

# Big Data-Zentren – Vorstellung und Panel

Volker Markl (BBDC), Erhard Rahm & Wolfgang Lehner (ScaDS), Michael Beigl (SDIL)

Moderation: Thomas Seidl

## 1 Abstract

Zur Erforschung der verschiedenen Facetten von „Big Data“ wurden jüngst drei Zentren gegründet. Hierbei handelt es sich um die vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Kompetenzzentren BBDC (Berlin Big Data Center, Leitung TU Berlin) und ScaDS (Competence Center for Scalable Data Services and Solutions, Leitung TU Dresden und Uni Leipzig) sowie das in Zusammenarbeit von Industrie und Forschung eingerichtete SDIL (Smart Data Innovation Lab, Leitung KIT). Diese drei Zentren werden zunächst in Kurzvorträgen vorgestellt. Eine sich anschließende Panel-Diskussion arbeitet Gemeinsamkeiten, spezifische Ansprüche und Kooperationsmöglichkeiten heraus.

## 2 Berlin Big Data Center (BBDC)

Im Rahmen des vom BMBF geförderten Kompetenzzentrums „Berlin Big Data Center“ (BBDC) unter der Federführung der Technischen Universität Berlin werden die Konsortialpartner Konrad-Zuse-Zentrum für Informationstechnik (ZIB), Fritz-Haber-Institut der Max-Planck-Gesellschaft, Beuth Hochschule und Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) in den Jahren 2014 bis 2018 neuartige und automatisch skalierbare Technologien entwickeln, welche tiefgreifende Analysen von „Big Data“ ermöglichen. Das BBDC verfolgt das Ziel, die Forschungsgebiete des Maschinellen Lernens und des Datenmanagements zu verschmelzen und ein hochskalierbares Open Source-System zu entwickeln, welches in der Lage ist, die Spezifikation, automatische Optimierung, Parallelisierung und Hardwareadaption sowie die fehlertolerante, effiziente Ausführung verschiedener Datenanalysemethoden (z.B. des Maschinellen Lernens, der linearen Algebra, der Statistik, der Wahrscheinlichkeitstheorie, der Computerlinguistik sowie der Signalverarbeitung) durchzuführen. Als Beispielanwendungen werden die Bereiche Gesundheit, Materialforschung, Informationsmarktplätze, Logistik und Industrie 4.0 zu betrachtet. Neben Forschung und Innovation werden im Kontext des BBDC auch mehrere Aktivitäten der Lehre gebündelt.

**Volker Markl** leitet das Fachgebiet Datenbanksysteme und Informationsmanagement (DIMA) an der Technischen Universität Berlin. Gleichzeitig ist Dr. Markl Status-Professor an der Universität von Toronto und leitet die Forschungsgruppe „Intelligente Analyse von Massendaten Smart Data“ am Deutschen Forschungszentrum für künstliche Intelligenz (DFKI). Seine Forschungsinteressen beinhalten neue Hardwarearchitekturen für das Informationsmanagement (z.B. neue Speichertechnologien wie NUMA, remote Memory oder Flash/Phase Change Memory, neue Verarbeitungstechnologien wie SIMD, Manycore), skalierbare Datenverarbeitung (Programmiermodelle, Anfrageübersetzung, -optimierung und -parallelisierung sowie skalierbares Data Mining) sowie Anwendungen in den Bereichen Text Mining und Informationsmarktplätze. Volker Markl ist auch Sprecher des „Data Analytics & Cloud“ Labs der TU Berlin und Direktor des vom BMBF eingerichteten Big Data Centers Berlin.

### 3 Competence Center for Scalable Data Services and Solutions (ScaDS)

Im Big-Data-Zentrum ScaDS (Competence Center for Scalable Data Services and Solutions) Dresden/Leipzig ([www.scads.de](http://www.scads.de)) entwickeln international führende Forscher der TU Dresden und der Universität Leipzig neue Lösungen für datenintensive wissenschaftliche und wirtschaftliche Anwendungen. Profilbestimmende Forschungsschwerpunkte liegen in den Gebieten der Datenintegration, der Wissensextraktion sowie der visuellen Analyse. Aufgrund der erstklassigen Datenbankexpertise an beiden Standorten werden Techniken des Datenmanagement und der Datenanalyse umfassend vorangetrieben.

**Erhard Rahm** promovierte und habilitierte an der TU Kaiserslautern und leitet seit 1994 den Lehrstuhl für Datenbanken an der Universität Leipzig. Seit 2014 ist er wissenschaftlicher Ko-Koordinator des BMBF-geförderten Big-Data-Zentrums ScaDS (Competence Center for Scalable Data Services and Solutions) Dresden/Leipzig. Seine Forschungsinteressen liegen vor allem in den Gebieten der Datenintegration und Big Data. Er ist Autor von mehreren Büchern und über 200 wissenschaftlichen Publikationen. Seine Forschungsergebnisse im Gebiet der Datenintegration wurden bereits mehrfach international ausgezeichnet, u.a. durch zwei renommierte Test-of-Time Awards: den VLDB 10-Year Best Paper Award sowie den ICDE Influential Paper Award.

**Wolfgang Lehner** ist Direktor des Instituts für Systemarchitektur an der Fakultät Informatik der TU Dresden. Seine wissenschaftlichen Schwerpunkte liegen im Bereich der analytischen Auswertung großer Datenbestände, wobei stets versucht wird, eine Durchgängigkeit von der statisch-analytischen Modellierung bis zur effizienten Umsetzung auf Ebene der Systemarchitektur zu erreichen. Über 200 Veröffentlichungen und eine Vielzahl von Aktivitäten in der Community zeugen von einer regen wissenschaftlichen Aktivität und internationaler Vernetzung. Wolfgang Lehner vertritt als Fachkollegiat aktuell die Datenbank- und Informationssysteme innerhalb der DFG, ist Mitglied des VLDB Endowments und wurde 2014 in die Academy of Europe berufen. An der TU Dresden ist er neben dem Kompetenzzentrum ScaDS auch im Sfb „HAEC – Highly Adaptive Energy Efficient Computing“, im Exzellenz-Cluster „cfAED – Advancing Electronics Dresden“ und als Sprecher im Graduiertenkolleg „RoSI – Rollenbasierte Informationssysteme“ aktiv.

## 4 Smart Data Innovation Lab (SDIL)

Das Smart Data Innovation Lab dient der Förderung der Smart Data Spitzenforschung in Deutschland. Durch Strukturierung von Big Data entstehen Informationen in Form von Smart Data, die zu Wissensvorteilen und zur Entscheidungsunterstützung genutzt werden können. Im SDIL werden in enger Zusammenarbeit von Industrie und Wissenschaft verbesserte Voraussetzungen für eine Spitzenforschung im Bereich Data Engineering/Smart Data geschaffen.

**Michael Beigl** ist seit 2010 Professor (W3) für Pervasive Computing Systems und Leiter des TECO am Karlsruher Institut für Technologie (TECO). Seit 2012 ist er darüber hinaus Dekan der Fakultät für Informatik. Zuvor hatte er von 2006 bis 2010 den Lehrstuhl für Verteilte und Ubiquitäre Systeme an der Technischen Universität Braunschweig inne und war 2005 Gastprofessor an der Keio Universität in Japan. Von 2000 bis 2005 leitete er als wissenschaftlicher Direktor des TECO der Universität Karlsruhe, ein Forschungslabor für anwendungsorientierte Forschung. Herr Beigl absolvierte sein Studium der Informatik an der Universität Karlsruhe und promovierte dort zum Dr.-Ing.. Er hat über 100 Artikel im Themenfeld des Mobile, Wearable, Ubiquitous und Pervasive Computing, des Internet der Dinge sowie der Datenanalyse veröffentlicht. Zu diesen Themen ist er in zahlreichen Gremien und Editorial Boards tätig und hat eine Vielzahl von Forschungsprojekten insbesondere zusammen mit der Industrie durchgeführt sowie Start-Ups gegründet. Seit 2014 ist Herr Beigl Sprecher des Smart Data Innovation Lab (SDIL).