

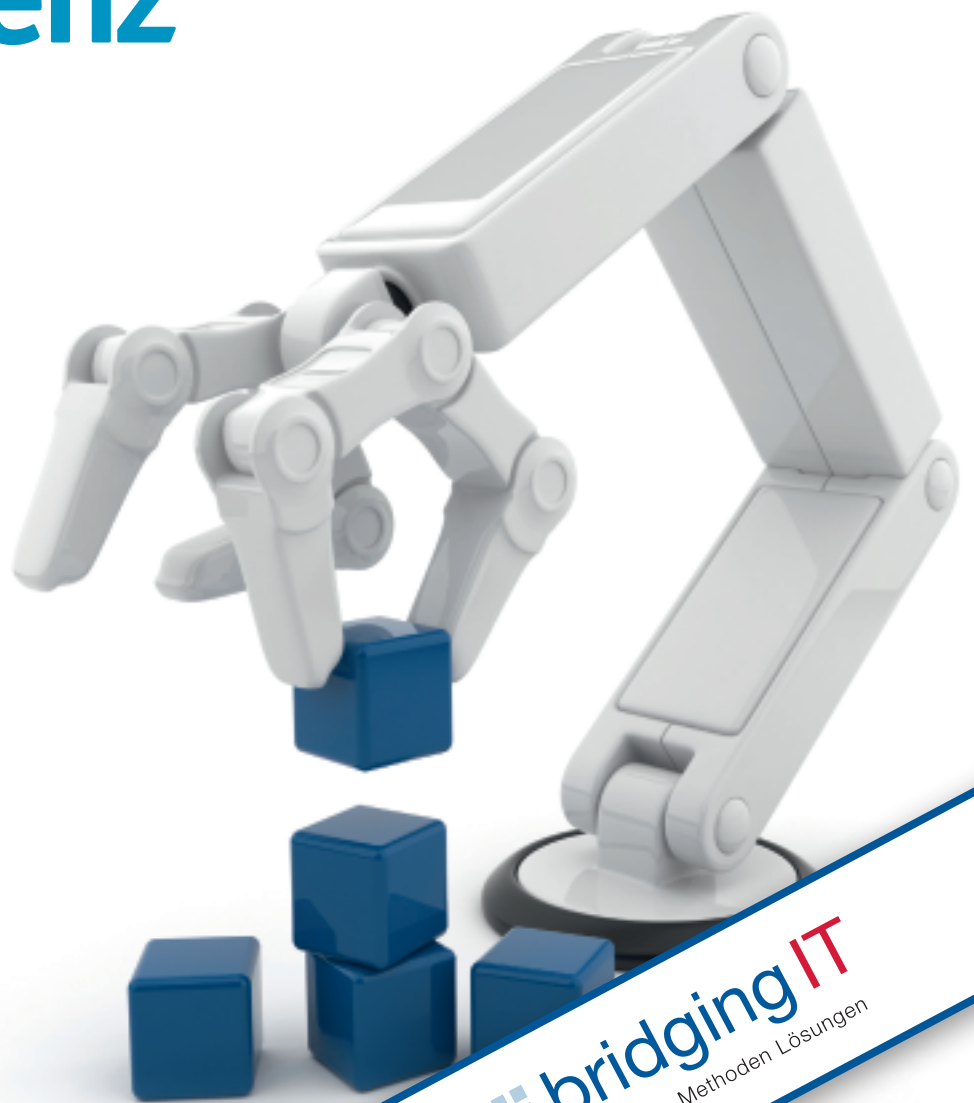
# JavaSPEKTRUM

Magazin für professionelle Entwicklung und Integration von Enterprise-Systemen

## Lernen und Künstliche Intelligenz

Digitale Vorahnung

Predictive Car Dispatching für eine autonome Fahrzeugflotte



### Interview

Interview mit Professorin Judith Simon über Ethik, Gerechtigkeit und Verantwortung in der IT

**bridging IT**  
Menschen Methoden Lösungen

**Sonderdruck für**

RESTful API  
Lambda-Ausdrücke mit Kotlin und Java

Lernen und Künstliche Intelligenz

G30325

## Digitale Vorahnung

# Predictive Car Dispatching für eine autonome Fahrzeugflotte

Kevin Berwind, Dirk Braun, Johannes Johannsen

Die Zukunft der individuellen Fortbewegung wird maßgeblich durch autonome Fahrzeuge beeinflusst werden. Seit Kurzem fährt beispielsweise der erste autonome Bus auf Deutschlands Straßen und Pkw werden bereits getestet und für den autonomen Einsatz fit gemacht. Unabhängig von der Technik wird es aber wichtig sein, dass seitens der Politik die entsprechenden Rahmenbedingungen insbesondere in Bezug auf Haftung und Eingriffsmöglichkeiten der Fahrgäste geschaffen werden, um die Mobilität der Zukunft voranzubringen. Dabei ist aktuell das große Ziel vieler Beteiligten, voll-autonome Fahrzeuge auf die Straßen zu bringen. Auf dem Weg dorthin gilt es aber, noch einige rechtliche und ethische Hürden zu überwinden und die Akzeptanz von alleine fahrenden Fahrzeugen bei den Nutzern zu schaffen und Ängste abzubauen.

Ist das passiert, werden voll-autonome Fahrzeuge dann in der Regel nicht mehr im Besitz des Einzelnen sein, sondern als eine durch künstliche Intelligenz gesteuerte Fahrzeugflotte auf unseren Straßen unterwegs sein und vor allem hinsichtlich der Sicherheit im Straßenverkehr viele Verbesserungen mit sich bringen, wie beispielsweise die Unfallzahlen signifikant zu reduzieren.

Eine autonome Car-Sharing-Flotte, die auf Basis einer künstlichen Intelligenz weiß, wann und wo welcher Bedarf herrschen wird, wird für viele Verkehrsteilnehmer die Notwendigkeit für einen eigenen Wagen obsolet machen. Die typischen Anwendungsfälle für die Nutzung solcher Flotten sind die normalen Pendlerfahrten. Diese lassen sich aufgrund ihrer Regelmäßigkeit vor allem bei Zeit und Strecke schon mit eher geringem Aufwand sehr gut planen, wodurch auch die Bedarfsprognose keine zu große Herausforderung dar-

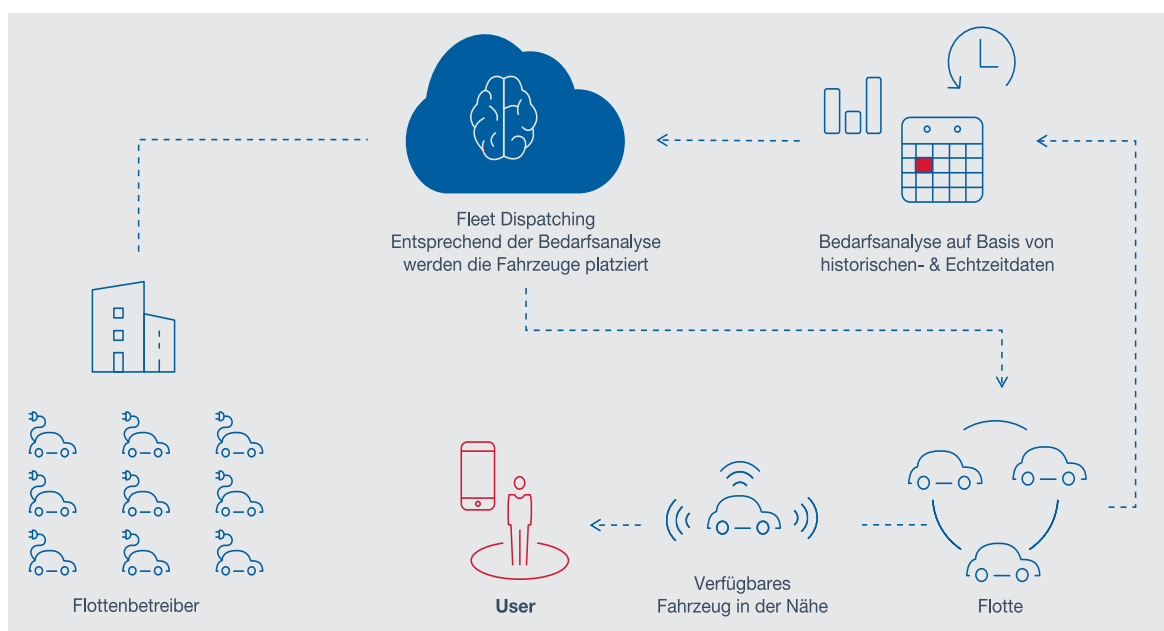
stellt. Abgeleitet aus den normalen Anwendungsfällen von heute werden die Fahrzeuge vornehmlich am Morgen aus den Wohngebieten zu den Unternehmen unterwegs sein, im Laufe des Tages die Fahrgäste bei den allgemeinen Erledigungsfahrten (Arztbesuche, Einkaufsfahrten, Amtstermin usw.) unterstützen, um dann am späten Nachmittag oder frühen Abend wieder in den Industriegebieten beziehungsweise bei den Firmen vor Ort zu sein, um die Arbeiter und Angestellten von dort nach Hause zu bringen. Am späteren Abend werden dann die üblichen Fahrten in und aus dem Stadtgebiet erfolgen, um die persönlichen Freizeitaktivitäten zu ermöglichen.

In diesen Fällen bestellt der Nutzer das von ihm gewünschte Fortbewegungsmittel in der für ihn geeigneten Qualität und Ausstattung, Nutzungshäufigkeit oder Dauer. Dieser Regelbetrieb lässt sich sehr gut anhand der jeweiligen Benutzerprofile prognostizieren. Dieses notwendige Fahrzeugangebot kann sehr gut skaliert werden, um zum richtigen Zeitpunkt die passende Anzahl von Fahrzeugen im Umfeld der Nutzer parat zu haben.

Aber was passiert bei unvorhergesehenen Ereignissen, also dann, wenn ein Nutzer einer anderen Mobilitätsvariante mit sehr geringer Vorlaufzeit auf eine verfügbare Alternative zum Beispiel zur Bahn oder zum Flugzeug angewiesen ist?

Wenn ein Zug aufgrund eines aktuellen Ereignisses mit großer Verspätung ankommt oder an einem Standort „strandet“, ist heute davon auszugehen, dass nicht für alle Mitfahrenden zeitnah eine alternative Reisemöglichkeit verfügbar ist, weil die Informationen zur Verspätung nur kurzfristig verfügbar sind und somit Fahrzeuge nicht oder nur nach entsprechender Umplanung verfügbar sein werden. Viele kennen sicher die Situation, an einem kleineren Bahnhof anzukommen, weil die Weiterfahrt nicht möglich ist, und

Abb. 1: Übersicht Car-Dispatching – Nutzerzentrierte Bedarfsermittlung





**Kevin Berwind** ist seit 2016 als Business Intelligence Berater bei der BridgingIT GmbH tätig. Er ist zudem im Center of Excellence Analytics engagiert und kümmert sich hier vorrangig um Big-Data-Analyseprozesse und Infrastrukturen. Kevin Berwind ist seit 2014 Dozent an der Hochschule Ludwigshafen am Rhein für Wirtschaftsinformatik. Zusätzlich promoviert er am Lehrstuhl für Multimedia und Internetanwendungen mit dem Schwerpunkt auf Big-Data-Analyseprozesse und Infrastrukturen. Studiert hat er Wirtschaftsinformatik (Bachelor) und Information Management & Consulting (Master).



**Dirk Braun** ist seit 2014 als Senior Consultant Innovation, Technologie und Trends bei der BridgingIT GmbH tätig. In dieser Rolle hat er unter anderem als Projektleiter und Flottenmanager das Thema eMobility des unternehmenseigenen Fuhrparks vorangetrieben. Als Berater gibt er seine Erfahrungen aus dem Umfeld Elektromobilität, Mobilitätskonzepte, Ladeinfrastruktur und Energie bei Kundenprojekten und auf Vorträgen weiter.



**Johannes Johannsen** berät seit 2011 nationale und internationale Kunden auf ihrem Weg in eine digitale Zukunft. Seit 2017 ist er bei der BridgingIT GmbH als Consultant für Geschäfts- und Organisationsberatung tätig. Er hat einen Bachelor of Laws (LL.B.) mit Schwerpunkt Corporate Management.

E-Mail: [smartmobility@bridging-it.de](mailto:smartmobility@bridging-it.de)

dort entweder Wartezeiten in Kauf nehmen zu müssen oder nur eine kleine Anzahl von Taxen vorzufinden, weil natürlich niemand so kurzfristig reagieren konnte, um den nötigen Bedarf zeitnah erkannt zu haben. Dann kommt die Zeit, wo künstliche Intelligenz und Prediction die bedarfsgerechte Steuerung einer Flotte übernehmen müssen.

Der Flottenbetreiber der Zukunft, sei es beispielsweise die Deutsche Bahn oder ein anderer Mobilitätsdienstleister, muss in der Lage sein, seinen Kunden dann einen umfassenden Service zu bieten. Und das setzt voraus, dass die steuernde künstliche Intelligenz die autonom fahrenden Fahrzeuge auf Basis von historischen Daten und Echtzeitinformationen so disponiert, dass diese bereits kurz nach Eintreten der Ausnahmesituation auf dem Weg zum Nutzer sind und diesen darüber informieren.

## Besserer Austausch von Verkehrsinformationen

Wie jeder Autofahrer werden auch die autonomen Pkw für bestimmte Routen noch etwas dazulernen, was bisher einem Navigationssystem mit seinen nahezu „statischen“ Daten nicht bekannt war. Also eine beispielsweise kurzfristig neue Verkehrsführung, eine Baustelle oder ein Hindernis, das bisher niemandem bekannt war, wie ein Baum, der gerade umgefallen ist und nun die Straße versperrt. Der Unterschied in einer solchen Situation zwischen Mensch und autonomen Autos liegt darin, dass sobald das autonome Fahrzeug das Hindernis dank der verbauten Kameras und der Technik mit Videoanalyse und Bilderkennung identifiziert hat, es seine neuen Erkenntnisse sofort mit dem Rest der Flotte oder sogar dem Rest der Welt teilen kann. Im Falle von Verkehrshindernissen kann zudem ein Bild an das Straßenverkehrsamt oder die Autobahnmeisterei übermittelt werden.

Heute wird ein Fahrer jedoch vom Navi weiterhin und trotz aller Verbotsschilder stur durch die Baustelle geschickt, und manchmal

ist es wirklich nicht leicht, dem Navigationssystem zu erklären, dass man nun eine andere Route benötigt. Es besteht schon heute die Möglichkeit, im Navigationssystem oder in einer der vielen Navigations-Apps eine Straßensperrung zu markieren, und die Anbieter können aufgrund der Bewegungsprofile teilweise schon erkennen, dass bestimmte Strecken aktuell nicht befahren werden sollten.

Man könnte auch ein Foto machen, teilweise auch direkt innerhalb der Navi-App, und das Hindernis dokumentieren. Das erfordert aber, dass man selbst aktiv wird, während man ja eigentlich damit beschäftigt ist, zu seinem Ziel zu fahren. So dauert es teilweise mehrere Monate, bis ein dauerhaftes Hindernis oder eine Straßenänderung in der Routenführung berücksichtigt wird. Heute weiß das System eben noch nicht, warum die anderen Fahrzeuge eine andere Strecke wählen. Aus der Ferne lassen sich Straßenschilder und Gegenstände auf der Fahrbahn eben nicht erkennen.

Diese Lücke könnte von den autonomen Autos geschlossen werden. Ein Bild sagt eben immer noch mehr als tausend Worte. Je nach Gestaltung könnte dies zu einem Alleinstellungsmerkmal für die Flotte werden oder es können auch die restlichen Verkehrsteilnehmer davon profitieren.

Zum einen durch eine zusätzliche Informationsquelle, zum anderen, weil Schäden und Hindernisse schneller behoben werden.

## Die Herausforderungen des Betreibers beim autonomen Flottenmanagement

Das Ziel für den Betreiber wird auch in Zukunft vornehmlich die Kosteneffizienz sein. Das heißt, die Flotte muss sich so steuern, dass die Nutzung und der Betrieb bezogen auf das zu erreichende Qualitätsziel des Betreibers, also die Kundenzufriedenheit, mit einem möglichst effektiven Mitteleinsatz (gefahrte Strecken, Treibstoffe, Werkstattkosten, Fahrzeit usw.) erreicht werden. Dass dies bei autonom fahrenden Fahrzeugen nicht mit der Unterstützung eines menschlichen Dispatchers erreicht werden kann, ist sicher klar, da hier alleine die Zusammenführung von mehreren Datenquellen und die Ableitung der entsprechenden Schlüsse schnell an ihre Grenze stoßen werden.

Dazu ist es notwendig, dass das KI-System auf einen möglichst großen Datenpool an Bestandsdaten zurückgreifen kann, aber auch in der Lage ist, persönliche Präferenzen der Nutzer und insbesondere aktuelle Informationsquellen ständig zu überwachen und in die Ermittlung des potenziellen Bedarfs und in das Fahrzeugangebot einfließen zu lassen. Ein konkretes Beispiel sind die Wetterdaten, denn hieraus lassen sich direkt Einflüsse auf die Verkehrssituation ableiten, und damit können die Routen- und Zeitplanung entsprechend optimiert werden.

Der Betreiber der Flotte muss sich bewusst werden, dass sein Erfolg maßgeblich von der intelligenten Verfügbarkeit seiner Fahrzeuge abhängt. Denn nur, wenn die Fahrzeuge in ausreichender Anzahl dort sind, wo sie benötigt werden, wird die Kundenzufriedenheit (und die damit verbundene Marktakzeptanz) entsprechend hoch sein, um eine wirtschaftliche Nutzung der Fahrzeuge zu ermöglichen. Wenn es aber aus Kundensicht zu lange dauert, bis die

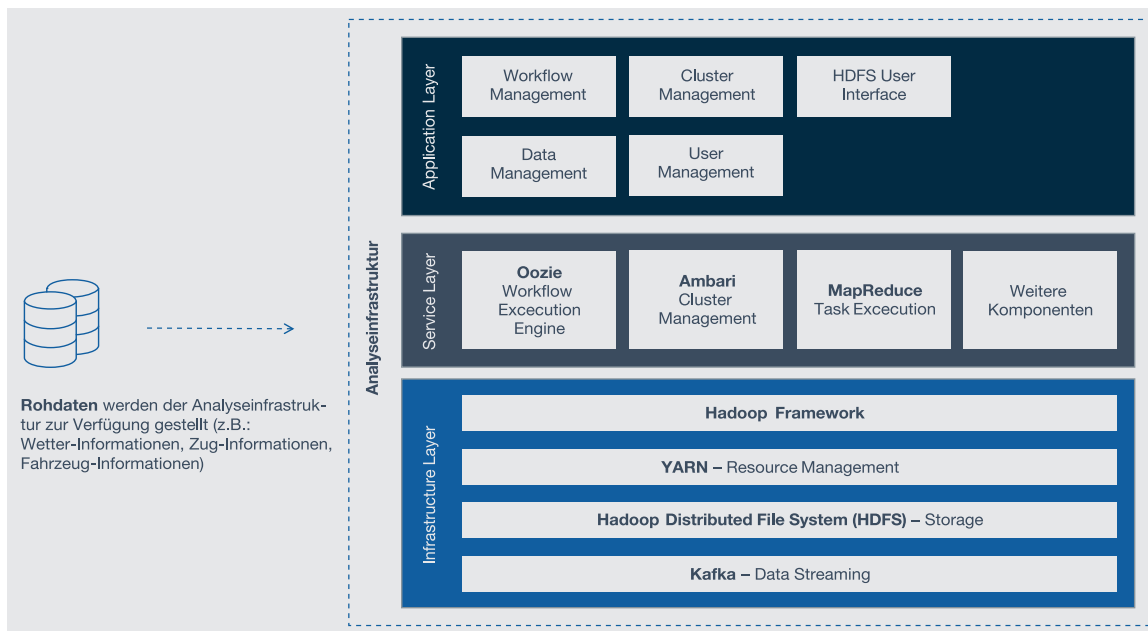


Abb. 2: „Predictive Car Dispatching“-Analyseinfrastruktur, angelehnt an Matthias Schneider [Ber17]

autonomen Fahrzeuge dort sind, wo sie benötigt werden, oder sogar der Fall eintritt, dass nicht genug Fahrzeuge verfügbar sind, werden in diesem Szenario sehr schnell Kunden abspringen und nach anderen Anbietern Ausschau halten.

## Softwarearchitektur für Predictive Car Dispatching

Die Voraussetzung für solch eine umfangreiche Analyse von Masendaten im Rahmen von Big Data wird erst durch eine geeignete (Analyse-) Softwarearchitektur ermöglicht. Hierzu soll im weiteren Verlauf des Artikels ein Architekturvorschlag auf Grundlage neuester wissenschaftlicher Erkenntnisse und bestehender Softwareprodukte der Apache Software Foundation (z. B. Hadoop, Kafka, Ambari) vorgestellt werden.

Die Rohdaten für das vorliegende Beispiel sind in externen, heterogenen Systemen abgelegt und werden über entsprechende (Web-)Services bereitgestellt. Die Infrastruktur nutzt hierbei beispielsweise aktuelle Wetterinformationen, Zug- oder Bahninformationen, Daten zur Verkehrslage und die aktuellen Fahrzeuginformationen (GPS-Daten, Reichweite, Fahrzeugstatus usw.), um über eine geeignete Analyse gezielt Fahrzeuge zu entsenden.

Die Architektur (s. Abb. 2) der „Predictive Car Dispatching“-Analyseinfrastruktur stützt sich auf die Architektur des Virtual Research Environments, die ein Analysesystem beziehungsweise eine Infrastruktur in den drei Schichten Application Layer, Service Layer und Infrastructure Layer organisiert.

Der *Infrastructure Layer* stellt die grundlegenden Funktionen der Infrastruktur zur Verfügung. Hier befinden sich Komponenten wie das Hadoop-Framework, YARN, das Hadoop Distributed File System (HDFS) und die Data-Streaming-Komponente Kafka. Das Framework Hadoop basiert auf Java und liefert die Basis für skalierbare und verteilte Softwareanwendungen. Das Resource Management der Analyse-Infrastruktur beziehungsweise des Hadoop-Clusters wird von YARN übernommen. Es ermöglicht die dynamische Zuteilung von Ressourcen für die verschiedenen Analysejobs.

Bei HDFS handelt es sich um ein hochperformantes und hochverfügbares Datensystem zur Speicherung von großen Datenmengen (Big Data). Die Daten liegen hierbei auf mehreren Knoten innerhalb des Hadoop-Clusters. Zum Verarbeiten und Speichern der

Daten nutzt das HDFS das Master-Slave-Prinzip. Der sogenannte NameNode (Master) beinhaltet alle Metadaten über Verzeichnisstrukturen und Dateien innerhalb des Dateisystems. Der DataNode (Slave) ist für die Verwaltung der Daten auf den einzelnen Knoten innerhalb des Clusters verantwortlich. Kafka ist die Softwarekomponente, die das Verarbeiten von Datenströmen ermöglicht. Diese sammelt hierzu die Rohdaten der Quellsysteme von den sogenannten Produzern und stellt diese im Kafka-Cluster innerhalb verschiedener Topics zur Verfügung. Die Topics werden wiederum in verschiedene Partitionen unterteilt und auf mehreren Clusterknoten redundant gespeichert. Die Daten können nun von den verschiedenen Zielsystemen (Consumern) gelesen werden.

Dem *Service Layer* werden die Softwarekomponenten zugeordnet, die dem Benutzer eine Funktionalität oder einen kompletten Service zur Verfügung stellen. Dieser Schicht werden die Workflow Execution Engine Oozie, der Cluster-Manager Ambari und die Task Execution MapReduce zugeordnet.

Die Workflow Execution Engine Oozie ermöglicht es, Workflows zu erstellen, die einen oder mehrere MapReduce-Jobs nach einem Workflow ablaufen lassen. Die Oozie-Workflows werden in hPDL (Hadoop Process Definition Language) definiert und ähneln der XML-Prozess-Definitionssprache JBOSS JBPM. Innerhalb der Oozie-Workflows können Software (z. B. Java-Programme) oder Services hinterlegt werden, die eine Auswertung der Daten vornehmen, um ein bestimmtes Analyseergebnis (z. B. Clustering, Prediction usw.) zu erhalten. Die Erstellung von definierten Oozie-Workflows kann direkt in der hPDL-Datei erfolgen oder aber über grafische User Interfaces (z. B. Hue). Der Workflow kann vor, während und nach der Laufzeit bestimmte Programme oder Services ansteuern, die dann wiederum die Verteilung der Fahrzeugflotte beauftragen.

Apache Ambari ist ein Werkzeug zum Cluster-Management, Provisioning und Monitoring für Hadoop-Cluster. Es ermöglicht ebenfalls die Installation von Hadoop-Services über den Cluster hinweg, bietet Start- und Stopp-Funktionalitäten zur Steuerung von Softwarekomponenten und Services und erlaubt ein Monitoring des gesamten Clusters über ein Dashboard (inkl. Ambari Metrics und Ambari Alert). MapReduce stellt das sogenannte MapReduce-Verfahren bereit, welches Daten (bis zu mehreren Petabytes) in den drei Phasen Map, Shuffle und Reduce verarbeitet.

Der *Application Layer* stellt die User Interfaces bereit. Diese Schicht ist abhängig von den jeweilig gewählten Softwarekomponenten. Apache Ambari liefert zum Beispiel eine komplette Web-Oberfläche (inkl. Dashboards), die es ermöglicht, den gesamten Hadoop-Cluster zu steuern, zu erweitern und zu monitoren. Innerhalb dieser Schicht könnten ebenfalls User Interfaces für User-, Data- und Workflow-Management (z. B. Hue) zur Verfügung stehen. Ebenfalls steht hier eine Web-Oberfläche zur Verfügung, die die Analyseergebnisse entsprechend für den User visualisiert. Denkbar wäre an dieser Stelle auch die Implementierung einer mobilen Applikation für Außendienstmitarbeiter. Beide Implementierungen können wiederum verschiedene Funktionalitäten für den Benutzer aufweisen.

Alle Schichten können natürlich durch weitere Softwarekomponenten der Apache Software Foundation oder Dritter erweitert werden. Auch eine Erweiterung durch eigene Implementierungen stellt kein Problem dar.

Letztlich wird es viele Lösungsansätze geben, welche Datenquellen durch welche Technologien miteinander verknüpft werden müssen, um den zukünftigen Kunden einer autonomen Fahrzeugflotte den Nutzervorteil zu bieten, den sie sich erwarten. Wichtig wird aber auch dann (wie schon heute) sein, wie schnell die Nachfrage der Nutzer erfüllt werden kann und wie sehr bei alledem die persönlichen Präferenzen der Fahrgäste berücksichtigt werden, um die Customer Journey für den Kunden zu einem vollen Erfolg zu machen.

## Literatur

[Ber17] K. Berwind u. a., Bigdata Reference Model, in: Big Data Analytics: Tools and Technology for Effective Planning, 2017

## Social Matching/Fahrgemeinschaften

„Hallo, deine Freundin Emma hat ebenfalls einen Wagen bestellt. Zeit und Strecke eurer Fahrten sind ähnlich. Wollt ihr zusammenfahren?“ So oder ähnlich könnten in Zukunft Fahrgemeinschaften entstehen. Gerade für Pendler könnten sich Fahrgemeinschaften in autonomen Autos als sehr hilfreich erweisen.

Das Matching geschieht im Hintergrund entsprechend der geplanten Strecke des Nutzers, seiner Präferenzen und der Strecken und Präferenzen anderer Fahrgäste. In diesen Präferenzen legt der Nutzer seine zeitliche Flexibilität und seine Bereitschaft für Umwege fest. Außerdem definiert er, mit wem er bereit wäre, einen Wagen zu teilen. Will er nur mit einem kleinen Kreis seiner engsten Freunde (Emma und Mathilda) zusammen fahren oder mit seinen Kontakten aus sozialen Netzwerken? Oder nur mit Teilmengen, wie beispielsweise Kollegen aus der Firma oder Menschen mit ähnlichen Hobbys, oder ist er auch bereit, mit Fremden den Wagen (und die Kosten) zu teilen?

Für den Nutzer bedeutet es, dass er basierend auf all diesen Informationen vom Mobilitätssystem Vorschläge erhält und über einen Single-Point-of-Contact (z. B. eine Mobilitäts-App) auf mehrere integrierte Mobilitätsalternativen zugreifen kann. Er wird von seinem Mobilitätsdienstleister über die besten Alternativen proaktiv informiert, um seine Reise mit möglichst geringer Verspätung und entsprechend seiner Präferenzen durchführen zu können. Nur dann wird die Customer Journey wirklich ein Erfolgserlebnis für den Nutzer.





# gemeinsam entwickeln

In unseren Communities entwickeln wir gemeinsam nachhaltige digitale Lösungen: mit unseren Kunden in konkreten Projekten, auf unseren Entwicklertagen oder übergreifend in unseren Centers of Excellence.

Mehr Informationen auf  
[www.bridging-it.de/gemeinsamentwickeln](http://www.bridging-it.de/gemeinsamentwickeln)