbgut vieler Zellen untersucht wird, kann bei der Sequenzierung auf Einzelzell-Ebene das Erbgut von einzelnen, individuellen Zellen bestimmt werden. Dadurch entstehen einerseits hochauflösende Informationen, andererseits aber auch gewaltige Mengen an Daten, die komplex und schwierig zu analysieren sind.

Das Projekt FASTGenomics leistet einen Beitrag zur Beschleunigung und Vereinfachung dieser aktuellen Genomforschung. Dafür wird eine spezielle Plattform zur Auswertung und Analyse von Einzelzell-Daten entwickelt, um Studien im Bereich der Transkriptomik (Untersuchung von RNA, der „Abschrift“ der DNA) zu erleichtern. So können Forscherinnen und Forscher und Medizinerinnen und Mediziner Big-Data-Technologien nutzen, um eigene Studien einfacher und schneller durchzuführen. Ein Prototyp der Plattform ist bereits online und stellt Werkzeuge zur Verfügung, mit der große Datensätze einfacher analysiert

werden können. Die Plattform ist zudem ein Forum zum Austausch mit anderen Wissenschaftlerinnen und Wissenschaftlern, da Forscherinnen und Forscher eigene Daten hochladen und ihre Ergebnisse mit der wissenschaftlichen Gemeinschaft teilen können. Dabei haben Datenschutz und -sicherheit einen hohen Stellenwert.

### Weg in die Praxis

Die Online-Plattform steht Wissenschaftlerinnen und Wissenschaftlern bereits als Prototyp zur Verfügung und wird für Anwendungen genutzt. Im weiteren Projektverlauf soll die Funktionalität der Plattform durch weitere Algorithmen, schnellere Rechenzeiten für noch größere Daten, sowie mehr Möglichkeiten zum Teilen von Daten und Studien-Ergebnissen kontinuierlich ausgebaut werden.

Neben dem Einsatz als Online-Plattform soll FASTGenomics künftig auch als Werkzeug für die wissenschaftliche Forschung in Unternehmen oder für die Auswertung von Einzelzell-Transkriptom-Studien z. B. in Laboren genutzt werden. Gemeinsam mit externen Partnern wollen die Projektbeteiligten den Nutzen von FASTGenomics im prototypischen Einsatz evaluieren und stehen zu diesem Zweck mit Unternehmen aus der Pharma-Branche im Austausch.

### Meilensteine im Projektverlauf



Die Online-Plattform von FASTGenomics soll in deutsche und europäische Forschungsnetzwerke integriert werden, um durch die Analyse der dort entstehenden Daten einen Nutzen für viele Forscherinnen und Forscher zu schaffen. Die Bekanntheit der Plattform bei den potenziellen Anwenderinnen und Anwendern soll erhöht werden, indem der aktuelle Stand der Plattform in einer wissenschaftlichen Publikation veröffentlicht wird.

Aufgrund der ermittelten Anforderungen der Nutzerinnen und Nutzer wird der Prototyp der Plattform komplett in der Cloud entwickelt, wodurch die Plattform mobil und auf verschiedenen Endgeräten genutzt werden kann, sowie regelmäßige Updates sichergestellt sind. Durch eine flexible Architektur soll das System künftig aber auch als Softwarelösung, zum Beispiel in einem Unternehmen, also auch ohne Verbindung zum Internet, lauffähig sein.

### Die Projekt-Ergebnisse auf einen Blick

Früher	Mit FASTGenomics
Einzelzell-Transkriptomik-Daten sind schwer handhabbar, aufgrund der großen Datenmengen und komplexer Algorithmen.	Mit fastgenomics.org entsteht eine nutzerfreundliche Plattform mithilfe von Cloud-Computing und vorgefertigten Algorithmen.
Studien aus wissenschaftlichen Publikationen sind schwer zu reproduzieren, da Originaldaten, Algorithmen und Konfigurationen nur schwer verfügbar oder transportierbar sind.	Studien können privat entwickelt und parallel zur Publikation auf der Plattform veröffentlicht werden.
Die großen Datenmengen, die durch Forschungen auf Einzelzell-Ebene entstehen, sind auch mit biologischem Wissen nur manuell und mühsam auszuwerten.	Visualisierungen und effiziente Filter zur Auswertung von Genomdaten können auf der Plattform durch biomedizinische Expertinnen und Experten erstellt werden.
Die Archivierung sowie das Erfassen von Änderungen und unterschiedlichen Forschungsständen bei Studienabläufen ist nicht möglich – dadurch ist die Reproduzierbarkeit von wissenschaftlichen Studien eingeschränkt.	Den Nutzerinnen und Nutzern werden archivierbare Apps in einem App Store zur Verfügung gestellt. Die Apps können für Studienabläufe kombiniert werden. Dokumentierte und archivierbare Studienabläufe zur Protokollierung von Ergebnissen werden möglich.



# Fachgruppe Rechtsrahmen

Bei der Entwicklung neuer Smart-Data-Technologien spielen rechtliche Fragestellungen eine zentrale Rolle: Damit die entwickelten Smart-Data-Innovationen möglichst reibungslos unter Beachtung der rechtlichen Rahmenbedingungen in die Praxis überführt werden können, wurde in der Smart-Data-Begleitforschung die Fachgruppe „Rechtsrahmen“ eingerichtet.

Mit der Unterstützung der Fachgruppe sollten die Projekte frühzeitig die für sie relevanten rechtlichen Vorgaben erkennen und bei der Entwicklung von Smart-Data-Innovationen berücksichtigen. Außerdem beschäftigte sich die Fachgruppe damit, bestehende rechtlichen Regelungen zu hinterfragen und Handlungsempfehlungen für eine Fortentwicklung des rechtlichen Rahmens in Deutschland und Europa abzugeben. Einer der zentralen Aspekte bei Smart-Data-Anwendungen ist der Schutz von personenbezogenen Daten. Darüber hinaus wurden viele Fragen aus dem Urheber- und Wettbewerbsrecht unter dem Gesichtspunkt des Spannungsfelds zwischen freiem Zugang zu Informationen und Schutzrechten in Bezug auf Daten diskutiert.

Die Fachgruppe hat mit zahlreichen Veröffentlichungen, Vorträgen und Workshops dazu beigetragen, wichtige rechtliche Themenkomplexe zum Datenschutz-, Wettbewerbs-, Urheber- und Leistungsschutzrecht im Kontext von Smart Data aufzuzeigen und die damit verbundenen Herausforderungen für die Entwicklung und Umsetzung der Projekte darzulegen.

## Datenschutz

Mit der EU-Datenschutz-Grundverordnung werden neue Regeln für die Verarbeitung von personenbezogenen Daten und damit entscheidende Vorgaben für die Smart-Data-Projekte anzuwenden sein. In Workshops wurden verschiedenste Aspekte, wie der Konflikt zwischen Datenminimierung versus Datenreichtum, die Reichweite des Personenbezugs sowie das Prinzip der Zweckbindung der Datenverarbeitung diskutiert. In Bezug auf viele Smart-Data-Anwendungsfälle stellte sich vor allem die Frage, wann Daten personenbezogen sind und damit das Datenschutzrecht angewendet werden muss, als eine der zentralsten Herausforderungen dar. Die Bewertung von Identifizierungsrisiken, also mit welcher Wahrscheinlichkeit ein Personenbezug (wieder)hergestellt werden kann, wurde als eine erste entscheidende Weichenstellung bei der Entwicklung von Smart-Data-Technologien identifiziert.

In der Arbeitsgruppe wurden Methoden und Verfahren zur Anonymisierung vorgestellt. Selbst wenn eine Re-Identifizierungswahrscheinlichkeit bestehen bleibt, kann der Einsatz die Datenminimierung ermöglichen, damit durch die Datenverarbeitung nur so viel Wissen über Betroffene generiert wird, wie für die jeweilige Anwendung zwingend erforderlich ist. Damit wurde gezeigt, dass rechtlichen Herausforderungen von Smart Data durch technische Lösungen begegnet werden kann. Hier bewies sich die interdisziplinäre Herangehensweise der Fachgruppe in Zusammenarbeit mit der Fachgruppe Sicherheit als sehr fruchtbar. Weitere technische Ansätze, wie die Anwendungen zur Datennutzungskontrolle, wurden präsentiert, um dem Zweckbindungsprinzip gerecht zu werden.

## Zugang zu Daten

Die Fachgruppe befasste sich auch mit der Rolle von Daten als Wirtschaftsgut. In Workshops und durch Expertinnen- und Expertenvorträge und Diskussionsrunden wurde die aktuelle Diskussion zum Aufbau einer europäischen Datenökonomie abgebildet. Die Rolle von Daten als essenzieller Bestandteil von Produkten oder Dienstleistungen ist eine rechtsgebietsübergreifende Herausforderung, bei der sich zwei Paradigmen gegenüberstehen: Einerseits die grundsätzlich freie Nutzbarkeit von Informationen als Grundlage neuer Wissensgenerierung, sowie andererseits die Schaffung von Eigentums- bzw. Schutzrechten, die dann die Einholung von Lizenzen bis hin zur Etablierung von Lizenzketten erfordern würden. Mit wissenschaftlichen Beiträgen näherte sich die Fachgruppe der Frage an, wie ein ausgeglichenes Verhältnis zwischen Schutz- und Zugangsrechten erreicht werden kann.

Der diskriminierungsfreie Zugang zu Daten sowie eine problemlose Zusammenarbeit über Systemgrenzen hinweg (Interoperabilität), etwa zwischen verschiedenen Unternehmen aber auch Ländern, sind dabei als wichtige Grundlage für die Entwicklung und Umsetzung weiterer Smart-Data-Technologien festgehalten worden. Die Fachgruppe plädiert daher für eine europäische Wettbewerbsregelung zum Zugang zu Daten, die gleichzeitig auch datenschutzrechtliche Aspekte mitberücksichtigt und für beides verbindliche Vorgaben bietet. Dafür wurden Handlungsempfehlungen in Form von Stellungnahmen zum EU-Konsultationsprozess „Building a European Data Economy“ formuliert sowie gegenüber dem BMWi in den politischen Diskurs eingebracht.

## Urheber- und Leistungsschutzrechte

Eine sich abzeichnende Problematik in Bezug auf Zugang und Nutzbarkeit von Daten, die Urheber-, Datenbankschutz- oder Leistungsschutzrechten unterliegen, ist die unterschiedlich praktizierte Gewährung von Nutzungsrechten – etwa bei wissenschaftlicher oder kommerzieller Nutzung. Eine zentrale Schlüsselfrage ist daher, wie durch rechtliche Vorgaben effektiv Diskriminierungsfreiheit hergestellt werden kann.

Die geplante EU-Urheberrechtsreform droht insoweit keinen positiven Einfluss auf die auch im Rahmen von Smart Data erfolgende kommerzielle Nutzung von Data-Mining-Technologien auszuüben, also die Analyse von vorhandenen Datenbeständen, um Muster, Trends oder Zusammenhänge zu erkennen. Eine Data-Mining-Erlaubnis, die auf wissenschaftliche Forschung beschränkt wäre, könnte im Umkehrschluss große Hürden für kommerzielle Smart-Data-Anwendungen bedeuten, wenn Rechte für eine Vielzahl von Texten und Textschnipseln von den jeweiligen Urheberinnen und Urhebern eingeholt werden müssen.

## Fazit

Durch die Arbeit der Fachgruppe sind die Projekte dafür sensibilisiert worden, die rechtlichen Anforderungen von Beginn an zu beachten und bei allen Entwicklungsschritten zu prüfen. Nur so wird sichergestellt, dass die Projektergebnisse später auch tatsächlich zum Einsatz kommen können. Beim Datenschutz zeigt sich, dass durch den Einsatz von technischen Schutzmechanismen rechtssichere Lösungen geschaffen werden können. Allerdings besteht hier noch viel Raum für die Entwicklung und den Einsatz weiterer technischer Innovationen, die gleichzeitig die rechtlichen Anforderungen berücksichtigen. Zudem bedarf es eines einheitlichen Gesamtkonzepts für die wettbewerbsrechtliche Regulierung zum Zugang und der wirtschaftlichen Verwendung von Daten, mit der Diskriminierungsfreiheit hergestellt sowie die Entwicklung von Standards angeregt wird, um langfristig eine Ende-zu-Ende-Interoperabilität zu erreichen.

# iTesa - Intelligent Traveller Early Situation Awareness

## Reisewarnungen in Echtzeit

### Das Projekt

In einer globalisierten Welt reisen nicht nur Touristinnen und Touristen durch ferne Länder. Auch Geschäftsleute fliegen zu auswärtigen Handelspartnerinnen und Handelspartnern oder fahren durch boomende Wirtschaftsregionen. Doch das Reisen birgt immer auch Risiken: Mit Naturkatastrophen, Epidemien oder Terroranschlägen muss in manchen Regionen gerechnet werden. Beim Ausbruch einer akuten Krise kann es lange dauern, bis es einen vollständigen Überblick über die genaue Lage vor Ort gibt. Dadurch verlieren Unternehmen wertvolle Zeit, ihre Mitarbeiterinnen und Mitarbeiter zu schützen und in Sicherheit zu bringen. Wie können also Unternehmen ihrer gesetzlichen Informations- und Fürsorgepflicht gegenüber ihren Mitarbeiterinnen und Mitarbeitern besser nachkommen und sie vor solchen Vorfällen schützen?

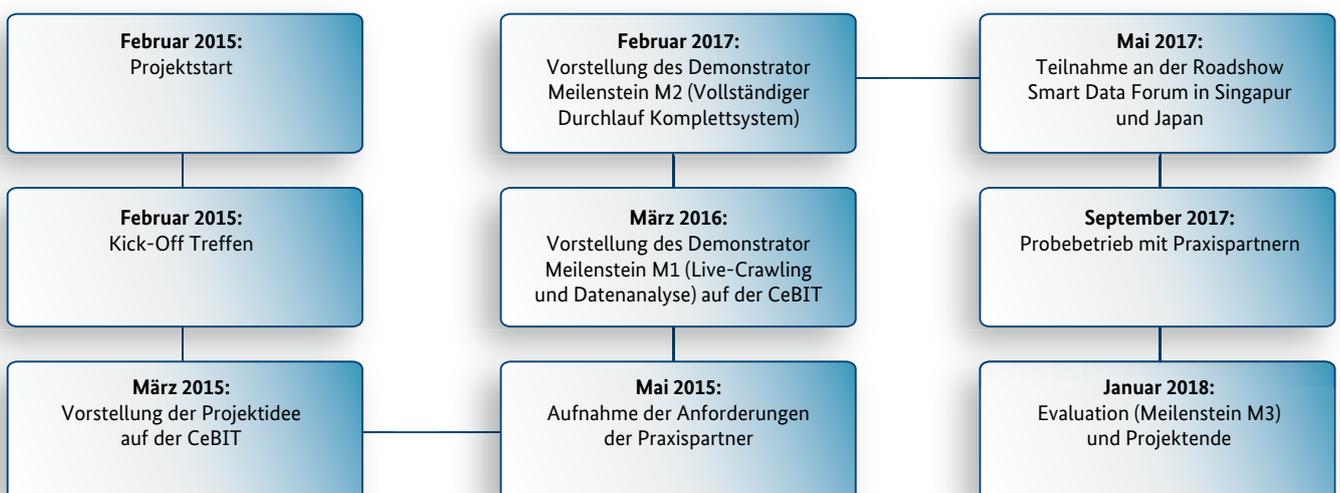
Das Smart-Data-Projekt „iTESA – Intelligent Traveller Early Situation Awareness“ möchte dieses Defizit beheben und Unternehmen und Reisenden ermöglichen, im Ernstfall schneller zu reagieren. Hierfür durchforstet iTESA mithilfe spezieller Datenanalysen und selbstlernender Algorithmen

weltweit zahlreiche Quellen wie Internetseiten, soziale Netzwerke, Agentur- und Pressemeldungen sowie Nachrichten und Informationen von Behörden nach möglichen Reiserisiken. Dies geschieht stets unter Berücksichtigung des deutschen Datenschutzes.

In der nächsten Stufe prüft iTESA die Meldungen auf ihre Glaubwürdigkeit sowie das jeweilige Risiko und ordnet einem Vorfall einen bestimmten Bedrohlichkeitsgrad zu. Auf diese Weise kann der nötige Handlungsbedarf realistisch eingeschätzt werden.

Neben der Analyse von Ereignissen stellt iTESA einen möglichst genauen geographischen Bezug her – bis hin zum betroffenen Straßenzug. Durch einen Abgleich mit Reisedaten aus Traveller-Systemen (Buchungssysteme der Global Distribution Systems oder Leistungsträger) ist iTESA in der Lage, zu ermitteln, welche Mitarbeiterinnen und Mitarbeiter oder Kunden sich in der Gefahrenzone befinden oder planen, dorthin zu reisen. Per App können diese anschließend gewarnt und mit passenden Alternativrouten versorgt werden.

### Meilensteine im Projektverlauf



## Weg in die Praxis

Um die Anbindung an die Praxis sicherzustellen, hat das Projekt iTESA von Beginn an mit Partnern zusammengearbeitet, die einen Querschnitt des zukünftigen Marktes für iTESA repräsentieren. So konnte beispielsweise im laufenden Projekt der deutsche Marktführer für die manuelle Reiserisikoermittlung, A3M GmbH, als Praxispartner hinzugewonnen werden.

Derzeit arbeitet das Projekt daran, den entwickelten Demonstrator qualitativ und quantitativ im Hinblick auf die Markterschließung zu bewerten, wobei sowohl eine inhaltliche als auch eine internationale Ausweitung diskutiert werden. Hierbei steht insbesondere die Ausweitung auf das Thema Sportevent im Fokus: So ist das Konsortium aktuell mit japanischen Partnern hinsichtlich einer möglichen Verwertbarkeit für die olympischen Sommerspiele 2020 in Tokio im Gespräch.

Weiterhin ist auch eine Verwendung von Teilergebnissen des Projekts über die iTESA-Applikation und Reisebranche hinaus angedacht. Dies betrifft beispielsweise die gesammelten Erfahrungen im Bereich dynamische Webrecherche

oder die datenschutz- und urheberrechtliche Bewertung der Verwendung öffentlich zugänglicher Internetdaten. Auf diese Weise ist es möglich, eine weitere Option in der Nachverwertung der Forschungsergebnisse zu erschließen.

Weiteren Forschungsbedarf stellte das Projektkonsortium unter anderem in der weltweiten vollautomatisierten semantischen Analyse von Sprachen fest, zum Beispiel für die in Asien verbreitete, auf Kanji-Schrift basierte Schriftform. Zudem gestaltete sich die Sammlung von Primärdaten sehr komplex, so dass die Mechanismen weiter verfeinert und regelmäßig an das Vorgehen der Webseitenbetreiber angepasst werden müssen. Damit das volle Potenzial von iTESA für die Reisebranche erschlossen werden kann, ist jedoch ein grundlegender Wandel bei der Generierung und Weitergabe digitaler Echtzeitreisedaten nötig. Hier ist sowohl die Verbindung von bestehenden IT-Systemen mit den innovativen Smart-Data-Techniken und -Verfahren auszubauen als auch die kontinuierliche Verarbeitung in Richtung der Analyse von Datenströmen (Streaming Analytics) weiter voranzutreiben.



## Die Projekt-Ergebnisse auf einen Blick

Früher	Mit iTesa
Reisewarnungen werden zeitverzögert ausgegeben, ohne dass Reisende oder Arbeitgeberinnen und Arbeitgeber automatisch darüber benachrichtigt werden.	Automatisches Alarmsystem kann Reisende und Arbeitgeberinnen und Arbeitgeber in Echtzeit vor Reiserisiken warnen.
Warnungen zu einzelnen Regionen oder Ländern müssen aus unterschiedlichsten Quellen manuell oder redaktionell zusammengesucht werden.	Öffentliche Datenquellen können automatisch miteinander verknüpft und für eine Analyse und mögliche Warnmeldung genutzt werden.
Die Reisebranche liefert Daten zu Reisebuchungen einmal täglich über Nacht.	Die Reisebranche kann Daten zu Reisebuchungen direkt nach Generierung liefern, wodurch diese aktueller zur Verfügung stehen.
Die Textanalyse erfolgt wesentlich über Schlüsselwörter.	Eine vollständige semantische Textanalyse ermöglicht eine bessere Bewertung der Relevanz.
Ein Redaktionsteam durchsucht Nachrichtenquellen wie Zeitungen manuell nach Hinweisen auf potenzielle Gefahren.	Redaktionsteams werden automatisiert auf bereits identifizierte Anomalien und Ereignisse hingewiesen.
Die Verwertung von öffentlich zugänglichen Daten aus dem Internet entspricht den Vorgaben des Bundesdatenschutzgesetzes.	Die Verwertung öffentlich zugänglicher Daten aus dem Internet entspricht den Vorgaben des neuen Bundesdatenschutzgesetzes, der Datenschutz-Grundverordnung und des Urheberrechtsgesetzes.

# SD4M – Smart Data for Mobility

## Serviceplattform für multimodale Smart-Mobility-Services

### Das Projekt

Mit dem Bus zum Bahnhof und mit dem Zug in die nächste Stadt, dann mit dem Mietwagen zum Reiseziel, anschließend mit dem Flugzeug und dem Taxi wieder zurück: Für eine Reise von der Haustür zum endgültigen Ziel sind meist mehrere Transportmittel nötig. Der Anspruch, eine kontinuierliche Reisekette zu gewährleisten, stellt den Mobilitätssektor jedoch vor diverse Herausforderungen. Unter anderem, weil die Nutzung von Mobilitätsangeboten oft stark schwankt und die verschiedenen Anbieter im Mobilitätsbereich nur selten aufeinander abgestimmt sind.

Das Projekt SD4M entwickelt deshalb eine branchenübergreifende Softwareplattform, die Daten von Mobilitätsanbietern mit offenen Datenquellen kombiniert. Zu diesen Datenquellen zählen unter anderem Newsseiten, Staumeldungen und Baustelleninformationen, offizielle Fahrplan- und Nahverkehrsdaten, sowie Social-Media-Kanäle wie Twitter. All diese Daten werden in Echtzeit extrahiert, aggregiert und miteinander in Zusammenhang gebracht.

Eine große Herausforderung ist neben der schieren Masse der Daten die Verlinkung zwischen strukturierten und unstrukturierten Daten, also das Herstellen einer semantischen Interoperabilität. So gilt es, etwa textliche Ausdrücke in valide Daten zu übersetzen. Was bedeuten zeitliche Formulierungen wie „eben“ oder „gerade“, oder räumliche Angaben wie „kurz vor Tegel“?

Ziel des Projekts ist es, ein Ökosystem für die Nutzung unterschiedlichster Daten zu schaffen, das umfangreiche

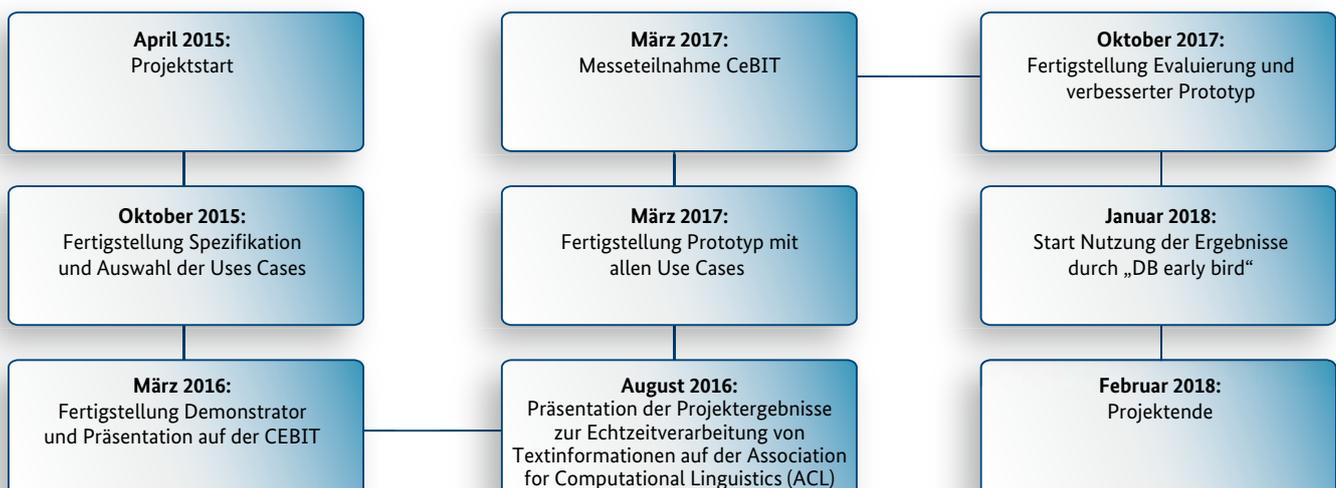
Prognosen erstellt und neue Optimierungspotenziale erschließt. Mobilitätsanbieter profitieren davon, da die intelligenten Daten helfen, Routen optimal auszulasten und Angebote besser anzupassen. Für Privatpersonen wird hingegen das Reisen schneller, preiswerter und bequemer. Darüber hinaus ist die SD4M-Plattform über standardisierte Schnittstellen für unabhängige Analyse-Spezialisten und Datenanalytik-Dienstleister zugänglich.

### Weg in die Praxis

Das Projekt SD4M legte von Beginn an einen starken Fokus auf die Praxisanwendung der Projektergebnisse. So nutzt der Konsortialpartner DB Systel GmbH die Referenzarchitektur der Softwareplattform als Vorlage für eine Anwendung zur Echtzeitanalyse von Big Data. Zudem sollen einzelne Komponenten der Software von DB Systel und DB Dialog für das Frühwarnsystem DB Early Bird genutzt werden, das relevante Informationen aus Social-Media-Kanälen automatisch herausfiltert. Die Erkenntnisse aus dem Projekt zur automatisierten Informationsgewinnung aus Freitexten kommen wiederum bei der DB Fernverkehr zum Einsatz, welche damit im Bereich des Kundenservices Prozesse optimiert.

Die entwickelte Plattform steht darüber hinaus auch Drittanbietern über standardisierte Schnittstellen offen. Dadurch können beispielsweise unabhängige Analyse-Spezialistinnen und -Spezialisten ohne eigene historische Datenbestände die Plattform als Integrations- und Datenzuliefersystem für ihre Dienstleistungen nutzen. Für Datenanalyse-Dienstleister könnten gesonderte Nutzungsvereinbarungen gelten (teil-

### Meilensteine im Projektverlauf

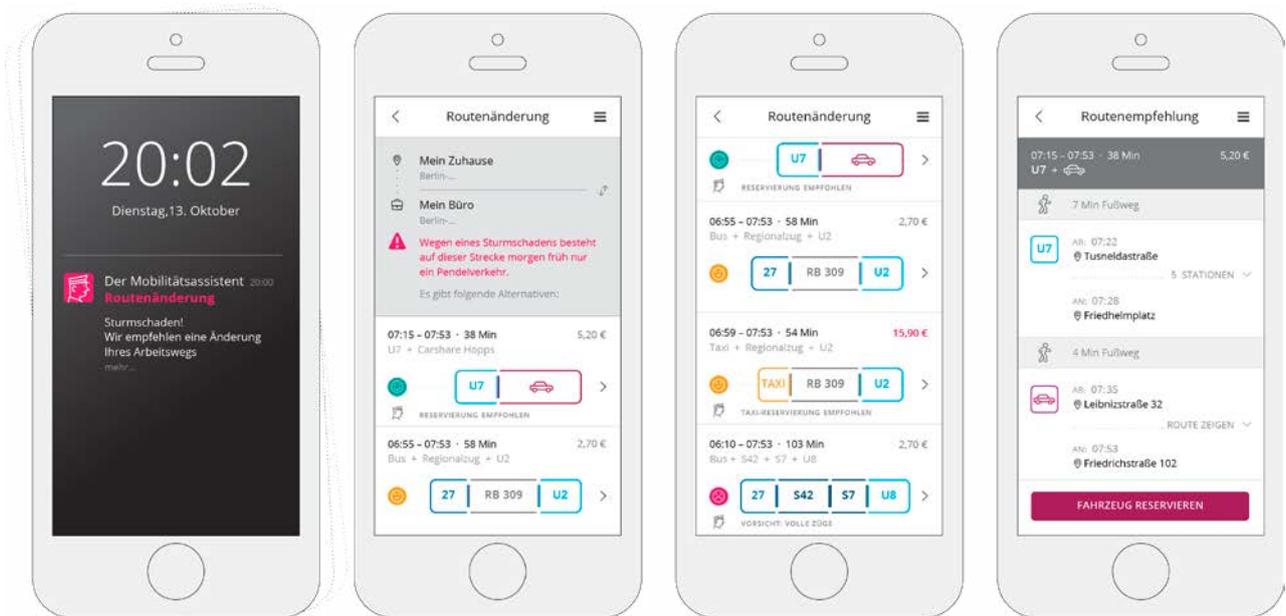


weise kostenpflichtig), die tiefere Eingriffe in das System zulassen.

Zudem hat das Projekt auch mehrere Ergebnisse öffentlich verfügbar gemacht: Es wurde beispielsweise ein Benchmarkdatensatz, also eine Art Datensatzvorlage, veröffentlicht, welcher die im Projekt entwickelten Mechanismen zur Ermittlung von Mobilitätsdaten aus Social-Media-Texten veranschaulicht. Ebenso veröffentlicht wurde eine Softwarekomponente, die es ermöglicht, Mobilitätsnachrichten aus

Social-Media-Quellen automatisch geographisch zu verorten und damit deren Zuordnung erleichtert.

Um das Potenzial der im Projekt entwickelten Ansätze voll auszuschöpfen, sind weitere Forschungen empfehlenswert. So könnte die Analyse der unterschiedlichen Textquellen noch auf andere Sprachen neben dem Deutschen ausgeweitet werden, um die Nutzung der Plattform über den deutschen Sprachraum hinaus zu ermöglichen.



**Die Projekt-Ergebnisse auf einen Blick**

Früher	Mit SD4M
<p>Mobilitätsrelevante Informationen sind nur für einzelne Verkehrsmittel/-regionen/-unternehmen verfügbar.</p>	<p>Mobilitätsrelevante Informationen stehen flächendeckend für alle Verkehrsmittel zur Verfügung. Datenwertschöpfungsketten enden nicht an Unternehmensgrenzen.</p>
<p>Informationen aus benutzergenerierten Quellen (z. B. Social Media) können nicht automatisiert verarbeitet und berücksichtigt werden.</p>	<p>Wissen aus allen relevanten Quellen kann für smarte Mobilitätsangebote verwendet werden.</p>
<p>Informationen zu mobilitätsrelevanten Ereignissen aus textbasierten Quellen können nur durch aufwändige Recherchetätigkeiten gefunden werden.</p>	<p>Textbasierte Informationen (z. B. aus Social Media) werden automatisiert verarbeitet und erweitern bestehendes strukturiertes Wissen zu Mobilität.</p>
<p>Für Big-Data-Analytikplattformen existieren ausschließlich herstellereigene Lösungen mit großen Barrieren für den Anbieterwechsel.</p>	<p>SD4M bietet eine herstellerunabhängige Referenzarchitektur und Umsetzung mittels Open-Source-Softwarekomponenten.</p>
<p>Methoden der Textanalytics und verschiedener Machine-Learning-Verfahren sind nicht oder in geringem Maß an die Mobilitätsdomäne angepasst.</p>	<p>Für die Anwendung von Textanalytics und Machine-Learning-Verfahren existieren mobilitätsspezifische Lösungen.</p>

# sd-kama – Smart-Data-Katastrophenmanagement

## Katastrophen effektiver managen

### Das Projekt

Bei Naturkatastrophen wie Überschwemmungen oder Hochwasser gibt es zumeist nur wenige oder unsichere Informationen über die aktuelle Lage vor Ort. Dabei sind verlässliche Daten zum Ausmaß und zur Intensität der Katastrophe, etwa zur Anzahl der bedrohten Personen, zu dem Zustand von Gebäuden oder Infrastrukturen essenziell, damit Krisenmanagerinnen und Krisenmanager und Rettungskräfte schnell und gezielt reagieren können.

Das Projekt sd-kama entwickelt daher ein echtzeitfähiges System, das die relevanten Informationen aus unterschiedlichen Datenströmen sammelt und analysiert. Dafür werden bereits vorliegende Daten zusammengeführt und durch Echtzeitinformatoren ergänzt, wie beispielsweise Informationen aus Satellitenbildern oder Bild- und Videoaufnahmen, die von freiwilligen Helferinnen und Helfern oder Einsatzkräften mittels einer im Rahmen des Projektes entwickelten App übermittelt werden.

Zudem werden mithilfe von Wearables und der App körperliche Messwerte von Einsatzkräften und Helferinnen und Helfern übermittelt, so dass Überlastungen rechtzeitig erkannt und Unterstützung bzw. eine Ablösung schneller organisiert werden können. Hinzu kommen Pegelstände und Wetterinformationen aus Diensten von Behörden. Moderne Plattformen für Echtzeitverkehrsdaten, zum Beispiel Navigationsprogramme wie HERE Maps, liefern darüber hinaus Informationen und Dienste für die Erreich-

barkeit von Einsatzorten. sd-kama trägt somit dazu bei, digitale Technologien, wie Wearables oder Bild- und Videoanalysen, wirkungsvoll in das Katastrophenmanagement zu integrieren.

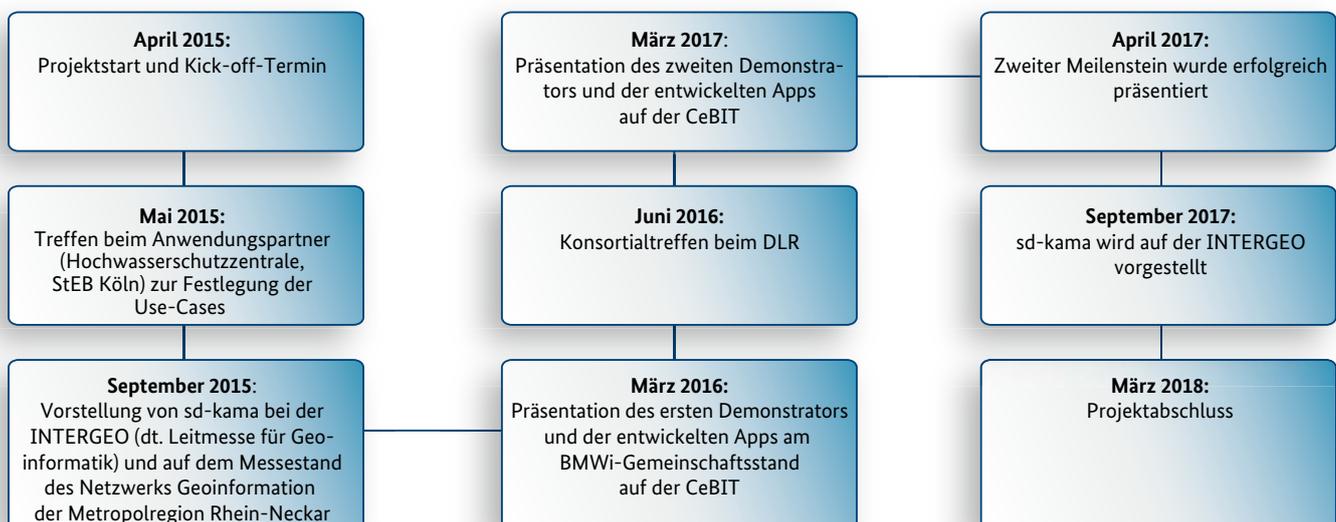
Durch die Verfügbarkeit dieser Fülle an Informationen entsteht ein umfassendes Bild von der Lage, wodurch die Einsatzzentren Ressourcen und Einsatzkräfte effizienter koordinieren, Gefahren besser einschätzen und effektive Entscheidungen treffen können. sd-kama ermöglicht damit ein zielgerichtetes, effektives Katastrophenmanagement mit bedarfsgerechtem Ressourceneinsatz – erprobt am Beispiel der Naturkatastrophe Hochwasser in Deutschland.

### Weg in die Praxis

Mit der Informationsplattform von sd-kama, einem Entscheidungsunterstützungssystem, ist ein Werkzeug entstanden, das auch nach der Laufzeit des Projektes von dem Anwendungspartner und den Projektpartnern genutzt werden kann. Darüber hinaus wird es einen Zugang für die interessierte Fachöffentlichkeit geben. Das modular und servicebasiert entwickelte System bietet nicht nur für den Hochwasserschutz wertvolle Möglichkeiten, um ein Lagebild in Echtzeit zu gewinnen.

Die ortsbezogenen Visualisierungen für die zahlreichen Echtzeit-Datenquellen eröffnen Optionen in unterschiedlichen Bereichen – so ist der Einsatz auch bei anderen Natur-

### Meilensteine im Projektverlauf



katastrophen denkbar. Da die Plattform Informationen aus unterschiedlichen Datenquellen sammelt und analysiert, kann auch der Ablauf von Großveranstaltungen beobachtet werden.

Das Ziel von sd-kama, in Echtzeit Informationen aus unterschiedlichen Datenströmen zu sammeln, zu analysieren

und diese Daten für das Katastrophenmanagement nutzbar zu machen, hat das Projekt erreicht. Eine große Herausforderung bestand darin, die Dienste durch eine rechtssichere Gestaltung attraktiv für Nutzerinnen und Nutzer und Institutionen zu machen, die großen Wert auf Datenschutz legen. Die Komponente der juristischen Begleitung hat deshalb wesentlich zum Erfolg des Projektes beigetragen.

### Die Projekt-Ergebnisse auf einen Blick

Früher	Mit sd-kama
Informationen von der Vor-Ort-Situation während eines Hochwassers sind nur per Telefon verfügbar.	Das Informationssystem von sd-kama liefert mithilfe von Echtzeitdaten Bilder und Videos direkt im System.
Überlastung von Einsatzkräften kann schwer festgestellt werden.	Überwachung zum Beispiel von Einsatzkräften in kritischen Situationen ist durch Wearables möglich.
Konsequenzen eines drohenden Deichbruchs sind nur sehr grob abzuschätzen.	Anwenderinnen und Anwender können die Konsequenzen von Ereignissen, wie einem Deichbruch, besser einschätzen. Durch Daten wie Bilder, Videos oder Informationen von Einsatzkräften haben sie ein umfassenderes Bild der Lage.
Informationen aus dem Einsatzgebiet kommen verzögert und auf diversen Kommunikationskanälen zu den Entscheidern.	Ein Eventsystem informiert in Echtzeit und gebündelt über relevante Ereignisse.
Aus Satellitenbildern abgeleitete Informationen müssen kompliziert über Datenimporte eingespielt werden.	Satellitenservices sind direkt im System sichtbar.



# ExCELL – Echtzeitanalyse und Crowdsourcing für eine selbstorganisierte City-Logistik

## Das Projekt in Kürze

Der zunehmende Verkehr in deutschen Städten sorgt nicht nur für lange Staus und eine hohe Feinstaubbelastung, sondern kostet Unternehmen auch viel Geld. Laut des britischen Forschungsinstituts Centre for Economics and Business Research (CEBR) belaufen sich die Kosten von Staus in Deutschland auf acht Milliarden Euro jährlich, 2030 soll dieser Wert gar auf 33 Milliarden Euro steigen. Um die Wettbewerbsfähigkeit zu erhalten, ist deshalb eine neue Mobilität gefragt.

Zur Steigerung der Effizienz im Straßenverkehr entwickelt das Projekt ExCELL eine Service- und Datenplattform, die verschiedene Mobilitätsdienstleistungen verknüpft und zur Verfügung stellt. Hierfür hat das Projekt Services für eine automatische Sammlung, Verarbeitung und Verknüpfung von geographischen und verkehrstechnischen Daten entwickelt. Dazu zählen ein Tracking Service, der den aktuellen Standort eines Mobilgerätes oder Fahrzeugs speichert, sowie ein Traffic Event Service, der eine Liste mit aktuellen

Verkehrereignissen bereitstellt. Da die Plattform modular aufgebaut ist, sind alle Elemente beliebig kombinierbar.

Ergänzt werden diese Daten anschließend durch per Crowdfunding gesammelte Daten, wie den Bewegungs- und Positionsdaten eines Nutzers oder aktiv von Nutzern übermittelte Daten wie Verkehrsstörungen oder Unfälle. Diese zusätzlichen Kontextdaten ermöglichen es, Analysen und Vorhersagen zu erweitern und zu verbessern. Durch die genauere Entscheidungsgrundlage kann die Plattform wiederum besser Verzögerungen berechnen, vor Staus warnen und alternative Wege empfehlen.

Neben Unternehmen profitieren alle Beteiligten von effizienteren Wegketten und Routen, die durch ExCELL ermöglicht werden: So können Mitarbeiterinnen und Mitarbeiter Stauzeiten vermeiden und Kunden besser und frühzeitig über Verspätungen informiert werden, beispielsweise per Nachricht auf ihr Smartphone.

## Meilensteine im Projektverlauf



## Weg in die Praxis

Um die Praxistauglichkeit sicherzustellen, hat sich das Projekt ExCELL von Beginn an auf konkrete Beispiele fokussiert. Der initiale Anwendungsfall sind dabei Handwerksbetriebe. Die Entscheidung für diese Branche fiel dabei unter anderem aufgrund des hohen Bedarfs nach Mobilität im Berufsalltag. Ausgewählte Handwerksbetriebe erprobten die Anwendung, um ihre Kundentermine und die damit verbundenen Fahrstrecken auf Basis der aktuellen Verkehrslage zu optimieren.

Ergebnis dieses Feldversuchs ist der Prototyp einer Lösung für kleine und mittelständische Handwerksbetriebe, die aus einer mobilen, webbasierten Komponente besteht. Diese Applikation optimiert die Termin- und Routenplanung beispielsweise unter Berücksichtigung der aktuellen Verkehrssituation sowie der Standorte der Mitarbeiter. Zudem werden dabei, wie bei allen im Projekt entwickelten Lösungen, die Möglichkeiten einer Ausgründung kontinuierlich geprüft und systematisch erprobt.

Als zweiten Anwendungsfall wird eine Lösung für Pflegedienstleister entwickelt. Derzeit werden systematisch Anforderungen erhoben, so dass 2018 erste Feldtests einer nutzerzentrierten Lösung erfolgen können.

Über die Anwendungsfälle hinaus wird die ExCELL-Service- und Daten-Plattform schrittweise für Entwickler von innovativen Anwendungen und Lösungen und zum anderen für Open-Source-Entwickler geöffnet. In diesem Zusammenhang soll das zunächst nur für Dresden entwickelte Modell auf weitere Städte und Kommunen übertragen werden. Eine zentrale Herausforderung ist es dabei, die entwickelte offene Plattform so zu stabilisieren, dass sie auch noch über das Ende der Projektlaufzeit fortbesteht.



### Die Projekt-Ergebnisse auf einen Blick

Früher	Mit ExCELL
Hohe Intransparenz, welche Daten die Algorithmen von kommerziellen Plattformen/Anbietern verwenden und wie diese verarbeitet werden.	Schaffung von Transparenz durch Offenlegung sämtlicher Algorithmen und Daten (Open Source und Open Data) innerhalb einer offenen Entwicklerplattform.
Hohe Intransparenz, wo Kunden- und Mitarbeiter- sowie Routen- und Verkehrsdaten gespeichert werden.	Alle erhobenen und verarbeiteten Daten werden anonymisiert und in Deutschland gespeichert/ gehostet.
Ausschließlich sensorische Geo- und Verkehrslagedaten werden zur Routen- und Tourenplanung sowie Verkehrsprognose genutzt.	Durch die zusätzliche Integration von durch Crowdsourcing erhobene Daten innerhalb der Plattform werden Echtzeitanalysen zur Routen- und Tourenplanung und Verkehrsprognose optimiert.
Dienstleister wie Handwerksbetriebe oder Pflegedienste planen ihre Kunden- und Patiententermine fallweise oder kurzfristig im Vorfeld.	Effizientere Planung durch Routen- und Tourenplanung unter Berücksichtigung der aktuellen und/oder zukünftigen Verkehrssituation.
Oftmals verspätete (und damit gestresste) Ankunft von Mitarbeiterinnen und Mitarbeiter bei Kunden, ohne dass Kunden/ Patientinnen und Patienten im Vorfeld darauf hingewiesen werden.	Kunden werden rechtzeitig auf mögliche Verspätungen von Dienstleistern hingewiesen. Damit verbessert sich die Kundenerfahrung (Customer Experience).

# VIRTUOSE-DE – Service-Plattform für echtzeitfähige Big-Data-Videoanalyse und -verarbeitung in der Cloud

## Durch intelligente Videoanalysen den Verkehr optimieren

### Das Projekt

Durch den zunehmenden Verkehr kommt es in zahlreichen Städten zu einem Mangel an Parkplätzen. Dabei belastet die Parkplatzzuche nicht nur die Fahrer, sondern auch andere Verkehrsteilnehmer, Anwohner und die Umwelt. Für eine schnelle und unkomplizierte Parkplatzzuche entwickelt das Projekt VIRTUOSE-DE eine cloudbasierte Service-Plattform für Big-Data-Videoanalysen und Datenverarbeitung in Echtzeit. Das Projekt beschäftigt sich unter anderem mit Herausforderungen bei der Analyse von komprimierten Datenströmen, um beispielsweise freie Parkplätze zu ermitteln.

Zwei Anwendungsszenarien stehen im Fokus des Projekts: Eines ist das intelligente Management von Parkflächen in Parkhäusern, wofür Systeme entwickelt werden, die dem Fahrer den Weg von der Parkhauseinfahrt bis zum nächstgelegenen Parkplatz zeigen. Zudem soll die Plattform in ähnlicher Weise auch bei der Parkplatzzuche auf der Straße eingesetzt werden. Ein weiteres Szenario betrifft die Sicherheit von Personen im öffentlichen Personennahverkehr (ÖPNV). VIRTUOSE-DE entwickelt eine Software, die mithilfe von cloudbasierten Echtzeit-Videoverarbeitung Situationen automatisch erkennen soll, die potenziell gefährlich sind.

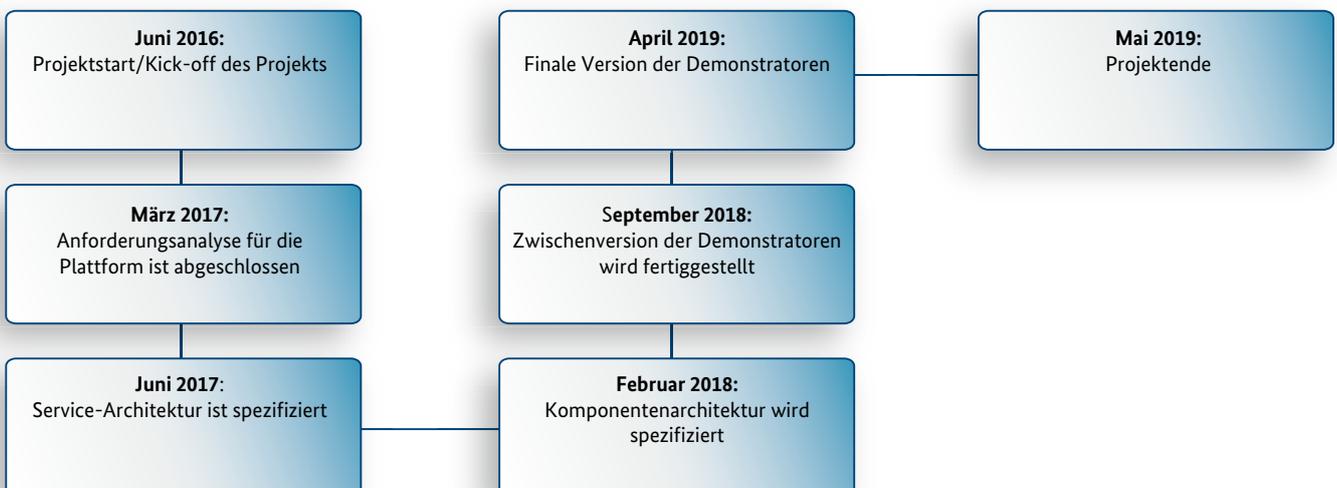
Für die Videoanalyse im Parkhaus, auf der Straße und im ÖPNV wird eine Service-Plattform entwickelt, die eine skalierbare Analyse, Verarbeitung und Speicherung großer Mengen von Bild- und Videodaten ermöglicht. Die Service-Plattform bietet darüber hinaus vielfältige Anwendungspotenziale, wie die videoüberwachte Sicherung von Großveranstaltungen, Gebäuden und öffentlichen Plätzen.

### Weg in die Praxis

Das Projekt VIRTUOSE-DE startet im Juni 2016. Um die Übertragbarkeit in die Praxis sicherzustellen, wurden bereits zwei Anwendungsszenarien definiert:

Das erste Szenario betrifft das Parkraummanagement, das zunächst in Parkhäusern und anschließend im öffentlichen Straßenverkehr umgesetzt werden soll. Dazu werden freie Parkplätze durch in den Parkhäusern angebrachte Kameras identifiziert und an das System von VIRTUOSE-DE übermittelt. Die Plattform ermittelt dann den Weg eines Fahrzeugs zum nächsten Parkplatz und sendet ihn an die Fahre-

### Meilensteine im Projektverlauf



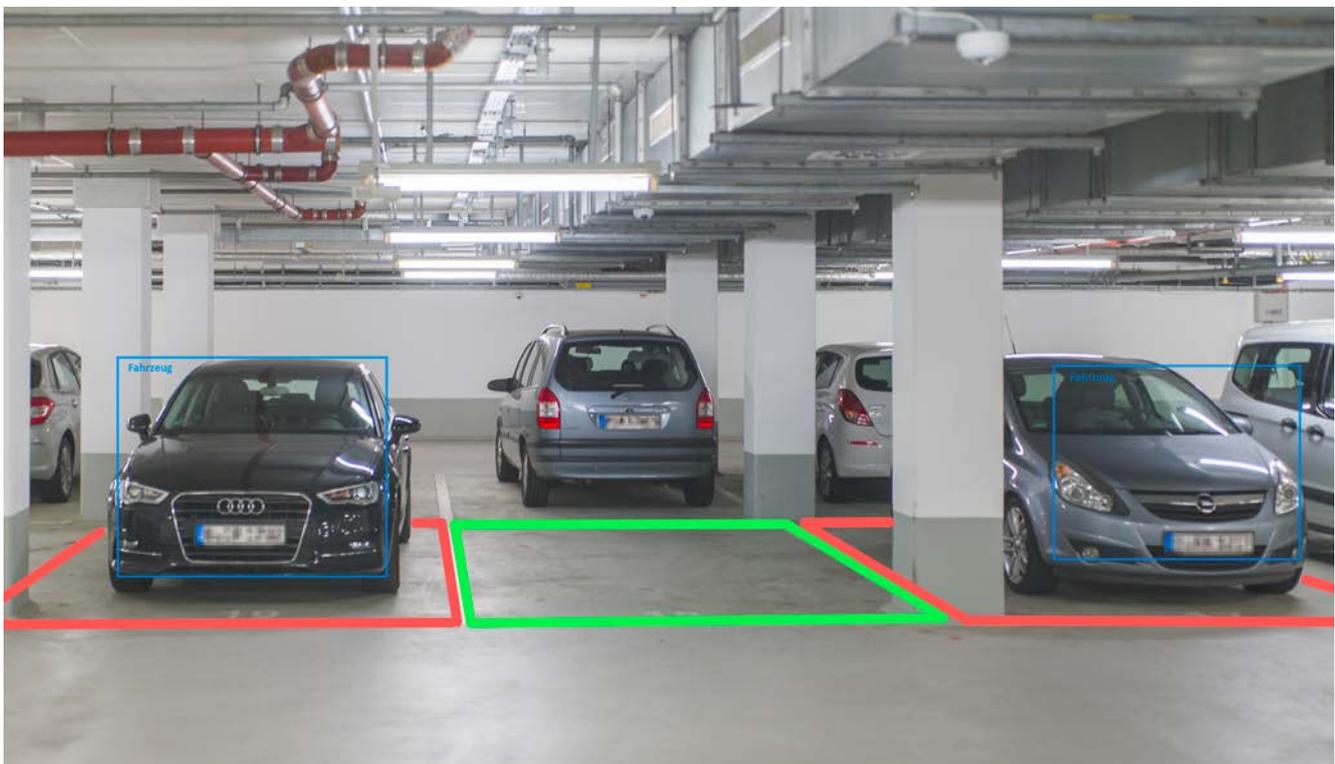
rin beziehungsweise den Fahrer. In einer künftigen Aufbau-  
stufe könnte das Fahrzeug auch selbständig zum nächsten  
freien Parkplatz fahren, wobei autonomes Fahren nicht  
während der Projektlaufzeit geplant ist. Neben der Position  
des freien Parkplatzes werden darüber hinaus auch durch  
Kameras erhobene Informationen zu weiteren Fahrzeugen  
und Fußgängerinnen und Fußgänger übermittelt. In der  
dritten Ausbaustufe des Projektes soll die Plattform auch  
die Parkplatzsuche im öffentlichen Straßenverkehr unter-  
stützen. Dazu sind Tests auf einem Parkplatz der TU Berlin  
geplant

Das zweite Anwendungsszenario betrifft videobasierte  
Sicherheitsvorkehrungen im ÖPNV. Durch die Echtzeitfä-  
higkeit der Technologie von VIRTUOSE-DE können kriti-  
sche Ereignisse wie medizinische Notfälle sehr schnell  
identifiziert werden. Diese Informationen stehen den Ein-  
satzzentralen zur Verfügung, somit können unmittelbar  
entsprechende Gegenmaßnahmen eingeleitet werden.  
Auch auf beobachtete Straftaten kann so Einfluss genom-  
men werden.

Tests zum Einsatz der Plattform werden zunächst in Innen-  
räumen von Gebäuden und anschließend in Bussen unter  
realen Bedingungen durchgeführt. VIRTUOSE-DE sucht  
dafür einen Partner und plant, einen Demonstrator mit  
echten Fahrzeugen und Systemen zu realisieren.

Die nächsten Schritte im Projekt beinhalten darüber hinaus  
die Entwicklung und den Test der einzelnen Komponenten  
des Systems, wie der Cloud-Technologie und der Analyse-  
Algorithmen. Im Anschluss werden Tests für die zwei zuvor  
skizzierten Anwendungsfälle vorgenommen und eine erste  
Version der Demonstratoren erstellt – die anschließend auf  
der Basis weiterer Tests optimiert werden.

Nach dem Projektende im Mai 2019 ist eine weitere Verbes-  
serung der Plattform geplant. So kann im Bereich des Park-  
raummanagements die Erweiterung auf autonomes Fahren  
im Parkhaus vorgenommen werden, im Bereich der video-  
basierten Sicherheit im ÖPNV ist ein flächendeckender  
Einsatz bei verschiedenen ÖPNV-Unternehmen und eine  
Einbindung in bestehende Systeme wie Videoüberwa-  
chungsanlagen denkbar.



## Die Projekt-Ergebnisse auf einen Blick

Früher	Mit VIRTUOSE-DE
<p>Autofahrer, die in einem Parkhaus auf der Suche nach einem freien Parkplatz sind, haben keine Informationen über den nächsten verfügbaren Parkplatz.</p>	<p>Autofahrer werden in einem Parkhaus zum nächsten freien Parkplatz geleitet. Dazu wird über die Infrastruktur des Parkhauses die Route zum Parkplatz an das Fahrzeug übermittelt. Auf längere Sicht kann das Fahrzeug autonom zum Parkplatz finden, sodass der Fahrer bereits an der Einfahrt zum Parkhaus das Fahrzeug verlassen kann.</p>
<p>Autofahrer, die in der Innenstadt parken wollen, sind heutzutage vor allem in den Großstädten längere Zeit auf der Suche nach einem freien Parkplatz.</p>	<p>Mit Angeboten basierend auf der VIRTUOSE-DE-Technologie können künftig Services bereitgestellt werden, die den Autofahrer direkt zum nächsten freien Parkplatz leiten.</p>
<p>Im ÖPNV kommt Videoüberwachung zu Dokumentationszwecken zum Einsatz. Die Aufzeichnungen werden bei einem Sicherheitsvorfall erst im Nachhinein analysiert.</p>	<p>Mit der VIRTUOSE-DE-Technologie können kritische Ereignisse nahezu in Echtzeit erkannt werden. Entsprechende Gegenmaßnahmen können dadurch sofort eingeleitet werden. Auf beobachtete Straftaten kann somit direkt Einfluss genommen werden.</p>
<p>Videoüberwachung basiert momentan besonders auf Installationen, die individuell gepflegt werden müssen. Zudem erfolgt die Auswertung der Videodaten nicht automatisch.</p>	<p>Die VIRTUOSE-DE-Plattform kann genutzt werden, um Services für Videoanalysen und -überwachung anzubieten („Video Surveillance as a Service“). Für die Videoanalyse werden unter anderem automatisierte Algorithmen verwendet.</p>

# Fachgruppe Sicherheit



Welches Maß an Sicherheit wird bei der Verwendung von Smart-Data-Technologien benötigt? Wo liegen besondere Sicherheits Herausforderungen in der intelligenten Auswertung großer Datenmengen? Welche Rolle spielt der Datenschutz bei Smart Data? Mit diesen grundlegenden Fragen beschäftigte sich die Fachgruppe Sicherheit. Die Arbeit der Fachgruppe behandelte neben Regeln und Leitlinien für den Umgang mit Daten, insbesondere mit personenbezogenen, ebenso die Themen Datenverarbeitung und -speicherung. Ein besonderer Fokus der Arbeit der Fachgruppe wurde frühzeitig auf Fragen der technischen Umsetzung von Datenschutzanforderungen gelegt.

## Sicherheit ist ein interdisziplinäres Thema

Von Beginn an verfolgte die Fachgruppe eine interdisziplinäre Herangehensweise in Kooperation mit der Fachgruppe Rechtsrahmen. Aus diesem Grund haben die beiden Fachgruppen gemeinsam mehrere interdisziplinäre Workshops zu Themen wie etwa der Anwendbarkeit datenschutzrechtlicher Vorgaben oder Maßnahmen im Sinne der Datenminimierung veranstaltet. Durch Expertinnen- und Expertenvorträge und Praxisbeispiele während der gemeinsamen Workshops konnten den Smart-Data-Projekten sowohl die Relevanz als auch die Herausforderungen beispielweise von Vorgaben wie der Anonymisierung oder Pseudonymisierung aufgezeigt werden.

Das Bundesdatenschutzgesetz (BDSG) schreibt eine Anonymisierung – oder Pseudonymisierung – personenbezogener

Daten vor, wann immer es möglich und verhältnismäßig ist. Die technische Umsetzung dieser zentralen Datenschutzanforderungen, war ein wichtiger Diskussionspunkt im ersten gemeinsamen Workshop beider Fachgruppen im Juli 2015. Im Speziellen diskutierten beide Fachgruppen die Frage, inwieweit Verschlüsselungsmechanismen für eine **Anonymisierung oder für eine Pseudonymisierung** von Smart Data genutzt werden können.

Um diese Frage besser beantworten zu können, haben beide Fachgruppen ein Begriffsverständnis von Anonymisierung diskutiert und in der Folge weiterentwickelt, welches gemeinsam und disziplinübergreifend verwendet werden kann. Zusätzlich zu diesem erarbeiteten Begriffsverständnis von Anonymisierung verfassten die beiden Fachgruppen ergänzend zum Thema eine neue Publikation mit dem Titel „Smart Data – Smart Privacy?“. Diese Publikation hat zusätzlich die Konzepte und Methoden der Anonymisierung und der Pseudonymisierung von Daten vorgestellt. Dabei wurden sieben Thesen entwickelt, als Impulse für eine interdisziplinäre rechtlich-technische Evaluation.

## Nutzung und Sicherheit von Daten

Die Fachgruppe Sicherheit beschäftigte sich auch mit den Themen **Anonymisierung, Datennutzungskontrolle, Authentifikation und Verschlüsselung**. Hierbei hat die Fachgruppe Sicherheit den Schwerpunkt auf die Kontrolle der Verwendung von Daten durch Dritte gelegt. Dazu referierte Denis Feth (Fraunhofer IESE), wie anhand des Frame-

works IND<sup>2</sup>UCE die Verwendung von Daten durch Dritte kontrolliert werden kann. Das Framework IND<sup>2</sup>UCE macht eine Datennutzungskontrolle möglich: Während eine Zugriffskontrolle lediglich regelt, wer auf welche Daten oder Dienste zugreifen darf, geht die Datennutzungskontrolle einen Schritt weiter. Mithilfe von bestimmten Regelsätzen, sogenannten Sicherheitsrichtlinien, kann sehr exakt und spezifisch geregelt werden, welcher Umgang mit den Daten erlaubt ist und welcher nicht. Auf diese Weise kann die Besitzerin oder der Besitzer der Daten festlegen, zu welchem Zweck die entsprechenden Daten genutzt werden dürfen.

Um persönlichkeitsrelevante Daten so sparsam wie möglich zu nutzen, können dafür zum einen Techniken für sichere Mehrparteienberechnungen eingesetzt werden, die aus der Kryptographie stammen. Zum anderen können auch Techniken zur Anonymisierung verwendet werden. Klassische Beispiele solcher Techniken sind das Verallgemeinern und Unterdrücken von Daten, aber auch das Veraschen. Bei Letzterem werden die Ursprungsdaten kontrolliert verfälscht. Das heißt, dass die Daten mithilfe einer „Rauschquelle“ mit einer definierten Verteilung verfälscht werden, aber ein Gesamtdatensatz für Statistikzwecke brauchbar bleibt. Dabei macht es einen Unterschied, an welcher Stelle im Prozess die Daten anonymisiert werden: Sowohl die erhobenen Daten, Zwischenergebnisse, aber auch Analyseergebnisse können anonymisiert werden. Anonymisiert man ausschließlich die Analyseergebnisse, behält man einen exakten Ausgangsdatenbestand, der beispielsweise für potentielle weitere Analysen bei einer vertrauenswürdigen Instanz lagern kann. Sebastian Probst Eide (Aircloak GmbH) stellte im Workshop eindrucksvoll dar, wie Datensätze durch Anonymisierungstechniken von personenbezogenen Angaben bereinigt werden.

## Modelle und Tools für Best Practice und Standards

Der Fachgruppe Sicherheit wurden die „**Best Practices und Vorgehensmodelle zur IT-Sicherheit und zum Datenschutz**“ das Standard-Datenschutzmodell (SDM) und der vom Bundesamt für Sicherheit in der Informationstechnik (BSI) entwickelte IT-Grundschutz als Orientierungshilfen präsentiert. Während sich der IT-Grundschutz des BSI auf grundlegende Sicherheitsmaßnahmen und dazu ergänzende Schutzprogramme fokussiert, verbindet das SDM Schutzziele der IT-Sicherheit mit Schutzziele, die von Datenschutzgrundsätzen abgeleitet wurden. Somit

können IT-Sicherheit und Datenschutz in einem Compliance-Prozess mitbedacht werden.

Im gemeinsamen Workshop der Fachgruppen Sicherheit und Rechtsrahmen standen neben dem Thema der Anonymisierung auch Methoden der Daten- und Risikominimierung im Vordergrund der Vorträge und Diskussionen. So stellte beispielsweise Herr Dr. Matthias Ihle (Averbis) das Averbis DeID-Tool zur Anonymisierung und De-Identifizierung von Freitexten vor. Herr Dr. Fabian Prasser (TU München) erläuterte ergänzend die Hintergründe, Methoden und Werkzeuge zur Anonymisierung (relationaler) Daten. Anschließend entwickelten die Teilnehmerinnen und Teilnehmer beider Fachgruppen in sogenannten „Breakout-Gruppen“ Thesen und Forderungen zu folgenden Fragen: Wie lassen sich Sach- von Personenbezügen in Maschinendaten abgrenzen? Wie können Gesundheits- und Bilddaten anonymisiert werden? Wie kann Text Mining auf öffentlich zugänglichen Daten erfolgen?

Neben diesen Fragen diskutierten die Teilnehmer beider Fachgruppe auch unter anderem Ansätze zum „**Privacy Preserving Data Mining**“. Beim Data Mining werden Daten automatisiert nach bestimmten Mustern durchsucht, um vorher nicht offensichtliche Zusammenhänge aufzudecken. Beim Privacy Preserving Data Mining werden diese Daten wiederum so verarbeitet, dass natürliche Personen im juristischen Sinne in den Ursprungsdatensätzen geschützt bleiben. Somit könnte durch die Technologien eine Annäherung an die Vorgabe des BDSG, möglichst wenige persönlichkeitsrelevante Daten zu verwenden, erreicht werden.

## Fazit der Fachgruppe Sicherheit

Die Arbeit der Fachgruppe Sicherheit und die dazugehörige Zusammenarbeit mit der Fachgruppe Rechtsrahmen haben deutlich erkennen lassen, dass die Fragen der technischen Umsetzung datenschutzrechtlicher Anforderungen für die Projekte im Förderprogramm eine hohe Relevanz besitzen. Im Laufe des Programmes setzen einige Projekte Anonymisierungs- sowie Datenminimierungskonzepte bereits sehr erfolgreich um. Weitere für die Projekte relevante Themen sind Authentifikation, Zugriffs- und Benutzungskontrolle und Verschlüsselung. Durch die Zusammenarbeit der Projekte mit den Fachgruppen Sicherheit und Rechtsrahmen konnte sowohl ein Problembewusstsein für die Wechselwirkungen von Recht und Technik geschaffen werden, aber auch Lösungen und Ansätze wurden präsentiert und diskutiert.

# SAKE – Semantische Analyse Komplexer Ereignisse

## Mit Semantik und Big Data zu effizienten Prozessen

### Das Projekt

Im Maschinenbau werden immer mehr Prozesse automatisiert. Damit einher geht eine wachsende Überwachung der Fertigungsschritte durch eine Vielzahl von Sensoren pro Maschine, wodurch täglich riesige Datenmengen entstehen. Durch die intelligente Verarbeitung dieser Datenmengen können Produktionsabläufe noch zuverlässiger und effektiver werden. Mit den gängigen Werkzeugen zur Datenauswertung ist eine Analyse jedoch nicht möglich, da diese nicht dafür geschaffen wurden.

Ziel von SAKE ist die Entwicklung einer passenden Plattform für die Integration und Verarbeitung von großen, heterogenen Datenströmen. Die Datenanalyse erfolgt mit Hilfe des Maschinellen Lernens. Das bedeutet, dass Daten modularisiert und in einem einheitlichen Format verarbeitet werden. Beim Einsatz des Maschinellen Lernens werden ausschließlich die für das Ziel der Analyse tatsächlich relevanten Daten herangezogen, so dass viel Zeit und Rechenleistung gespart wird.

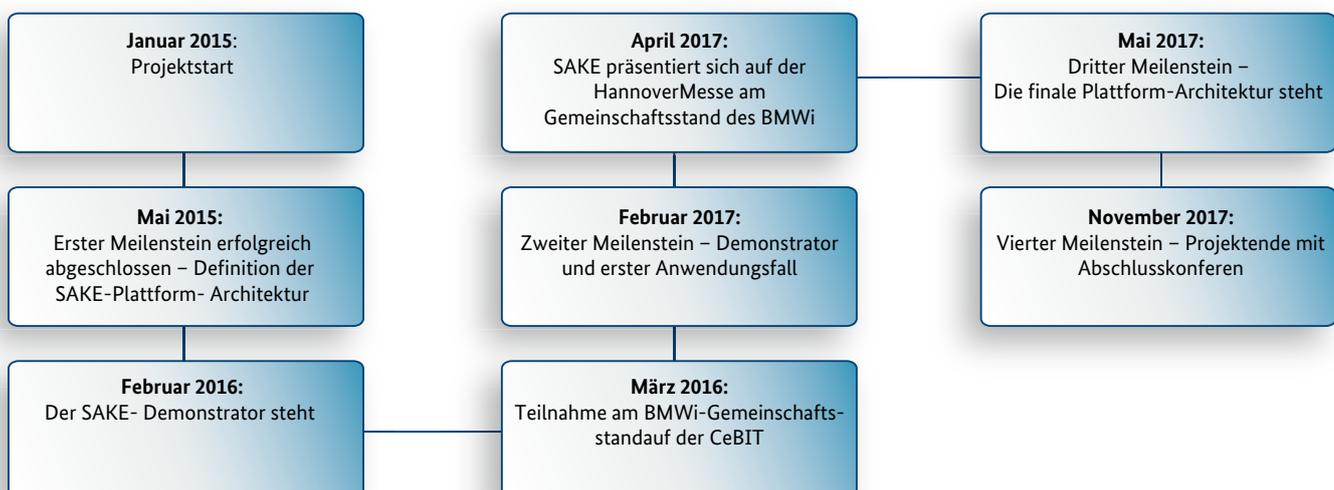
Durch die gewonnenen Analyse-Erkenntnisse können Ressourcen geschont und potenzielle Fehlerquellen erkannt und frühzeitig behoben werden. Zudem werden nicht nur Informationen darüber ausgegeben, ob ein Betriebsfehler vorliegt, sondern auch, welche Ursachen diesem Fehler zugrunde liegen. Durch diese Prozessoptimierung können Produktionskosten stark gesenkt werden. Die Anwendungsschwerpunkte im Projekt SAKE sind der Maschinenbau und das IT-Monitoring.

### Weg in die Praxis

In der dreijährigen Laufzeit von SAKE wurde im Austausch mit potenziellen Anwenderinnen und Anwendern auf Messen, Tagungen und Konferenzen sowie bei Kundengesprächen der beteiligten Unternehmen des Konsortiums der steigende Bedarf für Analysemethoden großer Datenbestände bestätigt. Dieser zeigt sich nicht nur bei den fokussierten Unternehmen aus dem Maschinen- und Anlagenbau, sondern auch aus den Bereichen Logistik und Automobilbau. Ein erstes Pilotprojekt konnte bereits im Automobilbau initiiert werden: Bei einem deutschen Automobilhersteller wurde mittels der zu dem Zeitpunkt vorhandenen SAKE-Ergebnisse ein Lernverfahren erstellt, mit dem sensorische Fehler in Messungen bei Testfahrten erkannt und automatisch korrigiert werden konnten. Dadurch konnten die Ergebnisse der sehr geld- und zeitaufwendigen Testfahrten zum einen nützlich verarbeitet und zum anderen deutlich schneller ausgewertet werden.

Aufgrund der positiven Projektergebnisse hat der Konsortialführer des Projekts, die USU Software AG, einen eigenen Geschäftsbereich gegründet, der sich mit der Analyse industrieller Daten beschäftigt. Das dabei eingesetzte Produkt KATANA (Karlsruhe Machine Tool Analytics) nutzt die Projektergebnisse dafür, intelligente Analysen für die frühzeitige Fehlererkennung und -behebung sowie die allgemeine Prozessoptimierung für Unternehmen aus verschiedenen Branchen anzubieten.

### Meilensteine im Projektverlauf



Auf Grundlage von SAKE-Ergebnissen konnte USU KATANA bereits für erste Kundenprojekte verwendet werden. Neben einem Unternehmen aus der Druckindustrie zählt ein Werkzeugmaschinenhersteller zu den Kunden von USU KATANA. Darüber hinaus laufen Gespräche mit internationalen Konzernen aus dem Maschinen- und Anlagenbau.

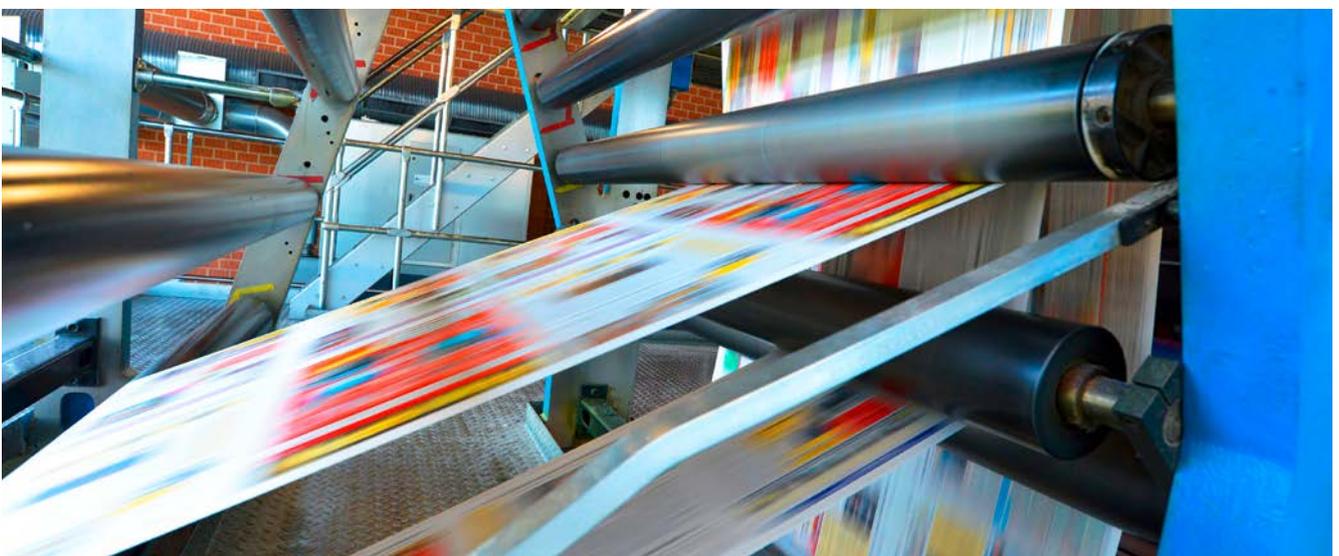
Das in SAKE entwickelte Konzept und die Ergebnisse legen wichtige Fundamente für weitere Projekte. Das Konzept diente als Grundlage für die erfolgreiche Bewerbung auf ein weiteres Forschungsprojekt, das sich mit produktions-

bezogenen Dienstleistungssystemen auf Basis von Big-Data-Analysen beschäftigt.

Das vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Projekt ProData ist im September gestartet. Auch das Projekt STEP, welches seit 2016 im Technologieprogramm Smart Service Welt des BMWi gefördert wird, baut auf den Ergebnissen von SAKE auf. Basierend auf den analysierten Daten der SAKE-Plattform entwickelt STEP smarte Dienste zur bedarfsgerechten und effizienten Planung von Techniker-Einsätzen.

### Die Projekt-Ergebnisse auf einen Blick

Früher	Mit SAKE
Im Schadensfall müssen Protokolldateien von Produktionsprozessen mühsam manuell ausgewertet werden, was Stunden oder auch Tage dauern kann.	Durch maschinelle Lernverfahren können Fehler in Datenströmen vorhergesagt werden. Fehleranalysen können auch bei großen Datenmengen in Sekunden oder Minuten durchgeführt werden.
Die Instandsetzung einer Maschine erfolgt erst, wenn diese ausfällt (reaktiv) oder aufgrund gesetzlicher oder sonstiger Auflagen in festen Zyklen (präventiv, zum Beispiel wegen TÜV-Tests). Dadurch kommt es zu Produktionsausfällen oder überflüssigen Wartungsarbeiten.	SAKE ermöglicht eine prädiktive Wartung: Wartungszeiträume und Dringlichkeit können anhand des tatsächlichen Zustandes vorhergesagt werden. Unnötige Wartungsarbeiten und zusätzliche Produktionsausfälle werden vermieden.
Die Analyse von Daten und der Einsatz maschineller Lernverfahren sind Spezialisten (Data Scientists) vorbehalten.	SAKE ermöglicht eine einfachere Interpretierbarkeit der Analyse-Ergebnisse. Maschinelles Lernen wird so auch für normale Anwender zugänglich.
Die Verarbeitung großer Datenmengen im Industriebereich erfordert teure Hardware.	Die SAKE-Architektur setzt auf skalierbare Komponenten, die auf Basis von Standard-Hardware arbeiten und bei Bedarf stückweise erweitert werden können. Hochwertige Datenverarbeitungslösungen können mit kostengünstiger Hardware realisiert werden und sind damit auch für den Mittelstand finanzierbar.



# PRO-OPT – Big-Data-Produktionsoptimierung in Smart Ecosystems

## Bessere Produktions- und Wertschöpfungsketten

### Das Projekt

Arbeitsteilung spielt in der industriellen Produktion, zum Beispiel der Automobilindustrie, eine große Rolle. So bilden Fahrzeuge ein äußerst komplexes Gesamtsystem, das aus Komponenten verschiedener Zulieferer besteht. Für die Wettbewerbsfähigkeit des Endprodukts muss die Produkt- und Prozessqualität über die gesamte Lieferkette hinweg optimiert werden. Dabei spielen die bei der Erstellung des Produkts anfallenden Daten eine zentrale Rolle. Die digitale Vernetzung aller beteiligten Unternehmen in dezentralen Datensystemen, sogenannten Smart Ecosystems, gewinnt deshalb an Bedeutung.

Eine zentrale Herausforderung ist die Wahrung der Datenhoheit der beteiligten Organisationen. Die einzelnen Unternehmen wollen die Kontrolle über ihre Daten behalten und zögern oft mit der Herausgabe. Der offene Austausch von Rohdaten birgt das Risiko des Verlusts von internem Geschäftswissen, das einen Erfolgsfaktor darstellt. Hinzu kommt, dass die Daten verteilt auf die unterschiedlichsten Datenquellen in verschiedenen Formaten vorliegen. Sie unterscheiden sich zudem bezüglich ihrer Qualität und Verlässlichkeit.

Das Ziel von PRO-OPT war es deshalb, die Wettbewerbsfähigkeit der gesamten Lieferkette durch organisationsübergreifende Datenanalyse sicherzustellen. Das Know-how und die Ziele der einzelnen Unternehmen, sowie deren Datenhoheit, sollen gewahrt bleiben. Hierzu hat PRO-OPT eine Smart-Data-Plattform entwickelt, die übergreifende Analysen erlaubt. Rohdaten verlassen dabei niemals das Unternehmen. Stattdessen werden Analysen jeweils in Teilen unternehmensintern ausgeführt und die Ergebnisse zentral zusammengeführt.

Die Plattform unterstützt sowohl die Katalogisierung von Daten und Analysemöglichkeiten als auch die Bewertung der Datenqualität und die Handhabung von verteilten großen Datenmengen (Big Data). Darauf aufbauend regelt sie, welche Analysen in welcher Situation für welches Unternehmen zulässig sind.

Die Smart-Data-Plattform von PRO-OPT wurde für die Verbesserung der Produktionsprozesse in der Automobilindustrie entwickelt – langfristig können aber auch andere Branchen wie die Medizintechnik oder die Energieversorgung von ihr profitieren.

### Meilensteine im Projektverlauf



## Weg in die Praxis

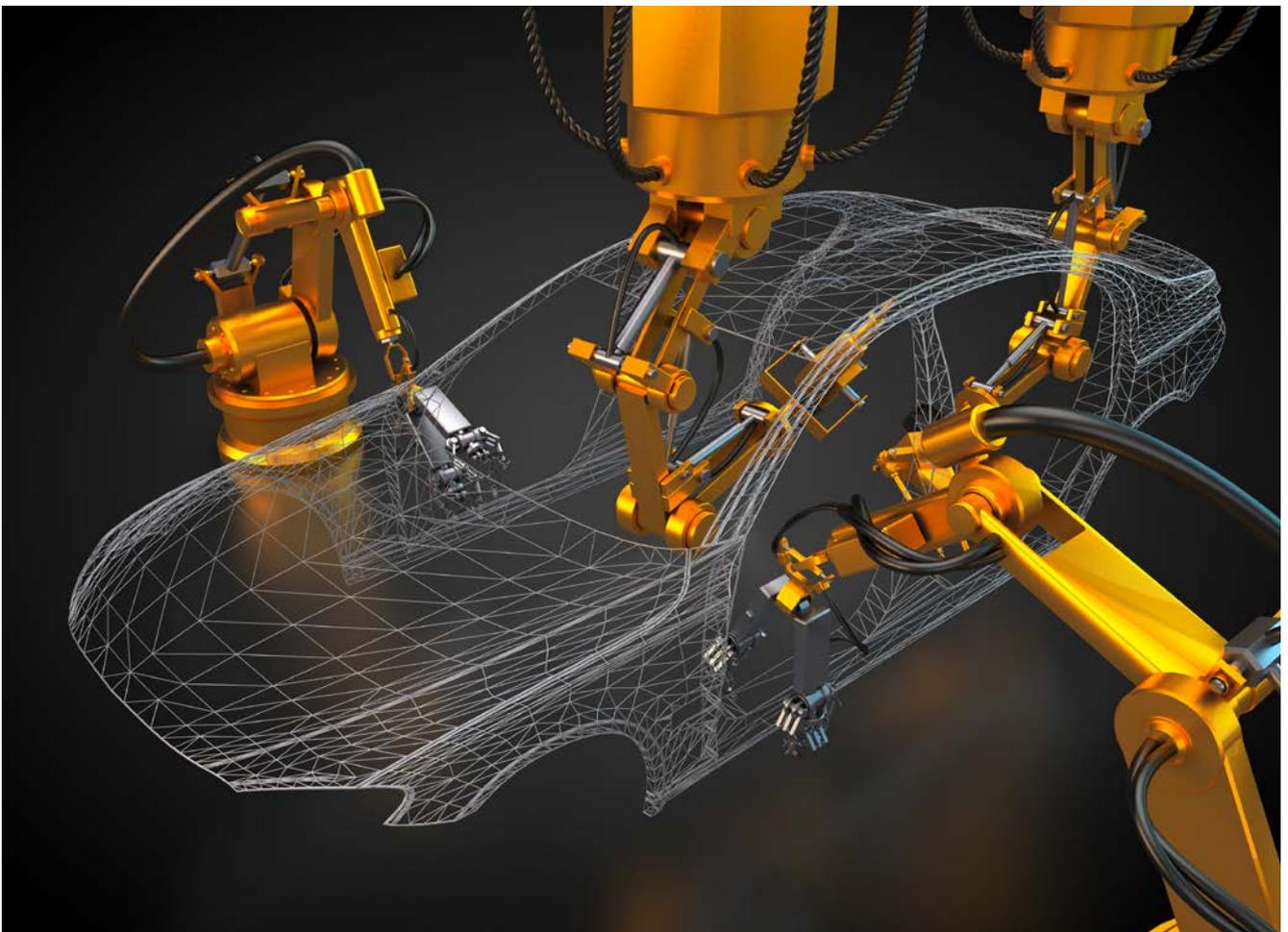
Während der knapp dreijährigen Projektlaufzeit von PRO-OPT konnten mehrere der entwickelten Projektbausteine und die Plattform selbst erfolgreich bei den beteiligten Projektpartnern erprobt werden. Das entwickelte Smart-Data-Mining, also die systematische Anwendung statistischer Methoden zur Analyse und Aufdeckung von Zusammenhängen zwischen großen Datenmengen, wurde erfolgreich mit Produktionsdaten getestet. Nach dem Projekt soll das Smart-Data-Mining als Teil einer Diagnoselösung des Projektpartners DSA (Daten- und Systemtechnik GmbH) für die Automobilproduktion eingesetzt werden. Ein späterer Einsatz beim Partner Audi ist angedacht.

Auch die prototypisch entwickelten Big-Data-Traceability-Analysen werden mittelfristig in Produkte des Projektpartners camLine einfließen. Damit können Prozessschritte rückverfolgt werden, sodass der Gesamtstatus eines Pro-

jekts oder beispielsweise Fehler einfacher und schneller erkannt werden können.

Der praktische Einsatz der gesamten PRO-OPT-Plattform bei Audi und beim assoziierten Partner Continental wird evaluiert. Zudem wird derzeit geplant, einzelne Komponenten der Plattform als Open-Source-Software zur Verfügung zu stellen. Das Konsortium wird zukünftig auch Beratungen zur Einführung von kollaborativer Datenanalyse in Smart Ecosystems, also der gemeinschaftlichen Datenauswertung von verschiedenen Unternehmen, anbieten.

Neben dem erfolgreichen Einsatz in der Praxis soll das Projekt auch auf der Forschungsebene weitergeführt werden. Ziel ist es, die Plattform so weiterzuentwickeln, dass die Ergebnisse von Datenanalysen noch stärker in den Entscheidungs- und Produktionsprozessen verankert werden können.



### Die Projekt-Ergebnisse auf einen Blick

Früher	Mit PRO-OPT
Probleme oder Fehler bei Produkten wurden von an der Produktion oder der Lieferkette beteiligten Partnern meist erst spät erkannt. Häufig kam es zu nachträglichen Garantieansprüchen.	Teilnehmer aus Produktion und Lieferkette können Analysen für andere Teilnehmer freigeben. So können diese das Verhalten ihrer Teilprodukte nachverfolgen. Zudem lassen sich auch die Wechselwirkungen der Teilprodukte untersuchen.
Die Problemanalyse ist ein aufwändiger Prozess: Benötigte Daten werden gar nicht oder nur durch unkontrollierten Dateiaustausch zur Verfügung gestellt.	Die PRO-OPT-Plattform ermöglicht den Zugriff auf Einzelergebnisse, unter Berücksichtigung der Schutzbedürfnisse aller beteiligten Unternehmen.
Voraussetzung für übergreifende Analysen ist die vorherige Zentralisierung von Daten.	PRO-OPT ermöglicht die Durchführung von Analysen in den einzelnen Unternehmen und die Zusammenführung von Ergebnissen. Das Unternehmen behält Hoheit über die Daten.
Die Analyse von Problemen führt oft zu Konflikten zwischen den Partnern.	Die Problemanalyse wird ein kooperativer Prozess.
Die Datenauswertung erfolgt manuell.	PRO-OPT ermöglicht die Anwendung statistischer Methoden zur Aufdeckung von Zusammenhängen („Smart-Data-Mining“)
Große Datenmengen in mehrstufigen Lieferketten können nur ausschnittsweise betrachtet werden.	Die übergreifende Traceability („Nachverfolgungs“-)Analyse erlaubt die Nutzung der Daten über die gesamte Lieferkette hinweg.

# SIDAP – Skalierbares Integrationskonzept zur Datenaggregation, -analyse, -aufbereitung von großen Datenmengen in der Prozessindustrie

## Optimaler Anlagenbetrieb in der Prozessindustrie

### Das Projekt

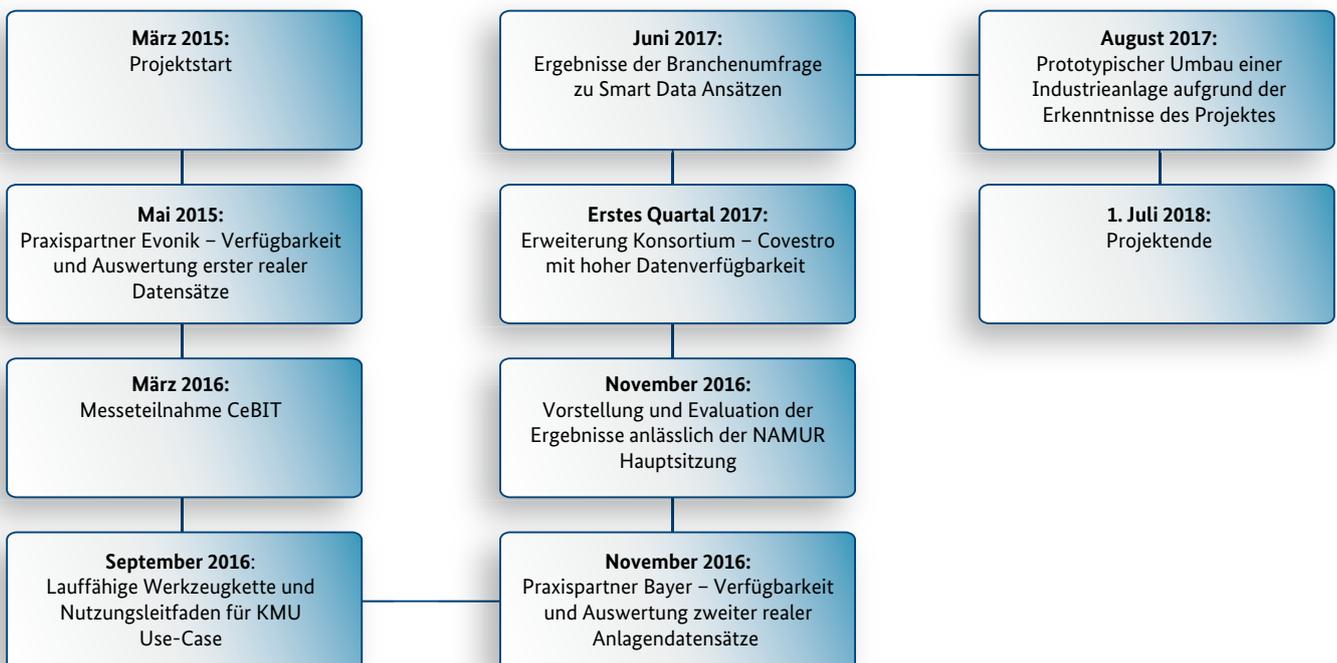
Die Prozessindustrie spielt am Industrie- und Technologiestandort Deutschland eine wichtige Rolle, da sie hochautomatisierte Anlagen – etwa für die Produktion von Arzneimitteln oder Spezialchemie – entwickelt und betreibt. In der Prozessindustrie kommen Geräte unterschiedlicher Hersteller zum Einsatz, welche die jeweiligen Prozesse zum Teil gemeinsam oder ganz unabhängig voneinander steuern. Während die Gerätehersteller sich auf die Auswertung von Daten zum Betrieb und zum Lebenszyklus ihrer Geräte konzentrieren, sammeln und verarbeiten die Anlagenbetreiber die entstehenden Daten der jeweiligen Teilanlagen und Teilprozesse. Trotz der vielen erhobenen Daten wird das Potenzial einer übergreifenden Analyse möglichst aller Prozessdaten nur selten oder überhaupt nicht genutzt.

Das Projekt SIDAP entwickelt deshalb eine Smart-Data-Lösung, welche die interaktive Analyse und Wartung der Systemtechnik- und Prozessleitsysteme ermöglicht und dadurch den Anlagebetrieb optimiert. Hierfür trägt die entwickelte Smart-Data-Lösung die Daten und Informationen

aus den unterschiedlichen IT-Systemen der einzelnen Prozesse zusammen, bereitet diese auf und setzt sie miteinander in Verbindung. Darauf aufbauend werden im Projekt neue Methoden und Modelle zur individuellen Analyse und sicheren Verarbeitung dieser erhobenen Prozessdaten entwickelt.

Ziel von SIDAP ist es, aus den unternehmensübergreifenden Nutzungsdaten neue Erkenntnisse über Ursachen und Zusammenhänge von Geräteausfällen zu gewinnen und konkrete Gegenmaßnahmen zur Behebung dieser Probleme zu entwickeln. Auf diese Weise können Unternehmen die Zahl ihrer Geräte- und Ausrüstungsausfälle reduzieren und die Leistung und Verfügbarkeit ihrer Anlagen verbessern. Von einer geringeren Fehlerquote profitiert letztendlich die Qualität der produzierten Produkte. Außerdem ist die verbesserte Verfügbarkeit der Anlagen die Basis für die weitergehende Automatisierung von Prozessen und die Fernüberwachung und -steuerung (Remote-Betrieb) von Anlagen im Kontext von Industrie 4.0.

### Meilensteine im Projektverlauf



## Weg in die Praxis

Um möglichst anwendungsbezogene Ergebnisse zu erzielen, verfolgt das Projekt SIDAP den Ansatz, zuerst konkrete Lösungen an Beispielen zu erarbeiten und ausgiebig zu testen, um sie in einem nächsten Schritt auf ähnliche Prozesse anwenden zu können. Hierfür werden innerhalb des Projekts insgesamt drei Anwendungsfälle (Use Cases) verfolgt: In einem Use Case steht die Fehlerdiagnose und Nutzungsoptimierung von Ventilen aus der chemischen Industrie im Mittelpunkt, während sich der zweite Anwendungsfall auf die Fernwartung von Anlagen konzentriert. In einem weiteren Use Case wird unter Leitung des Konsortienpartners Gefasoft GmbH nach Wegen gesucht, das im Projekt erworbene Wissen auf kleine- und mittelständische Unternehmen zu übertragen.

Eine große Herausforderung für die Projektteilnehmerinnen und Projektteilnehmer ist es, relevante Daten in ausreichender Qualität und Menge zu sammeln. Dies hat neben urheber- und datenschutzrechtlichen auch technische Gründe. Anlagen in der Prozessindustrie sind in der Regel historisch gewachsen und agieren getrennt voneinander. Zwar werden innerhalb der Anlagen große Mengen an Daten gesammelt, es fehlen jedoch oft die notwendigen Schnittstellen, die es erlauben würden, diese auch entsprechend zu extrahieren und zu verarbeiten. Eine komplette Umrüstung von Anlagen, um diese Verbindungsstellen zu erschaffen, ist jedoch äußerst kostspielig.

Als weitere Datenquelle nutzt das Projekt SIDAP wissenschaftliche Versuchsanlagen, um Daten zu sammeln. Ergänzend rüsten die Projektteilnehmer Evonik Industries AG, Covestro AG sowie die Bayer AG schrittweise bestehenden Anlagen mithilfe von Hardware-Erweiterungen nach. Um die einzelnen Geräte innerhalb einer Anlage wiederum miteinander zu verbinden und die Daten an den bestehenden Infrastrukturen vorbei zu schleusen, kommt im Projekt eine neuartige Systemarchitektur in Anlehnung an die NAMUR Open Architecture (NOA) zum Einsatz. Diese Plattform ermöglicht verschiedene IT-Komponenten und Sensorfunktionen einfach zu integrieren, ohne dabei deren Betrieb zu beeinträchtigen.

Aufbauend auf den dadurch erzielten Ergebnissen will SIDAP konkrete Lösungen für die Praxis entwickeln. Dies beinhaltet zum einen die feste Integration von computergestützten Methoden zur Erkennung von Mustern, Trends oder Zusammenhängen in Produktionsleitsystemen (Data Mining), um Daten direkt auswerten zu können. Zum anderen soll Anlagenbetreibern die Möglichkeit geboten werden, in einer neutralen Cloudumgebung ihre Daten mit den im Projekt entwickelten Analysemodellen zu analysieren. Mithilfe der Analyseergebnisse kann dann eine effiziente Instandhaltungsstrategie implementiert werden.



## Die Projekt-Ergebnisse auf einen Blick

Früher	Mit SIDAP
Schwer zugängliche Gerätedaten mit unbekannter Aussagekraft sind in den Datenbanken der Betreiber gespeichert.	Einfachere Identifikation der aussagekräftigen Daten für die Analyse von Gerätedaten.
Die Wartungsinformationen sind kaum zugänglich, unvollständig und nicht aussagekräftig.	Notwendige Informationen aus den Werkstätten zur Wartung und zur Instandsetzung der Regelventile sind bekannt. Eine App zur Erfassung von Wartungsinformationen unterstützt den Wartungsvorgang.
Anlagenbetreiber können ihre Gerätedaten nicht auswerten.	Anlagenbetreiber werten Gerätedaten aus und bauen auf Versuchsdaten (Fehlermustern) des Gerätelieferanten auf, um Fehler vorhersagen zu können.
Die Datenqualität der bestehenden Datensätze ist unbekannt. Die für Datenauswertung notwendige Vorverarbeitung fehlt oftmals.	Die Einschätzung der Datenqualität und der notwendigen Vorverarbeitung wird erleichtert. Die Datenqualität wird durch den Umbau auf bessere Regelventile in realen Industrieanlagen erhöht.
Es ist kein Modell und keine einheitliche Architektur zur Kommunikation zwischen Datenerhebung und -auswertung vorhanden.	Ein Modell für Ventile sowie Architekturvorschlag ist vorhanden. Die Datenauswertung wird somit stark vereinfacht und Ergebnisse sind schneller erzielbar.
Die Daten aus Produktionsleitsystemen werden in KMU nicht zur Datenauswertung genutzt.	Durch Data Mining in Produktionsleitsystem können Potenziale zur Produktionsoptimierung erschlossen werden. Ein Nutzerleitfaden ermöglicht Anwendern den schnellen Einstieg in die Thematik.

# Smart Data Web – Datenwertschöpfungsketten für industrielle Anwendungen

## Das Projekt

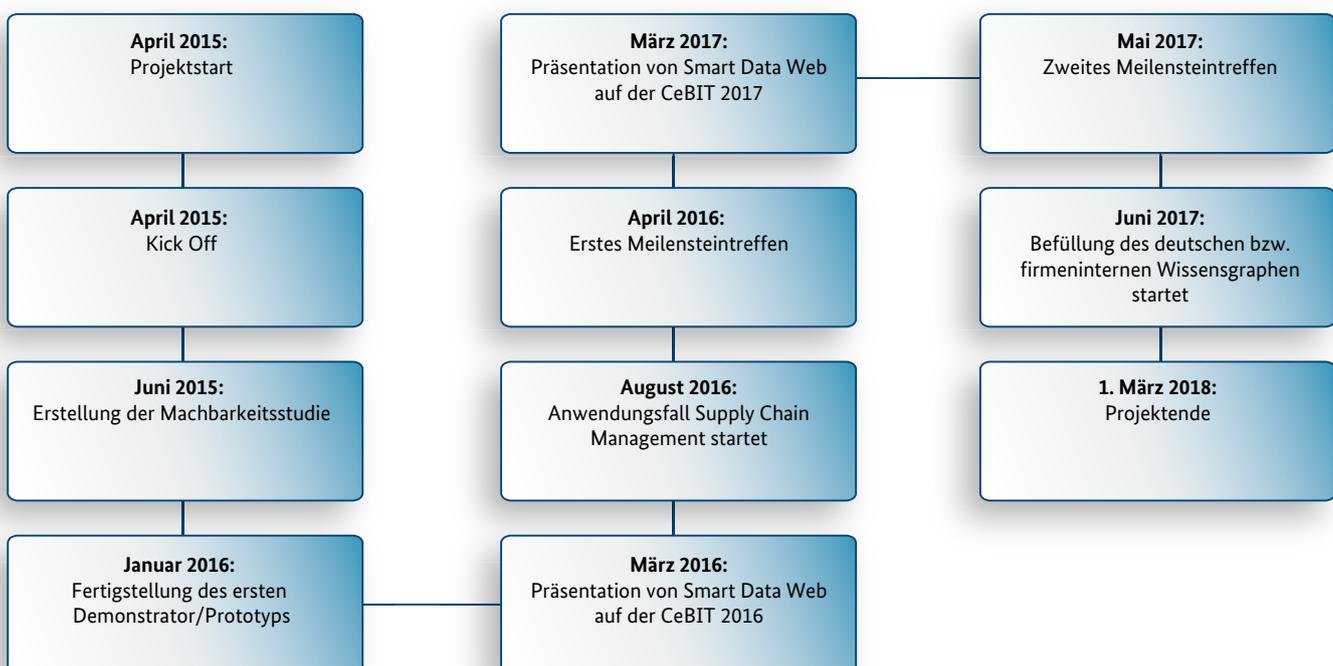
Ob Naturkatastrophen, Unruhen, Ausfälle von Zulieferern oder die Entwicklung neuer Produkte und Komponenten – es gibt eine Vielzahl von externen Ereignissen, die einen entscheidenden Einfluss auf die Wertschöpfungs- und Lieferketten eines Unternehmens haben. Informationen über diese Ereignisse sind auf Nachrichtenwebseiten oder in sozialen Medien zwar meist unmittelbar verfügbar, erreichen die Entscheider in den Unternehmen aber in der Regel spät und eher zufällig. Störungen sind dann oft nicht mehr zu vermeiden und führen zu enormen Kosten, die im Extremfall gerade für kleine und mittelständische Unternehmen existenzbedrohende Auswirkungen haben können.

Das Projekt Smart Data Web entwickelt ein smartes Wissensnetz, das Unternehmen schon früh die für sie relevanten Informationen zur Verfügung stellt. Ziel ist es, eine Brücke zwischen zwei bisher voneinander getrennten Datenwelten zu schlagen: dem öffentlich zugänglichen Internet und unternehmensinternen Informationsbeständen. Konkret werden bei Smart Data Web aus einer riesigen Menge an öffentlichen Datenströmen die für ein Unternehmen relevanten Daten – etwa über Kunden, Zulieferer, Transportwege oder Dienstleister – automatisch herausgefiltert, analysiert und aufbereitet.

All diese Informationen werden im nächsten Schritt intelligent mit den unternehmensinternen Daten verknüpft, sodass etwa Führungskräfte individuell auf ihr Themengebiet zugeschnittene Hintergrundinformationen erhalten. Diese dienen als Grundlage für bessere Entscheidungen: Ein global agierendes Technologieunternehmen erfährt beispielsweise mit Hilfe des Wissensnetzes schon früh von einem geplanten Streik bei einem seiner Zulieferer in Asien. Durch den zeitlichen Wissensvorsprung kann das Unternehmen dieser Situation proaktiv begegnen – etwa durch das Ausweichen auf einen anderen Anbieter – und so eine Unterbrechung der Produktion durch fehlende Bauteile verhindern.

Ein weiteres Einsatzszenario von Smart Data Web ist die Marktforschung einschließlich der Beobachtung von Kunden- und Anbietermärkten. Durch unternehmens- oder marktsegmentspezifische Analysen soll es Unternehmen künftig ermöglicht werden, sich schnell einen Überblick über die Wettbewerbssituation zu verschaffen, mögliche neue Kunden zu identifizieren oder günstigere und innovativere Lieferanten für eigene Produkte zu finden.

## Meilensteine im Projektverlauf



## Weg in die Praxis

Innerhalb des Projekts Smart Data Web wurden bereits mehrere Wege zur Überführung der Projektergebnisse in die Praxis identifiziert oder schon beschritten. Eine Schlüsselrolle spielen dabei die einzelnen Konsortialpartner. So führte der Konsortialpartner Beuth Hochschule für Technik Berlin mehrere Roadshows in ganz Deutschland durch, um möglichst viele Unternehmen von den Vorteilen der Smart-Data-Web-Technologien zu überzeugen.

Zudem haben die Konsortialpartner uberMetrics Technologies GmbH, VICO Research & Consulting GmbH und die Siemens AG die im Projekt entwickelten Technologien und Algorithmen bereits in bestehende Produkte und Technologien integriert, wie beispielsweise die entwickelten Methoden zur automatisierten Gewinnung von Informationen aus Texten oder der gezielten Beobachtung von Suchbegriffen im Internet. Darüber hinaus wurden erste Prototypen für das Lieferkettenmanagement und die Marktforschung entwickelt und von den Konsortialpartnern erfolgreich erprobt.

Anwendung im wissenschaftlichen Kontext finden die Technologien – etwa zur Texterkennung in Datenbanken oder Lernmethoden für Maschinen (Deep-Learning) – unter anderem in Projekt MACSS, das innerhalb des Technologieprogramms Smart Service Welt des BMWi gefördert wird. Dort soll die Kommunikation zwischen Ärztinnen und Ärzten, Patientinnen und Patienten und medizinischem Personal mit einer App verbessert werden. Diese Technologien finden auch im EU-Projekt H2020 Fashionbrain Verwendung, das an der Optimierung der Wertschöpfungskette in der Modebranche arbeitet. Zu den Anwendungsgebieten der Deep Learning-Methoden zählt darüber hinaus ein 2018 startendes Projekt zur automatisierten Erkennung von Hasssprache im Netz, das vom BMBF gefördert wird.

Als weiterer Baustein der Wissenstransformation in die Praxis wurde Mitte 2017 an der Beuth Hochschule für Technik Berlin das Start-up Qualification gegründet, das auf Smart-Data-Web-Technologien zur maschinellen Verarbeitung natürlicher Sprache (Natural Language Processing) basiert. Qualification will Unternehmen im Verlagswesen ermöglichen, automatisch Bestseller zu identifizieren.

## SDW Knowledge Graph



## Die Projekt-Ergebnisse auf einen Blick

Früher	Mit Smart Data Web
Unternehmensrelevante Nachrichten mit zeitkritischen Ereignissen (Streiks, drohende Insolvenz) können nur durch manuelle Recherche entdeckt werden.	Automatisches Web Monitoring filtert und visualisiert relevante Ereignisse für Entscheider.
Unternehmensrelevante Informationen (wie Umsatz, Standorte, etc.) sind über verschiedene, teils kommerzielle Quellen verteilt, die zudem in unterschiedlichen, oft nicht maschinenlesbaren Formaten vorliegen.	Datenquellen werden automatisch miteinander verknüpft und in einer standardisierten Darstellung über einen Wissensgraphen gespeichert und verfügbar gemacht.
Firmeninternes Wissen über Lieferanten ist in verschiedenen, heterogenen Datenbanken und Dokumentenablagensystemen gespeichert.	Firmeninternes Wissen kann geschützt und automatisiert mit externem Wissen über den SDW Knowledge Graph verknüpft, und abteilungsübergreifend nutzbar gemacht werden.
Sprachtechnologische Methoden zur Erkennung von industriebezogenen Fakten und Ereignissen in Webdaten sind nicht robust genug.	Sprachtechnologische Methoden sind weiterentwickelt und angepasst, sodass sie für das automatisierte Monitoring von industriebezogenen Fakten und Ereignissen eingesetzt werden können.
Es existiert kein offenes, allgemeines Wissensnetz für industrierelevante Daten.	Entwurf und prototypische Realisierung eines offenen, allgemeinen Wissensnetzes, das die Interessen industrieller Nutzer berücksichtigt.
Es existieren keine (frei verfügbaren) Big-Data-Analytik-Methoden für die automatische Textauswertung auf großen Mengen von Nachrichten.	Im Projekt werden Algorithmen zur Analyse von Texten für den Einsatz im Big-Data-Umfeld angepasst. INDREX, eine Methode für die Textanalyse innerhalb von Datenbanken, wird zu diesem Zweck mit den führenden Datenbankherstellern weiterentwickelt.

# Fachgruppe Wirtschaftliche Potenziale und gesellschaftliche Akzeptanz

In der Fachgruppe werden gemeinsam mit Projektvertreterinnen und -vertretern sowie externen Expertinnen und Experten themenübergreifende Fragen der wirtschaftlichen und gesellschaftlichen Verwertung von Daten diskutiert. Ziel dabei ist es, gemeinsame Positionen zu formulieren und so den politischen Diskurs über das Thema Smart Data mit besonderem Fokus auf wirtschaftliche Verwertungsfragen und gesellschaftliche Problemstellungen mitzugestalten. Zu den Themen der Fachgruppe zählen unter anderem Verwertungsstrategien, Geschäfts- und Abrechnungsmodelle, Markteintrittsstrategien sowie Aspekte und Chancen der Datenökonomie. Die Expertinnen und Experten der Fachgruppe stehen in engem Austausch mit den Smart-Data-Projekten, etwa durch regelmäßige Workshops und Vernetzungsveranstaltungen. Darüber hinaus unterstützt die Fachgruppe die Projektkonsortien durch umfangreiches, projektindividuelles Feedback.

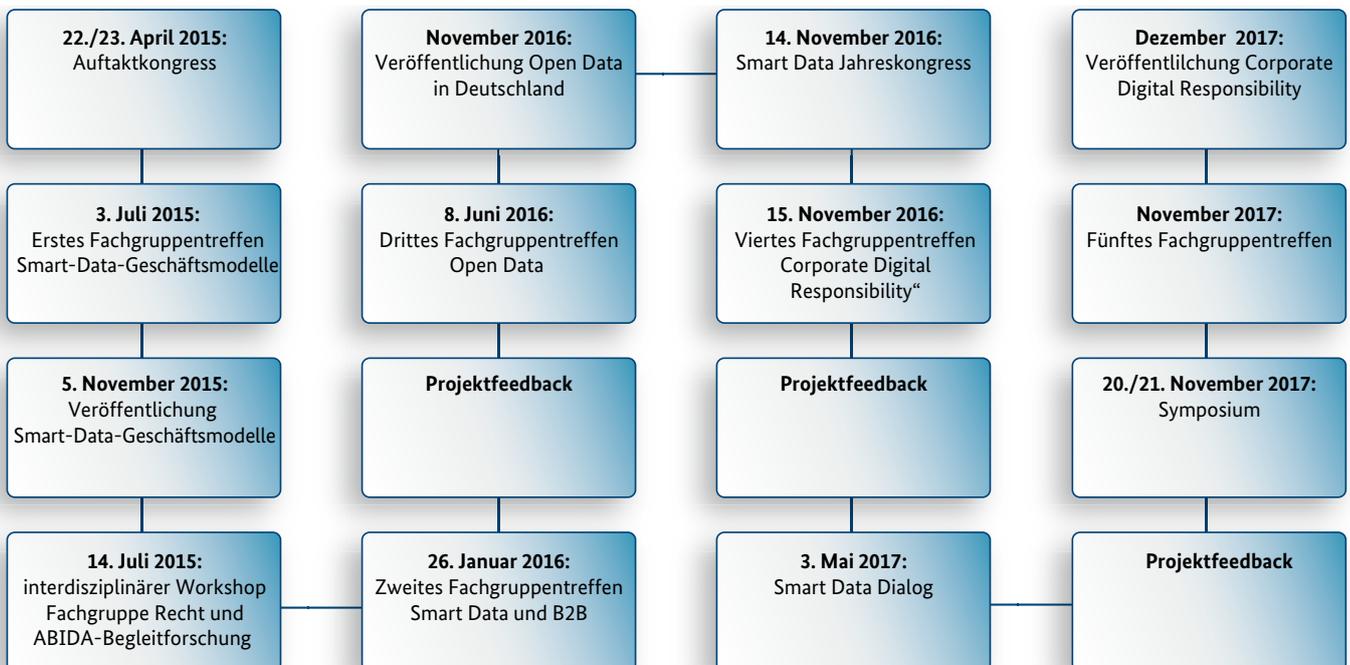
In der Arbeit der Fachgruppe zeigt sich der starke Zusammenhang zwischen der gesellschaftlichen Akzeptanz und

dem wirtschaftlichen Potenzial. Denn auch im Fall von Smart Data gilt: Ohne eine entsprechende Zustimmung zu mehrwertschaffenden Innovationen, kann sich deren wirtschaftliches Potenzial nicht oder nur beschränkt entfalten.

## Geschäftsmodelle im Kontext von Smart Data

Mit regelmäßigen Fachgruppentreffen fördert die Fachgruppe den projektübergreifenden Wissens- und Erfahrungstransfer und setzt gezielte Akzente zu aktuellen Themen und Entwicklungen. Das erste dieser Treffen fand im Juli 2015 zum Thema „Smart-Data-Geschäftsmodelle“ statt. Auf Basis der Ergebnisse entstand im November 2015 das Positionspapier „Smart Data Geschäftsmodelle“, das sich mit der Frage beschäftigt, wie Smart Data wirtschaftlichen Mehrwert schaffen kann. In der Publikation wurden sowohl verschiedene Ertragsmodelle vorgestellt, aber auch auf Hemmnisse und Herausforderungen im Umgang mit Smart Data eingegangen.

### Meilensteine im Projektverlauf



## Austausch und Vernetzung – Die Workshops der Fachgruppe

Nach der Gruppenbildungsphase des ersten Fachgruppentreffens öffnete die Fachgruppe ihre Treffen zunehmend für externe Experten. So fand im Juli 2015 ein gemeinsames Arbeitstreffen mit der Fachgruppe Rechtsrahmen und den Konsortialpartnern der Begleitforschung des vom BMBF geförderten Programms ABIDA statt. In diesem Zuge wurde ein kontinuierlicher Austausch zwischen den Begleitforschungen beschlossen, um Erkenntnisse zu teilen und Synergien zu nutzen.

Schwerpunktthema des zweiten Fachgruppentreffens war „Smart Data und B2B“. Dabei moderierten Vertreterinnen und Vertreter der Smart Data Begleitforschung einen Workshop für Projektteilnehmer und externe Expertinnen und Experten. In drei Arbeitsgruppen wurde zudem die Unterschiede von Smart-Data-Geschäfts- und Anreizmodellen in den Bereichen Business-to-Business (B2B) und Business-to-Consumer (B2C) erarbeitet.

Das dritte Treffen der Fachgruppe widmete sich dem Thema „Open Data“. Aufbauend auf drei Impulsvorträgen

von Vertretern aus der Wirtschaft erarbeiteten und diskutierten die Teilnehmerinnen und Teilnehmer bestehende Herausforderungen und Konflikte rund um die Verwendung von öffentlichen Daten. Die im Zuge dessen entwickelten Thesen bildeten die Grundlage für die Fachgruppenpublikation „Open Data in Deutschland“, die im Herbst 2016 veröffentlicht wurde. Sie zeigt den Handlungsbedarf im Open-Data-Umfeld auf und formuliert Forderungen an Verantwortliche in öffentlichen Institutionen und Unternehmen, um die wirtschaftlichen Potenziale von Open Data auch in Deutschland freizusetzen.

Unter der Überschrift „Corporate Digital Responsibility“ fand das vierte Treffen der Fachgruppe statt. Dieses Treffen wurde in Kooperation mit der Initiative D21 durchgeführt. Die Frage nach der digitalen Verantwortung ist ein sehr breit gefasstes Schnittstellenthema, das alle Smart Data-Projekte und Anwendungsbereiche einfängt. Derzeit ist ein Whitepaper der Fachgruppe zu „Corporate Digital Responsibility“ in Arbeit, das voraussichtlich im Herbst 2017 veröffentlicht werden wird.

# SmartRegio – Trend-Analysen auf Basis heterogener Massendaten

## Lokale Trends frühzeitig erkennen und nutzen

### Das Projekt

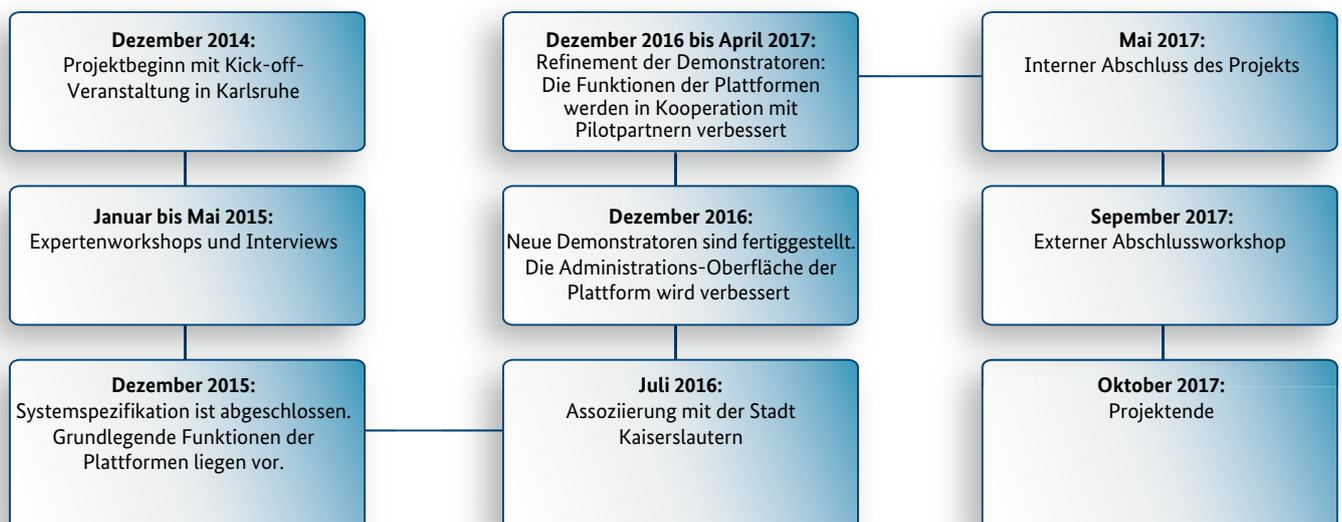
Öffentliche Einrichtungen sowie kleine und mittlere Unternehmen sind zumeist regional stark verwurzelt. Die Kenntnis von lokalen Entwicklungen sind deshalb für sie ein wichtiger Erfolgsfaktor: Neben sozioökonomischen und demografischen Trends, wie Migration oder die Energiewende, sind auch Haltungen und Meinungen zu technologischen und gesellschaftlichen Entwicklungen, lokalen Ereignissen und zum Produkt- und Leistungsangebot der Einrichtungen und Unternehmen von großer Bedeutung. Nur so können Firmen die Erwartungen erfüllen, dass lokale Unternehmen gezielter auf Kundenwünsche eingehen. Bislang beruht das Erkennen lokaler Veränderungen vor allem auf menschlicher Erfahrung und Beobachtung – und ist somit nicht objektiv. Eine Alternative stellen methodische Datenanalysen dar. Diese basieren zumeist auf teuren, speziell erhobenen Informationen, weshalb sie oftmals gerade für kleine Akteure unerschwinglich sind. Durch die Digitalisierung in Unternehmen und Einrichtungen sowie durch soziale Medien entstehen jedoch immer mehr Daten automatisch. Diese bereits vorhandenen Daten können künftig eine wichtige Informationsquelle im Hinblick auf lokale Entwicklungen darstellen.

SmartRegio entwickelte daher eine Lösung, um lokale Massendaten zu sammeln und aus diesen Informationen regionale Trends abzuleiten. Das Projekt setzte auf ständig entstehende Informationen aus unterschiedlichen Quellen wie Suchanfragen, Beiträge in sozialen Netzwerken oder vorliegende Statistiken, beispielsweise von Städten, Gemeinden oder Infrastrukturanbietern. Auch Daten der Kunden werden genutzt. Die Datengrundlage ist somit sehr breit angelegt. Folglich muss die Plattform von SmartRegio mit sehr heterogenen Formaten und Zugriffsmechanismen auf Daten funktionieren. Die modulare Plattform von SmartRegio kann die Informationen flexibel und in Echtzeit auswerten, räumlich-zeitliche Veränderungen aufdecken und die aufbereiteten Ergebnisse leicht verfügbar machen.

### Weg in die Praxis

Getestet wurde die Plattform von den Pilotanwendern, der Stadt Kaiserslautern und ihren Stadtwerken. Die Praxispartner sind am Einsatz und an der Weiterentwicklung der entwickelten Lösungen interessiert, da sich der bisherige Einsatz als sinnvoll erwies. So stellten die Stadtwerke bei-

### Meilensteine im Projektverlauf





## Die Projekt-Ergebnisse auf einen Blick

Früher	Mit SmartRegio
Die Stadt Kaiserslautern und deren Stadtwerke haben gemäß ihren Verpflichtungen operative Daten gesammelt, aber kaum genutzt.	Daten werden effektiv genutzt und können einfacher und schneller für strategische Planung, zum Beispiel neuer Infrastrukturen herangezogen werden.
Die Stadt Kaiserslautern hat Daten manuell zu statistischen Jahresberichten zusammengetragen.	Daten können weitgehend automatisiert und im Hinblick auf unterschiedliche Aufgaben ausgewertet werden, beispielsweise für Jahresberichte.
Die Stadtwerke Kaiserslautern konnten die Effektivität von Maßnahmen, wie Investitionen in die Infrastruktur oder in den Bildungsbereich, nur schwer oder gar nicht nachvollziehen.	Die Stadtwerke Kaiserslautern können mit SmartRegio auf Werkzeuge zurückgreifen, die automatisiert den Effekt von ihren Maßnahmen messen, indem relevante Daten zusammengeführt und analysiert werden.

# SmartEnergyHub – Datendrehscheibe für intelligente Energienutzung

## Das Projekt

Deutschlands Energiesystem befindet sich im Wandel: Während früher wenige Kraftwerke den Energiebedarf deckten, verschiebt sich das System hin zu einer dezentralen Energieversorgung, die zunehmend auch immer größere Mengen an erneuerbaren Energien integriert. Die Folgen sind unter anderem schwankende Energiepreise und ein steigender Bedarf an Speicherkapazitäten.

Wichtige und teilweise kritische Infrastrukturen wie Flughäfen, Seehäfen, Industrie- und Chemieparks, Fabriken und Fertigungsanlagen oder Bürogebäude sind bedeutende Verbraucher und gleichzeitig Erzeuger von Energie. Darüber hinaus besitzen sie erhebliche Speicherkapazitäten und verfügen über nennenswerte Potenziale zur flexiblen Anpassung des eigenen Energiebedarfs (Lastverschiebung). Die Betreiber dieser Infrastrukturen sind vor dem Hintergrund des sich wandelnden Energiemarktes bestrebt, ihre Angebote weiterhin zuverlässig, stabil und kostengünstig zu erbringen.

Das Projekt SmartEnergyHub möchte diese Akteure bei ihren Bemühungen unterstützen, das Energiemanagement ihrer Infrastrukturen weiter zu optimieren und damit unmittelbar zum Gelingen der Energiewende beitragen. Pilotanwender ist der Flughafen Stuttgart, der zur Vorzeige-

Infrastrukturzelle umgebaut werden soll. Dabei wird ein plattformbasierter Ansatz verfolgt, bei dem die Daten unterschiedlicher Sensoren im ersten Schritt miteinander verknüpft und anschließend mit weiteren Daten, zum Beispiel dem Fluggastaufkommen oder Wetterprognosen, intelligent zusammengeführt werden. Auf diese Weise sollen Infrastrukturbetreiber Synergieeffekte identifizieren und den gesamten Energiefluss optimieren können.

Neu an SmartEnergyHub ist vor allem die ganzheitliche medienübergreifende Betrachtung und Überwachung des Energiesystems mit einem Referenz-Energie-System (RES) unter Einbeziehung von Bedarfs-, Preis- und Wetterprognosen sowie weiterer externer Datenquellen. Eine intelligente Nutzung der Daten soll so eine Optimierung des Energiesystems in Echtzeit ermöglichen. Neu ist darüber hinaus die Berücksichtigung von Flexibilitäten auf der Bedarfsseite. Durch präzise Bedarfsprognosen werden zusätzliche Lastverschiebepotenziale erschlossen und für die Optimierung genutzt.

Am Ende des Projekts entsteht ein IT-Toolkit für Infrastrukturbetreiber, welches in der Lage ist, die bestehenden Systeme und bereits erfasste Daten des Infrastrukturbetreibers zu integrieren.

## Meilensteine im Projektverlauf



## Weg in die Praxis

Um eine spätere praktische Verwertung der Ergebnisse des Projekts SmartEnergyHub sicherzustellen, hat das Projekt einen Anwenderkreis aus Infrastrukturbetreibern gegründet. Das Ziel des Anwenderkreises ist es zum einen, weitere interessierte Anwender wie Flughäfen, kleine und mittelständische Industrieunternehmen oder auch Stadtwerke und Netzbetreiber auf das Projekt aufmerksam zu machen und ein Netzwerk für die weitere Verwertung aufzubauen.

Zum anderen werden so frühzeitig Anforderungen aus dem erweiterten Anwenderkreis an SmartEnergyHub aufgenommen und bei der Entwicklung berücksichtigt. Insgesamt werden während der Projektlaufzeit sechs Anwenderkreise bei großen Infrastrukturbetreibern wie Flughafen Stuttgart, ZF Friedrichshafen, Daimler AG oder Audi AG durchgeführt.

Seit etwa Anfang 2017 werden die im Projekt entwickelten Kernbausteine bereits praktisch umgesetzt. Hierfür wurde ein eigenes Workshop-Format für Stadtwerke und Industrieunternehmen entwickelt. Die Anwenderinnen und Anwender werden dabei mit der SmartEnergyHub-Metho-

dik vertraut gemacht, so dass sie diese im nächsten Schritt für sie relevanten Elemente in ihren Unternehmen anpassen und einsetzen können.

Ein Beispiel hierfür ist die RES-Methode zur visuellen Darstellung und Überwachung komplexer Energieinfrastrukturen. Die Überwachung und Steuerung erfolgt dabei durch eine Web-Oberfläche. Bereits diese Methode hat bei vielen Anwendern vor allem Erkenntnisgewinne in Hinblick auf die Kontrolle hochkomplexer Infrastruktursystemen und deren Optimierung in Echtzeit gebracht und zu signifikanten Einsparungen geführt.

Mit der Praxisumsetzung sind denkbare Schritte für die weitere Verwertung der Projektergebnisse skizziert: Neben dem Anwenderkreis, der in ähnlicher Form fortgesetzt werden soll, hat sich das Workshop Format „Zukunft des Energiemanagements“ und daraus resultierende sehr konkrete Verwertungsansätze für einzelne Funktionsmodelle von SmartEnergyHub als vielversprechendes Vorgehen herauskristallisiert.



## Die Projekt-Ergebnisse auf einen Blick

Früher	Mit SmartEnergyHub
<p>Unterschiedliche Datentypen (zum Beispiel aus Leitsystemen oder von externen Anbietern) werden nicht zentral gehalten, nicht in Relation zueinander gesehen, nicht auf ihre Qualität geprüft und es besteht keine Möglichkeit eines gesteuerten Ablaufs (automatisierter Workflow).</p>	<p>Zentrale Datenhaltung ermöglicht es, unterschiedliche Datentypen zusammenzuführen und aufeinander abzustimmen, zu überwachen und auf Qualität zu prüfen, durch eine Prozesssteuerung für weitergehende Prozesse wie Prognose- und Optimierungsrechnungen automatisiert zu verwenden und für Visualisierungszwecke bereitzustellen.</p>
<p>Die Wartung von kritischen Anlagen einer Infrastruktur, beispielsweise BHKWs oder Netzersatzanlagen (NEA), könnte nur unzureichend in den laufenden Betrieb eingeplant werden.</p>	<p>Die zukünftige Fahrweise einer Anlage (Prognose in Form einer Zeitreihe) wird regelmäßig aktualisiert und im System angezeigt. Durch die Prognosen sind Wartungen unter Berücksichtigung des Gesamtsystems planbar.</p>
<p>Auswertungen erfolgen häufig manuell im Nachhinein auf Basis von historischen Daten in EXCEL oder vergleichbaren Werkzeugen. Fehlentwicklungen oder Zusammenhänge werden entweder nicht oder zu spät erkannt und eine aktive Beeinflussung ist nur unzureichend möglich.</p>	<p>Der Betreiber und Entscheider gewinnt über die SmartEnergyHub-Anwendung und RES-Methode mit leistungsfähiger und flexibel einsetzbarer Visualisierung von Realwerten, Vergangenheitswerten und Prognosen einen konsistenten Gesamtüberblick über das Energiesystem.</p>
<p>Die Bereitstellung von Informationen aus dem Realbetrieb für Entscheider sowie die Steuerung von Anlagen aufgrund von Prognose-/ Optimierungsergebnissen ist nicht möglich.</p>	<p>Die Steuerung von Anlagen ist auf Basis von Prognose-/ Optimierungsergebnissen und Live-Informationen aus dem Realbetrieb während des Betriebs möglich. Zukünftig sind automatisierte Abläufe und Schalthandlungen denkbar.</p>
<p>Der Aufbau und die Bereitstellung von IT-Systemen, die vorgenannte Funktionen unterstützen, war bisher nicht, oder nur in Verbindung mit vergleichsweise hohen Investitionskosten möglich.</p>	<p>Standardisierte Lösungen können als Software-as-a-Service (SaaS) mandantenfähig aus der Cloud kostengünstig bei geringem Risiko bereitgestellt werden.</p>

# EDV – Einfaches Digitales Vergessen

## Selbstbestimmter Austausch von sensiblen Daten

### Das Projekt

In Unternehmen erfolgt der Austausch von sensiblen Informationen wie Projektergebnissen, Informationen zur Geschäftsstrategie oder auch Bewerbungsunterlagen zunehmend per E-Mail oder über externe Speicherbereiche wie Cloud-Lösungen. Allerdings sollen diese Informationen häufig nur temporär oder in einem bestimmten Kontext verfügbar sein, um die Gefahr des Missbrauchs oder der unabsichtlichen Weitergabe der Daten im Unternehmen zu verringern.

Das Projekt EDV entwickelt deshalb ein aus Soft- und Hardwarekomponenten bestehendes System, das einen selbstbestimmten Austausch sensibler Informationen sowie ein automatisches Löschen von Daten ermöglicht. Die eingesetzte Software soll einen sicheren Datenaustausch gewährleisten. So können durch eine Ende-zu-Ende-Verschlüsselung die übertragenen Daten auf Senderseite verschlüsselt und beim Empfänger wieder entschlüsselt werden.

Mit der Software-Lösung des Projekts EDV behält der Sender auch nach dem Versenden die Kontrolle über seine Daten. Auf diese Weise können Nutzungsrechte beispielsweise auch nachträglich geändert oder entzogen sowie Daten nach Ablauf einer bestimmten Frist gelöscht werden. Anwendungsbeispiele sind etwa digitale Bewerbungsunterlagen oder Zugriffsrechte bei Projektarbeiten, die nach einer festgelegten Frist entzogen werden.

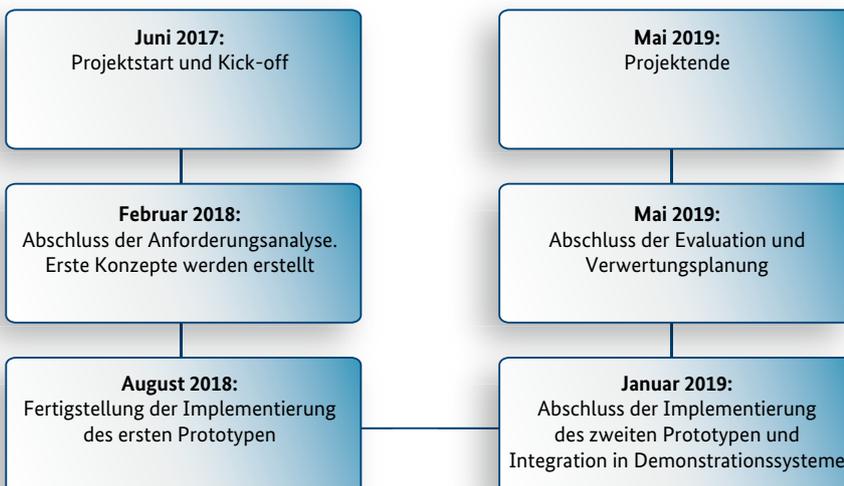
Die dadurch erhöhte Sicherheit kann die Bereitschaft steigern, sensible Dokumente und Informationen digital auszutauschen.

### Weg in die Praxis

Das Projekt EDV begann im Juni 2017. Im Projektkonsortium arbeiten die Forschungspartner des Forschungszentrums Informatik (FZI) am Karlsruher Institut für Technologie (KIT) und die Hochschule Pforzheim sowie von Wirtschaftsseite die CAS Software AG und die esentri AG zusammen. Von Beginn an hat das Projekt auf einen starken Praxisbezug großen Wert gelegt.

Bereits in der Analysephase des Projektes werden Anwendungsfälle skizziert und anschließend mit assoziierten Partnern ausgearbeitet. Nach der Entwicklung eines Prototypen für einen senderbestimmten Austausch von Daten wird der Lösungsansatz in einer Testphase evaluiert. Mögliche Kernanwendungsfälle bilden der Austausch von sensiblen Projekt- oder Produktinformationen zwischen Unternehmen sowie das Einbinden von EDV in einen digitalen Bewerbungsprozess zum sicheren und gesetzeskonformen Austausch von Unterlagen. Mit den assoziierten Partnern iPLON, Facility Care AG und dem Netzwerk Bundesverband IT-Mittelstand e. V. wird der Einsatz von EDV im deutschen Mittelstand geprüft. Mit den aus der Einsatzprüfung gewonnenen Erkenntnissen sollen im Anschluss Geschäfts- und Betreibermodelle für System- und Softwarelösungen entwickelt werden, die auf dem System von EDV basieren. Die Konsortialpartner des Projektes streben zudem einen Einsatz von auf EDV basierenden Lösungen in ihren eigenen Produkten an.

### Meilensteine im Projektverlauf



## Die Projekt-Ergebnisse auf einen Blick

Früher	Mit EDV
Sensible Daten stehen dem Empfänger auch nach Ablauf von gesetzlichen Fristen oder vertraglich geregelten Gültigkeitszeiträumen zur Verfügung.	EDV verhindert, dass nach dem Ablauf von Fristen beziehungsweise dem Überschreiten eines Zugriffsdatums der Empfänger die Daten einsehen und bearbeiten kann.
Der Sender von Daten muss bei dem Versand einen Kontrollverlust über die Informationen hinnehmen. Nach dem Verschicken beispielsweise von Dokumenten besteht keine Möglichkeit mehr, die Zugriffsrechte des Empfängers zu verändern, also zu entziehen oder zu erweitern.	Nach dem Versenden von Daten behält der Absender durch EDV die Möglichkeit, Zugriffsrechte des Empfängers zu verändern.
Der Empfänger von sensiblen Daten ist ohne Weiteres in der Lage, empfangene Daten (auch unabsichtlich) an Unbefugte weiterzuleiten.	Eine versehentliche oder unbeabsichtigte Weiterleitung von Daten an Unbefugte wird verhindert.
Sensible Daten werden oft über unverschlüsselte Kanäle oder ungesicherte Systeme ausgetauscht, sodass ein Abfangen der Informationen von Unbefugten relativ einfach möglich ist.	EDV bietet den Nutzern eine Ende-zu-Ende Verschlüsselung, so dass Daten in lesbarer Form nur beim Absender und beim berechtigten Empfänger vorliegen.
Bei vielen Datenaustauschverfahren (zum Beispiel Emails ohne digitale Signatur) ist es nicht möglich, die Identität des Absenders zu überprüfen.	EDV stellt sicher, dass der Empfänger der Daten die Authentizität des Absenders überprüfen kann.
Viele Sicherheitsverfahren beim Datenaustausch erfordern ein gewisses technisches Verständnis oder bedeuten einen Zusatzaufwand bei der Verwendung.	EDV will den sicheren Datenaustausch durch eine einfach zu benutzende Software und unkomplizierte Integrationsmöglichkeiten in bestehende Datenaustauschprozesse fördern.
Viele gängige Datenaustauschverfahren sind nicht in bestehende IT-Systeme integrierbar.	Das System wird durch offene Schnittstellen und ggf. durch Plugin-Konzepte die Integration in bestehende Softwaresysteme und Prozesse ermöglichen.

