

# 1 Einleitung

Kapitel 1 beschreibt die Zielstellung und Einordnung dieser Arbeit. Die Ausgangssituation und die daraus abgeleitete Problemstellung der Dissertationsschrift wird in Abschnitt 1.1 erläutert. Die Zielsetzung und Forschungsfrage der vorliegenden Arbeit wird in Abschnitt 1.2 formuliert. Die Wissenschaftstheoretische Verortung der Arbeit in Abschnitt 1.3. Die Forschungskonzeption wird in Abschnitt 1.4 dargelegt, Abschnitt 1.5 erläutert den Aufbau der Arbeit.

## 1.1 Ausgangssituation und Problemstellung

*„Ein halbes Jahrtausend nach Gutenberg ist nicht der Mangel, sondern der Überfluss an Information unser größtes Problem.“*

- DORN 1994, S. 13 f.

Die **Ressource Information** und ihre richtige Verwendung sind von enormer Bedeutung für Unternehmen. Dabei ist die Qualität dieser Information, gerade in der Flut an zur Verfügung stehenden Information, unklar. Im Zeitalter der Informationsgesellschaft ist der Einsatz von Informationstechnologie (IT) alltäglich geworden. Digitalisierung, cyber-physische Systeme, das Internet der Dinge und Industrie 4.0 sind Begriffe, welche die unterschiedlichen Ansätze der Nutzung von Information, Daten, IT und Sensorik beschreiben. Diese Ansätze erschließen neue Bereiche und Prozesse bzw. intensivieren den Technologieeinsatz in vorhandenen Domänen. Dadurch werden Unternehmen in die Lage versetzt, Entscheidungen und Handlungen besser zu fundieren (s. SCHUH, BIRKMEIER, BLUM, RESCHKE 2016a, S. 50). Prozesse und Entscheidungen im Unternehmen werden vereinfacht, beschleunigt und teilweise automatisierbar. Es mangelt aber an der strukturierten Beurteilung der eingehenden Information und einer Betrachtung der Änderungen der Information während sie entlang der Prozesse durch die Systeme verarbeitet wird.

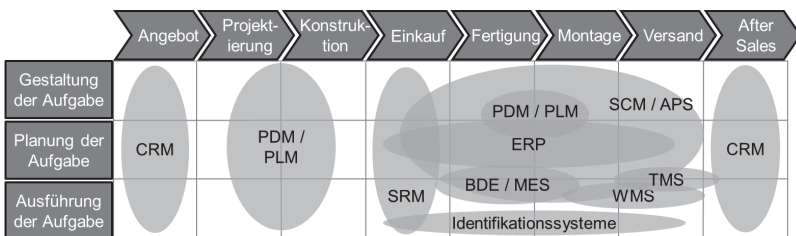
Verstärkt wird dies, durch das **Wachstum der Informationsmenge**. Die Menge an Daten und Informationen, die Unternehmen zur Verfügung stehen und verarbeitet werden müssen, wächst nämlich erheblich. RUDOLPH U. LINZMAJER (s. RUDOLPH U. LINZMAJER 2014, S. 13) bestätigten unlängst, dass sich das Datenvolumen analog zum Moorsche Gesetz entwickelt. Zusammen mit dem exponentiellen Wachstum der Rechenkapazität vergrößert sich das Datenvolumen exponentiell. Das Datenvolumen verdoppelt sich dabei ca. alle zwei Jahre (s. DORSCHER 2015, S. 7).

**Daten und Informationen werden intensiver und vernetzter genutzt.** 90 Prozent aller industriellen Produktionsprozesse sind heute bereits durch Informations- und Kommunikationstechnik (IKT) gestützt (s. KAGERMANN ET AL. 2013b, S. 17). Trends, wie Industrie 4.0, das Internet der Dinge und Big Data beschreiben die jüngsten technologischen und gesellschaftlichen Auswirkungen dieser Entwicklung und verstärken die Vernetzung (s. KAGERMANN ET AL. 2013b, S. 18). ORCUTT, vom Massachusetts Institute of Technology, prognostizierte im Jahr 2013, dass die Datenmenge, die von

Maschinen, Fahrzeugen, Sensoren, etc. erzeugt wird bald die Menge an Daten, die von Menschen erzeugt wird übersteigen wird (s. ORCUTT 2013). GARTNER prognostiziert, dass im Jahr 2020 26 Milliarden Geräte im Internet der Dinge miteinander verbunden sein werden und damit ein vielfaches mehr als Menschen (s. RIVERA U. VAN DER MEULEN 2014). Durch diese automatisierte Produktion und Nutzung von Information, **entfällt das regulierende Element des Menschen**, der neben der Verarbeitungsgeschwindigkeit auch die Plausibilität einschätzen kann und die angebotene Information gegenüber seinem eigenen Wissen verifiziert.

Informationssysteme werden im Unternehmen stark integriert eingesetzt. Dies geschieht heute standort- und unternehmensübergreifend. Insbesondere bei der Informationsbereitstellung für hochmoderne Produktion, beim Einsatz von cyber-physischen Systemen und der Smart Factory (s. KAGERMANN ET AL. 2013b, S. 6). Dabei ist das Vertrauen in die Korrektheit der Daten und Informationen ein grundlegender Aspekt für die Akzeptanz und die Entwicklung von Industrie 4.0 (s. KAGERMANN ET AL. 2013b, S. 51).

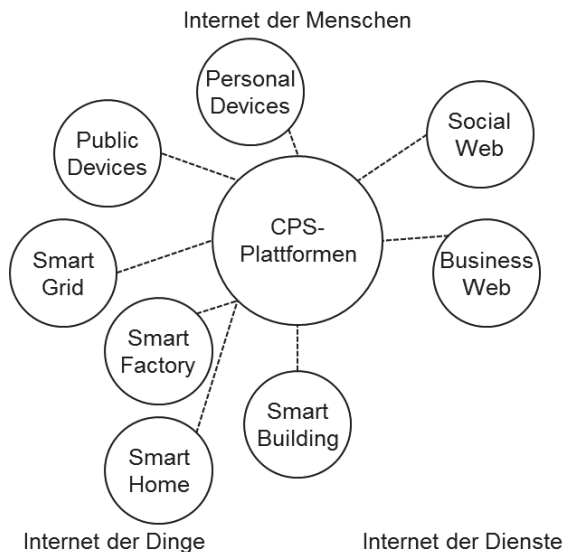
Die bisherige hohe Verzahnung der IT im Unternehmen (siehe Abbildung 1-1) wird in cyber-physischen Systemen mit der Anbindung an weitere Domänen (Internet der Menschen, Internet der Dienste und Internet der Dinge) und in der Entwicklung der Industrie 4.0 weiter steigen (siehe Abbildung 1-2). Dabei ist der durchgängige Informationsfluss von entscheidender Bedeutung für zukünftige Produktionssysteme (SPUR 1997, S. 23). Die notwendige Transparenz aus der Vielzahl an Daten wird als digitaler Schatten bzw. digitaler Zwilling bezeichnet. Relevant ist die Durchgängigkeit entlang der Unternehmensprozesse, weniger die hohe Auflösung. Ziel ist es alle notwendigen Datensätze aufgabenspezifisch zu erfassen (SCHUH, BIRKMEIER, WALENDZIK, LUCKERT, WEBER, BLUM 2016b, S. 745).



**Abbildung 1-1: Darstellung der Verzahnung von IT-Systemen in der Auftragsabwicklung, eigene Darstellung nach OEDEKOVEN 2011 - Zugl.: Aachen, Techn. Hochsch., Diss 2011, S. 29 i. A. a. STRAUBE U. DOCH 2007, S. 48**

Durch diese fortschreitende globale Vernetzung steigt die organisatorische Komplexität. Dabei führen die Produktionsprozesse, durch ihre hohe Dynamik, zu höheren Anforderungen an die Planung und Steuerung (vgl. SCHEIBMAYER 2014 - Zugl.: Aachen, Techn. Hochsch., Diss 2014, S. 1; WESTKÄMPER U. ZAHN 2009, 7 ff.; NYHUIS 2010, 3 f.; KAGERMANN ET AL. 2013a, S. 11; SCHUH ET AL. 2011, S. 797).

Diese höheren Anforderungen sind implizit gestiegenen Anforderungen an die Güte der eingehenden Informationen. Planung und Steuerung kann nur erfolgen, wenn die zu beeinflussenden Größen auch messbar sind und damit die Auswirkungen der Regelung bekannt sind (vgl. WESTKÄMPER U. ZAHN 2009, S. 8 f.). **Bezogen auf wissensintensive Arbeitsprozesse erfordert dies die Messung und Steuerung der Qualität der Daten**, der Informationen und des Wissens (vgl. EPPLER 2006, S. 27). ALBRECHT betont ebenso, dass ein Modell für die ganzheitliche Informationsqualität, die Qualität der Daten, der Information und des Wissens berücksichtigen muss. Darüber hinaus bewertet er den Schutz der Daten gegen Korruption und Informationsverlust durch Störungen ebenso wie die Berücksichtigung des Informationsverhalten der Benutzer als wichtige Aspekte (vgl. ALBRECHT 1999, S. 2 ff.; vgl. MICHELSON U. RIEKERT 2001, S. 67). Dies wird aber in vielen Unternehmen nur teilweise umgesetzt. Einerseits fehlt das Bewusstsein für die Thematik, andererseits mangelt es an geeigneten Mitteln und Werkzeugen und **die vorhandenen Methoden werden aufgrund des hohen Aufwands nicht eingesetzt** (vgl. EPPLER 2006, S. 13 f.), denn die vorhandenen Ansätze erfordern eine intensive Anpassung an die Prozesse, die Situation und Begebenheiten, damit aussagekräftige Ergebnisse bei Messungen erzeugt werden können (vgl. PIPINO ET AL. 2002, S. 218).



**Abbildung 1-2: Vernetzung im Internet der Dinge, eigene Darstellung nach KAGERMANN ET AL. 2013b, S. 28, mit Bezug auf FERBER U. DEBEVOISE 2012, S. 13**

Die Kombination einer Vielzahl von Systemen und ihre Vernetzung in verzweigten Informationsflüssen schafft ein Netzwerk zum Informationsaustausch. Dabei werden In-

formationen über verschiedene Systeme weitergereicht bzw. übertragen. Diese Übertragungen sind aber oft nicht verlustfrei, denn Daten werden verändert, angepasst, zusammengefasst oder reduziert. Für den Nutzer bzw. konsumierenden Prozess ist die Veränderung und der Pfad der Informationsbereitstellung intransparent. Informationen über die Quelle von Daten, den Ablauf der Informationsübertragung oder etwaige Veränderungen sind oft nicht vorhanden oder werden nicht dokumentiert. Metadaten werden häufig verändert, nicht erfasst oder bewusst gelöscht, die eine Rekonstruktion ermöglichen würden. **In der Folge ist die Qualität der Informationen und Daten in den Systemen und in den Prozessen nicht belegt oder gesichert** (vgl. WANG U. STRONG 1996, S. 5ff. ). **Die Informationsqualität ist für den Nutzer intransparent.** KAGERMANN ET AL. betonen die Bedeutung der Orchestrierung der Systeme und Prozesse. Deren Akzeptanz und Erfolg sind maßgeblich von dem Vertrauen, der Verlässlichkeit und der Nutzbarkeit der Daten und Informationen abhängig (vgl. KAGERMANN ET AL. 2013b, S. 28 f.; KAGERMANN ET AL. 2013b, S. 51; LINDNER ET AL. 2013, S. 3–7).

Der **wirtschaftliche Schaden**, basierend auf schlechter Daten- und Informationsqualität, ist aufgrund der Komplexität des kausalen Nachweises der Ursachen, nicht fassbar. WIETHAUS benennt die Größenordnung der „Total Cost of poor Data Quality“ zwischen 8 und 12 Prozent des Unternehmensumsatzes und die durchschnittliche Fehlerrate von Kundendatenbanken auf 15 – 20 Prozent (WIETHAUS 2001, S. 30). Basierend auf den Projekterfahrungen des Autors und den zur Validierung konsultierten Experten ist eine durchschnittliche Fehlerrate bei Kundendatenbanken von 15 – 20 Prozent in Unternehmen sehr plausibel, meist liegen die Fehlerraten sogar bei ca. 30 Prozent oder mehr.

SCHEIBMAYER ET AL. beschreiben in ihrer Studie von 136 Unternehmen, dass 40 Prozent ihre Produkt- und Anlagedaten als unterdurchschnittlich bewerten und knapp 2 Prozent die Produktdaten als schlecht oder sehr schlecht bewerten (s. SCHEIBMAYER, BIRKMEIER, KNAPP 2014, S. 16). Diese Ergebnisse entsprechen den Ergebnissen der Studie von GÖTZE U. LEIDICH und GÖTZE ET. AL., die im Rahmen des BMWi Projekts „eBEN“ 134 Unternehmen untersuchten. Ihre Befragung ergab zwar in der Selbsteinschätzung zunächst offenbar bessere Werte der Datenqualität: 77 Prozent der Unternehmen schätzten ihre Stammdaten als gut bzw. ausgezeichnet ein und ca. 23 Prozent als befriedigend (16 Prozent) oder mangelhaft (7 Prozent) (s. GÖTZE ET AL. 2014, S. 17; GÖTZE U. LEIDICH 2014, S. 7). GÖTZE U. LEIDICH untersuchten intensiv die Diskrepanz zwischen subjektiver, wahrgenommener Datenqualität gegenüber der realen Qualität und stellten fest, dass real nur 36 Prozent der Datenqualität gut (27 Prozent) bzw. ausgezeichnet (9 Prozent) waren. Dagegen waren 46 Prozent der Daten von befriedigender und 18 Prozent von mangelhafter Qualität (vgl. GÖTZE U. LEIDICH 2014, S. 7; GÖTZE ET AL. 2014, S. 11 ff.). Zusätzlich zur Intransparenz über die Daten- und Informationsqualität ist die **Selbsteinschätzung der eigenen Datenqualität durch Unternehmen deutlich positiver als sie in der Realität ist.**

Auch andere Fachdisziplinen können keinen konstruktiven Lösungsweg vorschlagen. Betrachtet man die Ausgangssituation eines Entscheiders bei einer wissensintensiven (betrieblichen) Aufgabe im Sinne der Spieltheorie, ist das primäre Problem das der