



NEGZ

Kompetenznetzwerk
Digitale Verwaltung

Kurzstudie

38

Drei Fallstudien zum Process Mining

**Gestaltung geeigneter Voraussetzungen für das Erschließen von
verborgenem Prozesswissen in der öffentlichen Verwaltung**

Oliver Gutermuth · Alessandro Benke · Peter Fettke · Wolfgang Reisig



Inhaltlicher Ansprechpartner

Oliver Gutermuth	oliver.gutermuth@dfki.de
Alessandro Benke	a.e.benke@outlook.de
Peter Fettke	peter.fettke@dfki.de
Wolfgang Reisig	reisig@informatik.hu-berlin.de

Über das NEGZ

Das NEGZ · Kompetenznetzwerk Digitale Verwaltung ist Fachnetzwerk und Denkfabrik zur Digitalen Verwaltung.

Wir bündeln die Expertise von Unternehmen, Forschungseinrichtungen, öffentlichen Körperschaften und Verbänden, um die Digitalisierung der deutschen Verwaltung zu unterstützen und voranzutreiben.

Wir veröffentlichen Studien und Impulse, veranstalten Austauschformate, vermitteln Kompetenzen und bringen uns in die Fachdiskussion ein.

Impressum

Erscheinungsjahr 2025

ISSN 2626-6032

DOI 10.30418/2626-6032.2025.38



Dieses Werk ist nach „Creative Commons Namensnennung 4.0 International“ lizenziert. Sie dürfen das Werk bei Nennung der Urheberinnen und der Lizenz teilen und bearbeiten.

<https://creativecommons.org/licenses/by/4.0>

Herausgeber

NEGZ e.V.

Markgrafendamm 24 / Haus 16 · 10245 Berlin

030 7543 89 55

office@negz.org · www.negz.org

Gestaltung: Nicole Mank

Titelbild: Freepik

Kurzdarstellung des Themas

In Verwaltungsabläufen sind zunehmend Informationssysteme beteiligt, aus denen man mit Hilfe von Process-Mining-Techniken wertvolles Prozesswissen gewinnen kann. Eine zentrale Voraussetzung für die Anwendung von Process-Mining-Methoden ist die Extraktion und Aufbereitung relevanter Prozessdaten aus den beteiligten Anwendungssystemen. Hierbei existieren insbesondere bei der Analyse von Verwaltungsabläufen zahlreiche Herausforderungen, die das Erschließen vorhandener Potenziale erschweren. Die vorliegende Studie zielt darauf ab, die Voraussetzungen für erfolgreiches Process Mining in der öffentlichen Verwaltung zu erforschen und umsetzbare Handlungsempfehlungen für die Akteure zu entwickeln.

Schlagworte: Process Mining; öffentliche Verwaltung; Logdaten; Prozessmodellierung; HERAKLIT

Inhaltsverzeichnis

Zusammenfassende Empfehlungen	6
1 Einleitung	7
1.1 Motivation und Ausgangslage	7
1.2 Problemstellung	8
1.3 Zielsetzung, Methodik und Aufbau	9
2 Wissenschaftlicher und praktischer Hintergrund	11
2.1 Geschäftsprozessmanagement	11
2.2 Dokumentation von Prozessinformationen	11
2.3 Process Mining	12
2.4 HERAKLIT	14
3 Ergebnisse	16
3.1 Erhebung der Verwaltungsverfahren	16
3.1.1 KoPers.	16
3.1.2 IT-Beschaffungen BUKEA / BSW	20
3.1.3 JUS-IT	25
3.2 Beurteilung der Voraussetzungen für Process Mining	29
3.2.1 KoPers.	29
3.2.2 IT-Beschaffungen BUKEA / BSW	30
3.2.3 JUS-IT	31
3.3 Herausforderungen für Process Mining	31
3.3.1 Vorgehensweise zur Ermittlung von Herausforderungen	31
3.3.2 Identifizierte Herausforderungen	32
3.3.3 Zusammenfassung für untersuchte Fachverfahren	33

3.4 Maßnahmen für gute Voraussetzungen für das Process Mining.	36
3.4.1 Erläuterungen zu den vorgeschlagenen Maßnahmen	36
3.4.2 M1 – Abläufe vollständig mit IT unterstützen.	37
3.4.3 M2 – Standards festlegen (adressiert H2 und H7).	37
3.4.4 M3 – Sinnvoller Detailgrad für Logdaten	37
3.4.5 M4 – Gute Auswahl an Attributen.	37
3.4.6 M5 – Geeignete Menge an Attributsausprägungen	38
3.4.7 M6 – Gute Verständlichkeit erzeugen	38
3.5 Diskussion	38
4 Handlungsempfehlungen	42
5 Zusammenfassung	45
Literatur	47
Über die Autoren	49

Zusammenfassende Empfehlungen

Im Zuge der Digitalisierung entstehen in Organisationen immer mehr Daten, die Details zu den mithilfe von IT durchgeführten Aktivitäten festhalten. Seit einigen Jahren werden diese Daten zunehmend genutzt, um Prozesswissen zu generieren und Abläufe zu verbessern. Dies schließt auch die öffentliche Verwaltung ein. Techniken, die dies ermöglichen, werden unter dem Begriff *Process Mining* zusammengefasst. Um möglichst viele Potenziale zu erschließen und den Nutzen aus der Verfügbarkeit entsprechender Techniken zu erhöhen, sind Anwender bestrebt, optimale Voraussetzungen für den Einsatz von Process Mining zu schaffen. Wie dies gelingen kann und welche weiteren Aspekte zu berücksichtigen sind, wird in der vorliegenden Studie behandelt. Folgende Empfehlungen können als übergreifende Leitlinie dienen:

Status quo erheben

Damit die Voraussetzungen für den Einsatz von Process Mining in einem bestimmten Verfahren verbessert werden können, müssen diese zunächst bekannt sein. Somit muss der Analyse des Ist-Zustands besondere Bedeutung zukommen. Hierfür sind insbesondere die Kenntnisse der Prozessbeteiligten, dokumentierte Beschreibungen der Abläufe und die vorhandenen Logdaten zu berücksichtigen.

Dokumentieren und kommunizieren

Um die gesammelten Kenntnisse zu den Abläufen geeignet zu dokumentieren, empfehlen sich Modelle, die neben Prozessschritten auch Informationen zu den eingesetzten Systemen und relevanten Daten abbilden können. Die Studie verwendet hierfür HERAKLIT. Mithilfe der Modelle kann dann ein gemeinsames Verständnis für die Abläufe erzeugt werden, um mit Prozessbeteiligten potenzielle Maßnahmen hinsichtlich der Verbesserung von Voraussetzungen für Process Mining zu diskutieren.

Verbesserung der Logdaten

Alle Maßnahmen, die dazu dienen, die Erzeugung von Logdaten für festgelegte Ziele des Process Mining qualitativ und quantitativ zu verbessern, begünstigen auch die Ergebnisse und das Erschließen von Potenzialen aus der Anwendung von Process-Mining-Techniken auf diesen Daten. Wesentliche Aspekte sind hierbei Vollständigkeit, Umfang, Struktur, Variation und Formate der Logdaten. Sie sind entscheidend, damit sich aus technischer Sicht Details zu Abläufen geeignet in Daten niederschlagen und mit entsprechenden Analysen verwertet werden können.

Einheitliche Anforderungen

Aus der Organisationsgröße und Komplexität der öffentlichen Verwaltung resultiert sowohl die heterogene Systemlandschaft mit unterschiedlichsten Voraussetzungen für Process Mining als auch eine Vielfalt von Potenzialen, die mit diesen Techniken erschlossen werden könnten. Um abseits konkreter Maßnahmen für einzelne Fachverfahren Voraussetzungen zu verbessern, sollten gewisse Standards vereinbart werden. Diese vereinfachen Analysen mit Process Mining insbesondere bei system- und behördenübergreifenden Verfahren.

1 Einleitung

1.1 Motivation und Ausgangslage

Das Management von Prozessen stellt für die meisten Organisationen einen zentralen Schwerpunkt zur ständigen Ausrichtung und Erwägung erforderlicher Anpassungen dar. Öffentliche Verwaltungen sind hier keine Ausnahme. Eine Studie zur Digitalisierung von Landesverwaltungen kam 2019 zu dem Ergebnis, dass erheblicher Handlungsbedarf bestünde, um ein ganzheitliches *Geschäftsprozessmanagement* zu etablieren, weil fast die Hälfte der befragten Organisationen ihre Geschäftsprozesse noch nicht oder nur in geringem Umfang dokumentieren.¹ Systematische Optimierungen und die Digitalisierung der Abläufe wurden sogar in nur 14% der untersuchten Fälle bestätigt.

Aufgaben im Rahmen des Geschäftsprozessmanagements werden meist zyklisch umgesetzt, wobei für vorhandene Prozesse ausgehend von der tatsächlichen Ausführung unter Berücksichtigung ausgewählter Anforderungen und Ziele ein angepasstes Vorgehen entworfen und implementiert wird.² In welchem Umfang der neu geplante Prozess dann die Anforderungen und Ziele erfüllt und ob die Ausführung weitere evtl. ungeplante Konsequenzen nach sich zieht, lässt sich anschließend analysieren und beurteilen. Grundlage für derartige Untersuchungen ist auch eine geeignete Repräsentation durch dokumentierte Informationen oder Modelle.

Um Prozesse zu untersuchen, werden immer häufiger auch Daten einbezogen, die von der beteiligten Informationstechnologie (IT) erzeugt werden. Techniken, die der Rekonstruktion, dem Vergleichen oder weiterführenden Analysen von Prozessen auf Basis der auf diesem Weg erzeugten Daten stattfinden, werden unter dem Begriff *Process Mining* subsummiert. Insgesamt lässt sich im Zeitverlauf beobachten,

dass die Voraussetzungen, Process Mining zielführend einzusetzen, tendenziell immer besser werden, weil Prozesse zunehmend von IT unterstützt werden und hierbei auch bewusst detaillierte Informationen zu den einzelnen Bearbeitungsschritten gespeichert werden. Viele Modernisierungsstrategien werden unterdessen auch sehr prozessorientiert ausgerichtet, um gezielt Entlastung für strapazierte Bereiche zu schaffen, ausgewählte Potenziale freizusetzen oder bestimmte Fehler zu vermeiden. Die Herangehensweise mit Fokus auf einzelne Verfahren ist besonders wichtig für große, komplexe Organisationen und ist deshalb auch für die öffentliche Verwaltung von entsprechender Bedeutung.

Protokollierte Prozessdaten, die IT-Anwendungen weitgehend automatisiert speichern, enthalten wertvolle Informationen. Der Einsatz von Process Mining soll diese Informationen zugänglich machen und Nutzen entfalten. Doch die Daten, auf denen Process-Mining-Techniken angewandt werden sollen, fallen im Hinblick auf Umfang, Detailgrad und verwendete Formate häufig sehr heterogen aus. Gerade in der öffentlichen Verwaltung, wo Geschäftsprozesse oft mehrere Organisationseinheiten und diverse Anwendungssysteme einschließen, existieren für den Einsatz von Process Mining zahlreiche Herausforderungen, mit denen sich die vorliegende Studie beschäftigt. Hierzu zählen insbesondere die Fragmentierung, Heterogenität und Unvollständigkeit der Daten aber auch unklare Definitionen und damit Interpretationsspielräume bei der Auswertung dieser protokollierten Prozessinformationen.

Obwohl in der öffentlichen Verwaltung diverse Hindernisse für die Anwendung von Process Mining existieren, zeigen sich hier gleichzeitig große Potenziale, von den Techniken zu profitieren. Gerade die hohe Kongruenz vieler Prozesse für gleiche oder ähnliche Aufgaben verschiedener Verwaltungseinheiten erlaubt eine gemeinsame Verwertung

1 (AIOS, BearingPoint, Cassini, CGI, GBTEC, PwC, 2019)

2 (Houy, Fettke & Loos, 2010)

gewonnener Erkenntnisse, ohne dass im Vergleich zur Privatwirtschaft Zusammenarbeit und Erkenntnistransfer durch Wettbewerbssituationen gehemmt werden. Aus den Anreizen, gemeinsam Prozesswissen aufzubauen, Referenzprozesse zu entwerfen und Best Practices zu entwickeln, geht für den Einsatz von Process Mining in der Verwaltung ein erhöhtes Interesse hervor.

1.2 Problemstellung

In der Verwaltungspraxis ergeben sich für die Anwendung von Process Mining verschiedene Herausforderungen. Eine grundlegende Voraussetzung für Process-Mining-Techniken sind protokollierte Prozessdaten, die von der eingesetzten IT erzeugt werden. Doch auch unter der Annahme, dass ein Prozess umfassend von IT unterstützt wird, ergeben sich im Detail Probleme, die anhand des in Abbildung 1 aufgezeigten Beispiels erläutert werden sollen.

Abbildung 1 zeigt im oberen Teil in vereinfachter Darstellung einen Beispielprozess zur digitalen Verarbeitung von Antragsformularen, der mithilfe diverser Anwendungssysteme abgewickelt wird. Heutzutage speichert der überwiegende Teil der in der Praxis eingesetzten Software-Anwendungen Informationen zu durchgeführten Prozessen. Doch hierzu existieren keine einheitlichen Richtlinien und Anforderungen an die erzeugten Daten. Sie gehen meist von der vorgesehenen Verwendung des jeweiligen Entwicklers aus.

Dies hat u. a. Implikationen auf die Datenstrukturen, die Vollständigkeit, den Detailgrad und die Lesbarkeit bzw. Zugänglichkeit der Daten im Dateisystem der Anwendung. Dass die gespeicherten Prozessdaten so unterschiedlich ausfallen können, stellt den Einsatz von Process Mining bereits vor der Interpretation dieser Informationen vor diverse Probleme (vgl. Abbildung 1, Punkt 1).

Wenn Prozesse nun mehrere Anwendungen einschließen, müssen die Daten aufbereitet werden, um den gesamten Ablauf analysieren zu können. Dies adressiert die Zusammenführung und Vereinheitlichung der Datenstrukturen und -formate, um bspw. ein gemeinsames Format für die Dokumentation festgehaltener Zeitpunkte („Zeitstempel“) zu erhalten. Ergänzend ist es besonders wichtig, anwendungsübergreifend Abläufe zu kennzeichnen und jede Durchführung eines Ablaufs unterscheiden zu können. Eine derartige Aufbereitung umfasst die drei Teilschritte, *Extrahieren*, *Transformieren* und *Laden* (kurz: ETL, vgl. Abbildung 1, Punkt 2).

Liegen nach erfolgreicher Aufbereitung homogenisierte Logdaten vor, können Process-Mining-Techniken eingesetzt werden, um Teilschritte und chronologische Ordnungen in Verwaltungsabläufen aufzudecken, diese in Anlehnung an bestimmte Vorgaben oder Referenzen zu prüfen oder die Prozessinformationen durch Ergänzung weiterer Daten und Metriken zu verbessern (vgl. Abbildung 1, Punkt 3). Gewonnene Erkenntnisse können dann dazu genutzt werden, Strategien zur Erschließung ausgewählter Potenziale zu ent-

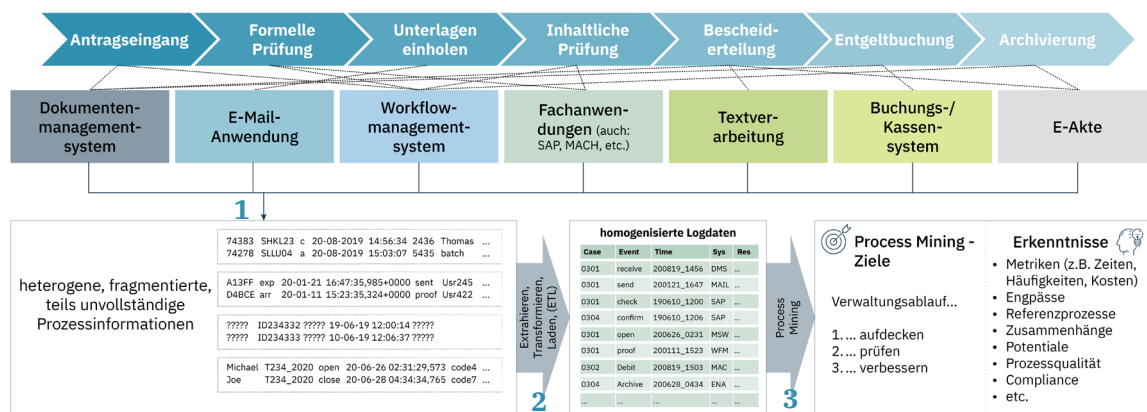


Abbildung 1: Herausforderungen für Process Mining in der Verwaltung

wickeln und diese bei einer entsprechenden Umsetzung auch nachzuhalten.

Zentrale Voraussetzung für den erfolgreichen Einsatz von Process Mining ist damit die Bereitstellung einer geeigneten Datenbasis. Entscheidend ist hier auch, in welchem Umfang ein bestimmtes Verwaltungsverfahren mit IT-Systemen unterstützt wird, da sich Vorgänge außerhalb solcher Systeme gar nicht in Daten niederschlagen. Die vorliegende Studie untersucht die herrschenden Voraussetzungen für die Anwendung von Process Mining anhand exemplarischer Fachverfahren der öffentlichen Verwaltung und stellt entwickelte Lösungsansätze zu identifizierte Herausforderungen vor.

1.3 Zielsetzung, Methodik und Aufbau

Das zentrale Ziel der Studie ist die Untersuchung der Gegebenheiten in der öffentlichen Verwaltung, um mit Hilfe von Process Mining verborgenes Prozesswissen aus Daten erschließen zu können. Hierbei sollen insbesondere Herausforderungen für den Einsatz dieser Techniken identifiziert werden. Anschließend werden Maßnahmen entwickelt und vorgeschlagen, die die Voraussetzungen für Process Mining verbessern und dadurch zusätzliche Potenziale in Aussicht stellen. Die Untersuchungen und Lösungskonzepte orientieren sich hierfür an drei exemplarischen Fachverfahren aus der Verwaltungspraxis:

- Bewerbungs- und Einstellungsverfahren für neue Mitarbeiter einschließlich der Personenneuanlage in **KoPers** (Kooperation Personaldienste)
- **IT-Beschaffungsvorgänge** in der Behörde für Umwelt, Klima, Energie und Agrarwirtschaft (BUKEA) sowie der Behörde für Stadtentwicklung und Wohnen (BSW)
- Bearbeitung von Fällen zur Heranziehung bei Unterhaltsvorschuss in **JUS-IT**

Zu diesen Verfahren werden im Rahmen von Experteninterviews sowohl fachliche Abläufe als auch Details zur Protokollierung der Prozessinformationen erhoben und analysiert. Die Fachverfahren werden mithilfe der auf Petrinetzen basierenden Modellierungsnotation HERAKLIT dokumentiert.³ Die Modellierungsnotation eignet sich, um die verschiedenen Dimensionen der Fachverfahren zu analysieren: Datenobjekte, deren dynamische Veränderungen im Verlauf des Verfahrens sowie die zugrundeliegende Systemarchitektur. Aus der Prozessdokumentation und bereitgestellten Logdaten können dann die Voraussetzungen für den Process-Mining-Einsatz abgeleitet werden. Identifizierte Herausforderungen und Probleme werden in passende Lösungsansätze und Handlungsempfehlungen überführt. Konkret stellen sich für die Studie folgende Forschungsfragen:

RQ1: Wie können die erhobenen Abläufe der drei Fachverfahren geeignet beschrieben werden?

RQ2: Welche Informationen der Abläufe in den drei erhobenen Fachverfahren werden als Logdaten gespeichert und stehen damit grundsätzlich für Process Mining Methoden zur Verfügung?

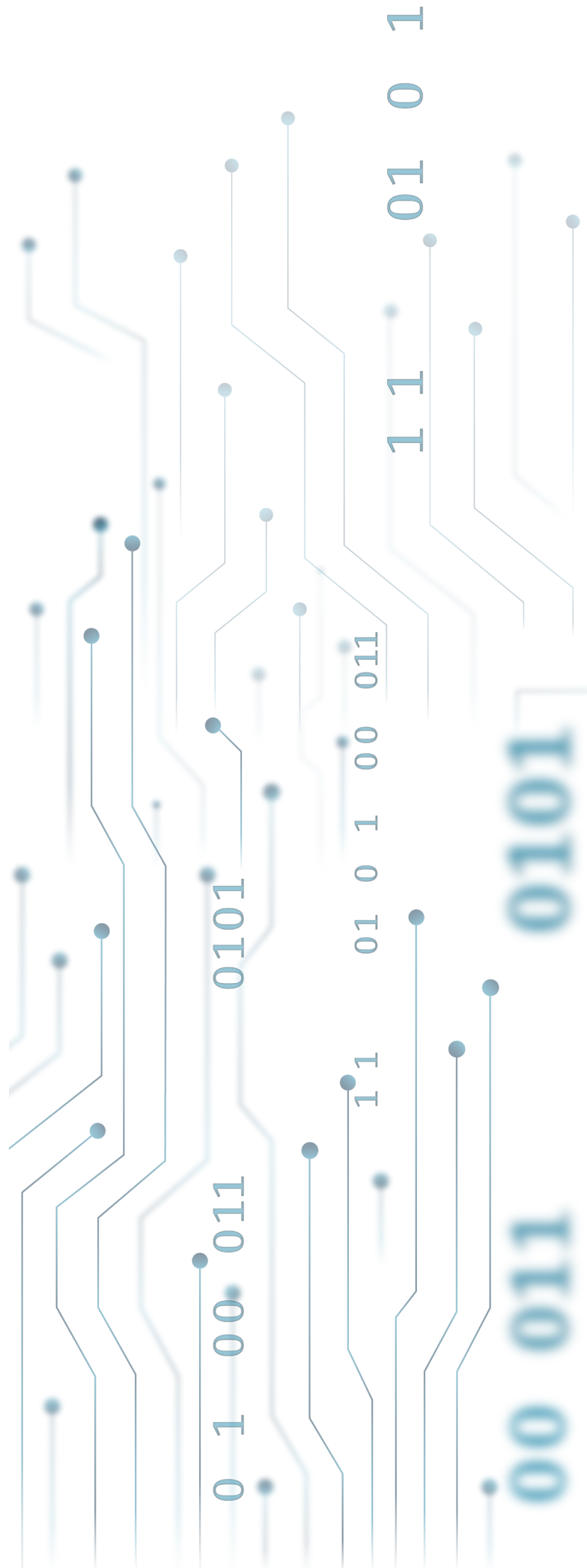
RQ3: Welche Herausforderungen für Analysen mit Process Mining gehen aus der Gegenüberstellung erhobener Abläufe und den jeweiligen Logdaten hervor?

RQ4: Welche Maßnahmen können den festgestellten Herausforderungen entgegenwirken und damit den Einsatz von Process Mining begünstigen?

Im Anschluss an den einleitenden ersten Abschnitt werden in Abschnitt 2 zentrale Terminologien erläutert, um ein grundlegendes Verständnis relevanter Themenfelder und Techniken zu vermitteln. Abschnitt 3.1 stellt die drei Fachverfahren aus der Praxis vor. Die Erhebung schließt dabei Experteninterviews, die Untersuchung schriftlicher Prozessdokumentationen sowie die Analyse exemplarischer Prozessdaten ein. Aus den

³ <https://heraklit.org>

Ergebnissen werden dann die jeweiligen Modelle entwickelt. Abschnitt 3.2 stellt eine konzeptionelle Analyse der Voraussetzungen vor, die bei der Anwendung von Process Mining im jeweiligen Fachverfahren relevant werden. Anschließend werden hieraus im Rahmen von Abschnitt 3.3 allgemeine Herausforderungen abgeleitet, die eine Übertragbarkeit auf weitere Verwaltungsbereiche zulassen. In Abschnitt 3.4 werden dann sinnvolle Maßnahmen erarbeitet und vorgestellt, um den identifizierten Herausforderungen zu begegnen. Sämtliche Ergebnisse werden in Abschnitt 3.5 diskutiert, woraus die in Abschnitt 4 vorgestellten zentralen Handlungsempfehlungen für die Verwaltung resultieren. Die Studie schließt dann in Abschnitt 5 mit einer Zusammenfassung und einem Fazit ab.



2 Wissenschaftlicher und praktischer Hintergrund

2.1 Geschäftsprozessmanagement

Geschäftsprozesse sollten sich regelmäßig an individuellen Erfordernissen einer Organisation ausrichten, um Leistungs- und Qualitätssteigerungen zu ermöglichen.⁴ Die Steuerung damit verbundener Maßnahmen wird als Geschäftsprozessmanagement bezeichnet. Es beschreibt den ganzheitlichen Ansatz zur Gestaltung, Einführung, Lenkung und Fortentwicklung der Geschäftsprozesse.⁵ Der Begriff umfasst Methoden, Konzepte und Techniken, mit denen sowohl operative als auch strategische Ziele erreicht werden sollen.⁶ Da sich Ziele von Organisationen regelmäßig ändern, wird Geschäftsprozessmanagement meist zyklisch beschrieben und unterscheidet mehrere Phasen, die iterativ durchlaufen werden (vgl. Abbildung 2).⁷



Abbildung 2: Regelkreis des kontinuierlichen Geschäftsprozessmanagements⁸

Nach Houy et al. (2010) ist die *Strategieentwicklung* der Ausgangspunkt eines jeden Zyklus

im Geschäftsprozessmanagement. Aus den entwickelten Strategien werden dann im Rahmen der *Definition und Modellierung* konkrete Maßnahmen abgeleitet und anschließend eingeführt (*Implementierung*) wird. Tangierte Geschäftsprozesse werden von nun an operativ umgesetzt (*Ausführung*), während das *Monitoring und Controlling* Informationen zwecks Evaluation sammelt, die zur *Optimierung und Weiterentwicklung* herangezogen werden können und dann meist einen neuen Zyklus einleiten.

2.2 Dokumentation von Prozessinformationen

Die Beschreibung von Prozessen kann sehr unterschiedlich gestaltet sein. Bestimmte Methoden besitzen dabei je nach Kontext Vor- und Nachteile. In komplexen Prozesslandschaften einer Organisation ist die Wahl eingesetzter Methoden zur Darstellung von Abläufen noch wichtiger, weil verschiedene Adressaten mit teils abweichenden Perspektiven auf die Prozessinformationen existieren. Im folgenden Teil sollen wichtige Aspekte bei der Beschreibung von Prozessen aufgegriffen werden und eine grobe Einordnung typischer Methoden vorgenommen werden.

Prozesse beschreiben strukturiert Zustände, zwischen denen kausale Zusammenhänge existieren. Optional können auch die Aktivitäten beschrieben werden, die den Wechsel der Zustände auslösen. Ein wesentlicher Aspekt bei der Darstellung eines Ablaufs, der unabhängig von der Methode wichtig ist, ist die Unterscheidung, ob ein einzelner, konkreter Ablauf oder ein allgemeines Schema für mehrere Varianten abgebildet wird. Während ein einzel-

4 (Becker, Mathas & Winkelmann, 2009)

5 (Bucher & Winter, 2009)

6 (Weske, 2007)

7 (Houy, Fettke & Loos, 2010)

8 (Houy, Fettke & Loos, 2010)

ner Ablauf genau eine Sequenz von Zuständen beschreibt, die auch nebenläufig auftreten können, fasst das allgemeine Schema mehrere Ablaufvarianten in einer gemeinsamen Beschreibung zusammen.

Sowohl ein einzelner Ablauf als auch ein allgemeines Schema kann unterschiedlich abgebildet werden. Besonders relevant hierfür ist der Verwendungszweck der Darstellungen in Verbindung mit den adressierten Zielgruppen. Verbreitete Möglichkeiten sind natürlichsprachliche Beschreibungen, Prozessmodelle und Logdaten.

Natürlichsprachliche Beschreibungen

erläutern Abläufe mithilfe von Text. Sie sind intuitiv zu verstehen. Einschränkend gilt jedoch, dass die Abläufe eine überschaubare Komplexität hinsichtlich enthaltener Details sowie insbesondere alternativer und nebenläufiger Teilprozesse aufweisen müssen, um Menschen bei der Interpretation der Informationen kognitiv nicht zu überfordern.

Für die Darstellung von **Prozessmodellen** existieren diverse Methoden, die sich im Detail stark unterscheiden können. Eine weitgehende Gemeinsamkeit ist die semiformale Abbildung des bezogenen Ablaufs, der mit Hilfe grafisch gekennzeichneten Objekttypen mit natürlichsprachlichen Labels in Zustände zerlegt wird und unter Verwendung gerichteter Kanten kausale Zusammenhänge kennzeichnet. Wenn ein Prozessmodell ein allgemeines Schema wiedergeben soll, enthält es ausreichend Freiheitsgrade, um verschiedene verteilte Abläufe abbilden zu können.

Logdaten sind strukturierte Daten, die Informationen zu Abläufen enthalten. Dabei werden eingetretene Ereignisse mit einem Zeitstempel gespeichert. Neben dem Ereignistyp und einem Zeitstempel umfassen Logeinträge in der Regel weitere Attribute wie bspw. eine Fallzuordnung (Case ID), welche die Zugehörigkeit des einzelnen Ereignisses zu genau einem

konkreten (einzelnen) Ablauf ermöglicht. Je nach Anwendung und Dokumentationszweck können weitere Attribute wie bspw. Angaben zu beteiligten Ressourcen (z. B. Mitarbeiter, Maschinen, etc.) vorhanden sein. Logdaten werden von Anwendungssoftware erzeugt und dienen in erster Linie der Protokollierung von Vorgängen. Sie können auch dazu verwendet werden, die protokollierten Abläufe zu analysieren. Die Datenstruktur schränkt jedoch die Möglichkeiten ein, allgemeine Ablaufschemen und komplexe Zusammenhänge ohne Weiteres zu erkennen. Um Menschen hier gewisse Einblicke zu ermöglichen, kann Process Mining eingesetzt werden.

2.3 Process Mining

Es existieren diverse Verfahren, um implizites Wissen in Datenbeständen zu entdecken und zu analysieren.⁹ So können mit Methoden des Data Minings nützliche, nicht-triviale Informationen aus Datenbeständen extrahiert werden.¹⁰ Ähnlich können auch prozessorientierte Daten aus der Protokollierung von Ereignissen untersucht werden und lassen dann u. a. eine gewisse Rekonstruktion von Vorgängen zu. Hierbei kommen je nach Fokus unterschiedliche Algorithmen zur Anwendung, die aus den Daten Prozessmodelle generieren können.¹¹ Nahezu jedes Informationssystem hält Prozessinformationen in Form protokollierter Log-Daten fest.¹² Werden solche transaktionsbezogenen Daten mit dem Ziel analysiert, Prozesse zu erkennen, zu überwachen und zu verbessern, wird dies als Process Mining bezeichnet.¹³

Die Analyse von Prozessen auf Basis von Ausführungsdaten legt ein allgemeines Vorgehen zu Grunde, das in 3 Phasen unterteilt werden kann und in Abbildung 3 gezeigt wird.¹⁴ Es wird verdeutlicht, dass die Ausführung von Geschäftsprozessen (*Execution Process*) bei entsprechender Anbindung an

⁹ (Düsing, 2006)

¹⁰ (Bissantz & Hagedorn, 1993)

¹¹ (Rozinat, Alves de Medeiros, Günther, Weijters & van der Aalst, 2008)

¹² (van der Aalst, Weijters & Maruster, 2004)

¹³ (van der Aalst, Adriansyah, de Medeiros & et. al., 2011)

¹⁴ (Cook & Wolf, 1995)

Informationssysteme Ereignisse protokolliert (*Collection*). Diese Daten können nun -ggf. auch unter Verwendung mehrerer Quellen- mittels Process Mining interpretiert werden (*Inference*), um Abläufe in einem Modell wiederzugeben (*Process model*).

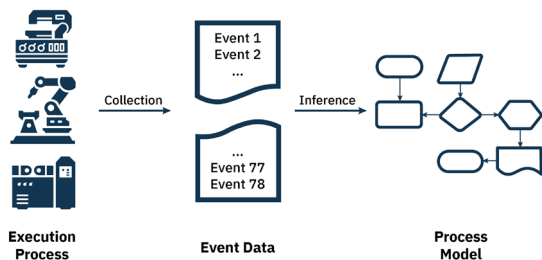


Abbildung 3: Leitfaden zum Process Mining¹⁵

Damit ein zielgerichtetes Process Mining betrieben werden kann, existieren Mindestanforderungen an die Protokollierung. So muss neben den zeitlichen Beziehungen zwischen Ereignissen auch eine Zuordnung jedes einzelnen Ereignisses zu einer konkreten Durchführung des Prozesses ersichtlich sein.¹⁶ Die Anwendung von Process Mining unterscheidet insbesondere zwischen den folgenden drei typischen Techniken:¹⁷

- **Process Discovery** erzeugt ein Modell aus Ereignis-Logs und ermöglicht somit eine Nachbildung untersuchter Vorgänge.
- **Conformance Checking** verwendet ein bereits vorhandenes Modell (bspw. Prozessmodell, Organigramm, Geschäftsregel, etc.), um dieses den Abläufen eines aus Ereignis-Logs resultierenden Prozesses gegenüberzustellen.
- **Process Enhancement** dient der Erweiterung oder Verbesserung eines vorhandenen Modells, indem es durch zusätzliche Informationen (bspw. die Ergänzung zeitlicher Beziehungen) angereichert oder verändert wird.

Weiterhin existieren für Process Mining bei der Verarbeitung von Log-Daten unterschiedliche Algorithmen, die signifikanten Einfluss auf die Interpretation der Daten und die Präsentation der Ergebnisse besitzen. Obwohl herausgestellt werden kann, dass jeder Algorithmus bestimmte Stärken und Schwächen besitzt, fällt eine direkte Gegenüberstellung schwer, da kein etabliertes Benchmarking zur Evaluation verfügbar ist.¹⁸ Aus diesem Grund empfiehlt sich je nach Anwendungskontext eine individuelle Auswahl.

Zur Umsetzung von Process Mining entwickelten sich im Zeitverlauf nach dem ersten Algorithmus, dem *Alpha Miner Algorithm* (van der Aalst et al. 2002) unter anderem der *Genetic Miner Algorithm* (Medeiros et al. 2005 und 2007), der *Fuzzy Miner Algorithm* (Bozkaya 2009; Gunther & van der Aalst 2007) sowie der *Heuristic Miner Algorithm* (Weijsters et al. 2006).¹⁹ Ein wesentliches Unterscheidungskriterium der Algorithmen stellt die Behandlung seltener Prozessvarianten dar. Für die Anwendung in der Praxis kann die Auswahl eines Algorithmus also beispielsweise davon abhängen, ob zu Gunsten der Übersichtlichkeit des Modells auf eine vollständige Abbildung aller aufgetretenen Varianten verzichtet werden kann.

Process Mining zur Ableitung von Prozessmodellen aus Logdaten verwendet in gängigen Softwareanwendungen der Praxis meist das Konzept des *Directly-Follows-Graphs* (DFG). Bei diesem Ansatz werden dokumentierte Ereignisse anhand ihres Zeitstempels chronologisch total geordnet. Diese Totalordnung trifft keine weiteren Aussagen zu Kausalitäten zwischen Aktivitäten und stellt eine pauschale Annahme dar, die wesentliche Aspekte der tatsächlichen Abläufe vernachlässigt. Entsprechend sind Prozessmodelle, die mit Hilfe entsprechender Discovery-Algorithmen des Process Minings erstellt wurden, kritisch zu hinterfragen und häufig verbesserungswürdig.

¹⁵ In Anlehnung an: (Cook & Wolf, 1995)

¹⁶ (van der Aalst, Adriansyah, de Medeiros & et. al., 2011)

¹⁷ (van der Aalst, Adriansyah, de Medeiros & et. al., 2011)

¹⁸ (Gupta, 2014)

¹⁹ (Mahendrawathi, Noval, Hanim, Renny & Rivia, 2018)

2.4 HERAKLIT

HERAKLIT ist eine Modellierungsinfrastruktur zur Beschreibung digitaler Systeme.²⁰ Sie erlaubt die Modellierung der *Statik*, *Dynamik* und *Architektur* von Systemen. Im Rahmen der *Statik* werden relevante Daten und Objekte einer algebraischen Struktur (Σ -Struktur) folgend spezifiziert. Die *Dynamik* bezeichnet Veränderungen lokaler Zustände eines Schemas, wodurch sich Vorgänge und die Transformation von Daten beschreiben lassen. Die Struktur von Daten wird in einer *Signatur* algebraisch beschrieben. Hierbei werden zwei wesentliche Darstellungsformen unterschieden. Während ein *allgemeines Schema* einen Prozess mit den erforderlichen Freiheitsgraden für alle relevanten Ablaufvarianten abbildet, beschreibt ein *verteilter Ablauf* genau eine Abfolge lokaler Zustände, die auch im Schema enthalten ist. Der objektorientierte Ansatz erlaubt die Definition von Modulen, die über Schnittstellen miteinander komponiert werden können. Diese Modulkomposition beschreibt die *Architektur* des betrachteten Systems.²¹

Mit HERAKLIT können aus protokollierten Logdaten eines Prozesses verteilte Abläufe erstellt werden. Abgeleitete allgemeine Schemata unterstellen im Gegensatz zu den Directly-Follows-Graphen klassischer Process-Mining-Ansätze keine totale Ordnung anhand der Chronologie aufgetretener Ereignisse und ermöglichen somit insbesondere die verbesserte Darstellung von Nebenläufigkeiten. Dies gelingt durch Einbeziehung zusätzlicher in Prozesslogs enthaltenen Kontextinformationen wie z. B. der beteiligten Ressourcen. Prozesslogeinträge werden also im Rahmen der Notation mit HERAKLIT nicht länger isolierten Prozessinstanzen zugeordnet, die dann ausschließlich mittels statistischer Methoden in ein gemeinsames Modell überführt werden können. Klassische Process-Mining-Ansätze nehmen hingegen bewusst hin, dass ein erzeugtes Prozessmodell vorhandene Loginformationen meist unvollständig verwertet. Dies äußert sich vor allem darin, dass

die zugrundeliegenden Systeme nur teilweise und die Ereignisse von Abläufen mangels Berücksichtigung von Kausalitäten sowie der unterstellten totalen Ordnung falsch abgebildet werden. Hier verfolgt HERAKLIT einen neuartigen ganzheitlichen Ansatz (vgl. Abbildung 4).

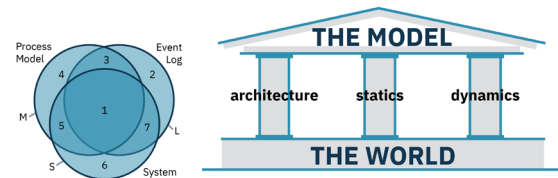


Abbildung 4: Konzepte klassischer Process-Mining-Techniken (links)²² und HERAKLIT (rechts)²³

Der zentrale Vorteil von HERAKLIT-Modellen wird in Abbildung 5 anhand eines Beispiels erläutert. HERAKLIT-Modelle beschreiben die Aktualisierung lokaler Zustände. Ein *Platz* (Ellipse) markiert dabei einen Zustand, eine *Transition* (Rechteck) einen lokalen Zustandsübergang.

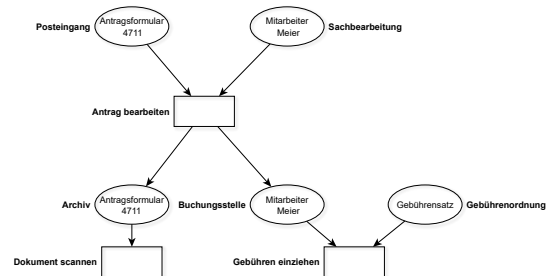


Abbildung 5: Verteilter Ablauf zur Demonstration lokaler Zustände.

Der in Abbildung 5 dargestellte Ablauf zeigt, dass der *Mitarbeiter Meier* im Rahmen des Prozessschrittes *Antrag bearbeiten* involviert ist, in dem der *Antrag 4711* bearbeitet wird. Zudem ist der gleiche Mitarbeiter im Rahmen des Prozessschrittes *Gebühren einziehen* beteiligt. Prozessschritte werden nicht als total geordnet verstanden. Dementsprechend werden im Rahmen der HERAKLIT-Methodik keine Symbolketten, sondern lokale Zustände auf der Ebene einzelner Prozessschritte abgebildet. Im klassischen Process Mining wäre

²⁰ (Fettke & Reising, 2024)

²¹ (Fettke & Reising, Breathing Life into Models: The Next Generation of Enterprise Modeling, 2022)

²² (Buijs, Dongen, & Aalst, 2014)

²³ (Fettke & Reising, Systems Mining with Heraklit: The Next Step, 2022)

es notwendig, eine statistische Auswertung über eine Menge von durch Symbolketten repräsentierten Prozessinstanzen durchzuführen, die Aufschluss über die Beteiligung des Mitarbeiters Meier an verschiedenen Schritten unterschiedlicher Fälle kennzeichnet. Der verteilte Ablauf in HERAKLIT stellt dies durch die Berücksichtigung beteiligter Ressourcen unmittelbar dar und zeigt entsprechend auf, dass jeweils eine kausale Beziehung zwischen *Antrag bearbeiten* und *Dokument scannen* sowie *Antrag bearbeiten* und *Gebühren einziehen* existiert.

Demgegenüber verwenden die Transitionen *Dokument scannen* und *Gebühren einziehen* keine gemeinsamen Ressourcen und lassen keine gegenseitigen Abhängigkeiten erkennen. Somit kann von einer Nebenläufigkeit der beiden Transitionen ausgegangen werden, die im Modell anders als bei den DFGs klassischer Process Mining Ansätze auch entsprechend abgebildet wird.



3 Ergebnisse

3.1 Erhebung der Verwaltungsverfahren

3.1.1 KoPers

3.1.1.1 Fachliche und organisatorische Einordnung

Einstellungsverfahren sind für die meisten Organisationen von hoher Relevanz. Ausgehend von einer vorhandenen oder geschaffenen Stelle werden über die Abwicklung des Bewerbungsverfahrens bis hin zur Einstellung einzelner Mitarbeiter zahlreiche Informationen verarbeitet und dokumentiert. Diese Vorgänge werden meist durch entsprechende Informationssysteme unterstützt. In Hamburg stellt das Zentrum für Personaldienste (ZPD) hierfür das Fachverfahren *Kooperatives Personalmanagement* (KoPers) bereit. Damit können angeschlossene Behörden Einstellungsprozesse IT-gestützt und einheitlich durchführen. Sie erhalten bei Bedarf seitens ZPD Hilfestellungen und technischen Support. Im

Rahmen der Erhebung wurde der Zusammenhang mit typischen Bewerbungsvorgängen erläutert und die Erfassung der Daten eines einzustellenden neuen Mitarbeiters exemplarisch demonstriert.

3.1.1.2 Modulstruktur

Abbildung 6 zeigt die für das Prozessmodell verwendeten Module und zugehörige Schnittstellen. Hier lässt sich erkennen, dass auf oberster Ebene zunächst zwischen *Kandidatenauslese* und *Onboarding* unterschieden wird. Die *Kandidatenauslese* umfasst die Vorgänge der ausschreibenden Organisationseinheit (*Abteilung Stellenausschreibung*) und die der Bewerber in jeweils separaten Modulen. Das Onboarding enthält Module zur Aktenanlage (*E-Akte erstellen*), die dazu eventuell erforderliche Beteiligung des neuen Mitarbeiters (*ausgewählter Bewerber*) und die Datenverarbeitung in KoPers. Hierbei wird innerhalb des Moduls KoPers zusätzlich zwischen der initialen Eingabe relevanter Information (*Datenerfassung*) und eventuell anfallenden Prüfungsvorgängen (*Prüfung*) unterschieden.

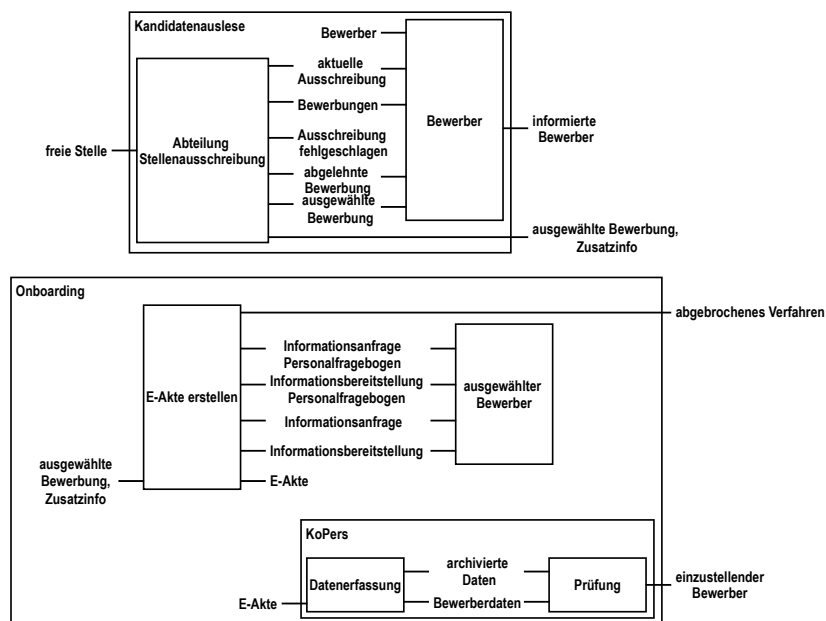


Abbildung 6: Architektur KoPers

3.1.1.3 Abläufe und zentrale Schnittstellen

Das Einstellungsverfahren wird dadurch ausgelöst, dass eine freie Stelle zu besetzen ist. Für diese freie Stelle wird von der einstellenden Behörde eine Ausschreibung erzeugt, woraufhin Bewerbungen eingereicht werden können. Eingehende Bewerbungen werden dann innerhalb einer bestimmten Frist gesammelt und ausgewertet. Gewöhnlich resultieren hieraus ein ausgewählter Kandidat und eine bestimmte Menge abgelehnter Bewerber. Diese werden von der ausschreibenden Organisationseinheit informiert.

Ausgehend von den Unterlagen des ausgewählten Bewerbers und den Zusatzinformationen zur adressierten Stelle werden mittels eines Personalfragebogens, den dieser Bewerber ausfüllen muss, weitere Bewerberinformationen in Erfahrung gebracht. Hierzu können eventuelle Rückfragen zur Klärung oder Ergänzung der Angaben erforderlich werden. Wenn alle erforderlichen Informationen eingeholt werden konnten, wird eine elektronische Akte erstellt. Sollten jedoch erforderliche Angaben fehlen und auch über das direkte Ansprechen des Bewerbers nicht einzuholen sein, kann dies zum Abbruch des Verfahrens führen.

Auf Basis der vollständigen E-Akte erfolgt die Datenerfassung in KoPers. Ein vollständiger Datensatz kann dann unter Umständen in Teilen zu prüfen sein. Zu prüfende Inhalte werden systemseitig anhand einer ermittelten Prüfwahrscheinlichkeit ausgewählt. Ausschlaggebend ist hier, wie kritisch bestimmte Angaben hinsichtlich der aus Fehlern resultierenden Konsequenzen ist und über welche Erfahrungen und Kompetenzen der erfassende Sachbearbeiter verfügt. Ein Prüfer evaluiert die zur Prüfung ausgewählten Informationen und nimmt gegebenenfalls Ergänzungen oder Korrekturen vor. Abschließend werden die Vorgänge aus Datenerfassung und eventueller Prüfung in KoPers archiviert und der einzustellende Bewerber kann hinsichtlich des weiteren Vorgehens informiert werden.

Abbildung 7 zeigt für KoPers das allgemeine Schema und Abbildung 8 die zugehörige Signatur. Abbildung 9 ergänzt eine exemplarische Instanziierung.

0101

01 0 1 00 011

0101

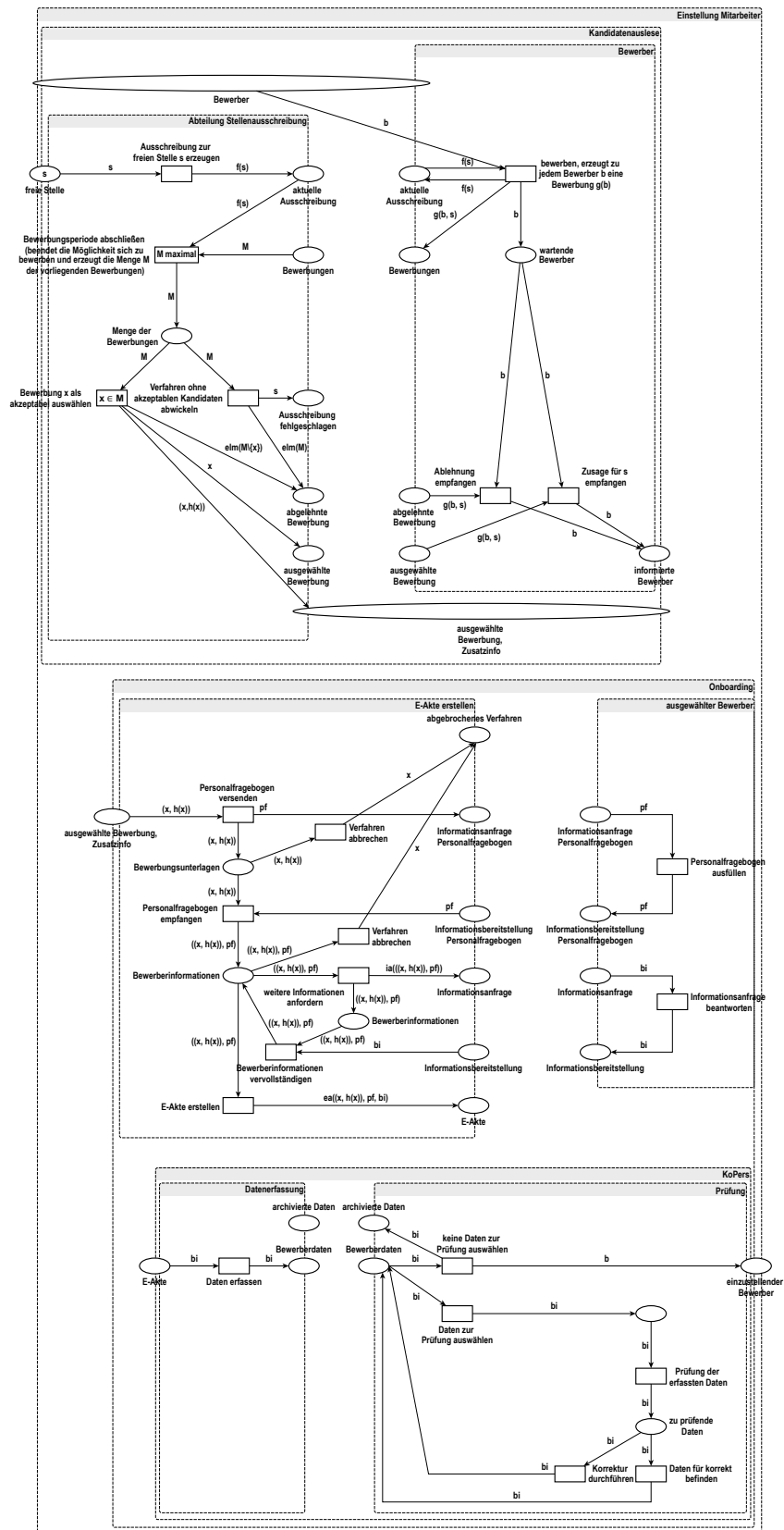


Abbildung 7: Allgemeines Schema (KoPers)

Signatur
<p>Grundsorten</p> <p>S Stellen</p> <p>A Ausschreibung</p> <p>B Bewerber</p> <p>U Bewerbungen</p> <p>Z Zusatzinformationen</p> <p>K Datenfeldbezeichnung</p> <p>V Datenfeldinhalt</p> <p>Abgeleitete Sorten</p> <p>$D \subseteq K \times V$ Datenfeld</p> <p>$BI \subseteq D \times \dots \times D$ Bewerberinformationen</p> <p>$PF \subseteq D \times \dots \times D$ Personalfragebogen</p> <p>Konstanten</p> <p>s: S</p> <p>Variablen</p> <p>d: D Datenelement</p> <p>b: B Bewerber</p> <p>x: U Bewerbung</p> <p>M: set of U</p> <p>p: P Prüfung</p> <p>pf: PF Personalfragebogen</p> <p>bi: BI Bewerberinformationen</p> <p>Funktionen</p> <p>f: $S \rightarrow A$</p> <p>g: $B \times S \rightarrow U$</p> <p>h: $U \rightarrow Z$</p> <p>ia: $U \times Z \times PF \rightarrow BI$</p> <p>ea: $U \times Z \times PF \times BI \rightarrow BI$</p>

Abbildung 8: Signatur des allgemeinen Schemas (KoPers)

Exemplarische Instanziierung
<p>Grundsorten</p> <p>S = {Planstelle_Nummer_53}</p> <p>A = {Ausschreibung_Sachbearbeiter_Bürgerbüro_Kennziffer_42}</p> <p>B = {Müller, Meier, Schmidt, Schulz}</p> <p>U = {Bewerbungsunterlagen_Müller, Bewerbungsunterlagen_Meier, Bewerbungsunterlagen_Schmidt, Bewerbungsunterlagen_Schulz}</p> <p>Z = {Einstellungstermin_01.01.2022}</p> <p>K = {Name, Straße, PLZ, Ort, Steuerklasse, E-Mail, IBAN }</p> <p>V = {Müller, Meier, Schmidt, Schulz, Hauptstr., Dorfstr., Bahnhofstr., Marktallee, 66121, 10115, 60306, 66862 Saarbrücken, Berlin, Frankfurt, Kaiserslautern, I, II, III, IV, mueller@web.de, meier@gmail.com, schmidt@yahoo.de, schulz@gmail.com, 1234, 4568, 8639, 0932}</p> <p>Abgeleitete Sorten</p> <p>D = {(Name, Müller), (Name, Meier), (Name, Schmidt), (Name, Schulz), (Straße, Hauptstr.), (Straße, Dorfstr.), (Straße, Bahnhofstr.), (Straße, Marktallee), (PLZ, 66121), (PLZ, 10115), (PLZ, 60306), (PLZ, 66862), (Ort, Saarbrücken), (Ort, Berlin), (Ort, Frankfurt), (Ort, Kaiserslautern), (Steuerklasse, I), (Steuerklasse, II), (Steuerklasse, III), (Steuerklasse, IV), (E-Mail, mueller@web.de), (E-Mail, meier@gmail.com), (E-Mail, schmidt@yahoo.de), (E-Mail, schulz@gmail.com), (IBAN, 1234), (IBAN, 4568), (IBAN, 8639), (IBAN, 0932) }</p> <p>PF = {(Name, Straße, PLZ, Ort, Steuerklasse, E-Mail, IBAN), ((Name, Müller), (Straße, Hauptstr.), (PLZ, 66121), (Ort, Saarbrücken), (Steuerklasse, I), (E-Mail, mueller@web.de), (IBAN, 1234)), (Mobilfunknummer, (Mobilfunknummer, 06131 123456)) }</p> <p>BI = PF</p> <p>Konstanten</p> <p>s = {Planstelle_Nummer_53}</p> <p>Funktionen</p> <p>f: $S \rightarrow A$</p> <p>f(Planstelle_Nummer_53) = Ausschreibung_Sachbearbeiter_Bürgerbüro_Kennziffer_42</p> <p>g: $B \times S \rightarrow U$</p> <p>g(Müller, Planstelle_Nummer_53) = Bewerbungsunterlagen_Müller</p> <p>g(Meier, Planstelle_Nummer_53) = Bewerbungsunterlagen_Meier</p> <p>g(Schmidt, Planstelle_Nummer_53) = Bewerbungsunterlagen_Schmidt</p> <p>g(Schulz, Planstelle_Nummer_53) = Bewerbungsunterlagen_Schulz</p> <p>h: $U \rightarrow Z$</p> <p>h(Bewerbungsunterlagen_Müller) = {Einstellungstermin_01.01.2022}</p> <p>h(Bewerbungsunterlagen_Meier) = {Einstellungstermin_01.01.2022}</p> <p>h(Bewerbungsunterlagen_Schmidt) = {Einstellungstermin_01.01.2022}</p> <p>h(Bewerbungsunterlagen_Schulz) = {Einstellungstermin_01.01.2022}</p> <p>ia: $U \times Z \times PF \rightarrow BI$</p> <p>ia((Bewerbungsunterlagen_Müller, Einstellungstermin_01.01.2022, ((Name, Müller), (Straße, Hauptstr.), (PLZ, 66121), (Ort, Saarbrücken), (Steuerklasse, I), (E-Mail, mueller@web.de), (IBAN, 1234)))) = (Mobilfunknummer, {})</p> <p>ea: $U \times Z \times PF \times BI \rightarrow BI$</p> <p>ea(((Bewerbungsunterlagen_Müller, Einstellungstermin_01.01.2022, ((Name, Müller), (Straße, Hauptstr.), (PLZ, 66121), (Ort, Saarbrücken), (Steuerklasse, I), (E-Mail, mueller@web.de), (IBAN, 1234))), (Mobilfunknummer, 06131 123456))) = ((Name, Müller), (Straße, Hauptstr.), (PLZ, 66121), (Ort, Saarbrücken), (Steuerklasse, I), (E-Mail, mueller@web.de), (IBAN, 1234), (Mobilfunknummer, 06131 123456))</p>

Abbildung 9: Exemplarische Instanziierung (KoPers)

3.1.1.4 Ablaufdokumentation und Logdaten

Da das genaue Vorgehen zur Ausschreibung einer Stelle und der Bewerberauswahl maßgeblich von der einstellenden Organisationseinheit abhängt, existieren hier keine einheitlichen Prozesse und damit auch keine gemeinsame Grundlage für die Erstellung von Logdaten. Gleiches gilt für den Dialog mit dem Bewerber zum Einholen relevanter Informationen und der Erstellung der E-Akte. Daten, die sich für eine Auswertung mit Process-Mining-Techniken eignen, entstehen erst im Rahmen der Anwendung von KoPers. Hier werden systemseitig alle Schritte der Erfassung und Prüfung sämtlicher Informationen feingranular dokumentiert. Dies umfasst die Unterscheidung von Benutzeraktivitäten in den einzelnen Datenfeldern der zu durchlaufenden Eingabemasken. Die erstellten Logdaten zeigen hierbei folgende Detailinformationen (vgl. Abbildung 10).

Signatur des Logs
<p> Grundsorten D Datum Z Uhrzeit NN Name BN Benutzername VN Vorname N natürliche Zahl B = {Tarifbeschäftigte} Beschäftigungsart A = {Neuanlage, Betaetigung} Aktion DB = {Person BV Zulage/Vst. E0, Person Unfallversicherung, Person BV/Sachbearbeiterzuordnung, Person BV/Sozialversicherungsdaten, Person Krankenversicherung, Person Steuerdaten, Person BV/SV-Nummer+Meldeverfahren, Person BV/Zusatzversorgung, Person BV/Steuerdaten, Person Arbeitszeit, Person BVL-Daten, Person BVL/Entwicklungsstufen, Person Anschrift, Person Bankverbindungen, Person BV Stammorganisation, Person Beschäftigungsverhältnis, Person BV/Vertragschronologie, Person, Person Familienstand, Aktionsbutton} Datenbestand F Datenfeld GM = {Person, Nur Person} Geschäftsmodul FN = {Kontierung FHH Doppik, Unfallversicherung II, SB-Zuordnung Person, Sozialversicherung, Person Krankenkasse, Steuerdaten Person, Meldeverfahren FHH, Beitrag Zusatzversorgung (bAV) FHH, Steuer FHH, Arbeitszeit FHH, Eingruppierung FHH, Private Erreichbarkeit, Bankverbindung FHH, Stammdienststelle, Arbeitsverhältnis Beschäftigte, Neuanlage Person} Formatname Z Zeichenkette Abgeleitete Sorten SD \subseteq D Systemdatum GD \subseteq D Geburtsdatum AV \subseteq D Abrechnungsstand Versorger PN \subseteq N Personalnummer VNR \subseteq N Vertragsnummer ERF \subseteq BN Erfasser AW \subseteq Z alter Wert NW \subseteq Z neuer Wert </p>

Abbildung 10: Signatur des Logs (KoPers)

Bei der Erfassung von Daten in KoPers navigiert der Benutzer durch verschiedene Eingabemasken und Formulare. In den Prozesslogs zeigt sich also auch immer eine gewisse Bearbeitungsreihenfolge. Ob eine gewählte

Reihenfolge jedoch begründete Ursachen hat, willkürlich entstanden ist oder maßgeblich auf das Design der Benutzeroberfläche zurückzuführen ist, lässt sich mit klassischen Process Mining Techniken nicht untersuchen.

3.1.2 IT-Beschaffungen BUKEA / BSW

3.1.2.1 Fachliche und organisatorische Einordnung

IT-Beschaffungsprozesse sind Vorgänge, bei denen Hardware, Software oder IT-bezogene Services angefordert, evtl. eingekauft und Nutzern bereitgestellt werden. Hierbei sind häufig bestimmte Anforderungen wie beispielsweise Präferenzen hinsichtlich der Anbieter und Lieferanten aber auch Richtlinien zu Standards, Sicherheitsaspekten, Vertragsgestaltung, etc. zu berücksichtigen. Aus diesem Grund werden derartige Beschaffungsprozesse meist unter Einbeziehung einer darauf spezialisierten Organisationseinheit durchgeführt. Dies gilt auch für die öffentliche Verwaltung. Die IT-Beschaffungen der Behörde für Umwelt, Klima, Energie und Agrarwirtschaft (BUKEA) und der Behörde für Stadtentwicklung und Wohnen (BSW) der Hansestadt Hamburg wird von der BUKEA Abteilung Z4 (IT-Management und Digitalisierung) koordiniert. Bedarfe werden primär von Dataport gedeckt. Eine Besonderheit ist hierbei der von Dataport bereitgestellte Servicekatalog (ITSM), der den Nutzern für ausgewählte Leistungen die Möglichkeit bereitstellt, diese direkt von Dataport zu beschaffen, ohne Z4 einzubinden. In anderen, weniger standardisierten Fällen holt Z4 passende Angebote ein und beauftragt den jeweiligen Anbieter entsprechend.

3.1.2.2 Modulstruktur

Die Module des entwickelten HERAKLIT-Modells beschreiben in Abbildung 11 auf der linken Seite die *beschaffende Behörde* (BUKEA / BSW) und auf der rechten Seite mit *Dataport* und *Dritten* die Anbieter entsprechender IT-Leistungen. Innerhalb des Moduls *Beschaffende Behörde* wird zusätzlich zwischen *Abteilungen*, von denen der Bedarf für eine bestimmte IT-Leistung

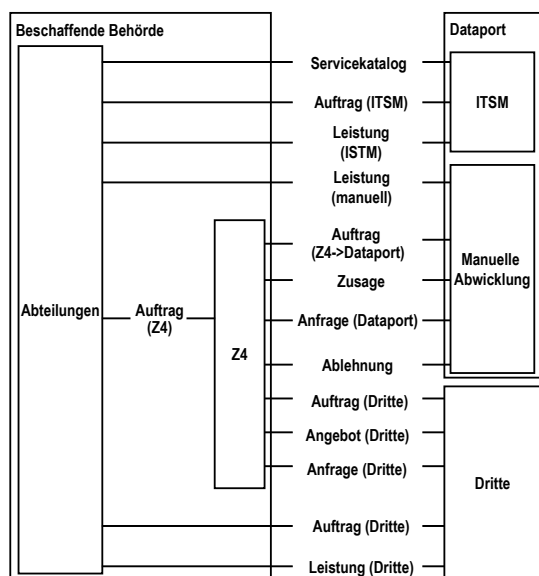


Abbildung 11: Architektur (IT-Beschaffungen BUKEA/BSW)

ausgeht, und Z4, die den Beschaffungsvorgang unter bestimmten Voraussetzungen unterstützen, unterschieden. Des Weiteren wird auf Anbieterseite bei *Dataport* zwischen jeweils einem Modul zur Bedarfsdeckung über den weitgehend automatisierten Servicekatalog (*ITSM*) und der *manuellen Abwicklung* differenziert.

3.1.2.3 Abläufe und zentrale Schnittstellen

Ausgangspunkt des allgemeinen Schemas ist ein existierender Bedarf in einer am Verfahren beteiligten *Abteilung*. Hieraus ergeben sich drei grundsätzliche Möglichkeiten, die Abwicklung des Vorgangs zur Deckung des Bedarfs einzuleiten. Im einfachsten Fall handelt es sich um eine Leistung, die der Dataport Servicekatalog (*ITSM*) umfasst. Dies sind insbesondere regelmäßig anfallende, standardisierte Leistungen wie Hardware oder Software-Lizenzen zur Arbeitsplatzausstattung, die bedarfsverursachende Abteilungen selbstständig beauftragen können. Wenn eine Abteilung eine derartige Leistung über *ITSM* anfordert, wird diese von Dataport disponiert und dem Auftraggeber ohne Umwege bereitgestellt.

Wenn die zu beschaffende Leistung jedoch nicht im *ITSM*-Servicekatalog enthalten ist oder Besonderheiten im Zusammenhang mit der Beauftragung eine Beteiligung von Z4

sinnvoll erscheinen lassen, erstellt die bedarfsverursachende Abteilung eine Bedarfsmeldung und sendet diese an Z4. In Abhängigkeit dessen, ob Z4 die Deckung des Bedarfs über Dataport für möglich hält, kommuniziert sie eine Anfrage an Dataport oder bereitet die Beauftragung Dritter vor. Dataport wird eingehende Anfragen prüfen und in Abhängigkeit von der eigenen Leistungsfähigkeit zusagen oder ablehnen. Eine erbrachte Zusage wird von Z4 dann in eine Beauftragung an Dataport umgewandelt. Abschließend disponiert Dataport die beauftragte Leistung und stellt diese in der Regel direkt der bedarfsverursachenden Abteilung bereit.

Wird die Beauftragung Dritter erforderlich, weil sich Z4 entweder direkt an einen Drittanbieter wendet oder Dataport ablehnt, generiert Z4 eine entsprechende Anfrage, die von Dritten mit einem Angebot beantwortet werden kann. Dieses Angebot kann Z4 ablehnen und weitere Anfragen an alternative Drittanbieter generieren oder annehmen und in einen Auftrag überführen. Der Auftrag wird dann dem jeweiligen Drittanbieter erteilt, der daraufhin die beauftragte Leistung disponiert und der bedarfsverursachenden Abteilung bereitstellt.

Unter bestimmten Voraussetzungen kann die Abteilung der beschaffenden Behörde, in der der Bedarf anfällt, entgegen der geltenden Beschaffungsrichtlinien (Bedarfe primär über Dataport zu decken) eigenmächtig handeln und sich zur Bedarfsdeckung direkt an Dritte wenden. In diesen Fällen, wo es sich beispielsweise um spezielle Bedarfe mit vergleichsweise geringen Kosten handelt, wird von der Abteilung ein Auftrag an einen Drittanbieter gerichtet, der die entsprechende Leistung disponiert und dem Auftraggeber direkt bereitstellt.

Das allgemeine Schema zum Verfahren wird in Abbildung 12 sowie die zugehörige Signatur in Abbildung 13 gezeigt. Eine exemplarische Instanziierung geht aus Abbildung 14 hervor.

3.1.2.4 Ablaufdokumentation und Log- daten

Die Dokumentation der Beschaffungsvorgänge variiert in Abhängigkeit der an einem Prozess beteiligten Organisationen bzw. Organisationseinheiten. Wenn eine Abteilung einen Bedarf selbstständig über den Servicekatalog von Dataport (ITSM) decken kann, werden die Anforderung, Bearbeitung und Bereitstellung der jeweiligen Leistung über die Systeme von Dataport abgebildet. Durch die enge Kooperation mit Dataport sind die hier

dokumentierten Vorgangsdaten grundsätzlich erreichbar, dennoch handelt es sich aus Sicht der Verwaltung um ein externes System, auf das nur bedingt zugegriffen werden kann.

Wenn Z4 eingeschaltet wird, dann wird der Bedarf über ein standardisiertes Formular kommuniziert. Hierbei handelt es sich um ein Microsoft Word-Formular, das Informationen zur Bearbeitung aller am Beschaffungsprozess beteiligten Organisationseinheiten vereint. Das heißt, die bedarfsverursachende Abteilung definiert den Bedarf, Z4 dokumentiert

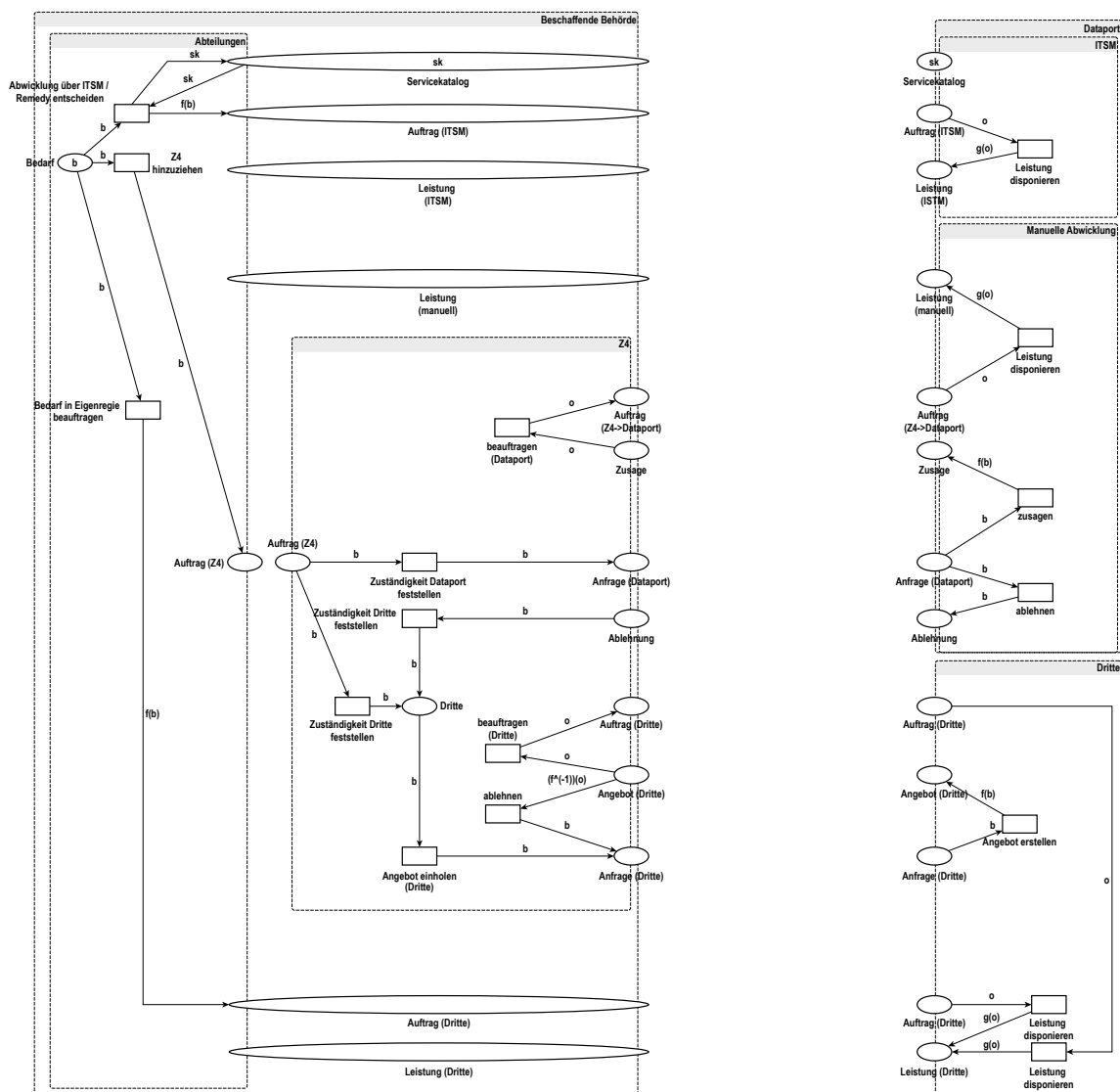


Abbildung 12: Allgemeines Schema (IT-Beschaffungen BUKEA/BSW)

Signatur
<p>Grundsorten</p> <p>I Leistung</p> <p>P Preis</p> <p>ANZ Anzahl</p>
<p>Abgeleitete Sorten</p> <p>$B \subseteq I \times \text{ANZ}$ (Bedarf)</p> <p>$O \subseteq B \times P$ (Angebot / Auftrag)</p> <p>$IX \subseteq I \times \dots \times I$ (erbrachte Leistung)</p>
<p>Konstanten</p> <p>b:B Bedarf</p>
<p>Variablen</p> <p>o:O Angebot</p> <p>i:I Leistung</p> <p>ix:IX erbrachte Leistung</p> <p>sk: set of I</p>
<p>Funktion</p> <p>f: $B \rightarrow O$ injektiv</p> <p>g: $O \rightarrow IX$</p>

Abbildung 13: Signatur des allgemeinen Schemas (IT-Beschaffungen BUKEA/BSW)

Exemplarische Instanziierung der Signatur
<p>Grundsorten</p> <p>$I = \{\text{Tastatur, Bildschirm, Office-Lizenz}\}$</p> <p>$P = \{65\text{€}\}$ Preis</p> <p>$\text{ANZ} = \{2\}$ Anzahl</p>
<p>Abgeleitete Sorten</p> <p>$B = \{(\text{Tastatur}, 2)\}$</p> <p>$O = \{(\text{Tastatur}, 2, 65\text{€})\}$</p> <p>$IX \subseteq \{(\text{Tastatur}, \text{Tastatur})\}$</p>
<p>Konstanten</p> <p>$b = \{(\text{Tastatur}, 2)\}$</p>
<p>Funktion</p> <p>f: $B \rightarrow O$</p> <p>$f((\text{Tastatur}, 2)) = (\text{Tastatur}, 2, 65\text{€})$</p> <p>g: $O \rightarrow IX$</p> <p>$g((\text{Tastatur}, 2, 65\text{€})) = (\text{Tastatur}, \text{Tastatur})$</p>

Abbildung 14: Exemplarische Instanziierung (IT-Beschaffungen BUKEA/BSW)

Signatur BUKEA Logs
<p>Grundsorten</p> <p>ID WorkflowID</p> <p>S = {Mitzeichnung, Mittelfreigabe, Mittelbindung, Schlusszeichnung, Auftragserteilung, Verfügung Eldorado / Papierakte} Schritt</p> <p>N Name</p> <p>VN Vorname</p> <p>L Leitzeichen</p> <p>K Kommentar</p> <p>E Erläuterung</p> <p>D Datum</p> <p>Z Zeit</p> <p>E = {Datei angelegt, Workflow wurde gestartet, Aufgabe zugestimmt, Workflowverlauf wurde angepasst, Aufgabe erledigt, Datei geändert, Aufgabe zugestimmt, Workflow beendet} Event</p> <p>Abgeleitete Sorten</p> <p>$Z \subseteq N \times VN \times L$ (Zuweisung)</p> <p>$A \subseteq N \times VN \times L$ (Ausführer)</p> <p>$T \subseteq D \times Z$ (Zeitstempel)</p> <p>$WDE \subseteq S \times Z \times E$ (Workflowdefinitionseintrag)</p> <p>$WD \subseteq WDE \times \dots \times WDE$ (Workflowdefinition)</p> <p>$C \subseteq E \times T \times S \times Z \times A \times K$ Case</p>

Abbildung 15: Signatur der Logdaten (IT-Beschaffungen BUKEA/BSW)

Bearbeitungsvermerke und Dataport hält Details zur Leistungserbringung fest. Weitergabe und Speicherung dieses Dokuments wird über Microsoft SharePoint (HIM-Workflow) abgewickelt. Hierdurch ist es möglich, Zeitpunkte, Zustände und beteiligte Personen im Rahmen der Bearbeitung festzuhalten. Dabei werden im SharePoint auch weitere Details dokumentiert, die nicht unmittelbar aus der Bearbeitung des Word-Formulars resultieren. So werden u. a. Genehmigungsvorgänge, eingeholte Angebote und erteilte Aufträge als separate Ereignisse festgehalten und teilweise mit entsprechenden Anlagen belegt. Die Dokumentation auf dem SharePoint bietet also eine chronologische Übersicht über einen einzelnen Beschaffungsvorgang und zeigt auf, welchen Personen bestimmte Aufgaben zugewiesen wurden und wann diese erledigt wurden. Ergänzend werden Dateioperationen wie die Anlage oder Änderung relevanter Dokumente festgehalten.

Informationen über die Nutzeraktivitäten im Sharepoint werden für jede Beschaffung in einem strukturierten elektronischen Dokument protokolliert. Die Struktur die-

ser Dokumentation wird in der Signatur in Abbildung 15 beschrieben. Dabei entspricht jede abgeleitete Sorte sowie die Grundsorten ID (WorkflowID), S (Schritt), L (Leitzeichen), K (Kommentar), E (Event) einem Datenfeld der zugrundeliegenden Dokumente.

In Fällen, bei denen die bedarfsverursachende Abteilung den Bedarf ohne Zutun von Z4 über Dritte abwickelt, ist der vollständige Beschaffungsprozess in Abhängigkeit des Anbieters individuell. Eine Dokumentation bzgl. der Verursachung des Bedarfs, der Angebotsauswahl, der Auftragserteilung und dem Eingang der Leistung kann im Einzelfall durch die jeweilige Organisationseinheit erfolgen. Eine einheitliche Protokollierung, wie sie von Z4 im Microsoft SharePoint verwendet wird, ist jedoch nicht gewährleistet. Weitere Details insbesondere zu den Bearbeitungsschritten des Anbieters finden außerhalb der Systeme der Verwaltung statt und entziehen sich damit jeglichen Analysemöglichkeiten.

3.1.3 JUS-IT

3.1.3.1 Fachliche und organisatorische Einordnung

In Kooperation mit dem Amt für Arbeit, Gesundheit, Soziales, Familie und Integration wurde der Prozess, *Heranziehung bei Unterhaltsvorschuss*, erhoben. Das Verfahren wird über JUS-IT, eine Fachanwendung für die Sozial- und Jugendhilfe, gesteuert und überwacht. Das erhobene Verfahren befasst sich mit der Heranziehung von Unterhaltspflichtigen bei erfolgter Leistung durch die Behörde. Auslöser für das Verfahren ist der Antrag, die Bewilligung und die Zahlung eines Unterhaltsvorschusses, den die Sozialbehörde ersatzweise für einen Zahlungspflichtigen aus einem zugestandenem Unterhaltsanspruch geleistet hat. Entsprechend existiert ein Leistungsfall, aus dem vor allem der Unterhaltsanspruch, die begünstigte Person (Unterhaltsempfänger) und die zur Unterhaltszahlung verpflichtete Person hervorgehen. Ziel der Behörde ist es, die bevorschusste Leistung vom Schuldner einzufordern, sofern dieser leistungsfähig ist. Begleicht ein leistungsfähiger Schuldner seine

Forderung nicht oder kann seine Leistungsfähigkeit nicht geeignet widerlegen, wird ein dreistufiges Mahnverfahren initiiert. Bleiben die Zahlungsaufforderungen des Mahnverfahrens erfolglos, wird in der Regel eine Vollstreckung zur Begleichung der Forderung eingeleitet.

3.1.3.2 Modulstruktur

Mit *JUS-IT*, *SAP* und der Fachanwendung *avviso* existieren für die Abwicklung des Verfahrens drei IT-Komponenten zur Durchführung, die je nach Verlauf zum Einsatz kommen. Die Abläufe des Moduls *Heranziehung* werden mithilfe von *JUS-IT* gesteuert. *Avviso* ist eine Fachanwendung für das Forderungsmanagement, die im Rahmen der Vollstreckung einer nicht ausgeglichenen Forderung verwendet wird. Das Modul *Mitarbeiter Vollstreckung* ist vor allem ein Bindeglied zwischen dem Modul *Heranziehung* und *avviso*. Die Aktivitäten des *Schuldners* werden in einem separaten Modul abgebildet. Dieses zeigt dessen Interaktion mit der Behörde. Die Architektur des Modells wird in Abbildung 16 dargestellt.

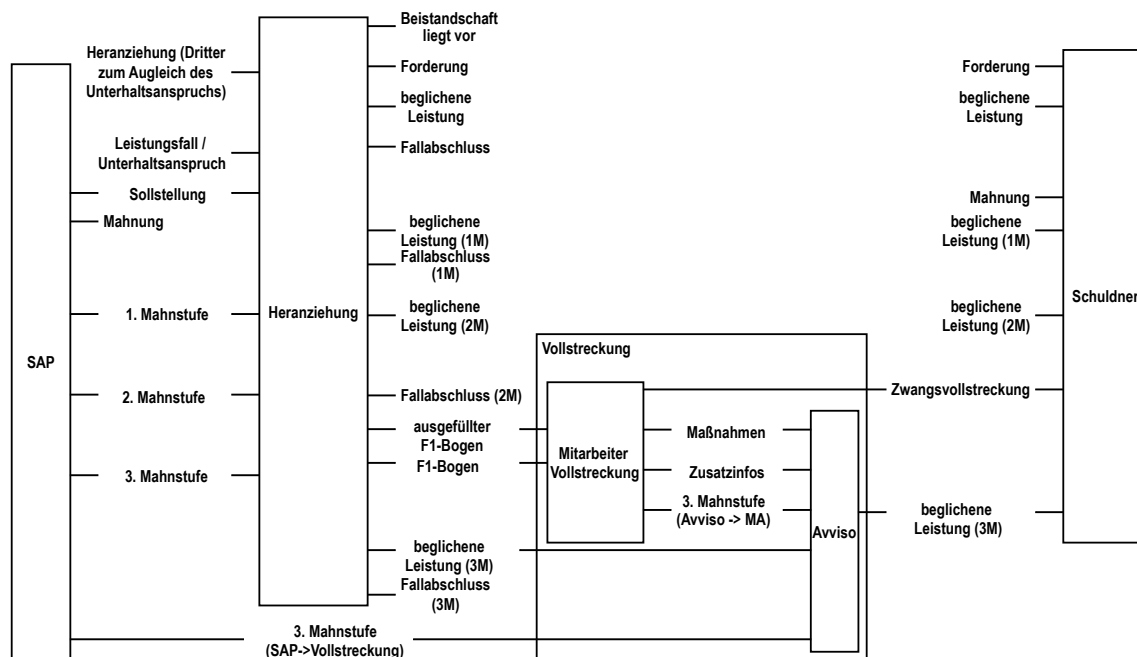


Abbildung 16: Architektur (JUS-IT)

3.1.3.3 Abläufe und zentrale Schnittstellen

Die Prozesslogik wird maßgeblich durch das Modul *Heranziehung* bestimmt. Hier startet der Ablauf mit dem Eingang eines Leistungsfalls bei erforderlicher Heranziehung. Sofern die Leistungsfähigkeit des Schuldners festgestellt wurde, erhält dieser eine Mitteilung über die Forderung und im *SAP*-System wird eine Sollstellung vorgenommen. Darauf kann der *Schuldner* reagieren und die Forderung begleichen. Mit der Begleichung der Forderung ist der Fall abgeschlossen. Wird die Forderung nicht beglichen, wird die erste Mahnstufe ausgelöst. In diesem Fall erhält der Schuldner eine schriftliche Mahnung, die aus *SAP* generiert wird, und erneut die Möglichkeit, die Leistung auszugleichen.

Andernfalls wird die zweite Mahnstufe ausgelöst. Zwischen *JUS-IT* und dem *SAP*-System werden Informationen über die Sollstellung der Forderung und die Mahnstufe abgeglichen.

Abbildung 17 zeigt dieses Zusammenspiel im Rahmen des allgemeinen Schemas, während Abbildung 18 die Signatur sowie Abbildung 19 eine exemplarische Instanziierung ergänzt.

SAP und *JUS-IT* sind über eine System-schnittstelle verbunden. Sollte die Leistung nach Ablauf der zweiten Mahnstufe nicht beglichen worden sein, wird der Fall in die dritte Mahnstufe eskaliert. Die Vollstreckung wird eingeleitet und der Fall an das *avviso*-System übermittelt. Daraufhin fordert ein *Mitarbeiter der Vollstreckung* mithilfe eines papierbasierten Formulars (F1-Bogen) zusätzliche Informationen von Sachbearbeitern der Leistungsbewilligung und Heranziehung an. Die angeforderten Informationen müssen durch einen Vollstreckungsmitarbeiter manuell vom F1-Bogen in *avviso* eingepflegt werden. Sobald diese Informationen vorliegen, werden Vollstreckungsmaßnahmen gewählt und umgesetzt. Wenn die Forderung dann ausgeglichen wurde, wird der Fall abgeschlossen.

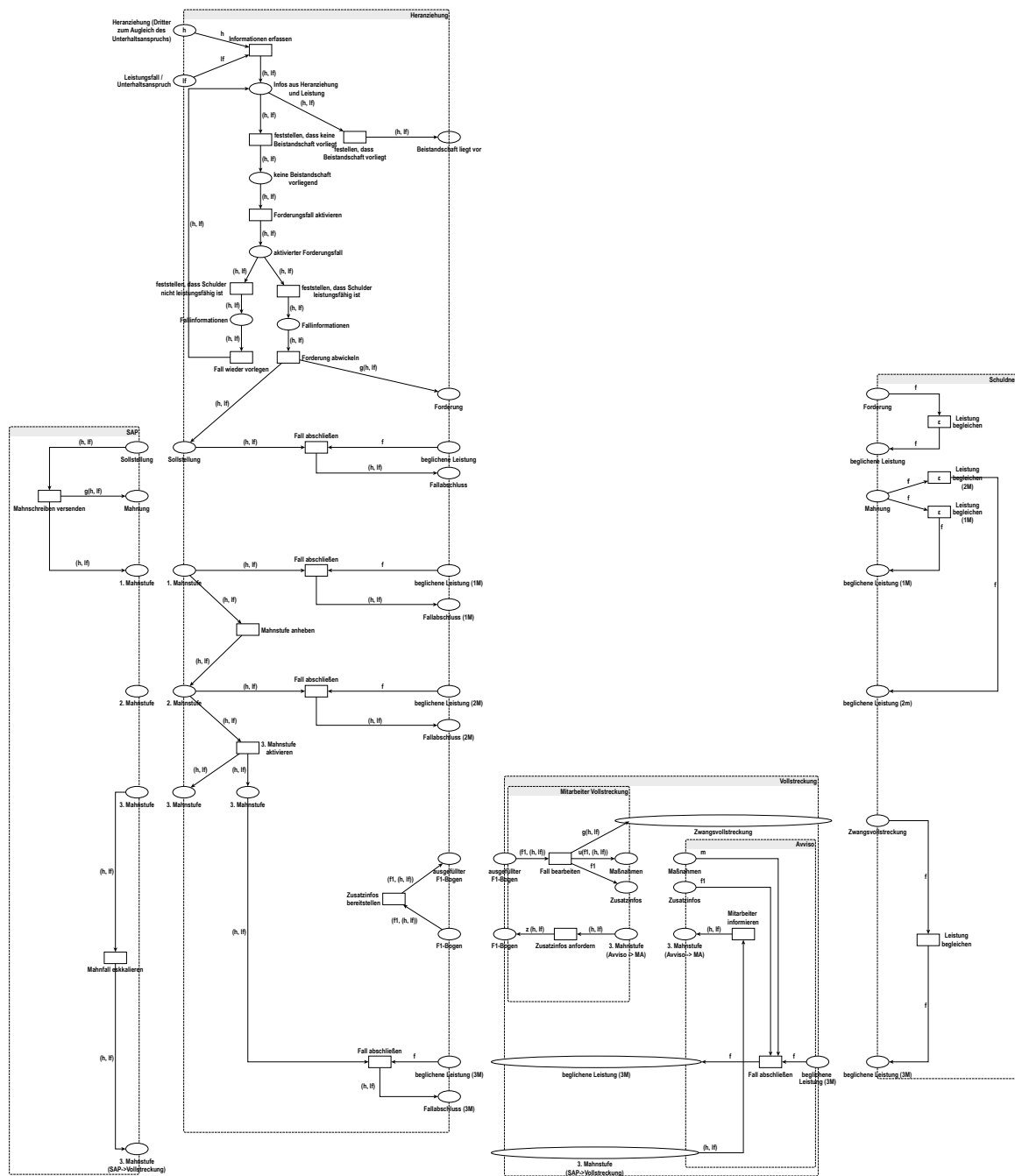


Abbildung 17: Allgemeines Schema (JUS-IT)

Signatur
<p>Grundsorten</p> <p>N Name</p> <p>F Forderung</p> <p>G monatl. Leistung</p> <p>T Leistungstyp</p> <p>M Vollstreckungsmaßnahme</p> <p>K Datenfeldbezeichnung</p> <p>V Datenfeldinhalt</p>
<p>Abgeleitete Sorten</p> <p>$D \subseteq K \times V$ Datenfeld</p> <p>$F1 \subseteq D \times \dots \times D$ F1-Bogen</p> <p>$H \subseteq N \times F$ Heranziehung</p> <p>$LF \subseteq N \times G \times T$ Leistungsfall</p>
<p>Variablen</p> <p>h: H Heranziehung</p> <p>lf: LF Leistungsfall</p> <p>f: F Forderung</p> <p>m: M Vollstreckungsmaßnahme</p> <p>f1: F1 F1-Bogen</p>
<p>Funktionen</p> <p>$g: H \times LF \rightarrow F$</p> <p>$u: F1 \times H \times LF \rightarrow M$</p>

Abbildung 18: Signatur zum allgemeinen Schema (JUS-IT)

Exemplarische Instanziierung der Signatur
<p>Grundsorten</p> <p>$N = \{\text{Horst Müller, Justin Müller}\}$</p> <p>$F = \{650\}$</p> <p>$G = \{325\}$</p> <p>$M = \{\text{Lohnpfändung}\}$</p> <p>$K = \{\text{pfändbares Einkommen, Arbeitgeber}\}$ Datenfeldbezeichnung</p> <p>$V = \{750, \text{Fleischerei Schmitz \& Söhne}\}$ Datenfeldinhalt</p>
<p>Abgeleitete Sorten</p> <p>$D = \{(\text{pfändbares Einkommen, 1150}), (\text{Arbeitgeber, Fleischerei Schmitz \& Söhne})\}$ Datenfeld</p> <p>$F1 = \{((\text{pfändbares Einkommen, 1150}), (\text{Arbeitgeber, Fleischerei Schmitz \& Söhne}))\}$ F1-Bogen</p> <p>$H = \{(\text{Horst Müller, 650})\}$</p> <p>$LF = \{(\text{Justin Müller, 325, Unterhalt})\}$</p>
<p>Funktionen</p> <p>$g: H \times LF \rightarrow F$</p> <p>$g((\text{Horst Müller, 650}), (\text{Justin Müller, 325, Unterhalt})) = 650$</p> <p>$u: F1 \times H \times LF \rightarrow M$</p> <p>$u(((\text{pfändbares Einkommen, 1150}), (\text{Arbeitgeber, Fleischerei Schmitz \& Söhne})), (\text{Horst Müller, 650}), (\text{Justin Müller, 325, Unterhalt})) = \text{Lohnpfändung}$</p> <p>$z: H \times LF \rightarrow H \times LF \times F1$</p> <p>$z(((\text{Horst Müller, 650}), (\text{Justin Müller, 325, Unterhalt}))) = ((\text{Horst Müller, 650}), (\text{Justin Müller, 325, Unterhalt}), ((\text{pfändbares Einkommen, 1150}), (\text{Arbeitgeber, Fleischerei Schmitz \& Söhne})))$</p>

Abbildung 19: Exemplarische Instanziierung (JUS-IT)

3.1.3.4 Automatische Ablaufdokumentation (Logging)

Im Rahmen der Studie wurden die von *JUS-IT* erzeugten Protokolldateien analysiert. Die enthaltenen Daten zeigen keine typische Log-Struktur, sondern beschreiben weitgehend zeitpunktbezogen die Ausführung von Datenbankoperationen. Abschnittsweise zeigen die Daten durchaus eine gewisse Ähnlichkeit zu typischen Ereignis-Logs. In diesen Abschnitten können bestimmte Aktivitäten mithilfe der jeweiligen Zeitstempel in eine chronologische Reihenfolge gebracht werden, sie beschreiben jedoch Systemvorgänge. Entsprechend wird nicht zwischen Benutzeraktivitäten, Batch-, Folge und Hintergrundoperationen unterschieden. Zusätzlich definieren sich derartige Aktivitäten meist durch einen konkreten Datenbank-Befehl („SQL-Statement“), der zwar eine bestimmte Datenbankoperation beschreibt, jedoch nur bedingt Interpretationen bzgl. intendierter Vorgehensweisen des Benutzers zulässt. Insgesamt kann also zusammengefasst werden, dass zwar eine gewisse automatische Dokumentation der Abläufe stattfindet, für Process-Mining-Techniken sind die erzeugten Protokolle jedoch nicht ohne Weiteres zu verwenden. Abbildung 20 zeigt die Signatur der protokollierten Informationen.

Signatur
Grundsorten RT Recordtyp AT = {Payment, Liability} Asset-Typ T Zeitstempel M Systemmodul D Dateiname SQL SQL-Statement SLE System-Log-Eintrag
Abgeleitete Sorten $AR \subseteq T \times RT \times AT \times D$ Asset Report $SL \subseteq RT \times T \times M \times SQL \cup SLE$ Systemlogs

Abbildung 20: Log-Signatur (JUS-IT)

Im dargestellten Prozess werden insbesondere Finanzinformationen geloggt. Eine besondere Rolle spielen dabei neben den Systemlogs, Asset Reports, die die Synchronisierung von Informationen zwischen *Leistungsbewilligung und Heranziehung* und *SAP* dokumentieren.

Die zur Synchronisation notwendigen Informationen sind in den referenzierten Dateien angegeben. Das *JUS-IT*-Anwendungssystem legt für jeden Leistungsfall einen separaten Case an und erzeugt eine Finanzkomponente. Hierin werden Informationen darüber festgehalten, in welchem Zeitraum Zahlungen (oder Forderungen) in welcher Höhe generiert werden sollen. Die Logs werden genutzt, um automatisiert und tagesaktuell jeweils anstehende Zahlungsströme zu identifizieren. Zahlungsrelevante Daten werden an ein SAP-System gesandt. Dazu werden die Daten an die *SAP*-Anwendung *Kassen- und Einnahmenmanagement (PSCD)* übermittelt. Diese generieren Ausgabedateien, die es ermöglichen, die in *JUS-IT* erstellten Informationen über Zahlungsströme in *SAP* zu erfassen.

3.2 Beurteilung der Voraussetzungen für Process Mining

3.2.1 KoPers

Eingaben, Änderungen und Bearbeitungsschritte, die in KoPers bei Neuanlage oder Anpassung eines Personaldatensatzes vorgenommen werden, werden feingranular mit entsprechendem Zeitstempel dokumentiert. Um die Änderungshistorie einzelner Datenfelder nachvollziehen zu können, enthält jeder Logeintrag, der ein bestimmtes Datenfeld referenziert, einen *alten Wert* und den *neuen (geänderten) Wert* (vgl. Abbildung 10: AW, NW). Bei jedem Bearbeitungsschritt wird auch der Nutzer festgehalten, der diesen ausführt. Entsprechend können diese Abläufe innerhalb des Moduls KoPers gut mit Process Mining analysiert werden. So ist u. a. erkennbar, welche Bearbeitungsschritte ügig aufeinander folgen konnten und welchen Schritten längere Pausen vorangingen. Hieraus kann bspw. abgeleitet werden, für welche Aktivitäten zusätzliche Informationen erforderlich sind, die nicht unmittelbar vorliegen und erst mit einem gewissen Aufwand aus weiteren elektronischen Quellen, Dokumenten oder von Prozessbeteiligten eingeholt werden müssen.

Derartige Analysen können dabei helfen, Verbesserungspotenziale für die Unterstützung der Mitarbeiter zu identifizieren und zu priorisieren. Auf diese Weise kann der Ablauf in Richtung verkürzter Durchlaufzeiten optimiert werden, woraus eine Entlastung des Personals und eine Verkürzung des Verfahrens für den neu einzustellenden Mitarbeiter resultiert.

Doch die Protokollierung in KoPers dokumentiert die Aktivitäten eines Anwenders nicht vollständig. Sie ist auf Ereignisse beschränkt, die zu einer Veränderung gespeicherter Mitarbeiterdaten führen. Dies ist zweckmäßig, um lückenlos die Entstehung und Veränderung der Datenbestände zu dokumentieren, lässt jedoch verschiedene Informationen vermissen, die bei einer Process-Mining-gestützten Analyse der Abläufe von Interesse wären. So wird beispielsweise die Navigation des Benutzers zwischen den zahlreichen Eingabemasken nicht explizit aufgezeichnet. Zwar wird durch die genauen Feldbezeichnungen bei entsprechenden Dateneingaben deutlich, auf welcher Maske jede einzelne Eingabe erfolgt. Dies erlaubt somit auch Rückschlüsse auf die gewählte Reihenfolge. Unklarheiten bspw. bei der Suche der richtigen Maske für eine bestimmte Eingabe können jedoch nicht zweifelsfrei identifiziert werden, weil das Durchschalten der Seiten nicht dokumentiert wird. Einige Verbesserungspotenziale wie bspw. Ansätze zur intuitiveren Gestaltung der Benutzeroberflächen bleiben somit verborgen.

Das Logging im untersuchten Fachverfahren ist auf die Datenerfassung in KoPers beschränkt und deckt somit nur einen Teil des in Abbildung 7 gezeigten Prozesses ab. Dies ist darauf zurückzuführen, dass Abläufe im Zusammenhang mit der Ausschreibung und Bewerberauswahl je nach einstellender Behörde recht individuell ausfallen und deshalb nur bedingt mit einem einheitlichen Informationssystem unterstützt werden können, aus denen die für Process-Mining-Techniken geeigneten Logdaten hervorgehen könnten. Im Verlauf der Studie wurde jedoch bekannt, dass zukünftig für die der Datenerfassung vorgelagerten Prozesse eine gemeinsam genutzte Recruiting-Software eingeführt werden soll. Durch diese Maßnahme wird in Aussicht gestellt, dass zukünftig weitere Teile des Einstellungsprozesses mit einem

einheitlichen Informationssystem unterstützt werden, bei dem die Erzeugung von Logdaten wahrscheinlich ist und damit für diese Workflows auch von einem nutzenstiftenden Einsatz ausgewählter Process-Mining-Techniken ausgegangen werden kann.

3.2.2 IT-Beschaffungen BUKEA / BSW

Die Beschaffungsvorgänge BUKEA / BSW werden von der Verwaltung nur dann zentral elektronisch dokumentiert, wenn die Abteilung Z4 eingebunden ist. Hier findet dann die Protokollierung relevanter Bearbeitungsschritte des Workflows in einem Microsoft SharePoint statt. Die darin enthaltenen Informationen bilden die wesentlichen Ereignisse im Verlauf jedes einzelnen Beschaffungsvorgangs ab und erlauben, diese geeignet zu rekonstruieren. Es handelt sich jedoch nicht um ein automatisches Logging im Hintergrund eines bestimmten Anwendungssystems. Die erzeugten Einträge werden durch Aktivitäten der Mitarbeiter auf dem SharePoint hervorgerufen. Die Vorgehensweise erscheint zweckmäßig, weil im Dialog zwischen Bedarfsverursacher, Z4 und Dataport bzw. alternativen Leistungserbringern kein gemeinsames Informationssystem genutzt wird. Vielmehr sind die Informationen der Bedarfsmeldungen, der kommunizierten Leistungs- und Anforderungsbeschreibungen, der Angebote und Aufträge relevant. Mitarbeiter können bei der Bearbeitung einzelner Schritte oder der Anlage vorgenannte Dokumente ergänzende Erläuterungen und Hinweise einbringen. Diese Hinweise sind wichtig, um Besonderheiten im einzelnen Vorgang – insbesondere beim Versuch, diesen nachzuvollziehen – erkennen zu können. Doch für konventionelle Process Mining Techniken sind derartige Informationen nur bedingt verwertbar. Ohne weitere Maßnahmen lassen sich hier nur die groben Schritte der Abwicklung und deren zeitliche Anordnung, die beteiligten Personen sowie einbezogene Dokumente auswerten.

Weiterhin fehlen für eine umfassende Analyse der Prozesse Aufzeichnungen des vorrangig einzuschaltenden Dienstleisters Dataport. Dies schließt auch Abstimmungen zwischen Z4 und Dataport ein, die über verschiedene

Kommunikationswege wie E-Mails oder Telefonate erfolgen können. Somit bleibt insbesondere verborgen, welche Prozessschritte zwischen dem Einholen von Angeboten für eine bestimmte Leistung bis zu deren Bereitstellung erfolgen. Dieser essentielle Prozessabschnitt kann u. a. durch die Art der Leistung, den zusätzlichen Abstimmungsbedarf, äußere Einflüsse, etc. beeinflusst werden und ist somit zur Bemessung von Gütekriterien der Abläufe von entscheidender Bedeutung. Um hier weitere Details einzubeziehen, müssten neben der organisationsübergreifenden Kommunikation auch die Logdaten von den bei Dataport genutzten Informationssystemen Berücksichtigung finden oder besser ein gemeinsames System genutzt werden.

3.2.3 JUS-IT

Anhand der Architektur des Fachverfahrens (Abbildung 16) wird deutlich, dass eine zentrale Herausforderung für den Einsatz von Process Mining in Abläufen zur *Heranziehung bei Unterhaltsvorschuss* darin besteht, dass bis zu drei IT-Systeme zum Einsatz kommen. Um also sämtliche protokollierten Informationen der Abläufe mit Process-Mining-Techniken analysieren zu können, müssten die Logs der beteiligten Systeme zusammengeführt und teilweise vereinheitlicht werden. Hierbei müsste der Schwerpunkt darauf liegen, die Log-Einträge der einzelnen Systeme derart aufzubereiten, dass einzelne Abläufe systemübergreifend nachvollzogen und ausgewertet werden können. Entsprechend müsste eine systemübergreifende Zuordnung der Logeinträge zum einzelnen Ablauf gekennzeichnet werden und ebenso sind für gemeinsam genutzte Attribute wie beispielsweise Zeitstempel einheitliche Datenformate zu wählen.

Für das untersuchte Verfahren zeigen sich hier einige Probleme. Besonders schwerwiegend ist die Erkenntnis, dass JUS-IT keine dedizierten Prozesslogs erzeugt. Die protokollierten Datenbankoperationen lassen sich zwar grundsätzlich in besser geeignete Attribute überführen, dies ist jedoch mit erheblichen Aufwänden verbunden. Einerseits setzt die Interpretation der dokumentierten SQLStatements bereits ein umfassendes Expertenwissen voraus und andererseits würde aufgrund der Vielzahl an

Einträgen eine automatisierte technische Lösung für die Transformation erforderlich werden.

Darüber hinaus existieren im Detail einzelne Schritte, die nicht von der unterstützenden IT erfasst werden. So werden im Verlauf des Verfahrens u. a. fallbezogene Informationen über einen papierbasierten Fragebogen (F1-Bogen) angefordert und manuell in das System avvisio übertragen. Eine Schnittstelle, die diesen Datenaustausch automatisiert ermöglichen und dokumentieren könnte, existiert nicht. Neben dem Problem, dass die papierbasierte Weitergabe der Daten nicht von Process Mining Techniken erfasst und analysiert werden kann, ist dieses Vorgehen auch aus fachlicher Sicht anfällig für Fehler und kann zu Inkonsistenzen führen.

3.3 Herausforderungen für Process Mining

3.3.1 Vorgehensweise zur Ermittlung von Herausforderungen

Anhand der im Rahmen der Studie durchgeführten Prozesserhebungen konnten in den untersuchten Fachverfahren unterschiedliche Voraussetzungen für Analysen mit Process Mining identifiziert werden. Die im folgenden Abschnitt beschriebenen Aspekte wurden induktiv aus der Betrachtung der erhobenen Anwendungsfälle abgeleitet und verallgemeinert. Dabei handelt es sich um Herausforderungen, die typische Probleme für Prozessanalysen insbesondere im Umfeld öffentlicher Verwaltungen beschreiben. Sie sind häufig auch auf andere Prozesse übertragbar. Somit können auch die daraus hervorgehenden und in Abschnitt 4 vorgestellten Handlungsempfehlungen für ein breites Spektrum von Prozessen in der öffentlichen Verwaltung Nutzen stiften. Dennoch sind die aufgeführten Herausforderungen (Abschnitt 3.3.2) und Maßnahmen (Abschnitt 3.4) aus der Analyse der drei untersuchten Fachverfahren hervorgegangen und bleiben deshalb hinsichtlich einer Übertragbarkeit auf die gesamte Verwaltungslandschaft grundsätz-

lich erweiterbar, decken jedoch typische und weitverbreitete Szenarien ab.

3.3.2 Identifizierte Herausforderungen

3.3.2.1 H1 – Abläufe werden nicht vollständig von den Logs abgebildet

Wenn Verwaltungsverfahren nicht oder nicht vollständig mithilfe von vorgangsprotokollierender IT durchgeführt werden, so fehlen verwertbare Logdaten und schränken die Anwendung von Process Mining entsprechend ein. Dieses Problem wird in Verwaltungsverfahren insbesondere dort sichtbar, wo Sachbearbeitung ohne den Einsatz von Informationssystemen erfolgt, Kommunikation zwischen Prozessbeteiligten außerhalb der dokumentierenden IT stattfindet, Medienbrüche existieren oder wesentliche Informationen, die Einfluss auf den Ablauf nehmen, ausschließlich in Dokumenten enthalten sind.

3.3.2.2 H2 – Logdaten verschiedener Systeme sind meist heterogen

Sobald ein Verwaltungsverfahren mithilfe mehrerer IT-Instrumente abgewickelt wird, existiert für Process Mining die Herausforderung, die protokollierten Informationen unterschiedlicher Systeme zu vereinen, um die Abläufe vollständig analysieren zu können. Dabei tritt meist das Problem auf, dass Systeme Logdaten weitgehend individuell protokollieren und diese dann hinsichtlich der Struktur, Formate und Attribute sehr unterschiedlich ausfallen. Somit sind die Logs mehrerer Systeme zunächst zu harmonisieren, wozu bereits umfassendes Know-how erforderlich ist. Zentrales Problem dieser Transformation ist das Zusammenführen der von den Systemen aufgezeichneten Attribute, so dass einerseits gekennzeichnet wird, welche Einträge der Systeme zu einem gemeinsamen Ablauf gehören und andererseits, welche Attribute der unterschiedlichen Systeme vergleichbare Informationen zum Prozess adressieren und deshalb zu einem Attributs-Typ vereint werden können.

3.3.2.3 H3 – Granularität des Logs

Erzeugte Logs können hinsichtlich ihres Detailgrads sehr unterschiedlich ausfallen. Einerseits kann es für Process Mining ein Problem darstellen, wenn Logs relativ wenige Ereignisse eines Ablaufs widerspiegeln und die Dokumentation damit vergleichsweise grob ausfällt. Dies hat zur Folge, dass die Aufzeichnung bestimmte Details vermissen lässt und interessante Ereignisse mit anderen konsolidiert werden. In der Folge fallen unter derartigen Voraussetzungen Analysen mit Process-Mining-Techniken recht oberflächlich aus und können nur bedingt Erkenntnisse zu detaillierten Fragestellungen liefern. Im umgekehrten Fall ist es ebenso nicht unproblematisch, wenn Logs sehr feingliedrig erstellt werden. Hier werden zwar keine Informationen konsolidiert und können somit auch nicht verwässern, es fällt jedoch oft schwer, die relevanten Ereignisse für bestimmte Fragestellungen auszumachen und für Analysen zu isolieren. Gerade für das Erstellen von Prozessmodellen sind sehr feingranulare Logs oft ein Problem und erfordern eine aufwendige Vorbereitung der Daten, damit die Modelle übersichtlich bleiben.

3.3.2.4 H4 – Menge protokollierter Attributstypen

Grundsätzlich existieren zwar keine festen Anforderungen an das Erstellen von Logs, in der Praxis hat sich jedoch durchgesetzt, dass Logs ausgewählte Attribute umfassen, die jeden einzelnen Logeintrag spezifizieren. Dabei ist zweckmäßig, dass mindestens ein Attribut das jeweilige Ereignis benennt und ein weiteres Attribut den Zeitpunkt des Eintretens festhält. Häufig wird den Einträgen eines Ablaufs außerdem eine einzigartige ID zugeordnet, damit Durchläufe separat betrachtet und gegenübergestellt werden können. Darüber hinaus können Logs jedoch eine Vielzahl weiterer Attribute umfassen. Da es in der Anwendung von Process Mining leichtfällt, Teilmengen der vorhandenen Attribute in die Analysen einzubeziehen, sind überdurchschnittlich viele Attributstypen selten ein echtes Problem. Doch auch hier kann die Bestimmung der Relevanz und das Treffen einer Auswahl mangels umfassender Kenntnisse Probleme bereiten.

Weitaus kritischer ist es jedoch, wenn wichtige Informationen, die bei Analysen von Interesse gewesen wären, nicht durch entsprechende Attribute repräsentiert werden und damit auch keine Erkenntnisse liefern können.

3.3.2.5 H5 – Variation der Attributs-Ausprägungen

Wenn für relevante Attribute eine Vielzahl von Ausprägungen verwendet wird, kann dies im Rahmen von Analysen zu Problemen führen, weil die Erkennung von Mustern auf statistische Verfahren angewiesen ist und deshalb vor allem Häufigkeiten ermittelt. Variieren die Ausprägungen betrachteter Attribute zu stark, wird die Bestimmung von Häufigkeiten und damit die Feststellung von Ähnlichkeiten gehemmt.

3.3.2.6 H6 – Intuitive Verständlichkeit der Logs

Es ist üblich, dass im Rahmen eines Loggings sowohl für bestimmte Attribute als auch deren Ausprägungen Schlagworte, Abkürzungen, besondere Formate oder codierte Bezeichner verwendet werden, die nicht zwingend selbsterklärend sind. Entsprechend wichtig sind die Interpretation der Attribute sowie deren Ausprägungen. Da das Logging in der Regel vom Entwickler eines IT-Instruments implementiert wird, Process-Mining-Analysen jedoch eher im Interesse der Anwender liegen, sind technische Erläuterungen meist nicht unmittelbar verfügbar, weshalb die möglichst intuitive Verständlichkeit der Attribute und ihrer Ausprägungen einen wichtigen Aspekt darstellt.

3.3.2.7 H7 – Format der Logdateien

Prozessinformationen können grundsätzlich sehr unterschiedlich gespeichert werden. Da es sich beim Logging um ein laufendes Protokollieren von Ereignissen handelt, ist es weit verbreitet, Einträge als Aufzählung der voneinander mittels Sonderzeichen getrennten

Attributausprägungen zu strukturieren und Einträge durch Zeilenumbruch voneinander zu trennen (z. B. im CSV-Format). Es existieren jedoch auch weitere Formate wie der auf der *Extensible Markup Language* (XML) basierende *eXtensible-Event-Stream-Standard* (XES)²⁴, bei dem den Daten Tags zuordnet werden, die Inhalte klassifizieren. Dennoch müssen Logdaten nicht zwingend in dieser Form vorliegen, sondern können bestimmte Formate, Verschlüsselungstechniken oder ungewöhnliche Strukturen aufweisen. In diesen Fällen ist eine Vorbereitung der Daten (ETL) erforderlich und kann entsprechend zusätzlichen Aufwand verursachen.

3.3.3 Zusammenfassung für untersuchte Fachverfahren

Die in Abschnitt 3.3.2 vorgestellten Herausforderungen haben sich im Verlauf der Untersuchung unterschiedlich in den drei erhobenen Fachverfahren widerspiegelt. Im folgenden Abschnitt wird erläutert, in welchem Umfang die jeweilige Herausforderung für jedes untersuchte Fachverfahren relevant ist. Im Anschluss an die Erläuterungen werden die Ergebnisse dieser Analyse in Tabelle 1 zusammengefasst.

Vollständigkeit

Abläufe spiegeln sich in generierten Logdaten oft nur unvollständig wider. Dieses Problem wurde in allen drei untersuchten Verfahren erkannt. Im Fachverfahren KoPers beschränkt sich das Logging auf die Datenerfassung bzw. -änderung. Alle weiteren Schritte des Recruitings und der Auswahl einzustellender Mitarbeiter wird von den beteiligten Behörden nicht einheitlich mit IT unterstützt. Bei den IT-Beschaffungen BUKEA / BSW findet eine Protokollierung der Prozessschritte nur unter Beteiligung von Z4 statt. Werden Bedarfe über den ITSM-Shop von Dataport oder in ausschließlicher Zusammenarbeit zwischen Bedarfsverursacher und Dritten gedeckt, entstehen auf Seiten der Verwaltung keine verwertbaren Logdaten. Der Einsatz von JUS-IT erzeugt detaillierte Logs und auch bei

²⁴ vgl. <https://xes-standard.org/>

der Verwendung der beiden mit dem Verfahren verknüpften Systeme, SAP und avviso werden Vorgänge der Benutzer in Form von Logdaten festgehalten. Neben der Nutzung von Schnittstellen zwischen den drei beteiligten Anwendungssystemen erfolgen jedoch auch bestimmte Prozessschritte außerhalb der IT. So werden u. a. Aktivitäten zum Datenaustausch mittels F1-Bogen nicht dokumentiert.

Heterogenität

Für KoPers wird derzeit nur ein einzelnes IT-System verwendet. Somit stellt sich die Herausforderung einer Zusammenführung von Logs unterschiedlicher Quellen bislang nicht. Dies könnte sich jedoch ändern, wenn zukünftig wie vorgesehen eine Recruiting-Software zum Einsatz kommen sollte. Bei ITBeschaffungen BUKEA / BSW wird zur Protokollierung einzelner Beschaffungsvorgänge eine manuelle Dokumentation auf dem Microsoft SharePoint vorgenommen. Neben ITSM und weiteren Systemen von Dataport oder Dritten, die im Rahmen der Bedarfsdeckung und im Verlauf der Prozesse eingesetzt werden, sich aber grundsätzlich außerhalb der Verfügungsgewalt durch die Verwaltung befinden, wurde bekannt, dass zur Abwicklung der Vorgänge auch Standardsoftware wie Microsoft Word und E-Mail-Kommunikation genutzt wird. Um sämtliche Logdaten der erhobenen Abläufe mit Process Mining untersuchen zu können, müssten auch diese Logdaten herangezogen werden, woraus sich grundsätzlich ein Bedarf zur Harmonisierung der Logs verschiedener Systeme ergeben würde. Dadurch dass neben JUS-IT zur Abwicklung der *Heranziehungen bei Unterhaltsvorschuss* verwaltungsseitig zwei weitere IT-Systeme eingesetzt werden, wird hier der Bedarf zur Zusammenführung von Logdaten am deutlichsten, um Abläufe möglichst vollständig analysieren zu können.

Granularität

Hinsichtlich der Granularität protokollierter Prozessschritte ist ein optimales Niveau immer auch von den genauen Zielen vorgesehener Analysen abhängig. Dennoch kann beurteilt werden, ob Abläufe einerseits mit den vorliegenden Logdaten ausreichend detailliert nachvollzogen werden können und andererseits nicht zu viele Einträge im Log enthalten

sind, die zur Rekonstruktion der Vorgänge keinen signifikanten Beitrag leisten. Im Fall von KoPers ist die Dichte der Einträge im Log ausgewogen, allerdings beschränken sich die Logeinträge auf Datenänderungen und schließen keine weiteren Nutzeroperationen wie bspw. die Navigation durch verschiedene Eingabemasken ein. Bei ITBeschaffungen BUKEA / BSW werden aufgrund der manuellen Dokumentation nur die wesentlichen Schritte des Verfahrens festgehalten. Dieses Granularitätsniveau ist für sehr detaillierte Analysen nicht hinreichend, erlaubt jedoch eine makroskopische Betrachtung der Abläufe. Im Fachverfahren JUS-IT liegen ausführliche Systemlogs vor. Diese zeigen zahlreiche Einträge auf, die nicht direkt auf Benutzerinteraktionen zurückzuführen sind, sondern ebenso Arbeitsschritte des Systems beschreiben. Auf diese Weise mischen sich Aktivitäten des Anwenders wie beispielsweise das Aufrufen eines bestimmten Falls mit Systemoperationen wie den Datenbankabfragen zur Anzeige des aufgerufenen Falls. Diese sehr detaillierte Protokollierung kann in Einzelfällen von Vorteil sein, stellt jedoch für das Nachvollziehen wesentlicher Prozessschritte eine große Herausforderung dar.

Menge der Attributstypen

Die Betrachtung des Umfangs protokollierter Attribute adressiert die Menge an Merkmalen, die jeder einzelne Logeintrag umfasst. Da sich Analysen mit Process Mining problemlos auf eine Teilmenge vorhandener Attribute beschränken können, wird hier ausschließlich beurteilt, ob in den Logdaten der einzelnen Fachverfahren offensichtlich wichtige Attribute fehlen. Bei KoPers und IT-Beschaffungen BUKEA / BSW erscheint die Auswahl verwendeter Attribute zweckmäßig. Die wesentlichen Informationen zur Identifikation des Ablaufs, der inhaltlichen und zeitlichen Einordnung festgehaltener Ereignisse sowie den Angaben zu den aus der Prozesserhebung bekannt gewordenen beteiligten Ressourcen sind jeweils durch entsprechende Attribute festgehalten. In JUS-IT wäre vor der Durchführung von Process-Mining-Analysen eine umfassende Aufbereitung der Logdaten vonnöten. Im Wesentlichen beschränken sich die verwendeten Attribute auf einen Zeitstempel und den in diesem Moment aus-

geführten Datenbankbefehl, wodurch sich zwar grundsätzlich Abläufe rekonstruieren lassen, wertvolle Hinweise auf eingebundene Ressourcen wie beispielsweise den ausführenden Anwender jedoch nicht eingeschlossen werden.

Variation der Attributs-Ausprägungen

Die Variation der Ausprägungen einzelner Attribute eines Logs müsste grundsätzlich für jedes Attribut individuell beurteilt werden. Aus diesem Grund wird hier nur auf offensichtliche Probleme eingegangen. In KoPers werden ausschließlich Ereignisse vom Typ Datenänderung festgehalten. Das Vorgehen wird dabei durch die Sequenz bearbeiteter Datenfelder und somit die Kombination der Attribute *Feldname* und *Datenbestand* bestimmt. Entsprechend lassen sich mit Process-Mining-Techniken innerhalb eines einzelnen Erfassungsvorgangs nur bedingt Muster erkennen. Bei IT-Beschaffungen BUKEA / BSW steht für das dokumentierte Ereignis eine zweckmäßige, überschaubare Auswahl an Attributen zur Verfügung. Logs erlauben somit u. a. auch die Identifikation sich wiederholender Teilprozesse innerhalb eines Einzelfalls. Problematisch fällt die Gestaltung des Logs hinsichtlich der Attributsausprägungen insbesondere bei JUS-IT auf. Da Ereignisse hier in der Regel durch sehr spezifische Datenbankbefehle (SQL-Statements) repräsentiert werden, existiert zunächst eine unübersichtliche Vielzahl von Ereignistypen, die Auswertungen mit Process Mining nur bedingt zulassen. Grundsätzlich ist es jedoch mit weiterführenden Techniken möglich, die Befehle zu klassifizieren und die Logdaten mit zusätzlichen, abgeleiteten Attributen anzureichern, wodurch dann auch bestimmte Analysen mit Process Mining ziel führend werden können.

Verständlichkeit

Die von KoPers erzeugte Logdatei ist intuitiv verständlich und leicht zu lesen. Dies ist jedoch zu einem erheblichen Teil darauf zurückzuführen, dass die protokollierten Ereignisse die Änderung von Daten in den jeweiligen Feldern betreffen. Deshalb fallen die in den Log-Einträgen festgehaltenen Informationen sehr homogen aus und lassen sich gut interpretieren. Die Logdaten, die im Rahmen von IT-Beschaffungen BUKEA / BSW generiert

werden, sind insbesondere deshalb gut zu verstehen, da Ereignistypen verständliche Namen tragen. Kleinere Probleme resultieren daraus, dass im Rahmen der Beschaffungsvorgänge bestimmte Prozessschritte durch Speicherung elektronischer Dokumente determiniert werden. Im Log tauchen diese Schritte mit dem Ereignis *Datei angelegt* auf. Wichtige Kontextinformationen sind jedoch nur im Dokument enthalten und lassen sich im Log bestenfalls noch vom dokumentierten Dateinamen ableiten. Die Interpretation der Logdaten von JUS-IT stellt eine tendenziell große Herausforderung dar, weil die aufgeführten Datenbankoperationen nur mit entsprechendem Know-how der Datenbanksprache zu verstehen sind und insbesondere die Aktivitäten des Anwenders nicht ohne Weiteres isoliert betrachtet werden können.

Datenformate

Die von KoPers erzeugten Logdaten liegen als Liste von Bearbeitungsschritten in einem Format vor, das sich für Analysen mit Process Mining eignet. Bei IT-Beschaffungen BUKEA / BSW lassen sich aus der Dokumentation der Vorgänge heraus einzelne Reports erstellen, die die protokollierten Ereignisse umfassen. Da die beteiligten Mitarbeiter die Protokollierung manuell auf einem MS SharePoint vornehmen, sind festgehaltene Daten nicht auf eine automatisierte Verarbeitung ausgerichtet, sondern liegen in einem für Menschen gut lesbaren Format vor. Im Fall von JUS-IT unterscheiden die Einträge zwar Attribute und geben insbesondere Zeitstempel und Datenbankoperation in einem Listenformat wieder. Die für viele Analysen erforderlichen Transformationen, bei denen die SQL-Statements in mehrere Attribute überführt werden, würden jedoch entsprechende Erweiterungen erforderlich machen.

Zusammenfassung

Die folgende Tabelle führt die identifizierten Herausforderungen, die bei der Untersuchung der Voraussetzungen für den Einsatz von Process Mining anhand der drei erhobenen Fachverfahren ermittelt wurden, in einer Übersicht zusammen. Hierbei wird anhand der verwendeten Symbole gekennzeichnet, ob die jeweilige Herausforderung im Rahmen

	KoPers	IT- Beschaffungen BUKEA/BSW	JUS-IT
H1 – Abläufe werden nicht vollständig von den Logs abgebildet	●	●	◐
H2 – Logdaten verschiedener Systeme sind meist heterogen	○	◐	●
H3 – Granularität des Logs	◐	●	●
H4 – Menge protokollierter Attribute	○	○	◐
H5 – Variation der Attributs-Ausprägungen	◐	○	●
H6 – Intuitive Verständlichkeit der Logs	○	◐	●
H7 – Format der Logdateien	○	◐	○

Tabelle 1: Fachverfahrensspezifische Relevanz identifizierter Herausforderungen

des einzelnen Fachverfahrens keine [○], mittlere[◐] oder hohe[●] Relevanz besitzt.

3.4 Maßnahmen für gute Voraussetzungen für das Process Mining

3.4.1 Erläuterungen zu den vorgeschlagenen Maßnahmen

Die im folgenden Teil vorgestellten Maßnahmen beziehen sich auf technische Aspekte, die gute Voraussetzungen für die Anwendung von Process Mining in der öffentlichen Verwaltung schaffen. Diese Empfehlungen gehen jedoch nicht auf juristische oder organisationspezifische Einschränkungen ein, die der Auswertung protokollierter Prozessinformationen grundsätzlich entgegenstehen könnten. Die aufgeführten Ansätze sind also gegebenenfalls zusätzlich mit relevanten Richtlinien wie dem Datenschutz oder besonderen Anforderungen (bspw. des Personalrats) abzustimmen und in Einklang zu bringen.

Weiterhin setzt die Gestaltung guter Voraussetzungen für Process Mining an einigen Stellen voraus, dass bereits ein gewisses Prozesswissen vorhanden ist, um bestimmte Anforderungen an die Logdaten formulieren und damit Potenziale aus Analysen erschließen zu können. Somit rücken Vorgaben zur Durchführung betrachteter Prozesse in den Fokus. Nur wenn die Vorgehensweise zur Durchführung adressierter Prozesse klar definiert ist und diese Informationen für ausführende Mitarbeiter zugänglich sind, lassen sich für den Einzelfall Maßnahmen ergreifen. Neben Dokumenten wie Arbeitsanweisungen und Prozessbeschreibungen sind Prozessmodelle wichtige Werkzeuge zur Dokumentation, Diskussion und Weiterentwicklung der Abläufe. Je nach Modellierungsmethode werden bestimmte Details in den Modellen vernachlässigt oder vereinfacht. Aus diesem Grund empfiehlt sich im Kontext von Process Mining eine Modellierungsmethode, die neben der Darstellung von Ablauflogiken prozessrelevante Informationen und Informationsobjekte formal beschreibt und einbezieht. Die im Rahmen der Studie vorgestellte und eingesetzte Modellierungsinfrastruktur HERAKLIT wird diesen Anforderungen gerecht.

Die vorliegende Studie sowie die vorgeschlagenen Maßnahmen und Hand-

lungsempfehlungen sind für die öffentliche Verwaltung erarbeitet worden. Viele der technischen Aspekte adressieren eingesetzte Softwarelösungen, die in den meisten Fällen keine Eigenentwicklungen sind, sondern von Softwareanbietern der Privatwirtschaft erworben werden. Entsprechend muss die Durchsetzung der Vorschläge im Zweifel indirekt erfolgen. Die Maßnahmen können der Verwaltung also helfen, neue oder weitere Anforderungen für eigenentwickelte Software zu formulieren, bestimmte Anforderungen bei der Auswahl und Anschaffung neuer Software zu berücksichtigen oder Entwickler aufzufordern, bestimmten Defiziten bei der bereits im Einsatz befindlichen Software nachträglich entgegenzuwirken.

3.4.2 M1 – Abläufe vollständig mit IT unterstützen

Um Abläufe vollständig in Logdaten zu dokumentieren und dadurch die Voraussetzungen zur Anwendung von Process Mining zu verbessern, muss eine umfassende Unterstützung durch IT gegeben sein. Dies schließt sowohl den Umfang der Prozessschritte als auch den Kreis aller Prozessbeteiligten sowie Vorgänge ein, die automatisiert abgewickelt werden. Besonders wichtig ist in diesem Zusammenhang die Aufnahme von Entscheidungen innerhalb betrachteter Abläufe. Das protokollierende System sollte in der Lage sein, die relevanten Aspekte sowie getroffene Absprachen und Abstimmungen transparent und nachvollziehbar festzuhalten. Entsprechend soll das eingesetzte IT-System den Workflow steuern und die Durchführung protokollieren. Falls hier relevante Informationen in einer ungeeigneten Form z. B. natürlichsprachlich einbezogen werden, ist es empfehlenswert, die wesentlichen Aspekte durchgeführter Schritte zusätzlich im Workflow manuell dokumentieren zu lassen.

3.4.3 M2 – Standards festlegen (adressiert H2 und H7)

Die Einführung und Verbreitung standardisierter Vorgaben für die Protokollierung von Logdaten können die Erschließung der darin verborgenen Potenziale erheblich steigern. Standards adressieren dabei neben dem

Dateiformat auch Richtlinien zur Struktur, der Namensgebung häufig genutzter Attribute und die Bezeichnungen wiederkehrender Attributausprägungen. Wenn Softwareanbieter derartigen Standards folgend Anwendungssoftware entwickeln, trägt dies in der Verwaltung zu einer Homogenisierung erzeugter Logdaten bei und steigert somit behördenübergreifend das Verständnis und die Möglichkeiten zur gemeinsamen Nutzung von Process Mining. Entscheidend ist insbesondere eine Strategie zur Kennzeichnung fachlich zusammengehöriger Abläufe. Nur wenn erkennbar ist, welche Vorgänge in Beziehung zueinander stehen, sich aber auf verschiedene Systeme erstrecken, lassen sich diese Abläufe vollumfänglich, system- und ggf. organisationsübergreifend analysieren.

3.4.4 M3 – Sinnvoller Detailgrad für Logdaten

Der Detailgrad der Logdaten wird insbesondere durch die Granularität der Einträge im Log bestimmt. Ein gut geeignetes Niveau ist hierbei von der Verwendung der Logdaten, also den geplanten Analysen, abhängig. Hierbei gilt, dass eine zu grobe Granularität tendenziell detailarm ausfällt und bestimmte Analysen nicht zulässt, während eine zu feine Granularität in der Regel zusätzlichen Aufwand verursacht, aber tendenziell umfänglichere Analysen erlaubt. Aus dieser Erkenntnis lässt sich ableiten, dass es grundsätzlich besser ist, die Ereignisse von Abläufen eher detailliert zu protokollieren. Eine sinnvolle Lösung kann hierbei durch die Struktur der Logdaten erreicht werden, indem zwar feingranulare Einträge erzeugt werden, diese sich jedoch ggf. mehrstufig verdichten lassen. So können Gruppen von Detailereignissen aggregiert werden und erlauben Process-Mining-Analysen auf mehreren Abstraktionsebenen.

3.4.5 M4 – Gute Auswahl an Attributen

Die Menge der in einem Log aufgegriffenen Attribute kann grundsätzlich großzügig gewählt werden und sollte das Ziel verfolgen, Merkmale möglichst *atomar* darzustellen. Jedoch ist wichtig, dass Klarheit herrscht,

welche Informationen mit einem bestimmten Attribut transportiert werden. Unter dieser Prämisse kann für Analysen mit Process Mining jederzeit selektiert werden, welche Attribute berücksichtigt werden. Essenziell sind *Zeitpunkt* eines bestimmten Ereignisses, *Ereignistyp* und die *Zuordnung* zu einer bestimmten Durchführung des protokollierten Ablaufs. Alle weiteren Attribute, die dabei helfen, Einträge im Log zu klassifizieren, können je nach Analysefokus nützlich sein. Umgekehrt kann ein zu großer Umfang enthaltener Attribute ohne entsprechende Kenntnisse über die jeweiligen Merkmale die Vorbereitung von Analysen erschweren.

3.4.6 M5 – Geeignete Menge an Attributsausprägungen

Die Menge möglicher Attributsausprägungen determiniert, inwiefern sich einzelne Logeinträge hinsichtlich eines bestimmten Attributs voneinander unterscheiden können. Da dies je nach Prozess hochgradig individuell ausfallen kann, sind Richtwerte nicht zielführend. So lässt sich nur allgemein ausdrücken, dass sich ein geeignetes Maß möglicher Ausprägungen dadurch auszeichnet, dass signifikante Unterschiede hinsichtlich eines Merkmals zu weiteren Ausprägungen führen, während irrelevante Abweichungen die Menge möglicher Merkmalsausprägungen nicht unnötig erhöhen sollten. Eine Voraussetzung für Ausgewogenheit hinsichtlich der Menge an Ausprägungen ist die in Abschnitt 3.3.2.4 aufgeführte gute Auswahl atomarer Attribute. Denn falls ein Attribut mehrere Merkmale kombiniert abbildet, erhöht sich die Menge möglicher Ausprägungen multiplikativ. Die daraus resultierende Beeinträchtigung der Analysemöglichkeiten gilt es, zu vermeiden.

3.4.7 M6 – Gute Verständlichkeit erzeugen

Um verständliche Logdaten zu erzeugen, herrschen mehrere, teilweise konkurrierende Zielstellungen. So sollen die Attributtypen und -ausprägungen intuitiv interpretierbar sein. Demgegenüber sollen sie idealerweise jedoch auch kompakt ausfallen und die Merkmale mit einer Zeichenfolge in beschränktem Umfang

eindeutig beschreiben. Wenn natürlichsprachliche Begriffe verwendet werden, müssen hinsichtlich des Log-Designs Entscheidungen wie beispielsweise die verwendete Sprache oder der Gebrauch von Abkürzungen getroffen werden, die je nach Adressaten auch Auswirkungen auf die Verständlichkeit haben können. Diesen Problemen kann und sollte eine ausführliche Dokumentation entscheidend entgegenwirken.

3.5 Diskussion

Im Rahmen der Studie werden drei ausgewählte Fachverfahren der öffentlichen Verwaltung aus zwei unterschiedlichen Perspektiven betrachtet. Zum einen wurden Experteninterviews geführt. Gewonnene Erkenntnisse wurden durch Anwendung der Modellierungsmethode HERAKLIT in die in Abschnitt 3.1 dargestellten Modelle überführt und mithilfe der Experten aus den jeweiligen Verwaltungsbereichen evaluiert. Diese Modelle beschreiben die untersuchten Abläufe und beantworten somit die Forschungsfrage RQ1. Zum anderen haben die Verfahrensverantwortlichen exemplarische Logdaten zu den betrachteten Abläufen bereitgestellt, um Einblicke in die für die Anwendung von Process Mining potenziell verfügbaren Daten zu gewähren. Durch die Untersuchung der Logdaten wurde deutlich, welche Informationen der betrachteten Abläufe von den eingesetzten IT-Systemen gespeichert werden und wie die Protokollierung gestaltet ist. Diese Erkenntnisse beantworten die Forschungsfrage RQ2.

Beide Perspektiven auf die Fachverfahren zeigen bestimmte Details auf und liefern Informationen zu den untersuchten Abläufen. Dabei existiert eine gewisse Schnittmenge, also Details, die sowohl aus den entwickelten HERAKLIT-Modellen als auch den bereitgestellten Logdaten hervorgehen. Um die Voraussetzungen und insbesondere Verbesserungspotenziale für die Anwendung von Process Mining zu betrachten, sind gerade jene Aspekte von Interesse, die zwar ausgehend von den Interviews in HERAKLIT-Modelle überführt werden konnten, bei reiner Betrachtung der Logdaten jedoch verborgen bleiben. Darüber hinaus wird beurteilt, welche technischen

Herausforderungen bei der Verwertung vorhandener Daten aus den Logs auszumachen sind. Aus den Erkenntnissen zu den in Logdaten fehlenden Informationen und den technischen Umständen, die eine Verwertung vorhandener Informationen erschweren, wurden in Abschnitt 3.3 sieben zentrale Herausforderungen für die Anwendung von Process Mining in den betrachteten Verfahren erarbeitet und beantworten damit die Forschungsfrage RQ3.

In Abschnitt 3.4 werden Vorschläge präsentiert, die die Voraussetzungen für die Anwendung von Process Mining in der öffentlichen Verwaltung verbessern können. Es handelt sich um abgeleitete Maßnahmen, die den im Rahmen der Fallstudie identifizierten Herausforderungen entgegenwirken. Hierbei werden hinsichtlich lösungsorientierter Handlungsempfehlungen vor allem die Aspekte, Vollständigkeit, Umfang, Struktur, Variation und Format von Logdaten adressiert. Auch wenn diese Empfehlungen aus nur drei untersuchten Fachverfahren resultieren und dort jeweils in unterschiedlichem Umfang aufgefallen sind, lassen sie sich auf einen Großteil der in der Verwaltung üblichen Verfahren und ein breites Spektrum möglicher Anwendungsszenarien übertragen. Mit den entwickelten Maßnahmen wird Forschungsfrage RQ4 beantwortet.

Die Ergebnisse der Studie bieten verschiedene Empfehlungen für die Verwaltung, um zukünftig gesteigerten Nutzen aus der Anwendung von Process Mining ziehen zu können. Die in Abschnitt 3.4 vorgestellten Maßnahmen und erläuterten Zusammenhänge halten hierfür gebündelt Erkenntnisse bereit. Es handelt sich hierbei vor allem um technische Maßnahmen zur Verbesserung der Logdatenbasis. Eine Umsetzung derartiger Empfehlungen besitzt grundsätzlich einen eher langfristigen Charakter, da vorgeschlagene Eingriffe signifikante Veränderungen der Logdaten nach sich ziehen. Die Verwendung der Logdaten, die vor entsprechenden Maßnahmen erzeugt wurden, gemeinsam mit denen, die nach den Maßnahmen entstanden sind, ist dann nur noch bedingt möglich. Neben den Maßnahmen aus Abschnitt 3.4 sind während der Untersuchungen weitere Aspekte aufgefallen, die zusätzliche Hinweise zur Vertiefung der Ansätze und weiterführende

Überlegungen beinhalten. Sie sollen im folgenden Teil schlaglichtartig aufgegriffen werden.

Das breite Aufgabenspektrum und die oft stark fragmentierten Organisationsstrukturen in der öffentlichen Verwaltung führen zu einer überaus heterogenen Systemlandschaft. Diese Ausgangslage belastet die Möglichkeiten zum Einsatz von Process Mining einerseits dadurch, dass Lösungen oft individuell zugeschnitten werden müssen und andererseits durch die Notwendigkeit, bei übergreifenden Prozessen verschiedene Systeme in die Analysen einbeziehen zu müssen. Hier kommt erschwerend hinzu, dass insbesondere im kommunalen Bereich für gleiche Aufgaben teilweise verschiedene Softwarelösungen eingesetzt werden. Vorgeschlagene Maßnahmen zur Standardisierung können dieses Problem zwar grundsätzlich abschwächen, erfordern jedoch eine umfassende Abstimmung und Verständigung auf gemeinsame Ziele. Gelingt es jedoch, mit Process Mining umfassendes Prozesswissen aufzubauen und bspw. unter Beteiligung diverser Verwaltungsbereiche Best Practices zu entwickeln, können davon auch zahlreiche Organisationseinheiten profitieren.

In allen drei untersuchten Verfahren wurde festgestellt, dass die analysierten Logdaten nur einen Ausschnitt betrachteter Abläufe wiedergeben. Dies ist eine für die öffentliche Verwaltung sehr typische Ausgangslage. Zwar können hier die voranschreitende Digitalisierung und die vorgeschlagenen Empfehlungen zur umfassenderen Unterstützung der Prozesse mit IT gegensteuern, auf absehbare Zeit wird dieses Problem jedoch in vielen Bereichen zunächst fortbestehen. Für die Anwendung von Process Mining bedeutet dies zwar eine Einschränkung, dennoch kann mit den vorhandenen Daten auch heute schon Nutzen generiert werden. So können für protokollierte Ausschnitte der Abläufe unter anderem häufige Vorgehensweisen aufgedeckt sowie Konformitätsprüfungen oder Bottleneck-Analysen durchgeführt werden, die wertvolle Erkenntnisse hervorbringen. Erhobene Modelle der Abläufe werden oft einen vollständigeren Überblick bieten, blenden jedoch auch bestimmte Details aus, um Komplexität zu reduzieren. Somit sind und bleiben beide Perspektiven auf die vorhandenen Abläufe für die Weiterentwicklung der Verwaltung wertvoll.

Die im Rahmen der untersuchten Fachverfahren analysierten Logdaten stellten sich hinsichtlich ihrer Granularität sehr unterschiedlich dar. Die Bandbreite reicht hier vom Festhalten der wichtigsten Ereignisse bis hin zur detaillierten Dokumentation sämtlicher Datenbankzugriffe. Ein optimales Niveau lässt sich dabei nicht pauschal festlegen, sondern ist abhängig von den Zielen geplanter Process-Mining-Analysen. So macht es einen großen Unterschied, ob mithilfe von Process-Mining-Techniken aus den Logdaten ein übersichtliches Prozessmodell abgeleitet werden soll oder ob das Ziel verfolgt wird, Erkenntnisse zur Usability (Gebrauchstauglichkeit) eines bestimmten Systems und damit mögliche Ansätze zur Verbesserung der Benutzerfreundlichkeit eingesetzter Softwarelösungen aufzudecken. Sinnvolle Anforderungen an den Umfang protokollierter Logdaten lassen sich also derart formulieren, dass Logdaten im Idealfall sehr detailliert festgehalten werden, sich jedoch anhand bestimmter Attribute und im Rahmen der Datenvorbereitung auch auf eine grobe Granularität reduzieren lassen.

Hinsichtlich der Verwendung von Attributen zur Strukturierung einer Log-Datei werden in Abschnitt 3.4 zielführende Maßnahmen vorgestellt. Dass Logdateien aus der Praxis diesen Empfehlungen häufig nicht entsprechen, ist meist darauf zurückzuführen, dass sie nicht primär darauf ausgerichtet sind, die Interessen des Anwenders hinsichtlich möglicher Analysen zu bedienen. Häufig erlauben die Logdaten bestimmte Funktionalitäten der Anwendungssoftware wie beispielsweise die Möglichkeit, durchgeführte Operationen rückgängig zu machen. Nicht selten adressieren erzeugte Logdaten unter Umständen auch den Entwickler, der mit deren Hilfe Fehler oder anomales Systemverhalten untersuchen kann. Insgesamt werden zwar einige Entwickler auch die Interessen der Anwender und den möglichen Nutzen der Logdaten für sie bei der Gestaltung und Arbeitsweise von Logging-Mechanismen berücksichtigen, dennoch sollte seitens der Verwaltung hier aktiv auf die verantwortlichen Entwickler zugegangen werden, um konkrete Anforderungen durchzusetzen.

Bezüglich des Logdaten-Designs werden gemeinsame Standards und Formate vorgeschlagen. Dies ist durch das breite Auf-

gabenspektrum der Verwaltung und die damit einhergehende IT-Systemlandschaft eine wichtige und zielführende Maßnahme, um den Einsatz von Process Mining zu begünstigen. Der Vorschlag stößt allerdings dort an gewisse Grenzen, wo Softwarelösungen eingesetzt werden, die nicht ausschließlich auf die Verwaltung zugeschnitten sind. Hier werden beispielsweise Office-Lösungen zur Textverarbeitung, Tabellenkalkulation oder E-Mail-Kommunikation adressiert. Derartige Standardsoftware wird in vielen Verfahren für Teilprozesse eingesetzt und erzeugt auch Logdaten. Folglich sollten diese Daten auch bei Process-Mining-Analysen eine gewisse Berücksichtigung finden. Allerdings wäre eine Umsetzung von Standards und gemeinsamen Formaten hier mit weiteren Branchen im internationalen Raum abzustimmen und ist deshalb nicht ohne Weiteres umsetzbar.

Die vorliegende Studie präsentiert auf Basis von Erkenntnissen, die in Zusammenarbeit mit der Verwaltungspraxis erarbeitet wurden, Handlungsempfehlungen zur Verbesserung der Voraussetzungen für die Anwendung von Process Mining. Dabei wurden aus der Vielzahl der in der Verwaltung existierenden Abläufe drei ausgewählte Fachverfahren untersucht. Viele Lösungsansätze lassen sich zwar grundsätzlich auf zahlreiche Verfahren der öffentlichen Verwaltung übertragen, weil die festgestellten Herausforderungen nicht an die vorgestellten Fachverfahren geknüpft sind, sondern in unterschiedlichen Kontexten auftreten und deshalb auch in vielen anderen Verwaltungsbereichen vermutet werden können. Dennoch existieren über die aufgegriffenen Aspekte hinaus weitere Herausforderungen für die Verwendung von Process Mining, die im Rahmen der Studie nicht deutlich in Erscheinung getreten sind. Dies betrifft u. a. Verfahren, die sowohl system- als auch behördenübergreifend ausgerichtet sind und deshalb einen erhöhten technischen sowie organisatorischen Abstimmungsbedarf aufweisen. Des Weiteren ist der Umfang der für die Untersuchung bereitgestellten Logdaten sehr unterschiedlich ausgefallen. So wurde bspw. in einem der drei betrachteten Verfahren nur ein einzelner Ablauf in Form protokollierter Daten übermittelt, während für ein anderes Verfahren eine Vielzahl an Logeinträgen verteilt auf diverse Dateien in

einer komplexen Ordnerstruktur bereitgestellt wurde. Somit ist im ersten Fall möglich, dass bestimmte Details in den Daten nicht enthalten waren und deshalb bestimmte Aspekte nicht erkannt werden konnten. Im anderen Fall können durch die Vielzahl der Einträge jedoch ebenso gut gewisse Auffälligkeiten unerkannt geblieben sein.

Das erarbeitete Ergebnis stellt sieben konkrete Maßnahmen vor, die den identifizierten Herausforderungen zur Anwendung von Process-Mining-Techniken sinnvoll entgegenwirken können. Daneben verdeutlicht die Untersuchung auch, dass die Analyse von Logdaten nur eine spezielle Perspektive auf die Prozesse der Verwaltung ermöglicht und gewissen Einschränkungen unterliegt. Hierbei werden insbesondere die Vollständigkeit und Interpretierbarkeit der Daten adressiert. Gleichzeitig wird erläutert, wie Analysewerkzeuge mit Process Mining die Logdaten auswerten und dass aus diesen Methoden weitere Einschränkungen resultieren. Insbesondere die Annahme einer totalen zeitlichen Ordnung betrachteter Einträge in einer Log-Datei blendet diverse Details aus und kann zu falschen Schlussfolgerungen führen. Somit verdeutlicht die Studie zwei weitere essenzielle Erkenntnisse: Zum einen lässt sich ein erheblicher Teil des wertvollen Prozesswissens zu den Abläufen in der öffentlichen Verwaltung nicht aus gespeicherten Logdaten gewinnen. Dieses Know-how liegt häufig bei den an den Verfahren beteiligten Mitarbeitern und nicht in gespeicherten Daten. Zum anderen wird deutlich, dass die Methodik zur Dokumentation von Prozesswissen – auch unter Einbeziehung von Logdaten – einen entscheidenden Einfluss auf deren Verwertung besitzt und deshalb besondere Anforderungen erfüllen sollte. Die im Rahmen der Studie verwendete Modellierungsmethode HERAKLIT kann die Architektur, Statik und Dynamik eines Systems beschreiben und eignet sich deshalb zur formalen Darstellung von Abläufen, die IT-Instrumente einbeziehen. Mit einem derartigen Ansatz kann es gelingen, verschiedene Perspektiven auf Verwaltungsverfahren zu konsolidieren und das Prozesswissen in und außerhalb gespeicherter Daten in nutzenstiftende Modelle zu überführen.

4 Handlungsempfehlungen

Ziele definieren

Um den Einsatz von Process Mining geeignet vorzubereiten, ist es wichtig, eine klare Vorstellung davon zu haben, welche Erkenntnisse mit dem Technologieeinsatz erzielt werden sollen. Zwar sind viele der in dieser Studie erarbeiteten Empfehlungen recht allgemeingültig und verbessern die Voraussetzungen für entsprechende Techniken unabhängig von konkreten Untersuchungsschwerpunkten. Dennoch ist es nützlich, bestimmte Maßnahmen bereits an den geplanten Analysen auszurichten und zu priorisieren, um die Anforderungen an die für Analysen relevanten Daten zu schärfen.

Vorhandenes Prozesswissen konsolidieren

Prozessorientierte Untersuchungen bauen in der Regel auf vorhandenem Know-how zu den adressierten Abläufen auf. Somit sind neben verfügbaren Dokumentationen insbesondere Mitarbeiter zu identifizieren, die wertvolles Wissen zur Vorbereitung von Process-Mining-Analysen bereitstellen können. Hierzu gehören sowohl Prozessbeteiligte, deren reguläres Aufgabenfeld die Umsetzung jeweiliger Verfahren einschließt, Prozessverantwortliche, die administrativ tätig sind und u.a. die strategische Ausrichtung der Vorgänge steuern sowie technisch versierte Personen, die mit den Anwendungen, Systemen und der IT-Infrastruktur vertraut sind. Die Zusammenführung dieser Quellen für wertvolles Know-how bezüglich der zu untersuchenden Prozesse ist essenziell, um die genaue Ausgangslage zur Vorbereitung von Process-Mining-Techniken zu beurteilen.

Prozesswissen geeignet dokumentieren

Wenn das vorhandene Know-how zu bestimmten Abläufen, die mit Process Mining weiterführenden Analysen unterzogen werden sollen, zusammengetragen wurde, ist es wichtig, dieses Wissen einheitlich, klar und verständlich zu dokumentieren. Damit soll eine gute Grundlage geschaffen werden, ein gemeinsames Verständnis adressierter

Prozesse zu schaffen, um Maßnahmen zur Prozessanalyse diskutieren und priorisieren zu können. Mit diesem Ziel werden in der Regel Prozessmodelle angefertigt. Sie besitzen die Eigenschaft, viele Details übersichtlich abbilden zu können und lassen sich kognitiv besser erfassen als beispielsweise eine natürlichsprachliche Beschreibung. Hinsichtlich der Auswahl einer Modellierungsmethode empfiehlt es sich, zur Vorbereitung von Process Mining, eine Notationsform zu wählen, bei der neben den betrachteten Abläufen auch Informationen zu relevanten Systemarchitekturen dokumentiert und formal beschrieben werden. In der vorliegenden Studie wird deshalb HERAKLIT verwendet.

Berücksichtigung ursprünglicher Logdaten

Neben dem Sammeln fachlicher und organisatorischer Details zu den betrachteten Abläufen ist es ebenso wichtig, die Ausgangslage hinsichtlich vorhandener Logdaten zu betrachten. Hierbei können technische Dokumentationen zur protokollierenden Software nützlich sein. Ebenso ist insbesondere bei fremdbezogenen IT-Instrumenten ein enger Austausch mit den jeweiligen Entwicklern anzustreben, damit diese bei der Interpretation vorhandener Logdaten unterstützen können. Hierbei geht es vor allem um Erläuterungen zu verwendeten Abkürzungen oder kryptischen Bezeichnungen für Attribute bzw. deren Ausprägungen. Ebenso kann der Entwickler unter Umständen erläutern, warum bestimmte Informationen nicht ersichtlich sind, dort jedoch vermutet wurden.

Vollständige Protokollierung der Abläufe

Um bestmögliche Ergebnisse aus Process-Mining-Analysen erzielen zu können, müssen sich alle Details, die im Rahmen der Analysen von Interesse sind, in den protokollierten Logdaten widerspiegeln. Somit ist von entscheidender Bedeutung, dass die betrachteten Abläufe möglichst umfassend von IT-Systemen unterstützt werden. Allgemein sollte also ein hoher Digitalisierungsgrad für alle Prozesse

in der öffentlichen Verwaltung angestrebt werden. Im Detail heißt dies für ein einzelnes Verfahren, dass Operationen, die den Zustand eines Ablaufs verändern, möglichst mit einer Anwendungssoftware ausgeführt und unter der Voraussetzung geeigneter Logging-Mechanismen protokolliert werden. Für Arbeitsschritte, die außerhalb der Software durchgeführt werden, sollte die Software passende Funktionalitäten anbieten, die dem Mitarbeiter eine manuelle Dokumentation mit wenig Aufwand erlauben. Dadurch können sich diese Prozessschritte ebenfalls in den Logdaten niederschlagen und werden ebenso aus fachlicher Sicht festgehalten. Werden für ein Verfahren mehrere IT-Instrumente mit jeweils separatem Logging eingesetzt, so ist eine Zusammenführung notwendig, die oft auch eine Transformation zwecks Harmonisierung der Daten erfordert.

Gestaltung der Logdaten

Eine optimale Ausrichtung der Inhalte einer Log-Datei adressiert maßgeblich den Umfang und die Struktur enthaltener Informationen. Bezüglich des Umfangs ist neben den im vorangegangenen Absatz beschriebenen Hinweisen zur Vollständigkeit auch der Detailgrad zu beachten, mit der die Ereignisse von Abläufen in den Logdaten festgehalten werden. Um gute Voraussetzungen für Process-Mining-Analysen zu schaffen, ist grundsätzlich eine feine Granularität der Einträge anzustreben, wobei sich idealerweise in den Daten auch Informationen zur Relevanz des einzelnen Ereignisses niederschlagen, sodass für den Verlauf des Verfahrens entscheidende Ereignisse von weniger relevanten unterschieden werden können. Diese Eigenschaft ermöglicht dann eine Betrachtung der Logdaten auf mehreren Detailgraden. Weiterhin sind auch die Menge der in einer Log-Datei enthaltenen Attribute sowie deren Ausprägungen zu berücksichtigen. Umfang und Struktur lassen jedoch aufgrund der sehr individuellen Anforderungen einzelner Verfahren kaum Richtlinien zu und müssen zweckmäßig gewählt werden. Essenziell ist jedoch, dass Klarheit bezüglich des Informationsgehalts jeder einzelnen

Ausprägung eines jeden Attributs herrscht. Sinnvoll ist in diesem Zusammenhang auch, Attributausprägungen so zu bezeichnen, dass sie möglichst intuitiv gelesen werden können.

Verständigung auf Standards

Aus der heterogenen Systemlandschaft der öffentlichen Verwaltung resultiert auch eine breite Vielfalt unterschiedlichster Logdaten. Das Ziel, eintretende Ereignisse von Abläufen zu benennen, mit einem Zeitstempel zu versehen, einem bestimmten Ablauf zuzuordnen und eingebundene Ressourcen zu annotieren, ist dabei vom Ansatz meist identisch, unterscheidet sich jedoch häufig gravierend in der Umsetzung. Somit ist es naheliegend, die Logdaten der einzelnen Systeme für die öffentliche Verwaltung zu vereinheitlichen, indem gewisse Standards eingeführt werden. Hierbei werden die Struktur von Logdaten, Formate und Bezeichnungen für Attribute bzw. deren Ausprägungen sowie die Verwendung bestimmter Dateiformate adressiert. Neben den Vorteilen für Process-Mining-Analysen von Abläufen, die system- oder behördenübergreifend ausgerichtet sind, fördert diese Maßnahme auch das gegenseitige Verständnis für Prozesse und begünstigt diesbezüglich die Zusammenarbeit und den Erfahrungsaustausch zwischen verschiedenen Organisationseinheiten.

Akzeptanz schaffen

Sowohl die Gestaltung günstiger Voraussetzungen für Process Mining als auch die spätere Durchführung und Auswertung bestimmter Analysen bedarf in der Regel die Mitwirkung prozessbeteiligter Mitarbeiter. Um deren Unterstützung sicherzustellen, ist es besonders wichtig, eine ausreichende Akzeptanz für geplante Maßnahmen zu schaffen. Nicht selten wird bei prozessorientierten Untersuchungen vermutet, dass die Ergebnisse für den Einzelnen unvorteilhaft ausgelegt werden könnten. Derartige Sorgen der Mitarbeiter sind nicht unberechtigt, weil Process Mining grundsätzlich Möglichkeiten bereithält, die Durchführung von Abläufen quantitativ und qualitativ auszuwerten. Aus diesem Grund ist

es wichtig, transparent darzustellen welche Untersuchungen mit welchen Zielen vorgesehen sind. Gelingt es hierbei, die Mitarbeiter von den Vorteilen geplanter Maßnahmen zu überzeugen, trägt dies in erheblichem Umfang positiv zur Verbesserung der Voraussetzungen für den Einsatz von Process Mining bei.

5 Zusammenfassung

Diese Studie verfolgt das Ziel, die Voraussetzungen für das Erschließen von verborgenem Prozesswissen in der öffentlichen Verwaltung mit Process Mining zu untersuchen und aus den Ergebnissen Handlungsempfehlungen für deren Gestaltung abzuleiten. Hierzu wurden mit Unterstützung des Amts für Digitalisierung (ITD) drei Fachverfahren der öffentlichen Verwaltung ausgewählt und analysiert. Die Untersuchung konzentrierte sich dabei insbesondere auf eine Prozess-erhebung, aus der verschiedene Modelle zur Darstellung der Abläufe jeweiliger Fachverfahren hervorgingen sowie die Betrachtung der bei der Durchführung der Verfahren entstehenden Logdaten. Diese Logdaten bilden die Grundlage für Analysen mit Process Mining. Entsprechend ist die Beurteilung, inwiefern diese Daten die erhobenen Prozesse wiedergeben, von essenzieller Bedeutung.

Anhand der drei Fachverfahren wurden insbesondere technische Unzulänglichkeiten herausgestellt, die potenzielle Analysen mit Process-Mining-Techniken belasten können. Diese konkreten Probleme wurden in sieben allgemeine Herausforderungen überführt und können somit auch auf viele weitere Fachverfahren der öffentlichen Verwaltung übertragen werden. Anschließend werden in der Studie Maßnahmen erläutert, die den sieben Herausforderungen gezielt entgegenwirken und damit die Voraussetzungen für den Einsatz von Process Mining verbessern. Die jeweiligen Ansätze finden sich auch in den herausgearbeiteten Handlungsempfehlungen des Abschnitts 4 wieder, wurden hierbei jedoch um zusätzliche Aspekte ergänzt, die außerhalb der technischen Betrachtung aufgefallen sind. Dabei handelt es sich vor allem um organisatorische und strategische Maßnahmen, die dazu beitragen können, Process-Mining-Analysen zu begünstigen.

Wesentliche Bestandteile der Studie sowohl für die Durchführung als auch die Erarbeitung von Ergebnissen sind die Erhebung, Modellierung und Evaluation der drei Fachverfahren. Erst die daraus hervorgegangenen Details

zu den betrachteten Abläufen erlauben Einschätzungen zur Güte der von den IT-Systemen erzeugten Logdaten und damit die Entwicklung von Ansätzen zur Verbesserung. Dieser Sachverhalt hat neben der primären Zielstellung der Studie, die sich aus den Forschungsfragen ergibt, weitere wichtige Erkenntnisse hervorgebracht, die die hohe Relevanz von dokumentiertem Prozesswissen unterstreichen.

Die meisten prozessorientierten Untersuchungen setzen bereits ein gewisses Know-how voraus. Selbst Ansätze, die speziell darauf ausgerichtet sind, Abläufe aufzudecken – wie die Discovery-Algorithmen des Process Minings – sind ohne eine geeignete Evaluationsgrundlage nur bedingt verwertbar. Die Studie zeigt auch, dass Fachverfahren in der Verwaltung häufig nur grob dokumentiert werden. Der Großteil detaillierter Prozesskenntnisse lag bei den Mitarbeitern. Mit Blick auf zukünftige Entwicklungen wie die fortschreitende Digitalisierung, die verstärkte Personalfuktuation, den zunehmenden Fachkräftemangel und den sich daraus ergebenden Bedarf, Abläufe immer wirtschaftlicher und effizienter auszurichten, muss die Erhebung und Dokumentation von Prozesswissen für die Verwaltung ein wesentliches Ziel darstellen.

Dokumentiertes Prozesswissen konserviert das gesammelte Know-how zu Abläufen und stellt eine wichtige Grundlage zwecks Diskussion, Analyse und Weiterentwicklung von Verfahren in jeglichen Organisationen dar. Dies gilt auch für die öffentliche Verwaltung und insbesondere im Zusammenhang mit Process-Mining-Ansätzen. Die vorliegende Studie stellt diesbezüglich mit der Methode HERAKLIT eine Notationsform vor, die sich eignet, um gewonnene Informationen zu Abläufen, Systemen und Daten vereint abzubilden. Im direkten Vergleich zu den zum Aufdecken von Prozessen verwendeten Process-Mining-Techniken, die Logdaten nur in zeitlich geordnete Graphen überführen, umfasst HERAKLIT viele Vorteile, die sich deutlich in den erstellten Modellen zeigen. Hier wird klar, dass Verwaltungsvorgänge erheblich komplexer abzubilden sind,

als es mit sequenziell verketteten Ereignissen (Symbolketten) der Discovery-Algorithmen des Process Mining gelingen kann.

Process-Mining-Techniken liefern vor allem Informationen zur chronologischen Ausführung bestimmter Prozessschritte. Zwar lassen sich diese Ergebnisse nur bedingt in allgemeine Prozessmodelle überführen (Process Discovery), weil kausale Zusammenhänge gerade bei nebenläufigen Teilprozessen unzureichend Berücksichtigung finden, dennoch spiegeln sie tatsächliche Ausführungen der Abläufe wider. Ein Nutzen resultiert dann insbesondere aus dem Vergleich der einzelnen Ausführung mit herrschenden Richtlinien (Conformance Checking) sowie der Ermittlung von Metriken und ausgewählten Indikatoren (Process Enhancement).

Gerade in der Verwaltung, wo mit dem breiten Aufgabenspektrum eine Vielzahl an Verfahren einhergeht, Prozesse oft miteinander verzahnt sind und von verschiedensten Mitarbeitern mit abweichenden Zielstellungen teils unterschiedlich umgesetzt werden, stellt die Analyse von Logdaten in Aussicht, viele Potenziale zu heben. Dies wird zusätzlich dadurch verstärkt, dass sich Aufgaben in der Verwaltung oft in teils unterschiedlichen Organisationseinheiten wiederfinden. Somit existiert eine gute Grundlage, verschiedene Abläufe gegenüberzustellen und gemeinsam Verbesserungen zu entwickeln, die dann in der Fläche kommuniziert und eingeführt werden können. Des Weiteren profitieren von verbesserten Verwaltungsprozessen indirekt auch Bürger und Unternehmen. Somit stellen die gewonnenen Erkenntnisse zur Verbesserung der Voraussetzungen für den Einsatz von Process Mining einen wichtigen Mehrwert bereit und können die Entwicklungen in Richtung moderner, digitaler Abläufe der öffentlichen Verwaltung entscheidend fördern.

Literatur

Aalst, W. v. (2016). *Process Mining. Data Science in Action*. Berlin, Heidelberg: Springer.

AIOS, BearingPoint, Cassini, CGI, GBTEC, PwC. (2019). *Digitalisierung der Landesverwaltung in Deutschland: Erfahrungen und Herausforderungen*. Retrieved 06 29, 2022, from <https://aios.de/wp-content/uploads/2019/10/Digitalisierung-der-Landesverwaltung-in-Deutschland.pdf>

Becker, J., Mathas, C., & Winkelmann, A. (2009). *Geschäftsprozessmanagement*. Berlin Heidelberg: Springer-Verlag.

Bissantz, N., & Hagedorn, J. (1993). Data Mining. *WIRTSCHAFTSINFORMATIK* 35(5), 481-487.

Bucher, T., & Winter, R. (2009). Geschäftsprozessmanagement - Einsatz, Weiterentwicklung und Anpassungsmöglichkeiten aus Methodik-sicht. (S. Reinheimer, Ed.) *HMD - Praxis der Wirtschaftsinformatik*(266).

Buijs, J., Dongen, B., & Aalst, W. (2014). Quality Dimensions in Process Discovery: The Importance of Fitness, Precision, Generalization and Simplicity. *International Journal of Cooperative Information Systems*(23).

Cook, J., & Wolf, A. (1995). Automating Process Discovery through Event-Data Analysis. In A. Wolf (Hrsg.), *17th International Conference on Software Engineering (ICSE'95)* (pp. 73-82). Seattle, Washington, USA.

Czarnecki, C., & Fettke, P. (2021). Robotic Process Automation. Positioning, Structuring, and Framing the Work. In C. Czarnecki, & P. Fettke, *Robotic Process Automation. Management, Technology, Applications* (pp. 3-24). De Gruyter.

Düsing, R. (2006). Knowledge Discovery in Databases. In P. Chamonì, & P. Gluchowski, *Analytische Informationssysteme* (pp. 241-262). Berlin Heidelberg: Springer.

Fettke, P., & Reisig, W. (2020). *Heraklit – die erkenntnistheoretisch motivierte Modellierung rechnerintegrierter Systeme*.

Fettke, P., & Reisig, W. (2022). Breathing Life into Models: The Next Generation of Enterprise Modeling. arXiv. doi:10.48550/ARXIV.2205.09591

Fettke, P., & Reisig, W. (2022). Systems Mining with Heraklit: The Next Step. doi:<https://doi.org/10.48550/arXiv.2202.01289>

Fettke, P., & Reisig, W. (2024). Understanding the Digital World. Modeling with Heraklit. Springer, 2024.

Gupta, E. (2014). PROCESS MINING ALGORITHMS. *International Journal of Advance Research In Science And Engineering*, 401-412.

Houy, C., Fettke, P., & Loos, P. (2010). Empirical research in business process management – analysis of an emerging field of research. *Business Process Management Journal*(16), pp. 619-661.

Mahendrawathi, E., Noval, A., Hanim, M., Renny, P., & Rivia, A. (2018). Analysis of Production Planning in a Global Manufacturing Company with Process Mining. *Journal of Enterprise Information Management, Vol. 31*, 317-337.

Reisig, W. (2013). *Petri Nets. Modeling Techniques, Analysis Methods, Case Studies*. Berlin Heidelberg: Springer.

Rozinat, A., Alves de Meldeiros, A., Günther, C., Weijters, A., & van der Aalst, W. (2008). The Need for a Process Mining Evaluation Framework in Research and Practice. In A. ter Hofstede, B. Benatallah, & H.-Y. Paik (Eds.), *Business Process Management Workshops* (pp. 84-92). Berlin Heidelberg: Springer.

van der Aalst, W. (2019). A practitioner's guide to process mining: Limitations of the directly-follows graph. *Procedia Computer Science*, 164(C), pp. 321-328. doi:<https://doi.org/10.1016/j.procs.2019.12.189>

van der Aalst, W., Adriansyah, A., de Medeiros, A., & et. al. (2011). Process Mining Manifesto. *Business Process Management Workshops* (pp. 169-194). Berlin Heidelberg: Springer.

van der Aalst, W., Joos C. A. M., B., & Boudewijn F., V. (2014). Quality Dimensions in Process Discovery: The Importance of Fitness, Precision, Generalization and Simplicity. *International Journal of Cooperative Information Systems*, 23(01). doi:10.1142/S0218843014400012

van der Aalst, W., Joos C. A. M., B., & Boudewijn F., V. (2014). Quality Dimensions in Process Discovery: The Importance of Fitness, Precision, Generalization and Simplicity. *International Journal of Cooperative Information Systems*, 23(01). doi:10.1142/S0218843014400012

van der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, 16(9), pp. 1128-1142.

Weske, M. (2007). *Business Process Management*. Berlin Heidelberg: Springer-Verlag.

Über die Autoren

Oliver Gutermuth ist seit 2018 wissenschaftlicher Mitarbeiter am Institut für Wirtschaftsinformatik (IWi) im Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI). Im Rahmen seiner Promotion beschäftigt er sich mit neuartigen Ansätzen und intelligenten Technologien im Public Sector. Strategien zur Unterstützung und Automatisierung von Geschäftsprozessen in dieser Domäne stellen dabei einen wichtigen Schwerpunkt da. Seine Hauptaufgaben bestehen in der Durchführung von Industrie- und Forschungsprojekten sowie der Vertretung des Instituts gegenüber Projektträgern und Konsortialpartnern aus Industrie und Wissenschaft. In diesen Projekten untersucht er unter anderem die Potenziale von neuartigen Mobilitätskonzepten, Robotic Process Automation sowie Process Mining und entwickelt Anwendungskonzepte unter Einbeziehung von Künstlicher Intelligenz.

Alessandro Benke ist IT-Architekt für digitale Identitäten und Web3. Er erstellt Software-Lösungen und Prozesse, schwerpunktmäßig im hoch regulierten Umfeld, insbesondere im öffentlichen Sektor, dem Finanzsektor sowie der Rechtsbranche. Sein Arbeitsschwerpunkt besteht in der technischen Vorbereitung von Organisationen und Unternehmen für den EU-weiten Rollout des EU Digital Identity Wallets. In diesem Zusammenhang gestaltet er sichere IT-Prozesse und Integrationsarchitekturen, die den Anforderungen der einschlägigen gesetzlichen Bestimmungen der Domäne genügen. Er war an der Entwicklung eines Identity Wallets sowie der Integration von Wallets mit IAM-Systemen beteiligt und hat Blockchain-Anwendungen zur Validierung von USt-IdNrn., zur manipulationssicheren Abwicklung innergemeinschaftlicher Lieferungen und zur Sicherung von Audit Trails entwickelt.

Prof. Dr. **Peter Fettke** ist Professor für Wirtschaftsinformatik an der Universität des Saarlandes und Forschungsgruppenleiter am Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) in Saarbrücken. In seiner Forschung befasst sich Professor Fettke zusammen mit seiner rund 30-köpfigen Forschungsgruppe insbesondere mit der Schnittstelle der Themenkomplexe Wirtschaftsinformatik und Künstlicher Intelligenz (KI). Insgesamt hat er mehr als 250 begutachtete Artikel publiziert. Seine Arbeiten zählen zu den meistzitierten Artikeln international führender Zeitschriften zur Wirtschaftsinformatik und er gehört zu den Top 10 der meistzitierten Wissenschaftler am DFKI. Ebenso ist er gefragter Gutachter renommierter nationaler und internationaler Konferenzen, Journale und Forschungsorganisationen.

Prof. Dr. **Wolfgang Reisig** ist emeritierter Professor für Softwaretechnik und Theorie der Programmierung an der Humboldt-Universität zu Berlin. In den 1970er und 1980er Jahren war er Assistent an der Universität Bonn und der RWTH Aachen, Vertretungsprofessor an der Universität Hamburg, Projektleiter bei der Gesellschaft für Mathematik und Datenverarbeitung (GMD), sowie Professor für Theoretische Informatik an der Technischen Universität München. Prof. Reisig war 1997 Senior Researcher am International Computer Science Institute (ICSI) in Berkeley, Kalifornien. Er erhielt die „Lady Davis Visiting Professorship“ am Technion, Haifa (Israel), den Beta Chair der Technischen Universität Eindhoven, und zwei mal einen IBM Faculty Award für seine Beiträge zur Analyse von Geschäftsprozessen und Service-Modellen.

www.negz.org