

Gabdouline, Gulnara; Kern, Christopher Julian; Krönung, Julia

Article — Published Version

Die ethischen Implikationen von Prozessdaten: Eine systematische Literaturanalyse von Datenethik und Process Mining

HMD Praxis der Wirtschaftsinformatik

Provided in Cooperation with:

Springer Nature

Suggested Citation: Gabdouline, Gulnara; Kern, Christopher Julian; Krönung, Julia (2023) : Die ethischen Implikationen von Prozessdaten: Eine systematische Literaturanalyse von Datenethik und Process Mining, HMD Praxis der Wirtschaftsinformatik, ISSN 2198-2775, Springer Fachmedien Wiesbaden GmbH, Wiesbaden, Vol. 61, Iss. 1, pp. 252-265, <https://doi.org/10.1365/s40702-023-01027-2>

This Version is available at:

<https://hdl.handle.net/10419/311043>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



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



Die ethischen Implikationen von Prozessdaten: Eine systematische Literaturanalyse von Datenethik und Process Mining

Gulnara Gabdoulline  · Christopher Julian Kern · Julia Krönung

Eingegangen: 25. August 2023 / Angenommen: 26. November 2023 / Online publiziert: 19. Dezember 2023

© The Author(s) 2023

Zusammenfassung Process Mining (PM) stellt eine wachsende Disziplin dar, die aufgrund ihres Potenzials zur Verbesserung von Geschäftsprozessen immer mehr Aufmerksamkeit von Forschern und Anwendern auf sich zieht. Wie jede neue Technologie gibt es jedoch auch im Kontext von PM-Bedenken hinsichtlich der ethischen Anwendung. Gerade bezogen auf Erhebung, Verarbeitung und Nutzung von Daten kann es hierbei zu Problemen kommen. Dieser Artikel zielt daher darauf ab, anhand einer Literaturanalyse ethische Implikationen im Process Mining herauszuarbeiten. Dabei wurden 39 Artikel aus sechs Zeitschriften im Bereich PM und 24 Artikel aus vier Zeitschriften im Bereich Datenethik analysiert. Die Ergebnisse zeigen das wachsende Interesse an der Datenethik und PM, aber es befasst sich nur ein geringer Anteil der analysierten PM-Artikel mit datenethischen Grundsätzen. Weitere Forschung ist in Bereichen bestimmter datenethischer Grundsätze, wie Datenqualität und der informierten Zustimmung, erforderlich. Insgesamt bietet diese Studie einen Ausgangspunkt für weitere Forschungen zur ethischen Nutzung von Daten bei der Anwendung von PM und verdeutlicht, dass diesem Bereich mehr Aufmerksamkeit gewidmet werden sollte.

Schlüsselwörter Datenethik · Prozessdaten · Process Mining · Literaturanalyse · Ethische Richtlinien

✉ Gulnara Gabdoulline · Christopher Julian Kern · Julia Krönung
Fernuniversität in Hagen, 58097 Hagen, Deutschland
E-Mail: gulnara.gabdoulline@yahoo.com

Christopher Julian Kern
E-Mail: christopher.kern@fernuni-hagen.de

Julia Krönung
E-Mail: julia.kroenung@fernuni-hagen.de

The Ethical Implications of Process Data: A Systematic Literature Review of Data Ethics and Process Mining

Abstract Process Mining (PM) is of increasing attention for researchers and practitioners alike, especially in regards of the potential to improve business processes in their efficiency. However, like any new data-driven technology, PM raises ethical concerns about the collection, processing, and use of data. The purpose of this paper is to demonstrate ethical implications in process mining by conducting a literature review. To do so, 39 articles from six journals in the field of PM and 24 articles from four journals in the field of data ethics were analyzed. The results show that despite there being growing interest in data ethics and PM, only a small percentage of the PM articles analyzed address data ethical principles. This highlights the need for clear ethical guidelines for use in PM. Further research is needed in areas of specific data ethics principles such as data quality and informed consent. As a result, this review provides a foundation for further research on the ethical use of data in the application of PM.

Keywords Data Ethics · Process Data · Process Mining · Literature Review · Ethical guidelines

1 Einleitung

Immer mehr Unternehmen und Organisationen verlassen sich zunehmend auf daten-gestützte Systeme und Geschäftsmodelle, um fundierte Entscheidungen zu treffen und einen Wettbewerbsvorteil zu erlangen (Elgendy und Elragal 2014). Der rechts-konforme Umgang mit Daten wird hierbei vorausgesetzt. Eine steigende Zahl an Initiativen zeigt, dass dieser nicht ausreicht, sondern ein ethischer Umgang mit Tech-nologie im Allgemeinen (bspw. Spiekermann 2015; Spiekermann et al. 2022; D. E. 2023) und mit Daten im Speziellen (Floridi und Taddeo 2016) auf gesellschaftliches Interesse stößt. Das neue Forschungsfeld „Digital Responsibility“ (Trier et al. 2023) befasst sich mit der verantwortungsvollen Gestaltung von digitalen Technologien und umfasst dabei auch deren ethische Implikationen. Gerade mit diesem Gedanken und dem der verantwortungsvollen Innovation (Stahl et al. 2014) ist der ethische Umgang mit Daten und datengetriebenen Geschäftsmodellen sowohl für Praktiker als auch für Wissenschaftler von Interesse.

Um aus den Daten der Geschäftsprozesse einen zusätzlichen Mehrwert generieren zu können, bietet sich Process Mining (PM), an – wodurch die Verbesserung von Geschäftsprozessen durch die Analyse von Ereignisprotokollen ermöglicht wird (van der Aalst 2011). Allerdings existieren auch hier ethische Bedenken: Unter-nahmen sollten bei den gesammelten Daten, die verarbeitet und gespeichert werden, Datenschutz, Datensicherheit und Transparenz berücksichtigen (van der Aalst 2016). Besonders Bedenken hinsichtlich des potenziellen Missbrauchs sensibler Daten wie bspw. persönlicher Informationen sind bei der Anwendung von PM präsent (van der Aalst 2016).

Dieser Artikel analysiert die ethischen Implikationen von PM-Anwendungen und leitet daraus Richtlinien und bewährte Verfahren zur Sicherstellung von verantwortungsvoller Datennutzung ab. Durch die ethische Nutzung von Daten kann Vertrauen geschaffen, die individuellen Rechte der Nutzer und deren Interessen geschützt werden (Alt et al. 2021). Um diesem Forschungsbereich zu erweitern wurde eine systematische Literaturanalyse durchgeführt. Hierfür wurden die folgenden Forschungsfragen gewählt:

FF 1: *Wie ist der aktuelle Forschungsstand ethischer Implikationen in Process Mining?*

FF 2: *Wie werden datenethische Fragen betrachtet und welche bewährten Verfahren können ethisches Verhalten im Zusammenhang mit Process Mining unterstützen?*

Dieser Artikel ist in mehrere Abschnitte unterteilt. Abschn. 1, die Einleitung betrachtet die wachsende Bedeutung datengestützter Systeme, insbesondere unter Berücksichtigung ethischer Aspekte. Der theoretische Hintergrund in Abschn. 2 beleuchtet Datenethik und Process Mining. Abschn. 3 erläutert die angewandte Forschungsmethode der Literaturanalyse. Die Ergebnisse werden in Abschn. 4, jeweils separat für Datenethik in Abschn. 4.1 und Process Mining in Abschn. 4.2 präsentiert. In der Diskussion in Abschn. 5, werden die ethische Implikationen gegenübergestellt. In der Zusammenfassung in Abschn. 6 werden die Hauptergebnisse zusammengefasst, und Impulse für die praxisnahe Implementierung von PM und Datenmanagement abgeleitet.

2 Theoretischer Hintergrund

2.1 Datenethik

Floridi und Taddeo (2016) definieren Datenethik als den Zweig der Ethik, der moralische Probleme im Zusammenhang mit Daten und Algorithmen untersucht und bewertet, um moralisch gute Lösungen zu formulieren. Die meisten ethischen Fragen in Zusammenhang mit Daten beziehen sich auf Daten über Menschen (Hand 2018). Datenethik bildet einen wesentlichen Bestandteil eines verantwortungsvollen Datenmanagements und stellt sicher, dass Daten zum Wohle der Allgemeinheit verwendet werden (Mittelstadt und Floridi 2016).

Zu den wichtigsten ethischen Grundsätzen gehören Datenschutz (Floridi 2014), Transparenz (Turilli und Floridi 2009), die Vermeidung von Voreingenommenheit, Verzerrung und Diskriminierung durch Daten (Mehrabi et al. 2022), Verantwortung (Slussareff 2022), Informierte Zustimmung (Mittelstadt und Floridi 2016), sowie Datenqualität (Cai und Zhu 2015).

2.2 Process Mining

PM ist ein datengesteuerter Ansatz zur Analyse und Optimierung von Geschäftsprozessen, bei dem Informationen aus Ereignisprotokollen extrahiert werden, um Mus-

ter in Aktivitäten zu erkennen und die Prozessleistung zu messen (van der Aalst und Weijters 2004). PM bietet Vorteile wie die Identifizierung von Prozessineffizienzen, die Kostenreduktion und die Verbesserung der Prozessqualität (Rozinat und van der Aalst 2006). Zu den drei Haupttechniken im PM gehören die Prozessentdeckung (Process Discovery), die Konformitätsprüfung (Conformance Checking) und die Verbesserung (Enhancement) (van der Aalst 2011).

PM weist verschiedene Perspektiven auf, darunter Kontrollfluss-, Fall-, Zeit-, Organisations-, Ressourcen-, Prozess- und Datenperspektiven. Jede Perspektive ermöglicht spezifische Analyse- und Visualisierungsmöglichkeiten (van der Aalst 2011). Bei der Anwendung von PM müssen bestimmte Datenanforderungen berücksichtigt werden (van der Aalst et al. 2010). Die Datenqualität kann sich auf die Genauigkeit und Gültigkeit der durch PM erzielten Ergebnisse auswirken, so dass es wichtig ist, sicherzustellen, dass die verwendeten Daten für den Zweck geeignet sind (Dustdar et al. 2006).

3 Methodik

Um die Forschungsfragen zu beantworten, wurde in dieser Arbeit eine Literaturanalyse nach Webster und Watson (2002) durchgeführt. Die Auswahl dieser Methodik erfolgte aufgrund ihrer strukturierten und systematischen Vorgehensweise, die eine adäquate Identifikation, Auswahl und kritische Bewertung relevanter Literatur ermöglicht. Für den Suchbegriff „Data Ethics“ wurden ausschließlich englischsprachige Zeitschriften im Bereich der Philosophie in Betracht gezogen. Die Auswahl erfolgte auf Basis des SCImago Journal Rank (SJR) (SCImago o.J.), um relevante Literatur zu identifizieren. Der SJR ist ein Indikator, der die wissenschaftliche Bedeutung von Zeitschriften anhand von Zitationsdaten bewertet. Die Top-26-Zeitschriften gemäß SJR-Ranking wurden mittels des Suchbegriffs „Data Ethics“, ohne zeitliche Einschränkungen, einer Volltextsuche unterzogen, was zu 33 Treffern führte. Durch sorgfältiges Screening der Volltexte wurden irrelevante Artikel eliminiert, sodass schließlich 24 Publikationen als relevant identifiziert wurden. Dabei wurde darauf geachtet, dass die ausgewählten Publikationen das Thema nicht nur erwähnen, sondern einen aktiven Forschungsansatz damit verfolgen, oder das Thema zumindest prominent behandeln. Exkludiert hingegen wurden solche Artikel, die das Thema Datenethik nur grob streifen, oder aber lediglich namentlich erwähnen Tab. 1.

Tab. 1 Dokumentation der analysierten Zeitschriften: Datenethik

Zeitschrift	Zeitraum ¹	Analysiert	Berücksichtigt
Minds and Machines	2019–2022	9	7
Philosophy & Technology	2017–2022	18	15
Science Technology and Human Values	2020	2	1
Synthese	2022	1	1
<i>Gesamt</i>	<i>2017–2022</i>	<i>33</i>	<i>24</i>

¹ Der Zeitraum beschreibt deskriptiv die zeitliche Abdeckung der gefundenen Artikel und stellt keine Einschränkung bei der Suche dar

Tab. 2 Dokumentation der analysierten Zeitschriften: Process Mining

	Zeitschrift	Zeitraum ¹	Analysiert	Berücksichtigt
TOP IS-Zeitschriften Tier 1 (Lowry et al. 2013)	MIS Quarterly (MISQ)	2016–2021	2	2
	Information Systems Research (ISR)	2014–2018	2	1
	Journal of Management Information Systems (JMIS)	2018	1	0
TOP IS-Zeitschriften Tier 2 (Lowry et al. 2013)	Decision Support Systems (DSS)	2006–2022	71	32
	European Journal of Information Systems (EJIS)	2006–2017	2	0
	Information & Management (I&M)	2019–2021	2	1
	International Journal of Electronic Commerce (IJEC)	2014	1	0
	Information Systems Journal (ISJ)	–	0	0
	Journal of the Association for Information Systems (JAIS)	2020–2022	6	2
	Journal of Information Technology (JIT)	2010	1	0
	Journal of Strategic Information Systems (JSIS)	2022	2	1
<i>Gesamt</i>	–	2006–2022	90	39

¹ Der Zeitraum beschreibt deskriptiv die zeitliche Abdeckung der gefundenen Artikel und stellt keine Einschränkung bei der Suche dar

Für den Suchbegriff „Process Mining“ wurden die vorgeschlagenen englischsprachigen Top-Tier Zeitschriften der Wirtschaftsinformatik von Lowry et al. (2013) bis Rang 2, ohne zeitliche Einschränkungen, untersucht. Lowry et al. verwenden Clusteranalysen, um ihre Stichprobe von 21 Top-Tier-Fachzeitschriften für den Bereich Information Systems in verschiedene Ebenen zu unterteilen, um deren Einfluss zu bewerten. Die Auswahl erfolgte aufgrund ihrer Reputation und ihres Einflusses in der Forschungsgemeinschaft. Für den Suchbegriff „Process Mining“ wurden 90 Publikationen identifiziert und nach Screening der Volltexte auf ihre Relevanz überprüft. Irrelevante Artikel wurden eliminiert, und nur Publikationen, die aktiv auf PM eingehen wurden inkludiert. Exkludiert wurden solche Artikel, die das Thema zwar nannten, sich aber nicht auf einer inhaltlichen Ebene damit befassen. Insgesamt wurden 39 relevante Publikationen für die Analyse berücksichtigt.

Eine vollständige Liste aller berücksichtigten Artikel ist auf Anfrage bei den Autoren verfügbar Tab. 2.

4 Ergebnisse der Literaturanalyse

4.1 Datenethik

Die vorliegende Konzeptmatrix in Tab. 3 präsentiert eine Übersicht über verschiedene datenethische Grundsätze, wie sie in ausgewählten relevanten Artikeln des Datenethik-Datensatzes identifiziert wurden. Im weiteren Verlauf erfolgt eine detaillierte Analyse bestimmter Grundsätze. Die in der Matrix aufgeführten Artikel

Tab. 3 Konzeptmatrix Datenethik

Publikation	Datenethische Grundsätze							
	Daten- qualität	Daten- schutz	Informierte Zustim- mung	Trans- parenz	Verant- wortung	Verzer- rung	Vorein- genom- menheit	Diskri- minie- rung
Aggarwal und Floridi (2020)	–	x	x	x	x	–	–	–
Chomanski (2021)	–	x	–	–	–	–	x	x
Floridi (2018)	x	x	–	–	–	–	–	x
Floridi et al. (2019)	–	x	–	x	–	–	–	x
Hagendorff (2020)	–	x	x	x	–	–	x	x
Laat (2021)	x	x	–	x	x	x	x	x
Lee und Floridi (2021)	–	–	–	x	–	x	x	x
Loi (2019)	–	x	–	–	–	x	–	x
Mökander et al. (2022)	x	x	–	x	–	–	–	–
Morley et al. (2021)	–	x	–	x	x	x	x	–
Ratti und Graves (2021)	x	–	–	x	–	–	x	–
Saetra und Danaher (2022)	–	x	x	x	x	–	–	–
Symons und Alvarado (2022)	–	x	–	–	x	x	x	x
Taylor (2021)	–	x	x	–	–	–	–	–

repräsentieren dabei diejenigen, die im nachfolgenden Text einer eingehenden Untersuchung unterzogen werden.

Die datenethischen Grundsätze wurden mittels einer induktiven Inhaltsanalyse der als relevant betrachteten Publikationen mit Fokus auf Datenethik ermittelt. Die Analyse zielte darauf ab, durch die Identifizierung von Schwerpunkten in den Publikationen Grundsätze zu erkennen, die für den adäquaten Einsatz datengetriebener Systeme von Bedeutung sind.

Datenqualität: Datenverständnis, -beschreibung und -bereinigung sind entscheidend, um Probleme in Bezug die Genauigkeit und Korrektheit von Aussagen zu vermeiden (Ratti und Graves 2021). Digital Governance ermöglicht Prozesse zur Verbesserung von Datenqualität, -zuverlässigkeit, -sicherheit und -zugänglichkeit

und fördert somit die Entwicklung effektiver Entscheidungsverfahren und Mechanismen zur Rechenschaftspflicht in datenbezogenen Prozessen (Floridi 2018).

Datenschutz: Die Nutzung von KI kann zu erheblichen Herausforderungen im Datenschutz, wie Datenlecks oder fehlerhaft trainierten Algorithmen, führen (Hagendorff 2020; Chomanski 2021; Morley et al. 2021). Die Datenschutzgrundverordnung (DSGVO) wird dabei als entsprechender rechtlicher Rahmen erachtet (Loi 2019; Saetra und Danaher 2022), wobei der bestehende Schutz personenbezogener Daten möglicherweise nicht ausreicht, um Diskriminierung von nicht identifizierten Personen zu verhindern. Resultierend daraus ergibt sich eine vertiefte Diskussion über die die Wahrung der Privatsphäre, der wissenschaftlichen Gültigkeit und den Eigentumsrechten (Loi 2019).

Informierte Zustimmung: Die ethische Verpflichtung zur Einholung der Zustimmung zur Veröffentlichung anonymisierter Daten besteht, trotz fehlender DSGVO-Vorschriften für anonymisierte oder pseudonymisierte Daten. Die Schwierigkeiten bei der Erlangung informierter Zustimmung werden ebenfalls untersucht (Aggarwal und Floridi 2020). Es wird in Frage gestellt, wie man sicherstellen kann, dass Personen den Zweck der gemeinsamen Datennutzung verstehen. Dabei wird betont, dass ohne Vertrauen informierte Zustimmung nicht möglich ist, was alternative Wege zur Sicherung der Rechtmäßigkeit erfordert (Taylor 2021).

Transparenz: Eine Möglichkeit für die verständliche Offenlegung zur Datenerfassung, -nutzung und -weitergabe besteht in der Durchführung von Audits, um das Vertrauen in KI-basierte Systeme zu stärken (Mökander et al. 2022). Gleichzeitig wird Kritik über mangelnde Transparenz in KI-Forschung und -Entwicklung, mit negativen Auswirkungen auf Nutzerzustimmung, Datenschutz und Rechenschaftspflicht, geäußert (Hagendorff 2020). Verschiedene Strategien zur Verbesserung der Transparenz werden vorgeschlagen, darunter die Veröffentlichung einer Liste von Verstößen durch Nutzer und technische Lösungen (Floridi et al. 2019).

Verantwortung: Eine der bedeutendsten ethischen Fragen im Zusammenhang mit dem Einsatz von Technologie betrifft die Verantwortung von Softwareentwicklern gegenüber den Personen, die von der Technologie beeinflusst werden (Saetra und Danaher 2022), wobei die Verantwortung für voreingenommene Daten nicht allein bei den Entwicklern der Algorithmen, sondern auch bei den für die Daten zuständigen Personen liegt (Symons und Alvarado 2022).

Verzerrung: Es wird vorgeschlagen, Datensatz-Datenblätter zu erstellen, um eine verantwortungsbewusste Datenverarbeitung zu gewährleisten. Diese sollen Informationen zur Datensatzqualität und möglichen Verzerrungen enthalten. Darüber hinaus wird auf die Notwendigkeit neuer Techniken und Software-Tools hingewiesen, die Verzerrungen, Erklärbarkeit und Robustheit berücksichtigen sollen (Laat 2021).

Voreingenommenheit: Die Forschung konzentriert sich darauf, die Auswirkungen von Voreingenommenheit auf rassische Minderheiten zu verstehen und Maßnahmen zur Beseitigung von Voreingenommenheit in Algorithmen zu ergreifen (Lee und Floridi 2021; Symons und Alvarado 2022).

Diskriminierung: Ein konkretes Beispiel für Diskriminierung ist der Cambridge-Analytica-Skandal, der die Problematik von Datenmissbrauch und schädlicher Vorurteile durch große Technologieunternehmen aufzeigt (Chomanski 2021). Eine mögliche Lösung hierfür ist die Einführung standardisierter Trainingsdatensätze (Hagendorff 2020).

4.2 Process Mining

Ähnlich wie in Abschn. 4.1 bei der Identifikation der datenethischen Grundsätze, wurden die für PM und Datenethik als relevant erachteten Publikationen mittels Inhaltsanalyse hinsichtlich ihrer Erwähnung datenethischer Grundsätze betrachtet. Zusätzlich wurde hier untersucht, ob die jeweilige Publikation eine Prozess- bzw. eine Datensichtweise einnimmt. Zur Darstellung des gegenwärtigen Forschungsstands bezüglich ethischer Implikationen im Bereich des PM wird hervorgehoben, dass mit 7 der 39 analysierten PM-Publikationen nur eine begrenzte Anzahl spezifische Grundsätze der Datenethik oder relevante datenethische Fragestellungen thematisiert.

Im Kontext des PM werden verschiedene Analyseansätze aus einer Datenperspektive betrachtet – darunter ein Qualitätsbewertungsansatz, der die Datenqualität aus zwei Hauptperspektiven bewertet: der Eignung der Daten für ihre Verwendung im PM und die Erkenntnisfähigkeit der Datenanalyse (Andrews et al. 2020). Weitere Perspektiven sind die Datenprozessperspektive (Caron et al. 2013), die die während eines Prozesses verwendeten, erzeugten oder manipulierten Informationen und ihre Beziehungen darstellt und in verschiedene Perspektiven unterteilt, sowie die Kontrollflussperspektive, die Datenflussperspektive und die organisatorische Perspektive (Suriadi et al. 2017).

In den Publikationen zu PM ließen sich einige datenethische Grundsätze identifizieren. So spielen Datenschutz und individuelle Rechte in Unternehmensanwendungen eine Rolle (Breuker et al. 2016). Einwilligung und Transparenz bei der Datenverarbeitung (Kratsch et al. 2022), sowie die Identifizierung von Datenqualitätsdimensionen und deren Bedeutung für PM (Andrews et al. 2020) werden als relevant angesehen. Datenschutzbedenken, insbesondere im Gesundheitswesen, werden bspw. im Fall von RFID-Etiketten hervorgehoben (Zhou und Piramuthu 2010). Die Erhebung ausreichender Daten für das PM ist mit Herausforderungen verbunden, insbesondere hinsichtlich Sicherheitsaspekten wie der Anonymisierung von Daten. Dabei ist die Wichtigkeit der Einhaltung datenethischer Grundsätze bei der Datenerhebung und -verwendung für die Analyse zu beachten (Cho et al. 2017).

Die Signifikanz der Nutzung von Simulationen basierend auf realistischen Patientenmodellen wird ebenfalls betont, insbesondere für die Untersuchung von Behandlungsstrategien, die in realen Studien ethische Bedenken aufwerfen könnten (Meyer et al. 2014).

5 Diskussion

Diese Literaturanalyse zeigt einerseits datenethische Grundsätze zur Betrachtung datengestützter Systeme auf und geht andererseits spezifisch auf PM als datengestützte Technologie ein. Die analysierten Artikel zeigen dabei, dass innerhalb der Wirtschaftsinformatik die ethischen Implikationen von PM noch nicht erschöpfend behandelt wurden. 7 der 39 analysierten Artikel zu PM befassen sich mit den identifizierten datenethischen Grundsätzen, während die restlichen Artikel des Datensatzes diese nur eingeschränkt oder gar nicht berücksichtigen.

Ein erheblicher Teil der betrachteten Literatur zu Datenethik behandelt hauptsächlich die funktionalen Aspekte von Technologien, wobei ein besonderes Augenmerk auf spezifischen Technologien liegt, oft im Kontext von Künstlicher Intelligenz (KI) (bspw. Hagendorff 2020; Ratti und Graves 2021; Sætra und Danaher 2022).

Datenqualität stellt einen eher funktionalen Aspekt der datenethischen Grundsätze dar, da dieser ebenso für die grundlegende Nutzung von datengetriebenen Systemen, wie PM-Systemen (vgl. van der Aalst et al. 2010), notwendig ist. In jedem Fall sind Daten von guter Qualität für informierte Entscheidungen notwendig, bei denen alle Aspekte und absehbaren Auswirkungen im Vorhinein berücksichtigt werden können (Floridi 2018; Ratti und Graves 2021).

Der Themenbereich Datenschutz wird oft aufgrund rechtlicher Normen berücksichtigt (bspw. Loi 2019; Sætra und Danaher 2022). Allerdings wird hierbei ethisches Verhalten als solches angesehen, das über die bloße Umsetzung von rechtlichen Vorgaben hinausgeht (Trier et al. 2023). Daher ist eine Ausrichtung notwendig, die sich nicht ausschließlich am rechtlichen Rahmen orientiert, sondern auch auf die Bedürfnisse der Personen eingeht, die im Fokus der Datenerfassung liegen.

Ein wichtiger Aspekt ist ein klares Verfahren für die Einholung der Zustimmung der Personen, deren Daten verwendet werden sollen (Kratsch et al. 2022). Eine transparente Gestaltung von Datenerhebungs- und -Verwendungspraktiken, in Form der Bereitstellung klarer Informationen über die Daten, die gesammelt werden und deren Verwendung, sowie der Personen, die Zugang zu den Daten haben, kann ebenfalls dabei helfen die Akzeptanz von PM-Anwendungen zu steigern. Dabei ist es wichtig, dass die einschlägigen Gesetze und Vorschriften eingehalten werden (Floridi 2018). Die Sicherstellung von Datenqualität gewährleistet, dass in einer Organisation überhaupt informierte Entscheidungen getroffen werden können, da PM-Anwendungen in hohem Maße von dieser abhängen (Dustdar et al. 2006). Zudem trägt sie zur Glaubwürdigkeit von Entscheidungen bei, was besonders wichtig ist, wenn die Ergebnisse einer Analyse erhebliche Auswirkungen auf Einzelpersonen oder Organisationseinheiten haben können (Hand 2018).

Transparenz ist ein datenethischer Grundsatz, der insbesondere bei der Erstellung und dem Betrieb von KI-Systemen vorherrscht (bspw. Ratti und Graves 2021; Mökander et al. 2022; Sætra und Danaher 2022). Dementsprechend findet er sich bei PM-Anwendungen auch in erster Linie beim Einsatz von KI wieder (Kratsch et al. 2022). Transparenz kann jedoch auch auf organisationale Strukturen und die Verwendungsweise der erhobenen Daten bezogen werden. In diesem Kontext kann in Zukunft untersucht werden, ob die klare Darlegung der Datenverwendung die Akzeptanz von beispielsweise PM-Maßnahmen in Organisationen erhöht.

Datenethische Grundsätze wie Voreingenommenheit, Verantwortung und Verzerrung finden in unserem Datensatz nur wenig Betrachtung (bspw. Laa 2021; Morley et al. 2021; Symons und Alvarado 2022). Hieraus lässt sich schließen, dass ein Bedarf an weiter Forschung, insbesondere im organisationalen Umfeld besteht. Management-Entscheidungen sind oft nachgelagert und ihre Auswirkungen sind nicht direkt spürbar (Crane und Matten 2016). Ebenso verhält es sich mit den Auswirkungen der Verwendung von datengestützten Technologien. Hier kann der verantwortungsvolle und ethische Umgang sowohl mit datengestützten Technologien als auch mit gesammelten Daten einen aktiven Beitrag zum gegenseitigen Verständnis bei der Erhebung und Verwendung von Daten leisten. Insbesondere der Aspekt der Verzerrung ist nicht immer offensichtlich (Laa 2021) und hängt sowohl von den Nutzern der Daten als auch von den initial verfügbaren Datensätzen ab. Daher ist weitere Forschung sowohl auf technischer Seite zur Minimierung von Verzerrungen als auch auf organisationaler Seite zum adäquaten Umgang mit gegebenenfalls verzerrten Datensätzen notwendig.

Die in Tab. 3 dargestellten datenethischen Grundsätze lassen sich zu weiten Teilen in den Publikationen zu PM wiederfinden. Datenschutz bildet dabei den am häufigsten genannten Grundsatz, gefolgt von Transparenz. Diskussionen über die DSGVO, welche Regelungen über den Datenschutz festlegt (Loi 2019), sowie aufkommende Bedenken über KI-Anwendungen (Hagendorff 2020) stellen Ausprägungen dieser Grundsätze dar. Informierte Zustimmung und Datenqualität wurden verhältnismäßig wenig betrachtet, was in einer schwierigen Umsetzung in komplexen in KI-Systemen begründet liegen kann (vgl. Mökander et al. 2022).

Weitere Faktoren umfassen die regelmäßige Überprüfung der vorgeschlagenen Data-Governance-Richtlinien, die Durchführung ethischer Prüfungen und die Sicherstellung, dass ethische Praktiken kontinuierlich verfolgt werden (Floridi et al. 2019). Der Umgang mit datenethischen Fragen im PM erfordert einen vielschichtigen Ansatz, der klare Grundsätze und Richtlinien, Maßnahmen zum Schutz der Privatsphäre, zur Transparenz, Rechenschaftspflicht, ethischen Überprüfung und kontinuierlichen Verbesserung umfasst. Unternehmen können durch die Übernahme dieser Best Practices sicherstellen, dass ihre PM-Praktiken ethisch, verantwortungsvoll und vertrauenswürdig sind. Eine zusätzliche Unterstützung könnte durch die Einbeziehung von Ethikausschüssen oder Experten erfolgen, um die möglichen Auswirkungen auf Einzelpersonen oder die Gesellschaft zu bewerten. Eine kontinuierliche Überwachung und Verbesserung der Anwendung von PM kann zudem die ethische Nutzung dieser Technologie stärken.

6 Zusammenfassung und Ausblick

6.1 Zusammenfassung

Der Artikel gibt einen Überblick über den aktuellen Stand der Forschung zu Datenethik und Process Mining. Dabei werden relevante Konzepte der Datenethik behandelt und bewährte Verfahren der Datenethik im Kontext von PM betrachtet. Im Zuge der hier durchgeführten Literaturanalyse wurden, Grundsätze und Richtlinien

der Datenethik, die Maßnahmen zum Schutz der Privatsphäre, Transparenz, Rechenschaftspflicht, Überprüfungen auf ethische Problemstellungen hin und kontinuierliche Verbesserungen umfassen, identifiziert. Aspekte wie Datenqualität, Fairness, informierte Zustimmung und Verantwortung bieten Potenzial für weitere Forschung. Die Literaturanalyse zeigt, dass sich ein vergleichsweise geringer Anteil der Artikel im Bereich PM mit datenethischen Grundsätzen befasst. Dies weist auf einen Bedarf an weiterer Forschung und gleichzeitig auf die Notwendigkeit einer verstärkten interdisziplinären Zusammenarbeit zwischen Philosophie und (Wirtschafts-) Informatik hin. Die Festlegung von Data-Governance-Richtlinien, die Anonymisierung und Pseudonymisierung von Daten sowie die Transparenz der Datenerhebungs- und -verwendungspraktiken sind mögliche Maßnahmen, um ethisches Verhalten im Rahmen des PM sicherzustellen.

6.2 Limitationen

Beim Verfassen dieses Artikels mussten leider einige Einschränkungen berücksichtigt werden. Die Literaturanalyse wurde mit einer begrenzten Auswahl von Zeitschriften durchgeführt, die sich vorrangig auf theoretische wissenschaftliche Publikationen konzentrieren und versuchen, diese für die Praxis nutzbar zu machen. Einerseits besteht die Möglichkeit, dass bei den nicht berücksichtigten Konferenzen weitere relevante Beiträge existieren, andererseits besteht auch die Chance, dass in der Praxis übliche Konzepte nicht berücksichtigt wurden.

6.3 Implikationen für Forschung und Praxis

Die Ergebnisse dieser Literaturanalyse liefern Impulse für die praktische Umsetzung von PM und Datenmanagement in Unternehmen und Organisationen. Hierdurch zeigt sich eine gewisse Notwendigkeit für weitere empirische Forschung, um ein tieferes Verständnis für die ethischen Aspekte des PM zu entwickeln. Die Erstellung praktischer Handlungsrichtlinien und Leitfäden ist von entscheidender Bedeutung, um ethisches Verhalten im PM zu fördern. In der Praxis ergeben sich verschiedene Handlungsempfehlungen: In erster Linie ist die Entwicklung und konsequente Umsetzung klar definierter Leitlinien von Bedeutung. Diese sollten sicherstellen, dass PM und Datenmanagement ethisch verantwortungsbewusst eingesetzt werden, indem sie klare Grundsätze und Praktiken im Umgang mit Daten festlegen. Zusätzlich ist eine aktive Identifikation ethischer Problemstellungen und die Entwicklung von Vorschlägen zu deren Lösung wichtig. Dabei können die in dieser Analyse präsentierten Themengebiete als Orientierung dienen.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- SCImago. <http://www.scimagojr.com>. Zugegriffen: 20. März 2023
- van der Aalst W (2011) Process mining: discovery, conformance and enhancement of business processes. EBL-Schweitzer. Springer (<http://swb.ebibli.com/patron/FullRecord.aspx?p=763874>)
- van der Aalst W (2016) Process mining: data science in action, 2. Aufl. Springer, Berlin Heidelberg <https://doi.org/10.1007/978-3-662-49851-4>
- van der Aalst WMP, Weijters AJMM (2004) Process mining: a research agenda. *Comput Ind* 53(3):231–244. <https://doi.org/10.1016/j.compind.2003.10.001>
- van der Aalst WMP, Rubin V, Verbeek HMW, van Dongen BF, Kindler E, Günther CW (2010) Process mining: a two-step approach to balance between underfitting and overfitting. *Softw Syst Model* 9(1):87–111. <https://doi.org/10.1007/s10270-008-0106-z>
- Aggarwal N, Floridi L (2020) Towards the ethical publication of country of origin information (COI) in the asylum process. *Minds Mach* 30(2):247–257. <https://doi.org/10.1007/s11023-020-09523-w>
- Alt R, Göldi A, Österle H, Portmann E, Spiekermann S (2021) Life engineering. *Bus Inf Syst Eng* 63(2):191–205. <https://doi.org/10.1007/s12599-020-00680-x>
- Andrews R, van Dun CGJ, Wynn MT, Kratsch W, Röglinger MKE, ter Hofstede AHM (2020) Quality-informed semi-automated event log generation for process mining. *Decis Support Syst* 132:113265. <https://doi.org/10.1016/j.dss.2020.113265>
- Breuker D, Matzner M, Delfmann P, Becker J (2016) Comprehensible predictive models for business processes. *MIS Q* 40(4):1009–1034. <https://doi.org/10.25300/MISQ/2016/40.4.10>
- Cai L, Zhu Y (2015) The challenges of data quality and data quality assessment in the big data era. *Data Sci J* 14(0):2. <https://doi.org/10.5334/dsj-2015-002>
- Caron F, Vanthienen J, Baesens B (2013) Comprehensive rule-based compliance checking and risk management with process mining. *Decis Support Syst* 54(3):1357–1369. <https://doi.org/10.1016/j.dss.2012.12.012>
- Cho M, Song M, Comuzzi M, Yoo S (2017) Evaluating the effect of best practices for business process redesign: An evidence-based approach based on process mining techniques. *Decis Support Syst* 104:92–103. <https://doi.org/10.1016/j.dss.2017.10.004>
- Chomanski B (2021) The missing ingredient in the case for regulating big tech. *Minds Mach* 31(2):257–275. <https://doi.org/10.1007/s11023-021-09562-x>
- Crane A, Matten D (2016) Business ethics. Oxford University Press
- Deutscher Ethikrat (2023) Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz. Deutscher Ethikrat, Berlin
- Dustdar S, Fiadeiro JL, Sheth AP (Hrsg) (2006) Business process management. 4th International Conference, BPM 2006, Vienna, September 5–7, 2006. SpringerLink Bücher, Bd. 4102. Springer, Berlin Heidelberg <https://doi.org/10.1007/11841760> (Proceedings)
- Elgendy N, Elragal A (2014) Big data Analytics: a literature review paper. In: Perner P (Hrsg) Advances in data mining: 14th industrial conference, ICDM 2014 St. Petersburg, July 16–20, 2014. Lecture Notes in Computer Science Ser. v.8557, Bd. 8557. Springer, Russia, S 214–227 https://doi.org/10.1007/978-3-319-08976-8_16 (Proceedings)
- Floridi L (2014) Open data, data protection, and group privacy. *Philos Technol* 27(1):1–3. <https://doi.org/10.1007/s13347-014-0157-8>
- Floridi L (2018) Soft ethics and the governance of the digital. *Philos Technol* 31(1):1–8. <https://doi.org/10.1007/s13347-018-0303-9>
- Floridi L, Taddeo M (2016) What is data ethics? *Philosophical transactions. Ser A Math Phys Eng Sci*. <https://doi.org/10.1098/rsta.2016.0360>

- Floridi L, Luetge C, Pagallo U, Schafer B, Valcke P, Vayena E, Addison J, Hughes N, Lea N, Sage C, Vannieuwenhuyse B, Kalra D (2019) Key ethical challenges in the European medical information framework. *Minds Mach* 29(3):355–371. <https://doi.org/10.1007/s11023-018-9467-4>
- Hagendorff T (2020) The ethics of AI ethics: an evaluation of guidelines. *Minds Mach* 30(1):99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Hand DJ (2018) Aspects of data ethics in a changing world: where are we now? *Big Data* 6(3):176–190. <https://doi.org/10.1089/big.2018.0083>
- Kratsch W, König F, Röglinger M (2022) Shedding light on blind spots—Developing a reference architecture to leverage video data for process mining. *Decis Support Syst* 158:113794. <https://doi.org/10.1016/j.dss.2022.113794>
- de Laat PB (2021) Companies committed to responsible AI: from principles towards implementation and regulation? *Philos Technol* 34(4):1135–1193. <https://doi.org/10.1007/s13347-021-00474-3>
- Lee MSA, Floridi L (2021) Algorithmic fairness in mortgage lending: from absolute conditions to relational trade-offs. *Minds Mach* 31(1):165–191. <https://doi.org/10.1007/s11023-020-09529-4>
- Loi M (2019) The digital phenotype: a philosophical and ethical exploration. *Philos Technol* 32(1):155–171. <https://doi.org/10.1007/s13347-018-0319-1>
- Lowry PB, Moody GD, Gaskin J, Galletta DF, Humpherys SL, Barlow JB, Wilson DW (2013) Evaluating journal quality and the association for information systems senior scholars' journal basket via bibliometric measures: do expert journal assessments add value? *MISQ* 37(4):993–1012. <https://doi.org/10.25300/misq/2013/37.4.01>
- Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A (2022) A survey on bias and fairness in machine learning. *ACM Comput Surv* 54(6):1–35. <https://doi.org/10.1145/3457607>
- Meyer G, Adamavicius G, Johnson PE, Elidrisi M, Rush WA, Sperl-Hillen JM, O'Connor PJ (2014) A machine learning approach to improving dynamic decision making. *Inf Syst Res* 25(2):239–263. <https://doi.org/10.1287/isre.2014.0513>
- Mittelstadt BD, Floridi L (2016) The ethics of big data: current and foreseeable issues in biomedical contexts. *Sci Eng Ethics* 22(2):303–341. <https://doi.org/10.1007/s11948-015-9652-2>
- Mökander J, Axente M, Casolari F, Floridi L (2022) Conformity assessments and post-market monitoring: a guide to the role of auditing in the proposed European AI regulation. *Minds Mach* 32(2):241–268. <https://doi.org/10.1007/s11023-021-09577-4>
- Morley J, Elhalal A, Garcia F, Kinsey L, Mökander J, Floridi L (2021) Ethics as a service: a pragmatic operationalisation of AI ethics. *Minds Mach* 31(2):239–256. <https://doi.org/10.1007/s11023-021-09563-w>
- Ratti E, Graves M (2021) Cultivating moral attention: a virtue-oriented approach to responsible data science in healthcare. *Philos Technol* 34(4):1819–1846. <https://doi.org/10.1007/s13347-021-00490-3>
- Rozinat A, van der Aalst WMP (2006) Decision mining in proM. In: Dustdar S, Fiadeiro JL, Sheth AP (Hrsg) Business process management: 4th international conference, BPM 2006 Vienna, September 5–7, 2006. Bd. 4102. Springer, Berlin Heidelberg, S 420–425 https://doi.org/10.1007/11841760_33 (Proceedings)
- Saetra HS, Danaher J (2022) To each technology its own ethics: the problem of ethical proliferation. *Philos Technol*. <https://doi.org/10.1007/s13347-022-00591-7>
- Slussareff M (2022) O'Neil, Cathy. 2016. Weapons of math destruction: how big data increases inequality and threatens democracy. Crown. *CyberOrient* 16(1):72–75. <https://doi.org/10.1002/cyo2.26>
- Spiekermann S (2015) Ethical IT innovation: a value-based system design approach. CRC press. <http://gbv.eblib.com/patron/FullRecord.aspx?p=4096994>
- Spiekermann S, Krasnova H, Hinz O, Baumann A, Benlian A, Gimpel H, Heimbach I, Köster A, Maedche A, Niehaves B, Risius M, Trenz M (2022) Values and ethics in information systems. *Bus Inf Syst Eng* 64(2):247–264. <https://doi.org/10.1007/s12599-021-00734-8>
- Stahl BC, Eden G, Jirotko M, Coeckelbergh M (2014) From computer ethics to responsible research and innovation in ICT. *Inf Manag* 51(6):810–818. <https://doi.org/10.1016/j.im.2014.01.001>
- Suriadi S, Wynn MT, Xu J, van der Aalst WMP, ter Hofstede AHM (2017) Discovering work prioritisation patterns from event logs. *Decis Support Syst* 100:77–92. <https://doi.org/10.1016/j.dss.2017.02.002>
- Symons J, Alvarado R (2022) Epistemic injustice and data science technologies. *Synthese*. <https://doi.org/10.1007/s11229-022-03631-z>
- Taylor L (2021) Public actors without public values: legitimacy, domination and the regulation of the technology sector. *Philos Technol* 34(4):897–922. <https://doi.org/10.1007/s13347-020-00441-4>
- Trier M, Kundisch D, Beverungen D, Müller O, Schryen G, Mirbabaie M, Trang S (2023) Digital responsibility. *Bus Inf Syst Eng* 65(4):463–474. <https://doi.org/10.1007/s12599-023-00822-x>

- Turilli M, Floridi L (2009) The ethics of information transparency. *Ethics Inf Technol* 11(2):105–112. <https://doi.org/10.1007/s10676-009-9187-9>
- Webster J, Watson RT (2002) Analyzing the past to prepare for the future: writing a literature review. *MISQ* 26(2):xiii–xxiii
- Zhou W, Piramuthu S (2010) Healthcare Process Mining with RFID. In: Rinderle-Ma S (Hrsg) *Business Process Management Workshops: BPM 2009 International Workshops Ulm, 7. Sept. 2009. Lecture Notes in Business Information Processing Ser: v.43., Bd. 43*. Springer, Berlin, Heidelberg, S 405–411 https://doi.org/10.1007/978-3-642-12186-9_38 (Revised Papers)

Hinweis des Verlags Der Verlag bleibt in Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutsadressen neutral.