

Robotergesteuerte Prozessautomatisierung (RPA): Reifegradmodell zur Identifizierung RPA geeigneter Prozesse hinsichtlich der Dimension Daten

Robotic process automation (RPA): maturity model for identifying RPA-suitable processes with regard to the dimension of data

Christian Kutzner¹, Kilian Kuitert¹, Stjepan Jurisic³, Malte Stonis¹, Peter Nyhuis^{1,2}

IPH – Institut für Integrierte Produktion Hannover gGmbH¹

Leibniz Universität Hannover²

IPRI – International Performance Research Institute³

Für die Einführung von RPA im Unternehmen müssen in einem ersten Schritt geeignete Unternehmensprozesse identifiziert werden. Hierzu werden die Prozesse hinsichtlich verschiedener Auswahlkriterien bewertet, wie z. B. Komplexität oder Fallhäufigkeit. Ein weiteres Kriterium zur Auswahl und Bewertung von Prozessen sind die dem Prozess zugrundeliegenden Daten. In diesem Paper wird ein Reifegradmodell vorgestellt, welches das Auswahlkriterium Daten in die 5 bewertbaren Kategorien Digitalisierungsgrad, Datenmenge, Datenvarianz, Datenformat und Datenverantwortung untergliedert.

[Robotic Process Automation, RPA, Reifegradmodell, Daten, Automatisierung]

For the introduction of RPA in a company, the first step is to identify suitable business processes. For this purpose, the processes are evaluated with regard to various selection criteria, such as complexity or case-sensitivity. Another criterion for the selection and evaluation of processes is the data on which the process is based. In this paper, a maturity model is presented, which subdivides the selection criterion data into the 5 assessable categories digitisation level, data quantity, data variance, data format and data responsibility.

[Robotic Process Automation, RPA, maturity model, data, automation]

1 ROBOTIC PROCESS AUTOMATION

1.1 ÜBERBLICK

RPA hat sich als effizienzsteigernde Automatisierungstechnologie in Unternehmen erwiesen, die viele Arbeitsschritte von Mitarbeitern am PC obsolet macht und damit deren Arbeitsweise stark verändert [Rol18]. Im Zusammenhang mit RPA bezeichnet der Begriff „Roboter“ keinen klassischen Industrieroboter, sondern eine

Software. Es handelt sich daher um Softwareroboter (SR), die in die bestehende IT-Landschaft integriert werden können, sodass es nur geringfügiger Änderung der bestehenden Systeme und Softwarelösungen bedarf. Definiert wird RPA als „software to automate tasks previously performed by humans that use rules to process structured data to produce deterministic outcomes“ [Lac18, Kut22].

Mithilfe von RPA können beispielsweise administrative Prozesse in ERP-Systemen, die einen hohen Standardisierungsgrad sowie eine hohe Wiederholhäufigkeit aufweisen, aufwandsarm automatisiert werden. Beispielsweise können standardisierte Lieferscheine automatisiert ausgelesen und die entsprechenden Informationen mithilfe von RPA direkt in das ERP-System übertragen werden. Die Vorteile, die RPA hier insbesondere mitbringt sind, dass die SR ohne große Auswirkungen auf die bestehende Infrastruktur aufgesetzt werden können und die Bearbeitungszeit der ausgewählten Prozesse, bei einer niedrigen Fehlerquote, wesentlich reduziert werden kann [Kün19].

Darüber hinaus bietet RPA auch die Möglichkeit, mit anderen Technologien kombiniert zu werden, um auch semi-strukturierte Daten, wie z. B. die Daten aus einem Lieferschein per Optical Character Recognition (OCR) auszulesen. Dementsprechend bietet RPA, insbesondere auch für KMU, eine kostengünstige Lösung innerhalb der eigenen Systemarchitektur, um den Digitalisierungsgrad in Unternehmen zu steigern und insbesondere repetitive Prozesse in den administrativen Unternehmensbereichen aufwandsarm zu automatisieren.

Die Umsetzung von RPA führt in logischer Konsequenz zu einer Verfügbarkeitserhöhung der administrativen Prozesse in einem Unternehmen, da der eingesetzte SR nicht an gesetzliche Arbeitszeitregelungen gebunden und somit theoretisch 24 Stunden pro Tag verfügbar und arbeitsfähig ist. Dieses Potenzial ist an ein hohes Maß an Wiederholbarkeit und eine entsprechende

Eingangsfrequenz der Prozessschritte gebunden. Durch diese Vorteile kann eine Produktivitätssteigerung hinsichtlich der Auslastung der Mitarbeiter und zum anderen eine Kostensenkung infolge einer Reduktion der FTE (Vollzeit-äquivalente von Personal) erzielt werden [Cap17].

Zusammenfassend lässt sich der Nutzen von RPA in den folgenden vier Punkten darstellen [Sme19]:

- Der Digitalisierungsgrad im Unternehmen wird sukzessive erhöht. Es besteht dabei kein Bedarf zur vollen Systemintegration. RPA nutzt bestehende Systeme übergreifend und ohne Systemveränderungen. → Sukzessive Digitalisierung
- Fehler bei der Datenübertragung werden reduziert. RPA arbeitet grundsätzlich fehlerfrei und revisions-sicher. → Reduzierte Fehlerrate
- Vergleichsweise günstige Lizenzkosten sowie ein überschaubarer Aufwand für den Aufbau von Anwendungswissen bei den Fachabteilungen [Cap17]. → Niedrige Investitionskosten
- Der manuelle Aufwand der Mitarbeiter für administrative Aufgaben wird reduziert. Mitarbeiter haben dadurch mehr Zeit für wertschöpfende Tätigkeiten. → Freie FTE (Vollzeitäquivalente von Personal) können in KMU direkt eingesetzt werden; mit freien FTE kann der akute Fachkräftemangel abgeschwächt werden [Sto19]. → Schneller Return on Investment

1.2 ALLGEMEINE ANFORDERUNGEN BEI DER PROZESSAUSWAHL

Die Einführung von RPA ist mit unterschiedlichen Anforderungen in den Bereichen Technologie, Datenquellen, Organisation und Mensch verbunden (siehe Abbildung 1). Im nachfolgenden werden die Dimensionen der Technologie, Organisation und Mensch genauer erläutert. Die Dimension Daten wird anschließend im Kapitel 2 zum Reifegradmodell (RGM) erläutert.

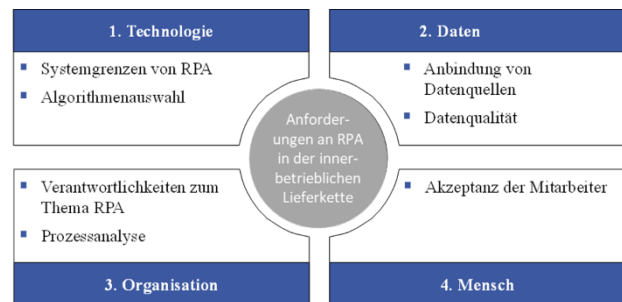


Abbildung 1: Anforderungsbereiche an RPA in der innerbetrieblichen Lieferkette

Technologie: Beim Einsatz von RPA können vier Umsetzungsstufen unterschieden werden: „Klassisches“ RPA, Cognitive Automation, Digital Assistants sowie Autonomous Agents (siehe Abbildung 2) [Kün19, Hor19, Ana17]. Der Grad an Automatisierung und abbildbarer Datenheterogenität vom klassischen RPA hin zu höher gelegenen Stufen steigt dabei sukzessive an.

Jede Stufe besitzt eigene Systemgrenzen in Hinblick auf die Einsatzmöglichkeiten der SR.

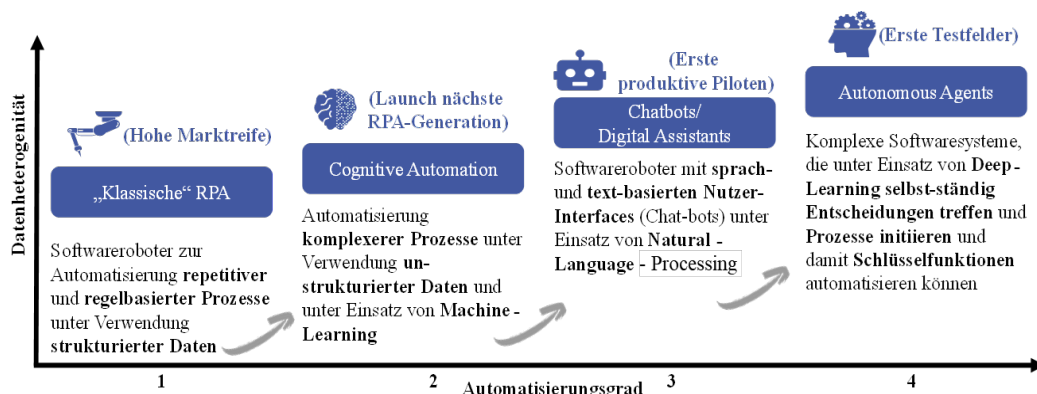


Abbildung 2: Entwicklungen von Softwarerobotern

Organisation: Für eine erfolgreiche RPA-Einführung ist es erforderlich, im Vorfeld eine RPA-Governance zu entwickeln. Sie erfüllt den Zweck, alle erforderlichen Richtlinien, Vorgaben und Verantwortlichkeiten für eine erfolgreiche Einführung für alle Bereiche des Unternehmens festzulegen [Lan21]. In der Governance sind Leitlinien verankert, die unter anderem die Organisation, die Strategie, die Prozesse, die Compliance und den Datenschutz betreffen [Lan21]. So werden auch die Vorgaben,

die Verantwortlichkeiten und die Techniken zur Identifikation geeigneter RPA-Prozesse von der Governance vorgegeben [Sme19]. Die IT-Governance deckt dabei die Rahmenbedingungen bezüglich der IT innerhalb des Unternehmens ab. Die RPA-Governance unter Einbezug der IT trägt dazu bei, die RPA-Technologie unternehmensweit zu etablieren und die vollen Nutzenpotenziale auszu-schöpfen [Pet20].

Mensch: Durch die digitale Transformation findet eine grundlegende Veränderung im Unternehmen statt. Eine Herausforderung ist diesbezüglich die fehlende Akzeptanz gegenüber einer neuen Technologie. Dies hängt insbesondere damit zusammen, dass die neue Technologie erstmal nicht greifbar ist und die Angst vor Arbeitsplatzverlust besteht [Buc20]. Die positive Wirkbeziehung zwischen der Akzeptanz von Technologien und deren Nutzung ist empirisch nachgewiesen [Sch16, Mer12, Lee04]. Hierbei hat sich insbesondere gezeigt, dass eine selbstbestimmende Motivation zur Nutzung der neuen Technologie entsteht, sobald die drei Grundbedürfnisse Autonomie, Kompetenz und Zugehörigkeit erfüllt sind [Buc20].

Fazit: Für eine zielgerichtete Einführung von RPA müssen zunächst geeignete Prozesse identifiziert und entsprechende technologische, datenbezogene, organisatorische und menschliche Anforderungen abgeleitet werden.

Im Folgenden erfolgt die genauere Betrachtung der Dimension Daten auf Basis des entwickelten RGMs.

2 REIFEGRADMODELL ZUR DIMENSION DATEN

Zur besseren Auswahl und Bewertung von Prozessen, die mit RPA automatisiert werden sollen, ist ein RGM hinsichtlich der Dimension Daten entwickelt worden. In diesem Kapitel wird in Abschnitt 2.1 ein Überblick über die in der bestehenden Literatur genannten Kategorien zur Bewertung von Daten gegeben. In Abschnitt 2.2 wird anschließend das ausgearbeitete RGM vorgestellt.

2.1 AUSWAHL GEEIGNETER DATENKATEGORIEN ZUM AUFBAU DES REIFEGRADMODELLS

Tabelle 1 bietet eine Übersicht über mögliche Kategorien zur Bewertung von Daten und stellt dar, wie oft diese in der untersuchten Literatur aufgegriffen und von RPA-Anbietern erwähnt werden. Die ermittelten Kategorien wurden auf ihre Anwendbarkeit zur Kategorisierung von RPA-Daten untersucht. Entscheidend war dabei nicht nur, wie häufig diese in der Literatur oder bei konkreten RPA-Softwareanbietern erwähnt wurden. Stattdessen wurde auch berücksichtigt, wie groß der Einfluss der jeweiligen Kategorien auf die Anwendung der Technologie RPA ist. Dabei wurden im Folgenden fünf zentrale Kategorien ausgewählt, die den größten Einfluss auf die Eignung der Daten eines Prozesses für die Automatisierung mit RPA haben. Diese werden in diesem Abschnitt nur kurz aufgeführt und im Kontext des RGMs in Abschnitt 2.2 mit Beispielen genauer beschrieben.

Digitalisierungsgrad: Der Digitalisierungsgrad ist von allen drei großen RPA-Anbietern aufgeführt worden, da die gesamte Automatisierung von Prozessen mittels RPA nur erfolgen kann, wenn die Daten digitalen Formaten vorliegen.

Datenmenge: Diese Kategorie ist nicht nur in der Literatur, sondern auch in der Praxis von großer Relevanz für RPA. Dies wird auch dadurch deutlich, dass diese Kategorie von der Literatur, den Softwareanbietern sowie in durchgeführten Interviews erwähnt wurde. Sie beeinflusst stark, wie gut sich Prozesse für die Automatisierung eignen. Die Datenmenge wird dabei auch in verschiedenen Kontexten betrachtet, wodurch die Kategorie „Frequenz“ mitbetrachtet wird.

Tabelle 1: Analyse bestehender Datenkategorien in der Literatur

Datenkategorie	Literatur					Anbieter		
	Laney et al. [Lan01]	IBM Corporation [IBM14]	Ishwarappa & Anuradha [IA15]	Stoffers, Karla & Kaufmann [SKK21]	Coito et al. [Co12]	IEC [IEC08] (25012)	UiPath	Blue Prism
Menge	●	●	●	●	○	●	●	○
Frequenz	●	●	●	○	●	○	○	○
Varianz	●	●	●	○	●	○	○	○
Wahrhaftigkeit	○	●	●	●	○	○	○	○
Nutzen	○	○	●	○	●	○	○	○
Verantwortlichkeit	○	○	○	●	●	○	○	○
Komplexität	○	○	○	○	●	○	●	○
Digitalisierungsgrad	○	○	○	○	○	○	●	●
Datentyp	○	○	○	○	○	○	○	●

Datenvarianz: Der Aspekt der Datenvarianz ist insbesondere im Kontext des Umfelds von Big Data zur Beschreibung der Datenbasis relevant. Im Kontext dieses Papers beschreibt dieser Begriff sowohl die in der Literatur oft erwähnte Varianz als auch der in der Praxis oft verwendete Begriff „Komplexität“. Weil die Komplexität der Daten und Datensysteme stark mit der Varianz in den Daten zusammenhängt, wurden beide Kategorien vereint.

Datenformat: Der Begriff Datenformat wird hier als Synonym für den von Softwareanbietern gewählten Begriff „Datentyp“ genutzt, um Verwechslungen mit der Begriffsdefinition aus der Programmierung zu vermeiden. Für das RGM beschreibt der Begriff verschiedene Dateiformate, welche einen starken Einfluss auf den für die Automatisierung notwendigen Aufwand haben. Unterscheidungen zwischen z. B. booleschen, ganzen oder natürlichen Zahlen, wie sie aus der Programmierung bekannt sind, besitzen hier keine Relevanz.

Datenverantwortlichkeit: Diese Kategorie wurde gewählt, um den notwendigen organisatorischen Hintergrund für die praktische Umsetzung des Modells zu beschreiben. Auch wenn theoretisch eine Umsetzung und Anwendung

durchführbar wäre, ist praktisch mindestens ein Ansprechpartner notwendig, der die Daten verwaltet und verwendet.

Im Folgenden wird das RGM sowie die fünf Kategorien näher erläutert und anhand von Beispielen verdeutlicht.

2.2 ÜBERSICHT REIFEGRADMODELL

Das entstandene RGM ist in Abbildung 3 dargestellt. Es besteht aus den zuvor beschriebenen fünf Kategorien Digitalisierungsgrad, Datenmenge, Datenvarianz, Datenformat und Datenverantwortung. Für jede Kategorie gibt es vier Stufen (Explorer, Tester, User, Leader (siehe Abbildung 3)), welche eine Einschätzung des Eignungsgrades eines Prozesses für die Anwendung von RPA ermöglichen.

Wenn sich ein Prozess in einer Kategorie in der ersten Stufe „Explorer“ befindet, ist eine Automatisierung mit RPA nicht möglich oder mit zu viel Aufwand verbunden, als dass es wirtschaftlich sinnvoll wäre. Zur Automatisierung mit RPA müssen die Daten gewisse Grundanforderungen erfüllen, die in dieser Phase nicht gegeben sind.

Wenn die Daten eines Prozesses in der zweiten Stufe „Tester“ einzuordnen sind, ist eine Umsetzung mit RPA meist theoretisch möglich. Eine praktische Umsetzung erfordert allerdings großen Aufwand und ist nur für Prozesse mit hohem Volumen und entsprechend großen Kosteneinsparungen wirtschaftlich. Zur Erhöhung der Wirtschaftlichkeit sollte auch hier zunächst der Fokus auf eine Verbesserung der Stufe gelegt oder ein Prozess aus einer höheren Stufe des RGMs ausgewählt werden.

Prozesse, bei denen sich die Daten in der dritten Stufe „User“ befinden, sind gut für die Automatisierung mit RPA geeignet. Sie bieten ein hohes Potenzial für Kosteneinsparungen und können aus Sicht der Daten mit relativ geringem zeitlichem Aufwand mithilfe von RPA automatisiert werden. Es kann sich trotzdem lohnen, die in den Prozessen verwendeten Daten weiter zu verbessern, um die letzte Stufe zu erreichen.

Die letzte Stufe „Leader“ des entwickelten RGMs enthält nur Prozesse, deren Daten sehr gut für die Automatisierung mit RPA geeignet sind. Diese Prozesse sind nicht nur schnell mithilfe eines SR automatisierbar, sondern im Vergleich zu Prozessen aus der Stufe „User“ auch zuverlässiger und einfacher zu pflegen sowie anzupassen. Ein Prozess, der sich hauptsächlich in dieser Stufe befindet, eignet sich besonders gut als erster Use Case, um Erfahrungen mit der Technologie aufzubauen und ein Gefühl für das theoretische Potenzial von RPA zu entwickeln.

In Abbildung 3 ist zudem eine kurze exemplarische Anwendung des RGMs dargestellt, in dem ein Prozess von drei zuständigen Mitarbeitern hinsichtlich der verschiedenen Kategorien bewertet und eingeordnet wurde. Die finale Stufe der jeweiligen Kategorie, in der Abbildung grün hinterlegt, richtet sich dabei immer nach der Konsensmeinung.

In den folgenden Abschnitten werden die einzelnen Kategorien sowie deren Stufen genauer erläutert.








Datenkategorien Robotic Process Automation	Explorer	Tester	User	Leader
Digitalisierungsgrad	Nur analoge Daten	Teilweise digitale Daten	Hauptsächlich digitale Daten, höchstens geringe Aufbereitung von Daten notwendig 	Nur digitale Daten, keine menschliche Interaktion notwendig
Datenmenge	Keine Daten vorhanden	Geringe Anzahl Daten vorhanden, mind. „Happy Case“ Keine Testdaten	Umfassende historische Daten, für alle Prozesspfade liegen Daten vor	Umfassende historische Daten für alle Prozesspfade Testdaten vorhanden 
Datenvarianz	Minimale Konsistenz in den Datenmengen, Datenformaten und Datenstrukturen	Geringe Konsistenz in den Datenmengen und der Datenstruktur, aber immer die gleiche Datenformate 	Gleiche Datenstruktur und Datenformate, Datenmengen variabel 	Datenstruktur, Datenformat und Datenmengen fast immer konsistent
Datenformat	Komplexe Videos, Bilder und Audiodateien	Strukturierte Bilder und Audiodateien Teilweise Zahlen und Texte	Hauptsächlich Zahlen und Texte, teilweise strukturierte Bilder 	Nur Zahlen und Texte Kopierbar und in festen Strukturen 
Datenverantwortung	Keine verantwortliche Person für den Prozess	Mehrere beteiligte und verantwortliche Personen	Mehrere beteiligte Personen, ein Verantwortlicher für den Prozess 	Datenmanagement oder Prozesseigner zuständig Alle Rechte zur Datenanpassung vorhanden

Abbildung 3: Reifegradmodell zur Auswahl von Prozessen für RPA hinsichtlich der Dimension Daten

2.2.1 KATEGORIE: DIGITALISIERUNGSGRAD

Datenkategorien Robotic Process Automation	Explorer	Tester	User	Leader
Digitalisierungsgrad	Nur analoge Daten	Teilweise digitale Daten	Hauptsächlich digitale Daten, höchstens geringe Aufbereitung von Daten notwendig	Nur digitale Daten, keine menschliche Interaktion notwendig

Die für einen Prozess zu verarbeitenden Daten können sowohl analog wie auch digital vorliegen. Wie jede reine Softwarelösung kann auch RPA nur Daten verwenden, die digital vorliegen. Daraus ergibt sich die Kategorie „Digitalisierungsgrad“ für die Bewertung von Daten. Der Digitalisierungsgrad bezeichnet folglich, wie weit die bei einem Prozess verwendeten Daten digitalisiert wurden.

Die niedrigste Stufe des Digitalisierungsgrades bilden dabei Prozesse, die nur analoge Daten verwenden. Prozesse ohne digitale Daten können nicht mithilfe von RPA automatisiert werden. Die zweite Stufe ist eine in Teilen fortgeschrittene Digitalisierung der Daten. In diesem Fall lassen sich die Teilprozesse, in denen digitale Daten verarbeitet und verwendet werden, möglicherweise mit RPA automatisieren. Die dritte Stufe ist von Prozessen geprägt, die hauptsächlich aus digitalen Daten bestehen. Wenn der Prozess zwischendurch von analogen Daten statt digitalen Daten unterbrochen wird, sinkt die Eignung deutlich, weil RPA keine analogen Daten verarbeiten kann. Idealerweise befinden sich höchstens am Anfang und am Ende des Prozesses analoge Daten, sodass die Teilprozesse nur auf digitalen Daten aufbauen. Die höchste Stufe des Digitalisierungsgrades bildet ein vollkommen digitaler Prozess. In diesem Prozess findet die Generierung der digitalen Daten direkt statt und es ist keine weitere menschliche Intervention notwendig. Weder am Anfang noch am Ende des Prozesses ist eine Umwandlung von analogen zu digitalen Daten notwendig. In diesem Kontext kann RPA die Rolle des Prozesseigners fast vollständig übernehmen.

2.2.2 KATEGORIE: DATENMENGE

Datenkategorien Robotic Process Automation	Explorer	Tester	User	Leader
Datenmenge	Keine Daten vorhanden	Geringe Anzahl Daten vorhanden, mind. „Happy Case“ Keine Testdaten	Umfassende historische Daten, für alle Prozesspfade liegen Daten vor	Umfassende historische Daten für alle Prozesspfade Testdaten vorhanden

Ein weiterer wichtiger Indikator für das Potenzial eines Prozesses zur Automatisierung mit RPA ist die Datenmenge, die im Prozess vorliegt. Der Begriff Datenmenge kann zunächst verschiedene Bedeutungen haben. Die verwendete Bedeutung im Bereich Big Data bezieht sich auf die Größe und die Anzahl der im Prozess verwendeten Daten [IBM14]. Es lässt sich folglich die Menge an Daten betrachten, die in einem Prozessdurchlauf verarbeitet werden. Je größer diese Menge ist, desto länger dauert ein Durchlauf eines SR. Andererseits lässt sich aber auch die Häufigkeit des Prozessdurchlaufs betrachten. So kann in einem Prozess z. B. nur eine geringe Menge an Daten verwendet werden, dieser Prozess muss aber sehr oft wiederholt werden, woraus schließlich große Datenmengen resultieren.

Zuletzt ist außerdem entscheidend, wie viele historische Daten zur Erstellung des SR vorliegen. Dieser Punkt ist wichtig, weil das Designen eines SR ohne Prüfen und Testen mithilfe von Real- oder Testdaten stark erschwert wird. Während große Datenmengen sich in vielen Situationen durch Skalierung mittels Parallelisierung von mehreren SR verarbeiten lassen, sind fehlende Daten für die Erstellung eines einzelnen SR schwer auszugleichen.

Die niedrigste Stufe für die Kategorie Datenmenge ist geprägt von fehlenden digitalen Daten. Ein SR kann nur mit sehr viel Aufwand ohne bereits vorhandene Daten oder bekannte Prozessabläufe erstellt und gepflegt werden. Aus diesem Grund eignen sich keine Prozesse, die neu sind und noch keine Daten generieren oder Testdaten besitzen. In der zweiten Stufe sind wenigstens für den problemfreien Standardfall Daten zur Modellierung des SR vorhanden. Idealerweise macht der Happy Case bereits einen großen Anteil der Prozessdurchläufe aus. Die alternativen Prozesspfade und notwendigen Daten in diesen Pfaden lassen sich dann langsam ergänzen. Die dritte Stufe ist durch die Anwesenheit umfassender Datenmengen zur Erstellung eines SR geprägt. Die Daten liegen hier für alle Prozesspfade vor und erlauben es theoretisch, den SR komplett zu designen, bevor er erstmals im operativen Betrieb eingesetzt wird. In der höchsten Stufe sind alle Daten vorhanden, die für eine fehlerfreie Implementierung von RPA nötig sind. Dazu gehören nicht nur umfassende historische Daten, die sämtliche Prozesspfade abdecken, sondern auch Testdaten, die sämtliche Abweichungsmöglichkeiten im Prozess beinhalten. Somit kann der SR nicht nur den Umgang mit allen bisher aufgetretenen Fehlern und Prozessabläufen lernen, sondern auch auf bereits antizipierte, theoretisch mögliche Fehler vorbereitet werden.

2.2.3 KATEGORIE: DATENVARIANZ

Datenkategorien Robotic Process Automation	Explorer	Tester	User	Leader
Datenvarianz	Minimale Konsistenz in den Datenmengen, Datenformaten und Datenstrukturen	Geringe Konsistenz in den Datenmengen und der Datenstruktur, aber immer die gleiche Datenformate	Gleiche Datenstruktur und Datenformate, Datenmengen variabel	Datenstruktur, Datenformat und Datenmengen fast immer konsistent

SR sind nicht intelligent. Sie folgen bestimmten, festgelegten Regeln, welche vorher klar definiert sein müssen. Abweichungen von diesen Regeln führen oftmals zum Scheitern eines SR. Wenn ein SR scheitert, können die Prozesseigner die gescheiterten Prozesse manuell übernehmen. Scheitert ein SR aber zu oft, lässt sich die Wirtschaftlichkeit der RPA-Lösung infrage stellen. Aus diesem Grund müssen sämtliche Abweichungen oder Varianzen der Daten möglichst gering gehalten werden. Die Varianz bezieht sich dabei auf alle Arten von Unterschieden, die bei den Daten innerhalb eines Prozessablaufs festgestellt werden können. Daten können in einem Prozess auf verschiedene Arten variieren. Ein Beispiel für die Varianz sind Unterschiede in den Datenmengen sowie Datenformaten je Prozessdurchlauf. Ein weiteres Beispiel wäre das Einpflegen von Besprechungsaufzeichnungen in eine Datenbank. Normalerweise liegen diese Aufzeichnungen im

Textformat vor und werden lediglich kopiert und übertragen. Wenn die Daten stattdessen im Bildformat vorliegen, z. B. als Zeichnung eines neuen Produkts, wird das Kopieren wahrscheinlich scheitern. Zuletzt kann auch die Struktur der Daten in sich stark variieren. Auch wenn immer dieselbe Menge an Daten im selben Datenformat vorliegt, müssen die Daten keine ähnliche Struktur aufweisen. Ein typisches Beispiel für diese Strukturprobleme ist die Verarbeitung von Angeboten im PDF-Format. Selbst wenn alle Angebote nur aus einer Kombination von Texten und Nummern bestehen, ist die Position und Anordnung dieser Datentypen auf den Seiten oftmals von Anbieter zu Anbieter sehr unterschiedlich. Eine reine RPA-Lösung ohne intelligente Komponente scheitert an solchen Strukturen.

In der niedrigsten Stufe der Kategorie Datenvarianz im RGM ist nur wenig Konsistenz über Datenmengen, Datenformate und Datenstrukturen vorhanden. Die verwendeten Daten im Prozess sind immer wieder anders aufgebaut, werden in anderen Formaten geliefert und weisen einen unterschiedlichen Umfang auf. Der Aufbau eines SR ist nicht möglich, da der Umfang aller möglichen Datenvariationen zu umfangreich ist und damit sämtliche Modellierungen entweder unmöglich oder zu aufwändig werden. Die zweite Stufe beinhaltet eine vorhandene Struktur in den Datenformaten. Die Daten müssen hier immer in denselben Datenformaten vorliegen, damit der SR auf die gleiche Art und Weise auf die Daten zugreifen kann. Außerdem muss der SR in der Lage sein, selbst die benötigten Informationen aus den zum Teil unstrukturierten Daten zu ermitteln. Dafür ist es notwendig, dass der Prozess zum Teil erweitert wird, z. B. muss der SR dann die Funktionalitäten der Software, welche die Daten darstellt, zum Auffinden der gewünschten Teildaten nutzen. In der dritten Stufe besitzen die Daten die gleiche Struktur und das gleiche Datenformat. Die Daten können vom SR konstant an den gleichen definierten Punkten abgegriffen werden und von der gleichen Software verarbeitet werden. Teilweise müssen noch komplexere Funktionalitäten in den SR eingebaut werden, damit dieser mit unterschiedlich großen Datenmengen je Prozessdurchlauf umgehen kann. Die höchste Stufe enthält nur konstante Datenstrukturen, Datenformate und Datenmengen. Der SR muss jetzt deutlich weniger Prozesspfade abdecken und kann schneller implementiert werden. Es gibt in dieser Stufe am wenigsten Möglichkeiten für den SR, auf Situationen zu treffen, in denen das einprogrammierte Vorgehen scheitert und die Aufgabe stattdessen von einem Menschen gelöst werden muss.

2.2.4 KATEGORIE: DATENFORMAT

Datenkategorien Robotic Process Automation	Explorer	Tester	User	Leader
Datenformat	Komplexe Videos, Bilder und Audiodateien	Strukturierte Bilder und Audiodateien Teilweise Zahlen und Texte	Hauptsächlich Zahlen und Texte, teilweise strukturierte Bilder	Nur Zahlen und Texte Kopierbar und in festen Strukturen

Daten können in vielen verschiedenen Datentypen bzw. Datenformaten auftreten. Zum Begriff Datenformat zählen hier zunächst fünf übergeordnete Medienformate: Zahlen,

Texte, Bilder, Audio und Video. SR ersetzen meist Prozesse, die auf einer Verknüpfung von Basisvorgängen beruhen. Beispiele für solche Basisvorgänge sind das Kopieren, Einfügen, Verschieben von Daten sowie das Ausführen von Vorgängen in Anwendungen durch Tastenkombinationen oder durch Interaktion mit der grafischen Benutzeroberfläche (GUI). Zentral für den SR ist dabei, ob er die Informationen aufwandsarm verarbeiten kann. Bei bestimmten Datenformaten ist dies fast immer gegeben. Ein Beispiel dafür sind Zahlen, die in einer TXT-Datei gespeichert sind. Diese können vom SR einfach markiert und kopiert werden, sowohl über die GUI als auch über Tastenkombinationen. Es gibt Datenformate, die so gut wie nie von SR verarbeitet werden können. Ein Beispiel dafür sind Videos, z. B. im MP4-Format. Während Menschen die Informationen aufnehmen, verarbeiten und mit dem gewonnenen Wissen andere Daten (z. B. eine Zusammenfassung in Textform) erzeugen können, hat ein SR aufgrund seiner fehlenden Intelligenz diese Möglichkeit nicht. Zuletzt sind auch Datenformate vorhanden, die nur teilweise mit SR verarbeitet werden können. Das Dateiformat PDF bietet ein gutes Beispiel. So könnte eine PDF-Datei nur aus Zahlen bestehen, die einzeln markiert und kopiert werden können. Alternativ könnten die gleichen Zahlen als Bild dargestellt werden, welches sich nicht kopieren lässt und nur mittels OCR vom SR verarbeitet werden könnte.

Die niedrigste Stufe des Datenformats im RGM ist gekennzeichnet von komplexen Video-, Audio- und Bilddateien. Diese Datenformate lassen sich nur sehr schwierig von SR verarbeiten. Die zweite Stufe enthält strukturierte Bilder und/oder Audiodateien, zum Teil möglicherweise auch noch Zahlen oder Texte. Strukturierte Bilder und Audiodateien können oft durch weitere Software oder Technologien wie Machine Learning analysiert werden, auch wenn die Auswertung fehleranfälliger und aufwändiger ist. In der dritten Stufe werden hauptsächlich Texte und Zahlen verwendet, strukturierte Bilder sind in einem geringen Anteil auch vorhanden. Die Auswertung der Texte und Zahlen läuft hier hauptsächlich über das Kopieren, aber auch Analysen von gescannten Dokumenten mithilfe von OCR sind möglich. In der höchsten Stufe gibt es keine Gefahr, dass der SR Daten fehlinterpretieren kann. Es werden nur Daten im Zahlenformat oder Textformat kopiert und sämtliche verwendeten Datenformate weisen eine feste Struktur auf.

2.2.5 KATEGORIE: DATENVERANTWORTUNG

Datenkategorien Robotic Process Automation	Explorer	Tester	User	Leader
Datenverantwortung	Keine verantwortliche Person für den Prozess	Mehrere beteiligte und verantwortliche Personen	Mehrere beteiligte Personen, ein Verantwortlicher für den Prozess	Datenmanagement oder Prozessleiger zuständig Alle Rechte zur Datenanpassung vorhanden

Nicht nur die Qualität der Daten selbst ist ausschlaggebend für die Eignung der Daten für RPA. Auch die Organisation, die den Prozess betreut und aktuell noch manuell umsetzt, ist entscheidend. RPA lebt von festen Regeln. Diese Regeln müssen allerdings auch bekannt sein oder definiert werden. Das ist nicht einfach, wenn die an den Prozessen beteiligten

Mitarbeiter sich nicht über die Regeln im Klaren sind, die sie zur Verarbeitung von Daten nutzen. Außerdem ist die Anwendung von RPA erschwert, wenn das Datenmanagement nicht vorhanden oder chaotisch ist.

geeignete Einführungsstrategie für RPA im Unternehmen aussehen kann, welche Prozessschritte hier insbesondere wichtig sind und was in diesem Kontext für die Unternehmen zu beachten ist.

Die niedrigste Stufe für die Datenverantwortung im RGM ist durch fehlende Zuständigkeiten gekennzeichnet. Ohne verantwortliche Personen oder geregelte Zuständigkeiten ist es schwer, mehr über den Prozess und die verwendeten Daten zu erfahren. Wenn doch sämtliche Informationen über die im Prozess verwendeten Daten gesammelt werden, ist immer noch nicht garantiert, dass die verwendeten Daten ihre Struktur erhalten und dauerhaft für den SR geeignet bleiben. Für die zweite Stufe müssen die Verantwortlichkeiten klar geregelt sein. Diese sind aber noch komplex und umfassend. So gibt es für die Daten eines Prozesses mehrere Verantwortliche, welche sich nur für die Daten eines Teilprozesses zuständig fühlen. Je mehr Personen Verantwortung für die Daten eines Prozesses tragen, desto mehr unterschiedliche Interessen und Meinungen müssen für jede Automatisierung und Anpassung berücksichtigt werden. In der dritten Stufe sind weiterhin viele Personen am Prozess beteiligt, aber es gibt nur eine sehr geringe Anzahl Personen, bestenfalls eine, verantwortliche Person. Diese Person hat nicht nur den Überblick über die im gesamten Prozess verwendeten Daten, sondern kann auch Veränderungen und Anpassungen der erstellten und verwendeten Daten für den SR autorisieren und falls nötig, den anderen Prozessbeteiligten gegenüber durchsetzen. Die höchste Stufe ist nicht nur von einem Prozessverantwortlichen, sondern auch von einem organisierten Datenmanagement für den Prozess gekennzeichnet. Dieses sollte nicht nur sämtliche im Prozess verwendeten Daten im Überblick behalten und bei der Erstellung des SR aushelfen, sondern auch wichtige Aufgaben wie die Bewahrung der Datensicherheit und Steigerung der Datenqualität vereinen.

3 ZUSAMMENFASSUNG UND AUSBLICK

In dem Paper sind verschiedene Kriterien (Technologie, Daten, Organisation, Mensch) erläutert worden, die bei der Einführung von RPA und der Auswahl geeigneter Prozesse eine entscheidende Rolle spielen. Zur genaueren Betrachtung und Bewertung der Prozesse hinsichtlich der Dimension Daten ist die Entwicklung eines RGMs sowie die Beschreibung hinsichtlich seiner einzelnen Einflussfaktoren erfolgt. Das RGM gliedert sich in die fünf Kategorien: Digitalisierungsgrad, Datenmenge, Datenvarianz, Datenformat und Datenverantwortung. Zudem ist zur einfachen Anwendung des RGMs die Einordnung hinsichtlich der einzelnen Kategorien anhand eines Beispielfalls erläutert worden.

Nachdem der Vorgang der Prozessauswahl für RPA in diesem Paper und der vorherigen Arbeit [Kut22] untersucht worden ist und geeignete Ansätze herausgearbeitet worden sind, soll in weiteren Arbeiten untersucht werden, wie eine

Literatur

- [Ana17] Anagnoste, S.: Robotic Automation Process - The next major revolution in terms of back office operations improvement, in: Proceedings of the International Conference on Business Excellence 11, 2017, 1, S. 676–686.
- [Buc20] Buchkremer, R.; Heupel, T.; Koch, O.: Künstliche Intelligenz in Wirtschaft & Gesellschaft, Springer Fachmedien Wiesbaden, Wiesbaden, 2020.
- [Cap17] Capgemini: Robotic Process Automation - Robots conquer business processes in back offices - A 2016 study conducted by Capgemini Consulting and Capgemini Business Services, 2017.
- [Coi21] Tiago Coito et al. „The impact of intelligent automation in internal supply chains“. In: International Journal of Integrated Supply Management 14.1 (2021), S. 1–27.
- [Her18] Hermann, K., Stoi, R., Wolf, B.: Robotic Process Automation im Finance & Controlling der MANN+HUMMEL Gruppe, in: Controlling: Zeitschrift für erfolgsorientierte Unternehmenssteuerung 30, 2018, 3, S. 28–34.
- [Hor19] Horváth & Partners: Next Generation Process Automation, 2019, <https://www.horvath-partners.com/media-center/studien/next-generation-process-automation/>.
- [IA15] Ishwarappa & J. Anuradha. „A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology“. In: Procedia Computer Science 48 (2015), S. 319–324.
- [IEC08] IEC. Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model. 2008. url: <https://www.iso.org/standard/35736.html>.
- [IBM14] IBM Corporation. Operational risk management in the world of big data: Unlocking the value of loss event data and driving the risk-aware enterprise. United States of America, 2014. url: <https://www.ibm.com/downloads/cas/716VKRPO> (besucht am 09. 11. 2021).
- [Kut22] Kutzner C, Jurisic S, Stonis M, Nyhuis P, Seiter M (2022). Robotergesteuerte Prozessautomatisierung zur softwarebasierten Automatisierung administrativer Prozesse der innerbetrieblichen Lieferkette (RPAlog). Logistics Journal : nicht referierte Veröffentlichungen, Vol. 2022.
- [Kün19] Künkele, S., Beukes, T.: RPA - Roboter im operativen Einkauf, Intelligente Prozessautomatisierung, 2019, <https://beschaffung-aktuell.industrie.de/e-procurement/intelligente-prozessautomatisierung/>.
- [Lac18] Lacity, M., Willcocks, L.: The next phase, Warwickshire, 2018.
- [Lan01] Doug Laney et al. „3D data management: Controlling data volume, velocity and variety“. In: META group research note 6.70 (2001), S. 1–4.
- [Lan21] Langmann, C.; Turi, D.: Robotic Process Automation (RPA) - Digitalisierung und Automatisierung von Prozessen, Springer Fachmedien Wiesbaden, Wiesbaden, 2021.
- [Lee04] Lee, J. D., See, K. A.: Trust in Automation, Designing for Appropriate Reliance, in: Hu-man Factors: The Journal of the Human Factors and Ergonomics Society 46, 2004, 1, S. 50–80.
- [Man18] Manutiu, S.: Digitalisierung im Controlling - Mehrwert durch Robotic Process Automation, in: Controlling 30, 2018, 3, S. 4–10.
- [Mer12] Mertens, P., Rässler, S.: Prognoserechnung, Siebte, wesentlich überarbeitete und erweiterte Auflage, Berlin, Heidelberg, 2012.
- [Pet20] Petersen, J.; Schröder, H.: Entwicklung einer Robotic Process Automation (RPA)-Governance. In: HMD. DOI: 10.1365/s40702-020-00659-y, Jg. 57 (2020), H. 6, S. 1130–1149.
- [Rol18] Roland Berger: The key to a successful RPA strategy, 2018, <https://www.rolandberger.com/de/Publications/Les-cl%C3%A9s-d%27une-strat%C3%A9gie-RPA-r%C3%A9ussie-Un-d%C3%A9fi-organisationnel-et-humain.html>.

- [Sch10] Schweizer-Ries, P., et al: Aktivität und Teilhabe – Akzeptanz Erneuerbarer Energien durch Beteiligung steigern. Projektabschlussbericht, Magdeburg & Berlin, 2010.
- [SKK21] Pascal Stoffers, Jürgen Karla & Jens Kaufmann. „Digitalisierung von Management-Reporting-Prozessen – Ein technologieorientiertes Reifegradmodell zum Einsatz in KMU“. In: HMD Praxis der Wirtschaftsinformatik (2021).
- [Sme19] Smeets, M., Erhard, R., Kaußler, T.: Robotic Process Automation (RPA) in der Finanzwirtschaft, Technologie – Implementierung – Erfolgsfaktoren für Entscheider und Anwender, 1. Auflage 2019, Wiesbaden, 2019.
- [Sto19] Stock-Homburg, R.: Personalmanagement, Theorien - konzeptinstrumente, 2019.
- [Wil17] Willcocks, L., Lacity, M. C., Sauer, C.: Outsourcing and offshoring business services, Cham, 2017.

Dr.-Ing. Malte Stonis, managing director, IPH – Institut für Integrierte Produktion Hannover gGmbH. Dr.-Ing. Malte Stonis, born in 1979, studied mechanical engineering at the Leibniz Universität Hannover. In 2006, he joined the IPH as a project engineer. In 2008, he became manager of the IPH process technology department. Malte Stonis has been coordinating managing director of the IPH since September 2016. His research focuses on increasing the efficiency and digitization of production, circular economy, disruption and humans in the digitized and automated production. Dr. Stonis is also a coach and mentor for startups in the field of production technology.

Prof. Dr.-Ing. habil. Peter Nyhuis, member of the management board of Institut für Integrierte Produktion Hannover gGmbH (IPH) and head of the Institute for Factory Plants and Logistics (IFA) at Leibniz University of Hannover.

Förderhinweis:

Das IGF-Vorhaben 21594 N der Bundesvereinigung Logistik (BVL) e.V. wird über die AiF im Rahmen des Programms zur Förderung der Industriellen Gemeinschaftsforschung (IGF) vom Bundesministerium für Wirtschaft und Klimaschutz (BMWK) aufgrund eines Beschlusses des Deutschen Bundestages gefördert.

M. Sc. Christian Kutzner, Project Engineer, IPH – Institut für Integrierte Produktion Hannover gGmbH. Christian Kutzner was born 1993 in Hanover, Germany. Between 2012 and 2018 he studied Industrial Engineering at the Leibniz University of Hanover.

Address: IPH – Institut für Integrierte Produktion Hannover gGmbH, Hollerithallee 6, 30419 Hannover, Germany, E-Mail: kutzner@iph-hannover.de

M. Sc. Kilian Kuiter, CoE Manager, Krone Business Center GmbH & Co. KG. Kilian Kuiter was born 1998 in Thuine, Germany. He studied Engineering and Business Administration at the Leibniz University of Hanover. Since April 2022 he works as a CoE Manager at Krone Business Center GmbH & Co. KG.

M. Sc. Stjepan Jurisic, Scientific Assistant, IPRI - International Performance Research Institute gGmbH. Between 2014 and 2020 he studied Business Administration and Economics at the University of Hohenheim and the Ruhr University in Bochum.

Address: IPRI - International Performance Research Institute gGmbH, Reuchlinstraße 27, 70176 Stuttgart, Germany, E-Mail: sjurisic@ipri-institute.com