

Abschlussbericht der Förderung zur internationalen Zusammenarbeit in der angewandten Forschung (PIZ) „CRISP-DM – ein Ansatz zur Systematisierung von Digitalisierungsprojekten in Produktionsumfeldern“

Berichtszeitraum und Förderzeitraum

Berichts- und Förderzeitraum ohne Vorbereitungen: 01.07.2018 bis 17.09.2018

Datum des Projektabschlusses: 17.09.2018



Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA

Projektgruppe Prozessinnovation

Universitätsstraße 9

95447 Bayreuth

Autor/Ansprechpartner/Stipendiat:

Christoph Schock, M.Sc.

Tel: 0921/78516-314

Email: christoph.schock@ipa.fraunhofer.de

1 Allgemeine Angaben

Thema des Forschungsvorhabens

Ziel der internationalen Zusammenarbeit ist es, eine Handlungshilfe für produzierende Unternehmen, insbesondere für kleine und mittlere Unternehmen (KMU) zu entwickeln, die diese dazu befähigt, eigenständig Digitalisierungsprojekte durchzuführen und nachhaltig zu verwerten.

Zu diesem Zweck soll eine neuartige Kombination etablierter Methoden des Data Mining sowie der Produkt- und Serviceentwicklung erforscht werden. Einerseits wird der Cross Industry Standard Process for Data Mining (CRISP-DM) auf Basis von Fallstudien eines laufenden Forschungsprojekts der Projektgruppe auf die Anforderungen produzierender Unternehmen angepasst und spezifiziert, um künftig als Handlungshilfe bei der Durchführung von Digitalisierungsprojekten zu dienen. Andererseits soll im Rahmen der internationalen Zusammenarbeit die Integration von Design Thinking (DT) Methoden durch den Einbezug aller Interessensgruppen in Digitalisierungsprojekte eine nachhaltige Verwertung erzielter Ergebnisse gewährleisten.

Inhaltsverzeichnis

1	Allgemeine Angaben	2
2	Ausgangssituation	4
3	Veränderung des Erkenntnisstandes seit der Förderung	5
4	Geleistete Arbeit und Ergebnisse	7
5	Verwertung der Ergebnisse	10
6	Literaturverzeichnis	11

2 Ausgangssituation

Zur Umsetzung von Digitalisierungsanwendungen bzw. datenbasierten Problemlösungen im Produktionsumfeld dienen in der Industrie und Forschung oft Methoden aus der Disziplin des Data Mining als Grundlage. Neben unternehmensinternen Methoden, dem Knowledge Discovery in Databases (KDD) und dem darauf aufbauenden Sample, Explore, Modify, Assess (SEMMA), kommt vor allem der CRISP-DM (Abbildung 1) zum Einsatz [KDN14].

Da der CRISP-DM im Vergleich zu KDD und SEMMA durch die Phasen Business Understanding und Deployment im Ansatz ein breiteres Spektrum an Handlungsfeldern abdeckt und die umfangreicheren Beschreibungen notwendiger Phasen bei Data Mining Projekten liefert, kann dieser im Data Mining als etabliert betrachtet werden.

Dennoch fehlen insbesondere in den Phasen Business und Data Understanding, Data Preparation und Deployment relevante Informationen für produzierende Unternehmen. Weiterhin wird aus Data Mining Sicht davon ausgegangen, dass alle zur Problemlösung notwendigen Daten verfügbar sind. Da dies in industriellen Produktionen nur selten der Fall ist, ist der CRISP-DM ohne weitere Anpassungen nur bedingt zur Durchführung von Digitalisierungsprojekten einsetzbar.

Deshalb wurde der CRISP-DM Prozess von der Projektgruppe Prozessinnovation zunächst um die Phasen Cross Linking und Data Synthesis erweitert. Die beiden Phasen beinhalten Aktivitäten, Modelle und Methoden sowohl zur Ableitung von Datenbedarfen, also auch zur sensorischen Nachvernetzung von Anlagen, um identifizierte Bedarfe zu decken. Weiterhin sind Aktivitäten zur Datensynthese definiert, um Bestandsdaten mit Daten aus der Nachvernetzung zu einem Datenrahmen bzw. einer Datenbasis zu homogenisieren.

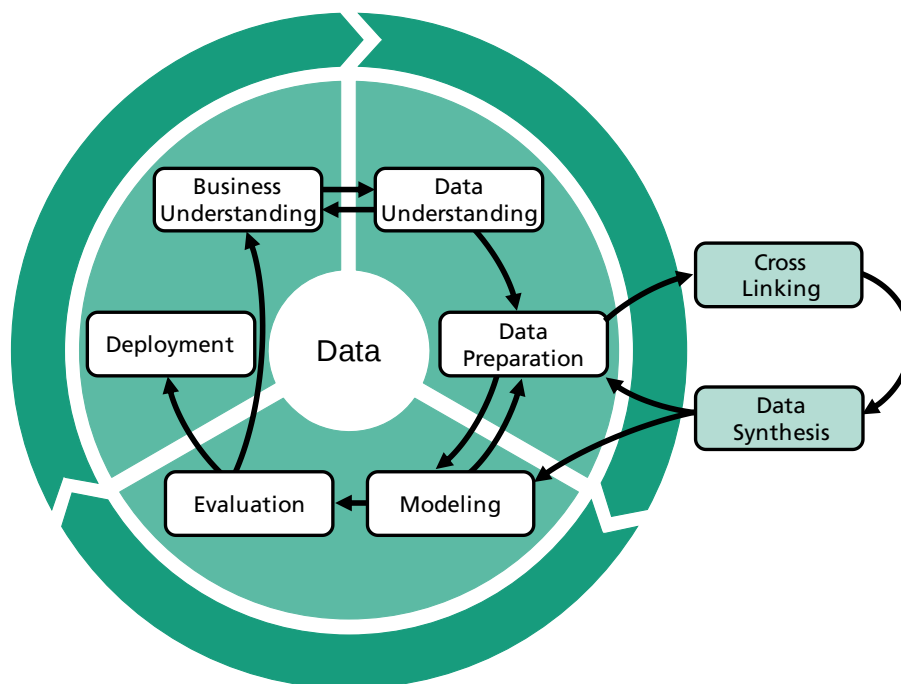


Abbildung 1: Erweiterter CRISP-DM Iteration 1

Das Design Thinking (DT) stellt im Vergleich zu Methoden der agilen Softwareentwicklung, bei denen durch User Stories und User Centered Design oftmals nur die Anforderungen des Product Owners erfüllt werden, einen innovativen Weg dar, um die Anforderungen und Wünsche von potentiellen Endnutzern abzudecken, auch wenn noch kein DT-Standardprozess etabliert ist. [IBM16; KRÜ13, UFL18] DT ist eine Methodik, die von Projektteams zur Befriedigung von Nutzerbedürfnissen angewandt wird. Während in der Wissenschaft der analytische Prozess die Lösung für ein gegebenes Problem darstellt, erlaubt der DT Ansatz die Erforschung von Problem und Lösungen gleichermaßen. Das DT bietet daher Potentiale bei der Durchführung von Digitalisierungsprojekten auf Basis des CRISP-DM.

3 Veränderung des Erkenntnisstandes seit der Förderung

DT verwendet Methoden aus verschiedenen Disziplinen und stellt einen iterativen Ansatz dar, bei dem es zunächst darum geht, das eigentliche Problem ganzheitlich zu verstehen, bevor mit der Lösungsfindung gestartet wird [GÜR13].

Die Methoden, die in den einzelnen Schritten des DT angewandt werden, stellen die Basis für die Erweiterung des CRISP-DM dar. Trotz der Vielzahl verschiedener DT Modelle [IDS14; GOO18; DFA18; DDT18] kann, wie in Abbildung 2 dargestellt, im Wesentlichen in vier Hauptschritte mit entsprechend zugeordneten Methoden unterschieden werden:

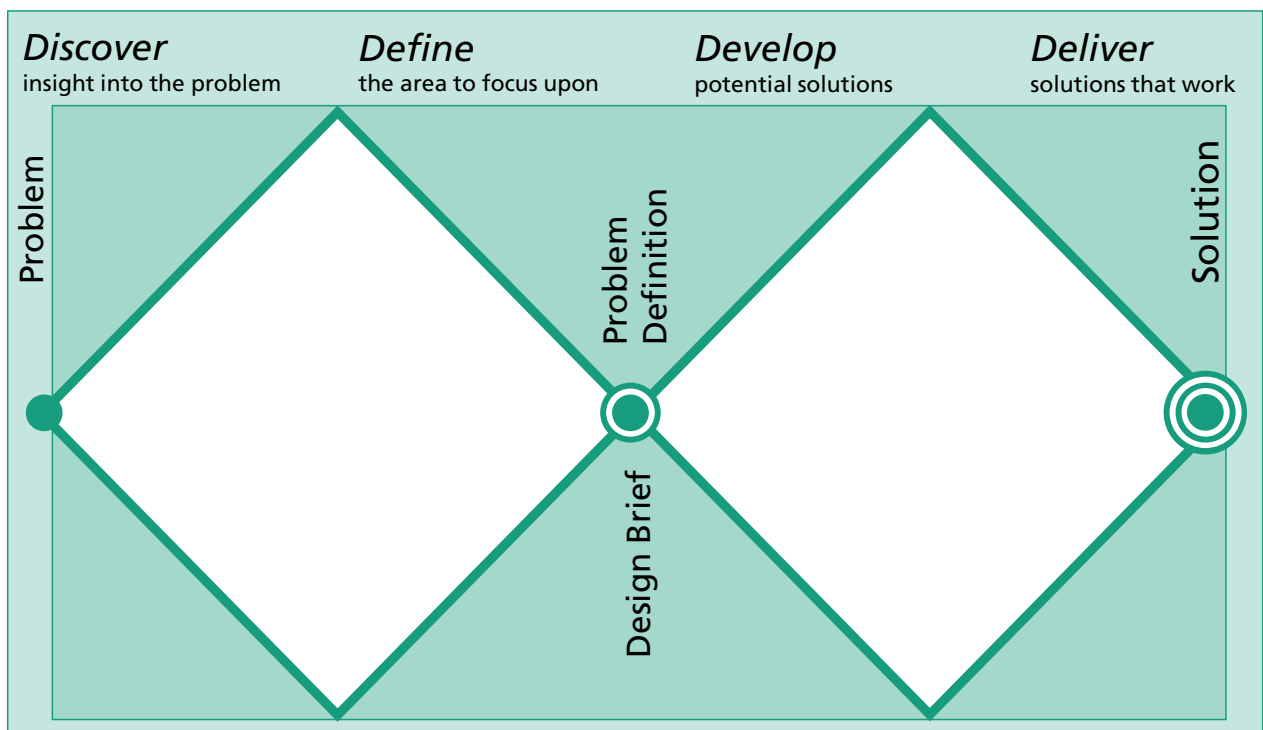


Abbildung 2: Double Diamond nach Design Council [DEG18]

Hier gilt es die Welt des Nutzers bzw. des „Problembesitzers“, die Hintergründe des Problems sowie potentiell existierende Perspektiven auf das Problem zu erfassen und zu verstehen. Ergebnis des Schritts sind verschiedene Problembeschreibungen.

2. Define

Hier gilt es die gesammelten Erkenntnisse aus Schritt Eins zusammenzuführen sowie zu analysieren, zu bewerten und zu priorisieren. Ergebnis des Schritts ist es, ein in den Folgeschritten zu lösendes Problem zu definieren.

3. Develop

Hier werden verschiedene Lösungskonzepte erarbeitet, falls möglich prototypisch umgesetzt sowie getestet und inkrementell optimiert. Ergebnis des Schritts sind verschiedene Lösungen (in Form von Prototypen) für das in Schritt Zwei definierte Problem.

4. Deliver

Hier gilt es die verschiedenen Prototypen zu testen, zu evaluieren sowie inkrementell in weiteren Iterationen zu optimieren.

Insbesondere für die Phasen der Divergenz (Discover und Develop) gibt es eine Vielzahl an Methoden, die in Tabelle 1 dargestellt sind. Für die Phasen der Konvergenz (Define und Deliver) werden vorrangig Evaluierungstechniken eingesetzt und Ergebnisse mit Hilfe von Kommunikationstechniken und Visualisierungen erzielt.

Tabelle 1: Methoden des DT

Discover	Define	Develop	Deliver
Stakeholder Analyse	Storytelling [DIG13]	Prototyping [VET12]	Evaluationstechniken
Case Study	Persona [DIG13]	Testing [VET12]	
Benchmarking	Use Cases	Scenario [DIG13]	
Fragebögen [DIG13]	Customer Journey [DIG13]	Storytelling [VET12]	
5 W Fragen [DIG13]		Co-creation [DIG13]	
5 Warums		Customer Journey [VET12]	
Abstraction ladder		Rollenspiele [DIG13]	

		[VET12]	
Interviews	[DIG13]		Bodystorming [DIG13]
[VET12]			
Einflussfaktoren Analyse		Brainstorming	[DIG13]
[VET12]		[VET12]	
Befragungen	[DIG13],	Storyboard	[DIG13]
Als Methode wenig		[VET12]	
hilfreich, da die Personen			
die eigenen Bedürfnisse			
in den Vordergrund			
stellen [VET12]			
Beobachtungen [DIG13],			
Laut Vetterli der			
Befragung methodisch			
überlegen, jedoch ist			
Unvoreingenommenheit			
essentiell [VET12]			

4 Geleistete Arbeit und Ergebnisse

Ziel der internationalen Zusammenarbeit ist es, eine Handlungshilfe für produzierende Unternehmen, insbesondere für kleine und mittlere Unternehmen (KMU) zu entwickeln, die diese dazu befähigt, eigenständig Digitalisierungsprojekte durchzuführen und nachhaltig zu verwerten.

Zu diesem Zweck wurde der CRISP-DM im Rahmen der drei Arbeitspakete sowie in zwei Iterationen erweitert:

1. Erweiterung der Business Understanding Phase
2. Erweiterung der Deployment Phase

Die Ergebnisse werden nachfolgend zusammengefasst dargestellt.

Ergebnisse Arbeitspaket 1:

Ein Auszug von „Methoden von Geschäfts-, Prozess- und Datenverständnis sowie Design

Thinking“ mit großer Relevanz für die Zielstellungen Steigerung der Effizienz (1), Steigerung der Effektivität (2) sowie Steigerung der Flexibilität (3) in Produktionsumfeldern ist in Tabelle 2 dargestellt.

Tabelle 2. Methoden von Geschäfts-, Prozess- und Datenverständnis sowie Design Thinking

Geschäfts-, Prozess- und Datenverständnis	Design Thinking
Wertstrommethode, Wertstrom 4.0 (Ergänzung um Daten- bzw. Informationsquellen)	Stakeholder Analyse
Fehler-Möglichkeiten-Einfluss-Analyse (FMEA)	Case Study
Input-Output Analyse (I-O Analyse)	Benchmarking
Ursache-Wirkungs-Analyse (U-W Analyse)	Fragebögen
Design of Experiments (DoE)	5 W Fragen
SIPOC Diagramm	5 Warums
Geschäftsmodelle, z. B. Osterwalder	Prototyping
Analyse externer und interner Datenquellen	Testing
Qualitätsreport Trainingsdaten	Scenario
Qualitätsreport Evaluationsdaten	Storytelling

Ergebnisse Arbeitspaket 2 und Arbeitspaket 3:

Im Sinne produktionsrelevanter Ziele und unter Berücksichtigung der Ergebnisse des ersten Arbeitspakets, ist in Abbildung 3 die „**Erweiterung und Spezifikation des CRISP-DM hinsichtlich produzierender Unternehmen**“ dargestellt sowie untenstehend zusammenfassend beschrieben.

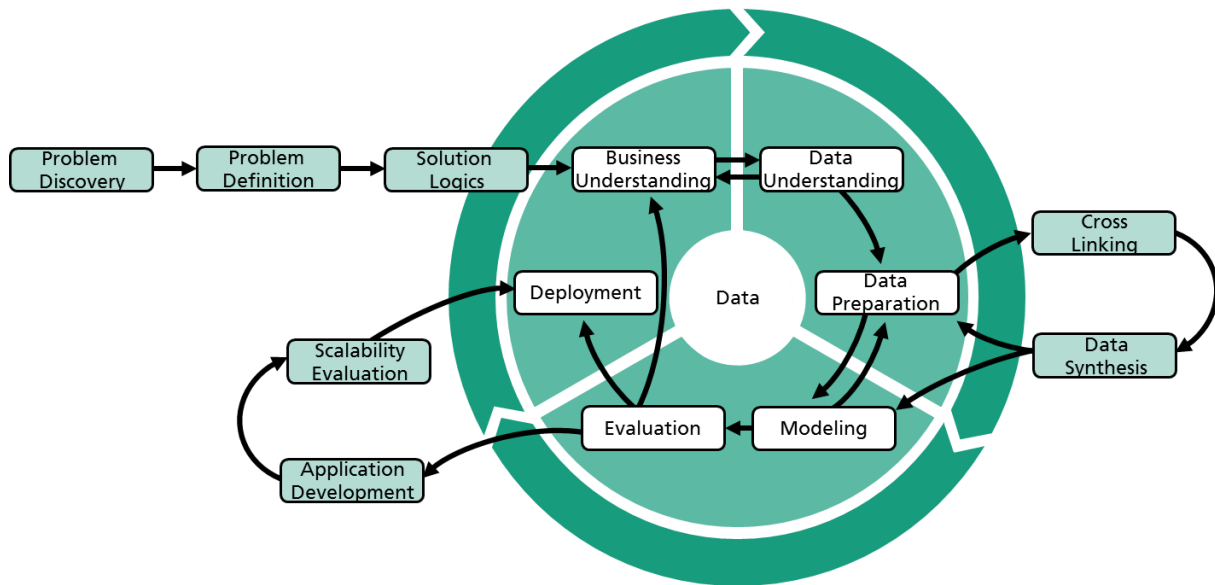


Abbildung 3: Erweiterter CRISP-DM Iteration 2

Zu Beginn eines Digitalisierungsprojekts findet in der Phase **Problem Discovery** ein divergenter Prozess zur Erforschung verschiedener Problemstellungen statt. Hier werden unterschiedliche Stakeholder hinsichtlich jeweiliger Perspektiven auf das Problem sowie möglichen Zielszenarien analysiert. So wird vermieden, dass lediglich die Problemstellung bzw. die Zielstellung einer einzigen Person bzw. Personengruppe einer Hierarchiestufe im Projekt verfolgt wird, ohne weitere Stakeholdergruppen (wie z.B. spätere Nutzer, Prozessbeteiligte, Informationsempfänger) zu vernachlässigen. In der Phase **Problem Definition** wird aus den verschiedenen Perspektiven aller Stakeholder eine Problemstellung abgeleitet und durch die Beschreibung einzelner Problemdimensionen ganzheitlich beschrieben. Für die einzelnen Problemdimensionen werden in der Phase **Solution Logics** jeweils mehrere Lösungsansätze erarbeitet und beschrieben. Hierbei ist es essentiell, Teams aus verschiedenen Stakeholdergruppen die Lösungsansätze erarbeiten zu lassen. Anschließend werden die Lösungsansätze der jeweiligen Problemdimension sowie deren Kombinationen evaluiert und zu einer ganzheitlichen Lösungslogik zusammengefügt. Darauf aufbauend startet mit der Phase **Business Understanding** der datenbasierte Teil des Digitalisierungsprojekts.

Nach der Definition der Lösungslogik sowie daraus hervorgehender Fragestellungen und Erfolgskriterien werden Data-Mining-Ziele abgeleitet, dokumentiert und kommuniziert. Nachfolgend werden in der Phase **Data Understanding** vorhandene Quellen für Trainings- und Evaluationsdaten identifiziert und sowohl die Qualität als auch die Granularität der Daten begutachtet sowie beschrieben. Mittels Standardoperationen der Datenbereinigung und Formatierung werden in der Phase **Data Preparation** die Daten in eine einheitliche Form gebracht. Nachfolgend werden in der Phase **Modelling** die Daten durch die Bildung verschiedener

Verarbeitungsmodelle analysiert. Die Analyseergebnisse werden in der Phase **Evaluation** hinsichtlich der in Phase Eins definierten Geschäfts- und Data-Mining-Ziele evaluiert. Löst ein Modell das definierte Problem in gewünschter Güte, findet in der letzten Phase **Deployment** die Integration in den Entscheidungsprozess statt. Zu diesem Zweck werden in der Phase **Application Development** alle Stakeholder eingebunden, um den Umfang und das Vorgehen der Implementierung der späteren Digitalisierungsanwendung (als Ergebnis des Digitalisierungsprojekts) zu definieren. Hierbei sind insbesondere die späteren Nutzer einzubinden. Anschließend findet in der Phase **Scalability Evaluation** eine Untersuchung aller Bestandteile der Anwendung hinsichtlich Skalierbarkeit und Übertragbarkeit statt.

5 Verwertung der Ergebnisse

Im Sinne der Verwertung der Ergebnisse sind in 2019 und 2020 detaillierte wissenschaftliche Veröffentlichungen der erarbeiteten Inhalte geplant. Die Ergebnisse können in ein Meta-Modell (oben schematisch dargestellt), ein Prozessmodell sowie ein Workflowmodell unterschieden werden. Zu jeder Phase bestehen zudem Methoden anderer Disziplinen (z.B. Lean Management, Business Canvas, etc.), die in Abhängigkeit der Freigabe der Geldmittelgeber veröffentlicht werden.

6 Literaturverzeichnis

Das Literaturverzeichnis enthält auch Quellen, die nicht unmittelbar im Bericht referenziert sind.

- [AZE08] Azevedo, A.; Santos, M. F.: KDD, SEMMA and CRISP-DM: a parallel overview, ISBN: 978-972-8924-63-8, 2008
- [KDN14] CRISP-DM, still the top methodology for analytics, data mining, or data science projects, online verfügbar unter: <https://www.kdnuggets.com/2014/10/crisp-dm-top-methodology-analytics-data-mining-data-science-projects.html>
- [GÜR13] 30 Minuten Design Thinking, ISBN 9783869364865
- [DIG13] Shared Ideas: Integration von Open-Innovation-Plattform-Methoden in Design-Thinking-Prozesse
- [VET12] Vetterli, C.; Brenner, W.; Uebernickel, F.; Berger, K.: Die Innovationsmethode Design Thinking, online verfügbar unter: <https://www.alexandria.unisg.ch/214442/1/ATTMMU9E.pdf>
- [BIT12] Bitkom, Big Data im Praxiseinsatz - Szenarien, Beispiele, Effekte, online verfügbar unter: <https://www.bitkom.org/Publikationen/2012/Leitfaden/Leitfaden-Big-Data-im-Praxiseinsatz-Szenarien-Beispiele-Effekte/BITKOM-LF-big-data-2012-online1.pdf>
- [BIT14] Bitkom, Big-Data-Technologien - Wissen für Entscheider: online verfügbar unter: <https://www.bitkom.org/Publikationen/2014/Leitfaden/Big-Data-Technologien-Wissen-fuer-Entscheider/140228-Big-Data-Technologien-Wissen-fuer-Entscheider.pdf>
- [BIT15] Bitkom - Big Data und Geschäftsmodell-Innovationen in der Praxis: 40+ Beispiele - Leitfaden
- [BRO08] Brown, T. et al.: Design thinking. Harvard business review, Juli 2009, online verfügbar unter: <https://hbr.org/2008/06/design-thinking>
- [CHA00] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C.; Wirth, R. (2000): CRISP-DM 1.0. Step-by-step data mining guide. Edited by CRISP-DM Consortium.
- [DDT18] Deep Design Thinking, online verfügbar unter: <http://deepdesignthinking.com/>
- [DEG18] Design Council: The Design Process: What is the Double Diamond?, online verfügbar unter: <https://www.designcouncil.org.uk/news-opinion/design-process-what-double-diamond>
- [DFA18] Design for America: We design tools for the network and beyond. Use these resources to apply creativity, design, and innovation to social challenges., online verfügbar unter: <http://designforamerica.com/resources/>
- [FAY96] Fayyad, U.; Shapiro-Piatetsky, G.; Smyth, P.: From Data Mining to Knowledge Discovery in Databases. AI Magazine Volume 17 Number 3 1996, online verfügbar unter:

<https://www.aaai.org/ojs/index.php/aimagazine/article/viewFile/1230/1131>

- [GOO18] Google: Design Sprint Kit, online verfügbar unter: <https://designsprintkit.withgoogle.com/methods/>
- [IBM16] Lucena, P.; Braz, A.; Tizzei, L.: IBM Design Thinking Software Development Framework; Conference Paper in Communications in Computer and Information Science, November 2016
- [IDS14] Hasso Plattner Institute of Design at Stanford: The Design Thinking Process, online verfügbar unter: <http://web.stanford.edu/group/cilab/cgi-bin/redesigningtheater/the-design-thinking-process/>
- [KRÜ13] Krüger, J.; Müller, J.: How can Design Thinking improve Software Engineering Processes? Online verfügbar unter: <https://hpi.de/dtrp/projekte/projekte-201213/how-can-design-thinking-improve-software-engineering-processes.html>
- [SAS17] Schebek, L.; Kannengießer, J.; Campitelli, A.; Fischer, J.; Abele, E.; Bauerdick, C.; Anderl, R.; Haag, S.; Sauer, A.; Mandel, J.; Lucke, D.; Bogdanov, I.; Nuffer, A.-K.; Steinhilper, R.; Böhner, J.; Lothes, G.; Zühlke, D.; Plociennik, C.; Bergweiler, S.: Ressourceneffizienz durch Industrie 4.0 – Potenziale für KMU des verarbeitenden Gewerbes. VDI Zentrum Ressourceneffizienz im Auftrag des Bundesministeriums für Umwelt, Naturschutz, Bau und Reaktorsicherheit
- [SAS17] SAS Institute Inc.: Data Mining Overview. Online verfügbar unter: <http://documentation.sas.com/?docsetId=emref&docsetTarget=p0pcscy5da8792n1agl4z85o9fnm.htm&docsetVersion=14.3&locale=e>
- [SCHRÖ16] Schröder. Herausforderungen von Industrie 4.0 für den Mittelstand, online verfügbar unter: <http://library.fes.de/pdf-files/wiso/12277.pdf>
- [SKH15] R. Steinhilper; F. Kübler; M. Hamacher: "Minimierung nicht wertschöpfender Energieaufwände durch energetische Anlagenoptimierung", Stuttgart, 2015.
- [SPA13] D. Hrsg. Spath: "Produktionsarbeit der Zukunft – Industrie 4.0", Fraunhofer IAO Leitstudie, Stuttgart, 2013.
- [UFL18] Uflacker, M.: Measuring the Impact of Design Thinking Software Development Processes and Products. Online verfügbar unter: <https://hpi.de/dtrp/projekte/projekte-201718/measuring-the-impact-of-design-thinking-on-software-development-processes-and-products.html>