

# Smart Data Innovation Lab - eine Forschungsplattform mit realen Daten

Dr. Harald Schöning

Head of Research  
Software AG  
Uhlandstr. 12  
64297 Darmstadt  
Harald.Schoening@softwareag.com

**Abstract:** Das Smart Data Innovation Lab ist eine nationale Forschungsplattform, die der Wissenschaft Zugang zu realen Industriedaten für Forschungszwecke auf einer leistungsfähigen IT-Infrastruktur verschafft und einen schnellen Ergebnistransfer zurück in die Industrie ermöglicht.

## 1 Hintergrund

Das Thema „Big Data“ hat in den letzten Jahren an Aufmerksamkeit gewonnen, nicht nur durch Fortschritte in Algorithmen und verfügbaren Softwarelösungen, sondern vor allem durch die rasant anwachsende Verfügbarkeit von Daten. Neben dem wachsenden Verkehr auf sozialen Medien ist die verstärkte Verbreitung von Sensoren und anderer vernetzter Objekte aus dem „Internet of Things“ auch im industriellen Einsatz ein wichtiger Treiber dieser Entwicklung. Mit der Verfügbarkeit dieser Daten geht der Bedarf einher, sie nutzbringend zu aggregieren und auszuwerten. Neue Anwendungen in Industrie, Städten, Medizin, Mobilität usw. werden möglich. Viele Forschergruppen haben den Bedarf erkannt und forschen an neuen Algorithmen und Lösungen. Insbesondere die Zusammenführung von Daten aus mehreren Quellen verspricht neue Erkenntnisse. Oft steht die Forschung allerdings vor dem Problem, dass sie ihre Ergebnisse nicht in realen Anwendungsszenarien mit realen Daten validieren kann, da diese Daten von deren Inhabern nicht frei zur Verfügung gestellt werden. Die Gründe hierfür sind vielfältig. So können solche Daten personenbezogene Informationen enthalten oder Rückschlüsse auf Produktionsverfahren und –geheimnisse zulassen. Andererseits sind die Dateninhaber aber durchaus daran interessiert, aus ihren Daten unter Mitwirkung von Forschern neue Erkenntnisse zu gewinnen.

## 2 Das Smart Data Innovation Lab

Hier setzt das Smart Data Innovation Lab (SDIL, sdil.de) an, das auf Initiative der AG6 des IT-Gipfels entstand. Es stellt eine nationale Forschungsplattform dar, durch die die

Wissenschaft Zugang zu realen Industriedaten für Forschungszwecke auf einer leistungsfähigen IT-Infrastruktur erhalten kann und ein schneller Ergebnistransfer zurück in die Industrie ermöglicht wird. Forschungsprojekte sollen insbesondere von der Zusammenführung verschiedener Datenquellen (Wissenschaft, Verwaltung, Industrie, kommerzielle Datenquellen) profitieren. Besonderes Augenmerk liegt auf geeigneten Nutzungsbedingungen und Sicherheitsmaßnahmen, die Risiken wie den Missbrauch sensibler Daten, Streitigkeiten bezüglich des Eigentums an abgeleiteten Daten und Forschungsergebnissen, Rechteverletzungen bezüglich eingelagerter Daten oder angebundener kommerzieller Datenquellen minimieren.

Partner im SDIL sind Dateninhaber, Forschungseinrichtungen und Software- und Hardwarehersteller. In z.Z. vier inhaltlichen Clustern (Innovation Communities, s. Abb. 1) werden unter gemeinsamer Leitung jeweils eines Unternehmens und einer Forschungseinrichtung die Akteure eines Themengebiets vernetzt. Eine eigene Arbeitsgruppe *Data Curation* ist für Datenpflege, -beschreibung und -verwaltung zuständig. Daten können dabei entweder auf der Plattform gespeichert sein oder dort nur registriert sein. In diesem Fall werden sie vom Dateninhaber direkt nach gesonderter Vereinbarung mit den jeweiligen Forschern verfügbar gemacht. SDIL fungiert dabei als Moderator. Die Arbeitsgruppe *Recht* entwirft entsprechende Nutzungsvereinbarungen.

Um eine leistungsfähige IT-Plattform zu gewährleisten, stellen verschiedene Software-Anbieter ihre Produkte zu Verfügung. Der Betrieb wird vom KIT sichergestellt, dazu haben Hardware-Partner Rechnerkapazität bereitgestellt. Abbildung 1 zeigt die Organisationsstruktur des SDIL. Forschungseinrichtungen können sich über den SDIL-Koordinator auf Plattform- und Datenverwendung bewerben.

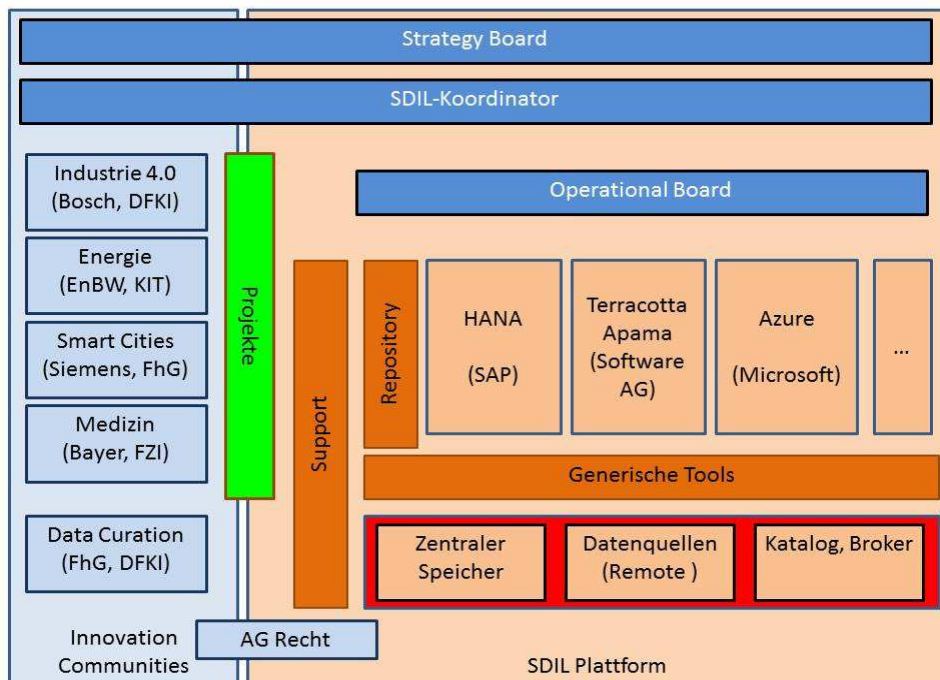


Abbildung 1: organisatorische Struktur des Smart Data Innovation Labs