

# Stücklistenprüfung mit Machine Learning und Data Science (Machbarkeitsstudie)



<b>Branche</b>	Kabelkonfektionierung
<b>Unternehmensgröße</b>	ca. 180 Mitarbeiter
<b>Projektkosten</b>	Personalkosten für 40 Tage
<b>Zeitaufwand</b>	10 Personentage innerhalb von 10 Monaten (Mikroprojekt).
<b>Mehrwert</b>	Personalkostenreduktion, Prozessoptimierung und IT-Integration

## Ausgangssituation

Die pema elektrotechnik gmbh fertigt Kabelkonfektionen, Kabelbäume und Baugruppen für unterschiedliche Branchen. Kunden erstellen dafür einen Kabelbaumbauplan, der u. A. aus einer technischen Zeichnung und Stücklisten dafür benötigter Teile besteht. Anhand dieser Dateien erstellt pema anschließend ein Angebot.

Damit einher gehen zwei Probleme:

1) Stücklisten unterschiedlicher Kunden werden in unterschiedlichen Formaten geliefert. Auch führen Kunden die Artikelnummern der Einzelteile auf den Stücklisten in ihrer eigenen Kodierung auf. Diese unterscheidet sich von pemas Einzelteilkodierung. Im Folgenden soll diese Herausforderung als **Übersetzungsproblem** bezeichnet werden.

2) Stücklisten sind fehlerhaft: Einzelteile fehlen oder falsche Einzelteile werden aufgelistet. Diese Problematik soll folgend **Ergänzungsproblem** genannt werden.

Um diesen Problemen entgegenzutreten, muss pema unter

Einsatz wertvoller Fachkräfte, Stücklisten händisch an das Format anpassen, umkodieren (die Einzelteilnummern der Kunden auf die pema-Einzelteilnummern ändern) und auf Richtigkeit und Vollständigkeit prüfen. Zwar werden für Formatanpassung und Umkodierung bereits Excel-Lösungen genutzt, weite Teile müssen aber trotzdem durch Erfahrung gelöst werden.

An dieser Stelle soll diese Machbarkeitsstudie ansetzen, wobei die zwei aufgeführten Problematiken gesondert betrachtet werden können. pema lieferte den Experten des Mittelstand 4.0-Kompetenzzentrum Stuttgart als Datengrundlage 173 Stücklisten, von denen jeweils sowohl das Kundenoriginal als auch die bearbeitete, verbesserte Version im pema-Standard vorlag. Unterschiede zwischen den zwei Stücklistenformaten sollten Informationen liefern. Die Stücklisten verteilten sich auf fünf Kunden. Der Kunde mit den wenigsten Stücklisten hatte dabei neun Stücklisten, der Kunde mit den meisten 80.

## Was ist die Lösung?

### Übersetzungsproblem

Kernstück der bisherigen Lösung ist eine sogenannte Artikelstammdatensdatei. In Form einer Exceldatei werden die Passungen zwischen Kundeneinzelteilcodes und der pema-internen Kodierung festgehalten. Pro Kunde besteht eine Tabelle, wobei i.d.R. eine Zeile folgende Informationen über das jeweilige Einzelteil enthält: Code beim Kunden, passender Code bei pema, Einzelteilbezeichnung, Lieferant der Einzelteile, uvm. Ziel ist, anhand der Artikelstammdatensdatei den Übersetzungsvorgang soweit wie möglich zu automatisieren – hin zu einer sauberen, standardisierten Stückliste.

Eine naheliegende Lösung dabei wäre, Kunden auf ein konkretes Stücklistenformat (inkl. standardisierter Codes) festzulegen. Leider erweist sich ein solches Vorgehen als schwer umzusetzen, da man einen direkten Marktnachteil erfährt: Kunden beauftragen andere Unternehmen, die beliebige Formate akzeptieren. Dennoch erwägt pema eine Vereinheitlichung auf Kundenseite, durch Anreize über bspw. den Preis.

Trotzdem kann zumindest die Umkodierung automatisiert ablaufen – benötigte Informationen sind schließlich vorhanden. Um das zu gewährleisten, muss die Artikelstammdatensdatei sauber gepflegt sein. Jede Ausnahme in der Datenbasis bedeutet eine verkomplizierende Fallunterscheidung im Algorithmus, die händisch hinzugefügt werden müsste. Um dem entgegenzutreten, bedeutet das konkret, dass **Data Governance Methoden** einzuführen sind:

- Ein/-e Artikelstammdatenverantwortliche/-r soll festgelegt werden. Ein kleiner Teil der Arbeitszeit soll sich der Instandhaltung und Standardisierung der Datei(en) widmen. Ebenso fungiert er/sie als Ansprechpartner/-in.
- Standardprozesse zum Hinzufügen neuer Daten in die

- Artikelstammdatensdateien sollen etabliert werden. Wer trägt was, wann, wie ein?
- Auch soll die Datendisziplin in den Fokus genommen werden: Es muss genug Zeit zum vollständigen Ausfüllen geben.

Gleichzeitig kann sogleich ein solides Datenmodell für die Artikelstammdatensdatei, das zukünftige KI-Potentiale bereits unterstützt, erarbeitet und implementiert werden:

- Zusätzliche Metadaten sollen erfasst werden, wie: *Teil\_hinzugefügt*, *Teil\_fehlerhaft*, *Bearbeitungszeit\_Korrektur*, etc.
- Anpassung an gängige Tabellenformate für KI-Algorithmen (Informationen werden ausschließlich zeilenweise gespeichert. In den vorliegenden Daten kam es vor, dass in einer Zeile mehrere Einzelteilcodes eingetragen waren).



#### Date Governance

Data Governance stellt das Rahmenwerk dar, welches die Grundlage für den Umgang mit und die Bewirtschaftung von Daten in einem Unternehmen für alle Stakeholder bildet. [1]

Diese Maßnahmen wurden anhand einer Ist-Zustandsanalyse erarbeitet und in einem Workshop kommuniziert. Sie sollen nicht nur eine Grundlage für das Übersetzungsproblem bieten, sondern gleichzeitig auch die Basis zukünftiger Innovationen durch KI festigen. Hierbei gilt die Devise: Zwar beansprucht saubere Datenhaltung kurzfristig eine Umgewöhnung und Zeit, auf lange Sicht zahlt sich diese Investition durch Automatisierung aber wieder aus.

Gefördert durch:



aufgrund eines Beschlusses  
des Deutschen Bundestages

## Ergänzungsproblem

Dieses Problem stellt aus Sicht des Machine Learnings (ML)/ Data Science eine interessante Herausforderung dar. Da sich die Datenbasis der vorliegenden Stücklisten auf eine für ML-Methoden zu geringe Anzahl beschränkt, wurden herkömmliche statistische Verfahren der Data Science genutzt, um die Daten auf die nötigen Informationen für die Lösung des Problems hin zu prüfen. An dieser Stelle wurde eine Analyse über gemeinsame Auftretenswahrscheinlichkeiten einzelner Stücklistenpositionen ausgewählt. Dazu werden sogenannte Bigramme erstellt: Einzelteile, die gemeinsam auf einer Liste vorkommen, erhalten eine Verbindung. Je öfter diese zusammen auf Listen vorkommen, desto stärker wird die Verbindung. Die darunterliegende Hypothese ist, dass eine Form funktionaler Untergruppen von Einzelteilen existiert – manche Teile treten fast immer mit gewissen anderen auf. Fehlt ein Gegenstück, ist die Wahrscheinlichkeit hoch, dass es sich um einen Fehler handelt.

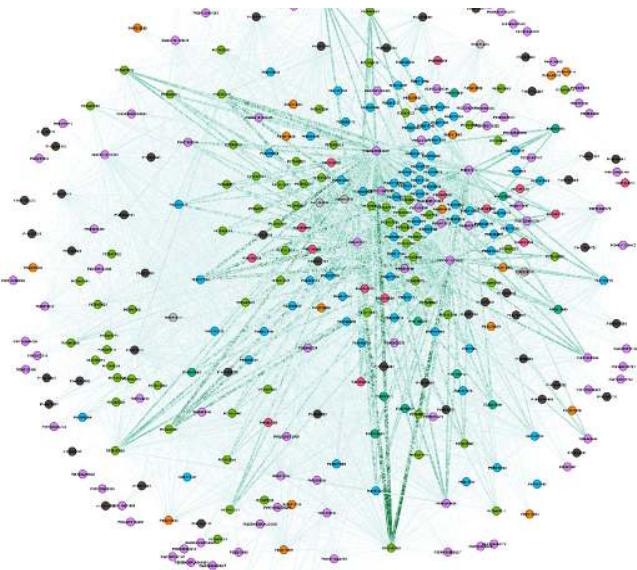


Abbildung 1: Erfahrungsgraph. Diese Visualisierung wurde durch den ForceAtlas2 Algorithmus [2] erstellt und soll in den Daten liegende Muster in Landkartenform illustrieren. Dabei ist jedes Einzelteil ein Kreis. Je dicker die Verbindungslinie, desto stärker die Verbindung zwischen dem Paar und desto öfter kam dieses Paar gemeinsam auf einer Stückliste vor.

Diese Zusammenhänge sind einem Spezialisten implizit bekannt und werden somit über die Zeit durch die Korrekturen der Stücklisten in die Datenbasis „hineinkodiert“. Diese Hypothese gilt es zu untersuchen. Abbildung 1 zeigt eine Landkarte der Bigramme. Je größer das Gewicht, desto näher zwei Einzelteile (Kreise). Und tatsächlich werden kleine „Inseln“ sichtbar. Nach Sichtung durch pema erweisen sich diese als tatsächlich deckungsgleich mit der Erfahrung. Die Hypothese funktionaler Untergruppen wird bestätigt. Hierbei muss unterstrichen werden, dass die Datenbasis von ca. 173 Stücklisten nur auf eine Tendenz und das Potential hinweist. Bei einer größeren Datenbasis ist davon auszugehen, dass sich mehr funktionale Untergruppen („Inseln“) abzeichnen und auch komplexere Zusammenhänge ersichtlich werden könnten. Weiterhin ist es auch möglich das Problem auf „N-Gramme“ zu erweitern – sich also nicht nur Paare, sondern auch Drillinge, Vierlinge, etc. anzusehen.

Da dieser Ansatz ausschließlich anhand der bereits korrigierten Versionen der Stücklisten funktioniert, kann pema zukünftig auf ihren kompletten Datenstamm von 17.000 Stücklisten im pema-Standard zurückgreifen (also Stücklisten aus der Vergangenheit, bei denen das Kundenoriginal verworfen wurde). Für dieses Projekt wurden nur die 173 Stücklisten betrachtet, bei denen Kundenoriginal und bearbeitete Version vorliegen.

Das bedeutet, dass durch diese große Datengrundlage eine Teilautomatisierung des Ergänzungsproblems durchaus möglich ist. Da in diesem Fall jedoch stets von Wahrscheinlichkeiten gesprochen wird, ist eine menschliche Überprüfung einer Meldung über fehlerhafte Stücklisten unumgänglich. Die angestrebte Lösung sollte also eine Form eines digitalen Assistenten sein, der die Fachkraft bei der Stücklistenprüfung bzw. der Fehlersuche unterstützt.

## Vorteile

Die angestrebte (Teil-) Automatisierung in beiden Aspekten der Stücklistenproblematik (Übersetzungs- und Ergänzungsproblem) würde Arbeitsstunden repetitiver Tätigkeiten einsparen. Durch Integration von Data Governance Methoden befähigt man außerdem das eigene Unternehmen dahingehend, für tiefere Digitalisierung empfänglich zu sein: Um erfolgreich KI-Methoden einsetzen zu können, müssen die Daten sauber sein.

Durch die aufgezeigten bereits bestehenden Potentiale in der kleinen Datenbasis (173 Stücklisten), kann durch eine Übertragung auf alle intern vorliegenden Daten (17.000 Stücklisten) ein Assistent umgesetzt werden. Dazu bietet sich bereits die hier angewendete Methode an.

## Kurz und Knapp

Anhand einer Ist-Zustandsanalyse wurden konkrete Handlungsschritte zur Implementierung von Data Governance Methoden erarbeitet. So können bei pema zukünftig Daten sauber erfasst und gehalten werden. Ebenso konnten Wege aufgezeigt werden, Formatstandards auf Kundenseite durch gezielte Anreize für die Kunden zu erreichen.

Die Machbarkeitsstudie konnte stabile Potenziale in der Datenbasis nachweisen. Diese unterstützen, bei ausreichender Menge an Daten, die Umsetzung eines intelligenten Assistenten. Dieser kann Fachkräfte auf potenzielle Fehler in Kabelbaum-Stücklisten hinweisen und damit zeitaufwändige, händische Überprüfung einsparen. Da der mögliche Aufwand zur Umsetzung überschaubar ist, sollten Effizienzgewinne auch nach Abzug der Umsetzungskosten für pema innerhalb relativ kurzer Zeit möglich sein.

*„Die Zusammenarbeit mit dem Mittelstand 4.0-Kompetenzzentrum Stuttgart gestaltete sich für uns sehr professionell. Die wissenschaftliche Herangehensweise an unsere Problemstellung veranschaulichte nicht erwartete Ergebnisse, die für uns einen enormen Mehrwert darstellen. Eine digitale Transformation in diesem Bereich könnte für uns einen weiteren Vorteil ggü. den Marktgleitern darstellen.“*

*- Nic Hahn, Mitglied der Geschäftsleitung*





Haben auch Sie Ideen oder Fragen zur Digitalisierung, dann wenden Sie sich an uns!

#### Ihre Ansprechpartner

Mittelstand 4.0-Kompetenzzentrum Stuttgart  
Jürgen Falkner / Themenfeldleitung Cloud-Computing & Smart Services  
Fraunhofer IAO  
Nobelstr. 12  
70569 Stuttgart  
[juergen.falkner@iao.fraunhofer.de](mailto:juergen.falkner@iao.fraunhofer.de)

Nic Hahn / Mitglied der Geschäftsleitung  
pema elektrotechnik gmbh  
Benzstraße 8  
72649 Wolfschlugen  
[Nic.Hahn@pema-elektrotechnik.com](mailto:Nic.Hahn@pema-elektrotechnik.com)  
Telefon 07022-957657  
<https://pema-elektrotechnik.com>

#### Was ist Mittelstand-Digital?

Mittelstand-Digital informiert kleine und mittlere Unternehmen über die Chancen und Herausforderungen der Digitalisierung. Die geförderten Kompetenzzentren helfen mit Expertenwissen, Demonstrationszentren, Best-Practice-Beispielen sowie Netzwerken, die dem Erfahrungsaustausch dienen. Das Bundesministerium für Wirtschaft und Energie (BMWi) ermöglicht die kostenfreie Nutzung aller Angebote von Mittelstand-Digital. Der DLR Projektträger begleitet im Auftrag des BMWi die Kompetenzzentren fachlich und sorgt für eine bedarfs- und mittelstandsgerechte Umsetzung der Angebote. Das Wissenschaftliche Institut für Infrastruktur und Kommunikationsdienste (WIK) unterstützt mit wissenschaftlicher Begleitung, Vernetzung und Öffentlichkeitsarbeit.

Weitere Informationen finden Sie unter [www.mittelstand-digital.de](http://www.mittelstand-digital.de)

Alle Praxisbeispiele des Mittelstand 4.0-Kompetenzzentrum Stuttgart finden Sie unter: <https://digitales-kompetenzzentrum-stuttgart.de/praxisinformationen/>

#### Impressum

##### Herausgeber und Redaktion

Mittelstand 4.0-Kompetenzzentrum Stuttgart c/o  
Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO  
Nobelstraße 12  
70569 Stuttgart  
Bildnachweis: pema elektrotechnik gmbh

Quellen: [1] Krotova A, Eppelsheimer J (2019) Was bedeutet Data Governance? Eine Clusteranalyse der wissenschaftlichen Literatur zu Data Governance. BMWi-Verbundprojekt: DEMAND - DATA ECONOMICS AND MANAGEMENT OF DATA DRIVEN BUSINESS  
[2] Jacomy M, Venturini T, Heymann S, Bastian M (2014) ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. PLoS ONE 9(6): e98679. <https://doi.org/10.1371/journal.pone.0098679>

Mittelstand-  
Digital 

##### Rechtsform

Das Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO ist eine rechtlich nicht selbstständige Einrichtung der Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.

Stand: Februar 2021

Gefördert durch:



aufgrund eines Beschlusses  
des Deutschen Bundestages