

Wetzel, Manuel; Koschmider, Agnes

Article — Published Version

Entwicklung einer VR-Umgebung zur Exploration von Process-Mining

HMD Praxis der Wirtschaftsinformatik

Provided in Cooperation with:

Springer Nature

Suggested Citation: Wetzel, Manuel; Koschmider, Agnes (2021) : Entwicklung einer VR-Umgebung zur Exploration von Process-Mining, HMD Praxis der Wirtschaftsinformatik, ISSN 2198-2775, Springer Fachmedien Wiesbaden, Wiesbaden, Vol. 59, Iss. 1, pp. 37-53, <https://doi.org/10.1365/s40702-021-00827-8>

This Version is available at:

<https://hdl.handle.net/10419/287643>

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/>



Entwicklung einer VR-Umgebung zur Exploration von Process-Mining

Manuel Wetzel · Agnes Koschmider 

Eingegangen: 28. Juli 2021 / Angenommen: 17. November 2021 / Online publiziert: 14. Dezember 2021
© Der/die Autor(en) 2021

Zusammenfassung VR-Umgebungen werden bereits in zahlreichen Anwendungsszenarien erfolgreich zur Visualisierung von Daten mit dem Ziel beispielsweise der Lernprozessunterstützung eingesetzt. Dieser Beitrag stellt eine VR-basierte Umgebung für Process-Mining mit dem Ziel der Prozessanalyseunterstützung vor. Die VR-basierte Umgebung ermöglicht, Prozessdaten aus Quellsystemen dynamisch anzubinden und zu laden und diese als ein dreidimensionales Prozessmodell zu visualisieren und zu analysieren. Die VR-Umgebung wurde systematisch basierend auf einer Anforderungsanalyse konzipiert, die aus kommerziellen (2D) Process-Mining Werkzeugen und verwandten Arbeiten aus der Literatur abgeleitet wurde. Bei der Implementierung der Umgebung wurde Wert auf Erweiterbarkeit und Offenheit der Umgebung gelegt. Im Gegensatz zu einer zweidimensionalen Visualisierung ermöglicht die VR-Umgebung für Process-Mining eine verbesserte Exploration der Zusammenhänge zwischen Prozessen und Daten (z. B. Prozesskennzahlen wie Ressourcenauslastung, Durchlaufzeit, Prozessabweichungen).

Schlüsselwörter Process-Mining · Immersive Analytics · Prozessmodell Visualisierung · Virtual Reality

Manuel Wetzel
Capgemini Deutschland, Lübecker Str. 128, 22087 Hamburg, Deutschland
E-Mail: manuel.wetzel@capgemini.com

Agnes Koschmider (✉)
AG Wirtschaftsinformatik (Process Analytics), Christian-Albrechts-Universität zu Kiel,
Hermann-Rodewald-Str. 3, 24118 Kiel, Deutschland
E-Mail: ak@informatik.uni-kiel.de

Design of a VR-based Environment for Process Mining Exploration

Abstract VR environments are already being used successfully in numerous application scenarios for the visualization of data with the aim of e.g., supporting learning processes. This paper presents a VR-based environment for process mining with the aim of supporting process analysis. The VR-based environment makes it possible to dynamically link and load process data from source systems and to visualize and analyze them as a three-dimensional process model. The VR environment was systematically designed on a requirements analysis derived from commercial (2D) process mining tools and related work from the literature. When implementing the environment, emphasis was placed on extensibility and openness of the environment. In contrast to a two-dimensional visualization, the VR environment for process mining supports an improved exploration of the relationships between processes and data (e.g., KPIs such as resource utilization, lead time, process deviations).

Keywords Process mining · Immersive analytics · Process model · Visualization · Virtual reality

1 Einleitung

Immersive Analytics ermöglicht durch VR- oder AR-Umgebungen Daten zu explorieren (Chandler et al. 2015). Zahlreiche Anwendungsszenarien für die immersive Datenexploration wurden bereits vorgeschlagen. Es existieren allerdings bisher kaum Vorarbeiten zur immersiven Analyse für das Process-Mining. Die Intention von Process-Mining ist die Erkennung von Prozessmodellen aus (digitalen) Daten mit dem Ziel Schwachstellen oder Abweichungen zwischen der Ist- und Soll-Situation zu finden (Laue et al., 2020). Die per Process-Mining gefundenen Prozesse werden anschließend anhand von Prozesskennzahlen wie Durchlaufzeit oder Prozessperformance durch Nutzer analysiert. Somit ist die zügige Erfassung der Abhängigkeiten und Ursache/Wirkungen zwischen dem Prozess und Prozesskennzahlen ein entscheidender Erfolgsfaktor von Process-Mining. In diesem Beitrag wird eine VR-basierte Umgebung für Process-Mining vorgestellt. Mit Hilfe der VR-Umgebung ist es möglich, im Gegensatz zu einer zweidimensionalen Visualisierung, beispielsweise die folgenden Fragestellungen auf einem Blick zu erfassen: Ist der Ressourceneinsatz für überproportionale Prozessdurchläufe effizient? Für welche Prozessabweichungen wird ein anderer Ressourceneinsatz benötigt? Warum brauchen bestimmte Prozessdurchläufe besonders lange trotz ausgelasteter Ressourcen? Die Beantwortung dieser Fragen über eine 2D-Visualisierung mit traditionellen Process-Mining Werkzeugen erforderten die Verknüpfung und Überlagerung von mehreren Fenstern. Im Gegensatz dazu ermöglicht die VR-basierte Umgebung eine verbesserte Exploration der Zusammenhänge zwischen Prozessen und Daten (z. B. Prozesskennzahlen wie Performance, Durchlaufzeit, Prozessabweichungen) und gibt zielführender einen Einblick in die Analyse im Vergleich zur 2D-basierten Process-Mining Analyse. Die VR-Umgebung bietet sowohl für unerfahrene als auch für erfahrene Nutzer einen Mehrwert. Unerfahrene Nutzer (d. h., Prozessmodellierungsanfänger oder Domänen-

experten) bietet die VR-Visualisierung einen leichteren Zugang zum Prozessmodell und deren Kennzahlen. Erfahrende Nutzer (d.h., Modellierungsexperten oder Prozessanalysten) bekommen durch den Wechsel zwischen Prozesssichten eine höhere Informationsdichte zu dem Prozessmodell. Damit die VR-Umgebung diese Mehrwerte stiftet, erfordert es einer Anforderungsanalyse zur Konzeption einer solchen Umgebung. Vor diesem Hintergrund adressiert dieser Beitrag die folgenden zwei Forschungsfragen:

- **Forschungsfrage 1:** Welche Anforderungen stellen sich an eine VR-basierte Umgebung zur Exploration von Process-Mining?
- **Forschungsfrage 2:** Wie kann eine solche VR-basierte Umgebung umgesetzt werden?

Der Beitrag ist wie folgt strukturiert. Das Kap. 2 fasst wichtige Begriffe des Process-Mining zusammen. Als Basis zur Gestaltung der VR-basierten Umgebung wurde eine Anforderungsanalyse mit Nutzern durchgeführt, die in Kap. 3 beschrieben wird. Auf der Grundlage der Nutzerstudie wurde die VR-basierte Umgebung implementiert, die in Kap. 4 vorgestellt wird. Verbesserungsmöglichkeiten der VR-basierten Umgebung werden in Kap. 5 skizziert. Der Beitrag endet mit einer Zusammenfassung und einem Ausblick.

2 Einführung in Process-Mining

Ziel von Process-Mining ist „die Analyse und Rekonstruktion von Prozessmodellen auf Basis von Ereignislogs (Protokollen) aus laufenden IT-Systemen“ (Drescher et al. 2017, S. 197). Die Terminologie von Process-Mining ist die folgende (Laue et al., 2020, S. 166ff.): Ein *Ereignisprotokoll* (engl. event log) ist Ausgangspunkt jeder Prozessanalyse im Process-Mining. Jeder *Arbeitsschritt* eines Prozesses wird durch das unterstützende Software-System als *Ereignis* im Protokoll aufgezeichnet. Ein Ereignis entspricht damit der Durchführung eines Arbeitsschrittes im Prozess. Dieser Arbeitsschritt ist im Modell als *Aktivität* beschrieben. Arbeitsschritte werden immer in einer bestimmten Prozessinstanz durchgeführt. Im Kontext des Process-Mining hat sich für „Prozessinstanz“ der Begriff *Fall* (engl. case) etabliert und meint beispielsweise eine konkrete Buchungsanfrage, eine konkrete Nutzung eines Röntgenscanners oder der Aufruf einer Website durch einen bestimmten Nutzer. Ein Prozessmodell beschreibt alle möglichen Ausführungsfolgen von Aktivitäten, die in Instanzen des Prozesses auftreten können. Abb. 1 fasst die Terminologie von Process-Mining zusammen.

Auf die Terminologie von Process-Mining wird im weiteren Verlauf zurückgegriffen und die Umsetzung der Process-Mining Konzepte in der VR-basierten Umgebung vorgestellt.

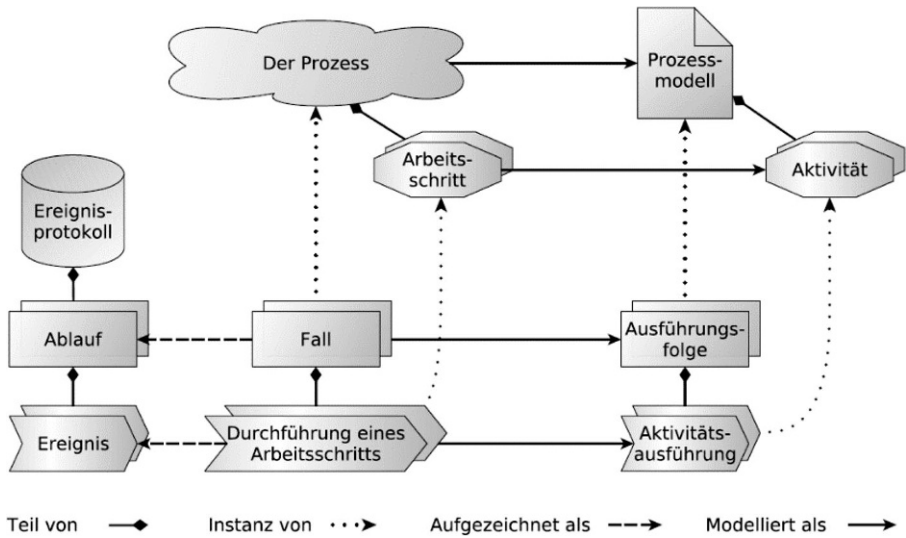


Abb. 1 Übersicht der Terminologie des Process-Mining [Laue et al., 2020, S. 167]

3 Immersive VR-basierte Datenvisualisierung

Verschiedene Formen der VR-basierten Datenvisualisierung wurden als Teilgebiet von Immersive Analytics bereits untersucht und entsprechende Lösungen implementiert. So existiert eine Umsetzung um zweidimensionale Diagrammtypen in einer VR-Umgebung gemeinsam kollaborativ zu betrachten (Lee et al. 2019). Es gibt Lösungen, die dreidimensionale Punktdiagramme erzeugen (Cordeil et al. 2019). Es existiert ebenfalls eine Implementierung zur Kombination von 2D-Visualisierungen auf einem Bildschirm mit einer VR-Umgebung. In dieser Lösung werden ein klassisches 2D-Dashboard mit einer VR-Umgebungen synchronisiert (Su et al. 2020).

Bisher behandeln aber nur wenige Arbeiten die immersive Analyse von Prozessen. In einer der wenigen Lösungen werden die räumlichen Positionen der Knoten im Prozessgraphen genutzt, um Attribute der Aktivitäten zu kodieren, die Zuordnung von Attributen zur räumlichen Dimension ist dabei frei konfigurierbar (Gall und Rinderle-Ma 2019). Allerdings setzen diese Lösungen ein Prozessmodell voraus und ermöglichen keine dynamische Datenanbindung.

4 Nutzerstudie als Grundlage zur Konzeption der VR-basierten Umgebung

Dieses Kapitel unterteilt sich wie folgt: das nächste Kapitel skizziert den Aufbau der Nutzerstudie. Abschn. 4.2. beschreibt die Ergebnisse der Nutzerstudie.

4.1 Aufbau der Nutzerstudie

Zur Erhebung von Anforderungen an die VR-Umgebung, und somit zur Beantwortung der Forschungsfrage 1 („*Welche Anforderungen stellen sich an eine VR-basierte Umgebung zur Exploration von Process-Mining?*“), wurden die Befragten zunächst gebeten, ihre Expertise im Bereich Data Science, Process-Mining, VR und User Experience Design auf einer Skala von 1 bis 3 anzugeben. 1 bedeutet: *Ich habe einmal davon gehört*, 2 bedeutet: *Ich habe erste Erfahrungen auf dem Gebiet gemacht und kenne relevante Konzepte*. 3 bedeutet: *Ich habe mehrjährige Erfahrungen in dem Gebiet und selbst Ergebnisse erzeugt*. Anschließend wurden die Befragten gebeten, Anforderungen an eine VR-basierte Process-Mining Umgebung in Form von User Stories auf einer Skala von –2 bis 2 zu priorisieren (2 – must; 1 – should; –1 – could; –2 not). Insgesamt wurden 20 Anforderungen aus kommerziellen (2D) Process Mining Werkzeugen und verwandten Arbeiten aus der Literatur abgeleitet und den Befragten vorgelegt. Zusätzlich konnten die Befragten in Freitextfeldern Kommentare (z. B. zusätzliche Anforderungen, Verbesserungsvorschläge, Meinungen) äußern. Die Nutzerstudie wurde im September 2020 durchgeführt. Die Nutzerstudie wurde von insgesamt neun Personen durchgeführt, von denen sechs Personen männlich und drei Personen weiblich waren. Die Anforderungen waren mindestens einem der abgefragten Fachgebiete (Data Science, Process-Mining, VR und User Experience Design) zugeordnet. So konnte in der Auswertung jede „Person-Anforderung-Bewertung“ mit der jeweiligen Expertise gewichtet werden. Dieses Vorgehen stellt sicher, dass jede Person alle Anforderungen unabhängig von seinem Fachgebiet bewerten konnte.

4.2 Ergebnisse der Nutzerstudie

An der Nutzerstudie haben neun Personen teilgenommen und ihre Expertise in den Fachbereichen wie in Tab. 1 dargestellt, angegeben.

Somit wurden in jedem Fachgebiet Experten befragt. Grundsätzlich ist die Anzahl der Befragten recht klein. Allerdings war der Fokus der Nutzerstudie auf dem grundsätzlichen Verständnis von Anforderungen an eine VR-basierte Process-Mining Umgebung, die erweiterbar und offen genug sein sollte, um weitere in Zukunft geforderte Anforderungen zu integrieren.

Tab. 2 zeigt einen Auszug der zehn hoch priorisierten von den insgesamt 20 Anforderungen geordnet nach der Priorisierung. Die vollständige Liste der Anforderungen

Tab. 1 Erfahrungen der Befragten

Fachgebiet	Geringe Erfahrung	Mittlere Erfahrung	Umfangreiche Erfahrung
Virtual Reality (inkl. Augmented und Mixed Reality)	5	3	1
Process Mining	4	1	4
Data Science	1	6	2
User Experience Design (inkl. UI Design)	2	5	2

Tab. 2 Auszug der höchst priorisierten Anforderungen aus der Gesamtliste

Funktion	Priorität	Referenz
Speichern und Laden von Analysen	1,7	z. B. Celonis SE
Prozessgraph Komponente	1,3	Ware und Mitchell 2008
Case-Cube Komponente	1,1	Garcia-Hernandez et al. 2016
Prozess Graph Komponente mit 3D Attributen	0,9	Gall und Rinderle-Ma 2019
Live-Daten	0,8	Marriott et al. 2018
Immersive Analytics Authoring	0,8	Lee et al. 2019
Details on Demand	0,6	Moran et al. 2015
Kollaboration	0,5	Billinghurst et al. 2018
Filtern	0,4	Moran et al. 2015
Linking & Brushing	0,3	Cordeil et al. 2019

findet sich in Anhang 1 (Tab. 3). Die Liste beantwortet damit auch die Forschungsfrage 1. Die Spalte *Funktion* beschreibt die Anforderung. Die Spalte *Priorität* gibt die Priorisierung der Anforderung an (möglich war ein Wert zwischen -2 und 2). Die Spalte *Referenz* bezieht sich auf die Quelle, die diese Funktion vorschlägt oder umgesetzt hat (z.B. wird bereits im Werkzeug wie Celonis genutzt oder wurde empirisch per Nutzerbefragung validiert).

Am höchsten priorisiert wurden das Speichern und Laden von Analysen (Priorisierung von 1,7). Diese Funktion soll Analysen vorbereiten, auf diese zu späteren Zeitpunkten zugreifen und zwischen mehreren Speicherständen wechseln zu können. Am zweit höchsten wurde die Prozessgraph Komponente mit einem Wert von 1,3 priorisiert. Die Prozessgraphvisualisierung ist ein Direkte-Nachfolger Graph (DFG), vgl. (Laue et al., 2020), was der Standardvisualisierung in kommerziellen Process-Mining Werkzeugen entspricht. Am dritthöchsten wurde die Funktion eines interaktiven Scatterplots bewertet (Priorität von 1,1), welche das Konstrukt des Falls (case) implementiert, vgl. Kap. 2. Die Prozessgraph Komponente mit 3D, d.h., Erweiterung des DFG um eine dritte Dimension, wurde mit 0,9 priorisiert. Konkret wird hierbei die Position der Graphknoten benutzt, um weitere Merkmale der Aktivität zu codieren, vgl. dazu (Gall, Rinderle-Ma (2019)). Hoch priorisiert wurde ebenfalls die Funktion „Live-Daten“ anzubinden. Damit sollen neue Datensätzen und neue Ereignis dynamisch integriert werden. Diese Anforderung wurde im Freitextfeld des Fragebogens nochmals als wertvoll bestärkt. Immersive Analytics Authoring, die mit einem Wert von 0,8 priorisiert wurde, bezeichnet verschiedene Interaktionsformen, um neue Visualisierungen zu erstellen, zu platzieren und die Datenauswahl zu konfigurieren. Diese Funktionen sind in zahlreichen Immersive Analytics Prototypen umgesetzt (vgl. z.B. Lee et al. 2019). Im Kontext der Kollaboration wurde vorgeschlagen, dass Nutzer virtuell und parallel gemeinsam an einer Analyse arbeiten sollen können. Diese Anforderung wurde mit 0,5 bewertet.

Basierend auf den Ergebnissen der Nutzerstudie wurde der VR-basierte Process Explorer implementiert, der im nächsten Kapitel vorgestellt wird.

5 Implementierung

Die Implementierung umfasst sowohl eine VR-Anwendung namens Process Explorer als auch eine Systemarchitektur in die der Process Explorer als Frontend eingebettet ist. Zunächst wird in Abschn. 5.1. die Systemarchitektur vorgestellt. Diese erfüllt die Anforderung „Live-Daten“ (vgl. Anhang 1; Tab. 3). Abschn. 5.2. beschreibt anschließend den Process Explorer und die Umsetzung zahlreicher hoch priorisierter Funktionen wie dem Speichern und Laden, dem immersiven Authoring, Details On Demand und der Case-Cube Komponente. Der Fokus des Kapitels liegt insbesondere auf der Prozessgraph Komponente. Denn der Kontrollfluss ist das zentrale Element eines Prozessgraphen. Das Kapitel endet mit einer Zusammenfassung in 5.3.

5.1 Architektur

Zunächst wird die Architektur betrachtet, die es ermöglicht Live-Daten in der VR Anwendung zu visualisieren. Die Architektur bildet das Fundament zur Implementierung des Process Explorers und beantwortet die Forschungsfrage 2 („Wie kann eine solche VR-basierte Umgebung umgesetzt werden?“). Abb. 2 gibt einen Überblick der Anwendungen von den Rohdatenquellen bis zur Visualisierung.

Das vorgeschlagene System besteht aus drei modularen Schichten:

1. Einem klassischen Process-Mining-Werkzeug für das exemplarisch Celonis angebunden wurde.
2. Einem Kommunikationsservice, der mit der Programmiersprache Python umgesetzt wurde und mit der Bibliothek Flask eine Schnittstelle (API) bereitstellt, um die Daten in aufbereiteter Form zu beziehen.

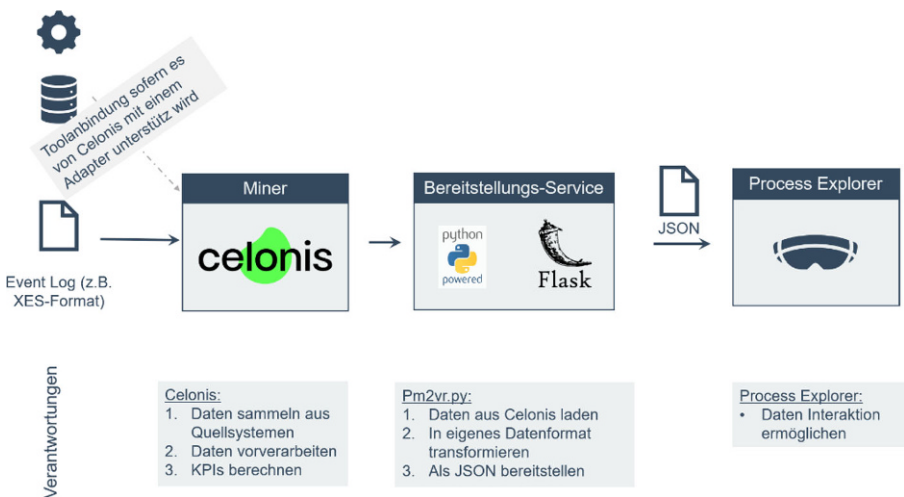


Abb. 2 Architekturübersicht der VR-basierten Umgebung für Process-Mining

3. Dem eigentlichen Frontend für die immersive Datenexploration, das als VR-Applikation implementiert wurde.

Die Einbindung eines etablierten Process-Mining Werkzeugs bietet Vorteile der prozessbezogenen Datensatzmanipulation und Bereinigung. So können über Adapter vielfältige Quellsysteme angebunden und deren Daten gefiltert, aufbereitet und um weitere Berechnungen angereichert werden. All diese Datenfilter, Aggregationen und Erweiterungen mit einem bereits vorhandenen Kontext von Prozessanalysen erlauben damit eine Datensatzvorbereitung. Ein wichtiger Teil dieser Datensatzvorbereitung ist die Berechnung von Prozessinstanzen für jeden Fall. Aus dieser Menge von Prozessinstanzen bzw. Traces erzeugt der Process Explorer später den Prozessgraphen.

Um die so aufbereiteten Daten strukturiert abgreifen zu können, müssen innerhalb von Celonis drei spezielle Tabellen vorbereitet werden: Die *Fall-Tabelle* umfasst sämtliche Fälle des Datensatzes und die zugehörigen Attribute. Die *Aktivitäts-Tabelle* umfasst die Namen der einzelnen Aktivitätsschritte. Die *Trace-Tabelle* umfasst die einzelnen Prozessinstanzen des Prozesses.

Auf Basis der drei Tabellen kann der Backend Service *PM2VR* (Process-Mining to Virtual Reality) die Daten via authentifiziertem API-Zugriff auslesen. Die auf diese Weise ausgelesenen Daten werden anschließend in ein JSON-Format überführt. Hervorzuheben ist dabei die flexible Struktur des Datenmodells, die es erlaubt eine beliebige Anzahl und heterogene Datentypen als Merkmale pro Fall oder Aktivität aufzunehmen. So können vielfältige, individuelle Kennzahlen oder abgeleitete Merkmale pro Fall, Aktivität oder Trace übergeben werden. Nach der Datentransformation sind die Daten für beliebige Clients an einer Schnittstelle abrufbar, die mit dem Flask Framework implementiert wurde. Technisch funktioniert die Kommunikation nach dem Pull-Prinzip. Die VR-App (d. h., der Process Explorer) sendet als Client einen parametrisierten http Get Request an den *PM2VR* Service. Aus diesen Parametern sowie dem Authentifizierungstoken liest der *PM2VR* Service den angefragten Datensatz aus Celonis aus, transformiert die Daten und sendet sie aufbereitet an den *Process Explorer*.

Durch die parametrisierte Anfrage können beliebig viele Datensätze aus Celonis an den *Process Explorer* angebunden werden. Der User kann sich innerhalb der Anwendung aussuchen, welcher Datensatz geladen werden soll. Die Verwendung der Zwischenschicht in Form des *PM2VR* Service ermöglicht die Abstraktion der Daten von dem Quellsystem. Die Aufbereitung in Form einer allgemeinen JSON Datenstruktur ermöglicht eine größtmögliche Unabhängigkeit und damit Modularität. Weitere Clients für verschiedene Endgeräte könnten die Daten aus dem *PM2VR* Service auslesen.

5.2 User Experience in der VR Prozesswelt

Nach dem Laden des Datensatzes in den *Process Explorer* kann der Prozess analysiert werden. Hierfür werden über ein Menü ein dreidimensionales Punktdiagramm und ein dreidimensionaler Prozessgraph erzeugt. Das dreidimensionale Punktdiagramm visualisiert die Fall-Attribute, sowie den dreidimensionalen Prozessgraphen.

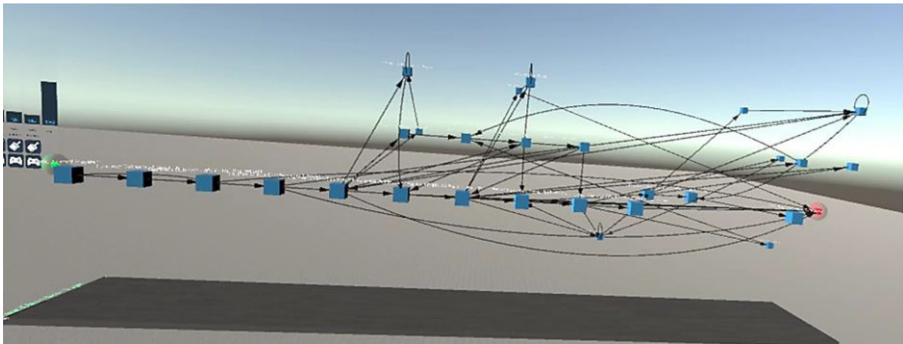


Abb. 3 Basisdarstellung des 3D-Prozessgraphen

Die Visualisierung und Interaktion des Prozessgraphen werden nachfolgend beispielhaft vorgestellt. Abb. 3 zeigt den Prozessgraphen für einen Beispieldatensatz.

Die Start- und Endknoten sind farblich hervorgehoben und die Aktivitätsnamen werden mit Elementnamen (Labels) über den Knoten angezeigt. Die Elementnamen sind derart programmiert, dass sie sich immer zu dem User drehen. Diese Basisdarstellung kann dem Analysezzweck entsprechend konfiguriert werden.

Der Umfang der dargestellten Traces kann durch Filter gesteuert werden, was die Analyse auf die relevanten Traces beschränkt. Zur Konfiguration können in einem Menü (siehe Abb. 4) Traces einzeln aus- oder abgewählt werden.

Dabei werden dem Analysten zusätzliche Angaben zu den Traces angezeigt. Die mittigen Textfelder zeigen wie viele Fälle und Traces aus dem Ereignisprotokoll durch die Selektion ausgewählt wurden. Außerdem ist für jede der Traces die durchschnittliche Durchlaufzeit angegeben. Die Anzahl der Fälle, die dieser Trace zugeordnet werden können, ist explizit angegeben und zum intuitiveren Verständnis zusätzlich als Balkendiagramm visualisiert.

Nach der Selektion von Traces kann auch die Form des Graphen konfiguriert werden. So kann die Position von Knoten (Aktivitäten) auf der Z-Achse verwendet werden, um zusätzliche Aktivitätsattribute zu codieren. Die Oberfläche, um diese Konfiguration vorzunehmen, zeigt Abb. 5. Die zur Verfügung stehenden Attribute sind diejenigen, die in Celonis als Spalten in der Aktivitäts-Tabelle angelegt und somit dynamisch geladen wurden. Diese Attribute können entweder der Z-Position oder der Knotengröße zugeordnet werden.

Weiterhin muss der Datentypen bei der Wahl der Knotengröße („Node Size“) der Attribute beachtet werden: kontinuierliche Attribute (z. B. Fallzahl der Aktivität) vs. diskrete Attribute (z. B. Nutzertyp). So sind für die Knotengröße nur Fallattribute vom numerischen Datentyp möglich. Während für die Position auf der Z-Achse numerische und diskrete Attributtypen wählbar sind. Bei Auswahl eines Attributs wird, wie in Abb. 5 dargestellt, der gesamte Wertebereich auf dem Panel angezeigt. Dieses ist je nach Datentyp (diskret oder kontinuierlich) entweder das Maximum und Minimum oder die Menge vorhandener unterschiedlicher Merkmalsausprägungen. Das Ergebnis zeigt Abb. 6.

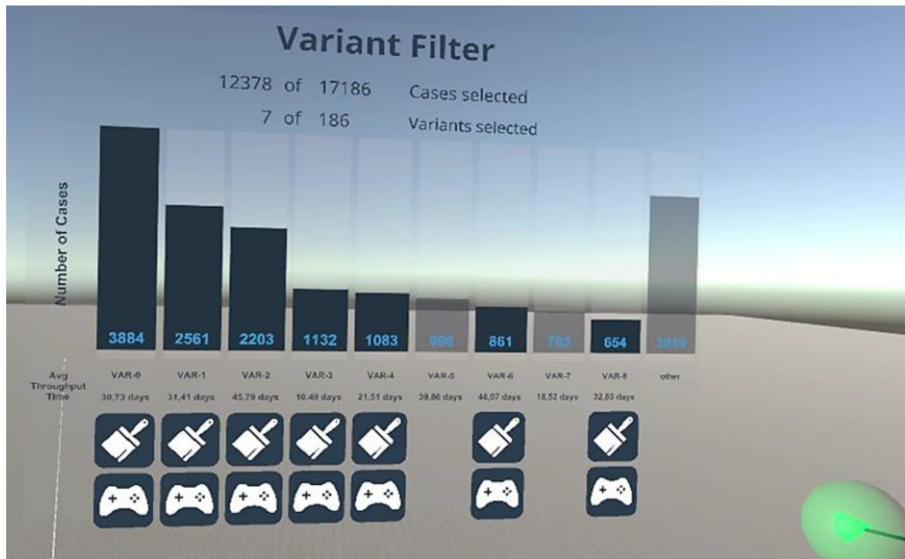


Abb. 4 Trace Filtering Menü

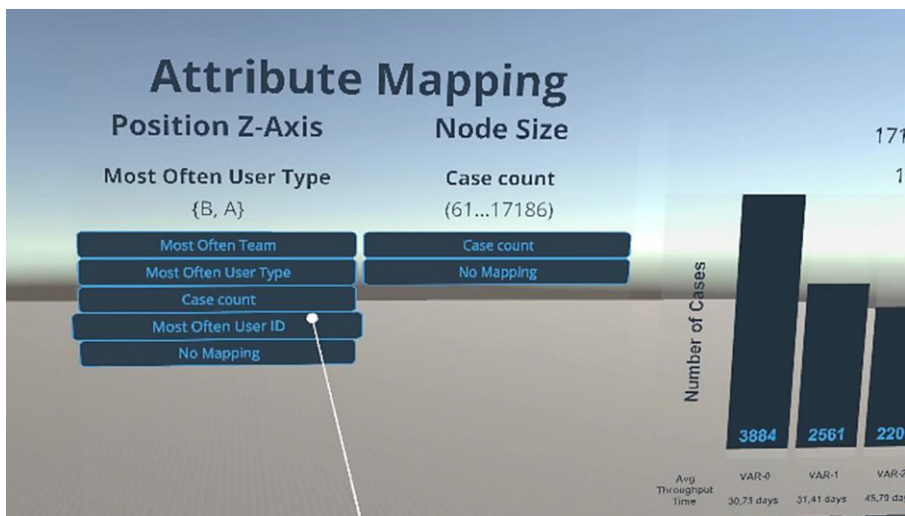


Abb. 5 Konfiguration weiter Darstellungsdimensionen

Abhängig von dem Nutzertypen wird dem Aktivitätsknoten eine unterschiedliche Z-Position zugewiesen (siehe Abb. 6a). Die Achsenbeschriftung auf dem Boden zeigt den Wert der jeweiligen Positionen (siehe Abb. 6b). So können im Beispiel Brüche zwischen automatisierter und manueller Bearbeitung aufgedeckt und Ineffizienzen des Prozesses festgestellt werden. Ein zusätzlicher Vorteil ist, dass durch die Nutzung der Z-Achse als dritte Dimension visuell weniger Kantenumwege und Kreuzungen vorhanden sind als in einer planaren 2D-Darstellung. So wirkt die gleiche Anzahl an

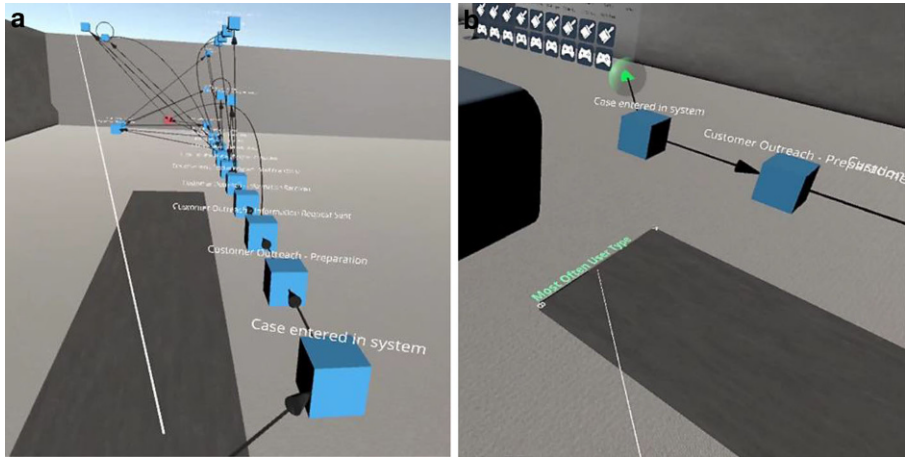


Abb. 6 Darstellung unter Anwendung der Attributcodierung. **a** Seitliche Ansicht. **b** Achsenbeschriftung

Verbindungen verständlicher und der Nutzer kann eine höhere Anzahl von Traces leichter kognitiv miteinander verglichen.

Werden bei der explorativen Analyse des Prozessgraphen mehr Informationen zu einer Aktivität benötigt, reicht ein Klick oder eine Berührung der Aktivität aus, um die Details-On-Demand Anzeige zu aktivieren. Alle zur Verfügung stehenden dynamisch geladenen Informationen zu der Aktivität werden so angezeigt (Abb. 7).

Eine weitere Form der Interaktion mit dem Graphen ist das *Einfärben von Traces*. Diese Funktion der VR-Umgebung wurde in der Freitextfunktion durch Befragte gefordert und ermöglicht, Abweichungen der Traces vom Hauptprozess zu untersuchen. Abb. 8a zeigt die Selektion im Menü.

Zur Identifizierung von Kanten in mehreren Traces wird durch die Farbe bestimmt, welche Traces betrachtet werden. Die Reihenfolge der Auswahl führt zu einer Priorisierung, die durch die Zahl 1 in Abb. 8a für den Nutzer dargestellt wird. So ergibt sich eine klare Reihenfolge der Prioritäten der Traces. Basierend darauf wird die Kante in derjenigen Farbe eingefärbt, die zu dem höchstpriorisierten beteiligten Trace gehört. All diejenigen Traces, die nicht eingefärbt wurden, sind im Gegensatz zu der Funktion des Trace Filtering (siehe Abb. 4) noch sichtbar. Die zugehörigen Knoten und Kanten werden mit einem transparenten Grau versehen, sodass sie optisch nicht von den gefärbten Traces ablenken.

Eine weitere Funktion, die Nutzer in der Studie priorisiert haben, ist die Bereitstellung von Conformance Checking (vgl. Anhang 1). Das Conformance Checking wurde im Werkzeug als Token-Replay umgesetzt (Laue et al., 2020). Durch diese Funktion kann der Prozess durchlaufen werden. Hierzu muss eine Trace ausgewählt werden, was den gesamten Graph grau einfärbt. Nur die erste Kante wird mit einer roten Färbung belegt. Sie zeigt auf die erste Aktivität der Trace. Aufgabe des Nutzers ist es, die erste Aktivität zu berühren, wodurch sich die graue Farbe zu blau verändert. Neben der sich öffnenden *Details-On-Demand-Anzeige* löst die Interaktion zusätzlich eine Animation aus und die nächste Kante wird eingefärbt (siehe Abb. 8).



Abb. 7 Details-On-Demand-Anzeige an Aktivitätsknoten

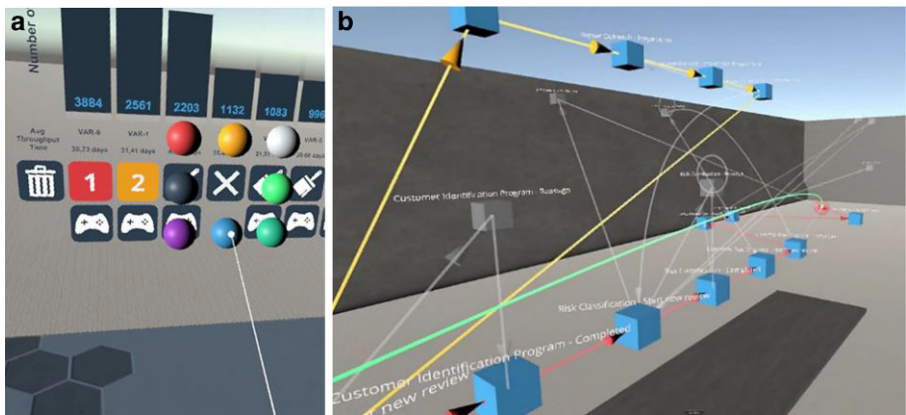


Abb. 8 Trace Einförbung zum Vergleich von Abweichungen. **a** UI zur Auswahl. **b** eingeförbte Traces in dem DFG

Auf diese Weise wird der Nutzer Aktivität für Aktivität durch den Prozess geföhrt. (Abb. 9).

Neben den Interaktionsmöglichkeiten mit dem Prozessgraphen, wurde als zweite Analysekomponente ebenfalls die Case-Cube Komponente (ein Punktdiagramm) zur Exploration der Fall-Attribute implementiert (z. B. zur Darstellung und Analyse von Fällen oder zur Analyse von Attributen mit langen Durchlaufzeit).

Neben den beschriebenen Komponenten bietet die Anwendung auch die Funktion beliebig viele Prozessinstanzen der beiden Komponenten frei in dem Raum zu platzieren und zu skalieren, was die Anforderung Immersive Analytics Authoring erföhlt. Auch die am höchsten priorisierte Funktion Speichern und Laden von Analysen wurde implementiert. Nach Abschluss der Analyse können die Konfigurationen

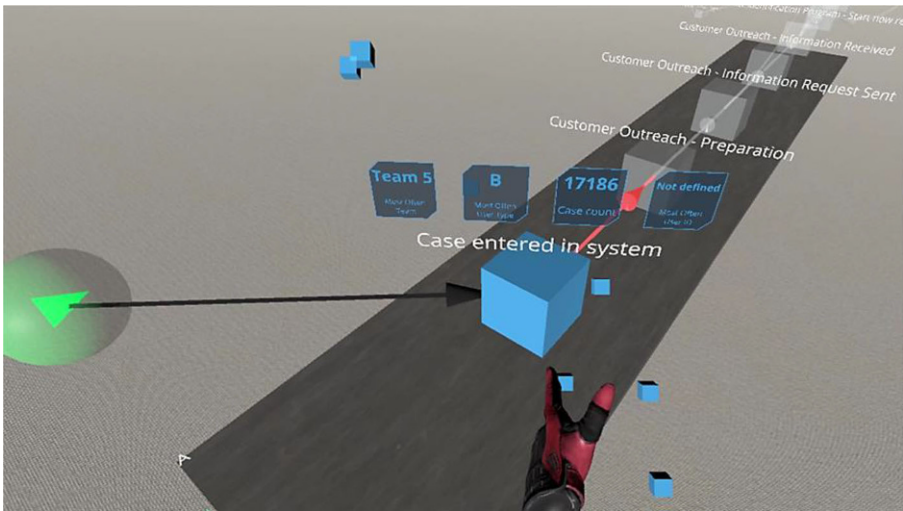


Abb. 9 Umsetzung von Conformance Checking als Trace Replay

gespeichert werden. Bei einem erneuten Laden des Datensatzes öffnet sich dann das gespeicherte virtuelle Dashboard mit den dazugehörigen Daten.

6 Zusammenfassung

Dieser Beitrag stellt eine VR-basierte Umgebung zur immersiven Exploration von Process-Mining vor. Im Rahmen einer Expertenbefragung wurden Anforderungen ermittelt und priorisiert. Zur Beantwortung der Forschungsfrage 1 („*Welche Anforderungen stellen sich an eine VR-basierte Umgebung zur Exploration von Process-Mining?*“) wurden 20 Anforderungen aus kommerziellen (2D) Process Mining Werkzeugen und verwandten Arbeiten aus der Literatur abgeleitet und im Rahmen einer Nutzerstudie untersucht. Die Analyse zeigt, dass das Speichern von Analyseergebnissen sowie die Darstellung des Prozessgraphen am höchsten priorisiert werden. Basierend auf den Ergebnissen der Anforderungsanalyse durch eine Nutzerstudie wurde eine entsprechende Architektur entwickelt und implementiert, die die Forschungsfrage 2 beantwortet („*Wie kann eine solche VR-basierte Umgebung umgesetzt werden?*“). Der VR Prototyp Process Explorer fokussiert insbesondere die Interaktion des Nutzers mit dem Prozessgraphen (d. h., die explorative Analyse des Prozessgraphen). Die Implementierung ist modular, sodass weitere Anforderungen bzw. Funktionen ergänzt werden können. Hervorzuheben ist außerdem die Datenflussanbindung an bestehende laufende Systeme. Zusammenfassend sind viele der hoch priorisierten Funktionen aus der Nutzerstudie in der VR-Umgebung umgesetzt worden.

Zukünftig planen wir folgendes umzusetzen: (1) das Kollaborationsfeature, (2) Erweiterungen der Filterfunktion, (3) datenorientierte, interorganisationale Sichten

zwischen einem Unternehmen und deren Lieferanten, (4) die Verbesserung der Performance, und (5) weitere Ausnutzung der dritten Dimension.

Mittels des *Kollaborationsfeature* soll eine kollaborative Umgebung (Mehrbenutzerbetrieb) ermöglicht werden. Die Modellierung als Mehrbenutzerbetrieb lässt zu, dass Geschäftsprozessen von Stakeholdern und Fachexperten gemeinsam analysiert werden. In der derzeitigen Version des Process Explorers können lediglich in Celonis Filter angelegt und geteilt werden und erlauben damit nur eine rudimentäre Filterung. Für den Anwender wäre es komfortabler, im Prozessmodell selbst zu filtern (z. B. nach der Durchlaufzeit oder nach Fällen mit einer überdurchschnittlich langen Laufzeit). Durch die Visualisierung von interorganisationalen Prozessen könnten Datenflüsse wie beispielsweise Lieferanten mit den jeweiligen Umsatzvolumen und Lieferzeiten analysiert werden.

Im Hinblick auf Performance könnte die VR-Umgebung optimiert werden. Derzeit benötigt der Download des vollständigen Beispiel-Datensatzes (ca. 17000 Fälle) etwa 90s von der Anfrage des Clients bis zum Empfangen der JSON-Antwort. Bisher war der Fokus der VR-Umgebung die Prozessgraphkomponente (d. h., eine kontrollflussorientierte Prozessanalyse).

Die Effektivität der dritten Dimension für Prozessmodellsimulationen wurde bereits gezeigt (Eichhorn et al. 2009). Zukünftig planen wir auch diesen Einsatz im Process Explorer zu prüfen und auch die Nutzung der dritten Dimension zur Modellierung von Prozesshierarchien und Ähnlichkeitsanalysen zwischen Geschäftsprozessen (Betz et al., 2008) auszunutzen.

Anhang

Tab. 3 Vollständige Liste der Anforderungen

Funktion	Langbeschreibung (User Story)	Priorität
Speichern und Laden von Analysen	Als User möchte ich erstellte Analysen weiterbearbeiten oder zu einem späteren Zeitpunkt anderen Menschen präsentieren. Deshalb sollen erstellte Analysen (Kombination aus Datensatz und konfigurierten Visualisierungen) gespeichert und geladen werden können	1,7
Trennung Analyst vs. Betrachter UI	Um die Interaktionen intuitiver zu gestalten und überladen zu können, soll es eine Auswahl geben, ob gerade die Analysen betrachtet werden „Betrachter-Mode“, oder ob neue hinzugefügt & konfiguriert werden sollen „Analyst-Mode“. Dadurch werden nur die kontextabhängigen relevanten Interaktionen und UI-Elemente angezeigt	1,3
Prozess Graph Komponente	Als Betrachter möchte ich in der Lage sein, den Verlauf des Prozesses zu verstehen und Hypothesen zu Schwachstellen zu entwickeln. Hierbei möchte ich von der verbesserten kognitiven Aufnahmekapazität in 3D profitieren und komplexere Verhalten erfassen können. Deshalb soll eine dreidimensionale Variante des Directly-Follows Graphen vom Analysten hinzugefügt werden können	1,3
Ladesymbol	Als User möchte ich ein visuelles Feedback bekommen, wenn etwas lange dauert. Deshalb soll, während Daten aus dem Backend geladen werden, eine visuelle Darstellung für den Download angezeigt werden	1,2

Tab. 3 (Fortsetzung)

Funktion	Langbeschreibung (User Story)	Priorität
Case-Cube Komponente	Als Betrachter möchte ich Muster, Cluster und Ausreißer in den Cases untersuchen. Deshalb soll es dem Analysten möglich sein, einen 3D Scatterplot hinzuzufügen, dieser zeigt die Cases und ihre Attribute an. Das Matching der Attribute auf die Dimensionen unterliegt hierbei dem Analysten	1,1
3D Attribute	Als Betrachter möchte ich noch mehr Informationen zur Verfügung haben, um Schwachstellen im Prozess zu identifizieren bzw. Hypothesen darüber zu entwickeln. Deshalb sollen Aktivitätsattribute (z.B. Abteilung, durchschnittliche Bearbeitungszeit, ausführende Rolle, automatisch/manuell) als Position der Knoten in einem 3D Koordinatensystem dargestellt werden	0,9
Teleporting	Als User möchte ich in der Lage sein, weite Entfernungen in der Virtuellen Welt schnell zurückzulegen. Deshalb soll eine Form des Teleportings ergänzend zum „normalen“ Laufen implementiert werden	0,9
Live-Daten	Als User möchte ich Veränderungen der zugrundeliegenden Daten einfach eingespeist bekommen. Deshalb soll die App eine Anbindung an eine „lebendige“ Datenquelle haben, sodass sich die Visualisierungen beim Laden neuer Basisdaten automatisch aktualisieren	0,8
Authoring	Als Analyst möchte ich in der Lage sein, selbst gewählte Visualisierungen nebeneinander zu betrachten. Deshalb soll es über ein Menü möglich sein, beliebig viele Komponenten (Scatterplot, Prozessgraph, usw.) einer Analyse hinzuzufügen und diese zu konfigurieren (z.B. Achsenparameter auszuwählen)	0,8
Details on Demand	Als Betrachter möchte ich bei der Analyse nicht von Informationen überladen werden, aber dennoch Zugriff auf alle Attribute haben. Deshalb soll bei Auswahl eines einzelnen Cases eine Einblendung aller zugehörigen Attribute erfolgen	0,6
Kollaboration (remote)	Als User möchte ich in der Lage sein, die Datenvisualisierungen mit anderen Usern zu teilen, um meine Gedanken in Echtzeit zu diskutieren und zu präsentieren. Deshalb sollen mehrere User zu einer Session joinen können, ihre Position sehen können und miteinander kommunizieren können	0,5
Zoomen, Platzieren, Rotieren	Als Betrachter möchte ich mir die Daten aus verschiedenen Perspektiven ansehen können und dabei Spaß haben. Deshalb soll es dem Betrachter möglich sein, die einzelnen Komponenten beliebig groß zu skalieren um „rein zu laufen“. Sowie „in den Händen zu halten“ und zu rotieren	0,5
3D Balkendiagramm	Als Betrachter möchte ich auch in der Lage sein, eine geeignete Visualisierung diskreter Attribute zu wählen. Deshalb soll eine 3D Variante des Balkendiagramms als Komponente hinzufügbare sein	0,5
„Echtes“ Filtern	Im Vergleich zu Linking and Brushing soll der Betrachter in der Lage sein, echte Filter auf den Datensatz anzuwenden, wodurch z.B. aggregierende KPIs neu berechnet werden – also alles was nicht durch einen einfachen visuellen Anzeigenwechsel erreicht werden kann, sondern im Backend neu berechnet werden muss	0,4
Linking & Brushing	Als Betrachter möchte ich mir Subsets oder einzelne Daten des Event Logs genauer betrachten. Die Auswahl eines (oder mehrerer) Datenpunktes in dem Scatterplot führt dazu, dass alle vorhandenen Visualisierungen sich synchronisieren und das ausgewählte Subset gehighlighted wird	0,3
Simulation „Token-Replay“	Als Betrachter möchte ich die zeitliche Komponente des Prozesses besser verstehen. Deshalb soll der Prozessgraph über eine Animation verfügen, in der die Cases im Zeitraffer entsprechend ihrer tatsächlichen realen Bearbeitungszeiten durch die einzelnen Knoten laufen	0,1

Tab. 3 (Fortsetzung)

Funktion	Langbeschreibung (User Story)	Priorität
Inhouse Fabrik Prozess Landkarte	Als Betrachter möchte ich bei physischen Prozessen den Materialfluss innerhalb eines Gebäudes sehen können, um lange Laufwege im Wertstromfluss zu identifizieren. Deshalb soll es eine Möglichkeit geben, Aktivitäten einem Standort innerhalb eines Fabrikmodells zuzuordnen und den Prozessgraphen über ein 3D Modell einer Fabrik zu legen	-0,1
3D Umgebung	Als Betrachter möchte ich mich in der virtuellen Welt gut orientieren können und meine Position bzw. Beobachtungen präziser kommunizieren können. Deshalb sollen die Analysen in einem visualisierten 3D Raum platziert werden (z. B. virtueller Konferenzraum oder Kunstgalerie anstatt in einem schwarzen „Nichts“)	-0,4
Geographische Prozess-Landkarte	Als Betrachter möchte ich Geopositionen und Entfernungen zwischen Prozessschritten begreifen. Deshalb soll es eine Komponente geben für Datensätze mit großen räumlichen Entfernungen zwischen den Prozessschritten. Der Prozessgraph wird dann als Erhebung einer Karte auf dem Boden angezeigt (quasi als Spezialisierung der Prozess Graph Komponente wobei X- und Y-Achse für Koordinaten verwendet werden)	-0,5
Money Stack	Als User möchte ich große Geldbeträge greifbarer präsentieren und Spaß haben. Deshalb soll der Analyst in der Lage sein, einen Stapel Geldscheine als Komponente hinzuzufügen und mit einem Geldwert zu kombinieren. Der Stapel wächst dann zur realen physischen Repräsentation des Betrages	-1,4
Individuelle KPIs	Es sollen eigene entwickelte KPIs von dem System unterstützt werden	-
2D-Tabellen Komponente	Als Analyst möchte ich auch Daten integrieren, für die noch keine 3D Lösung implementiert wurde	-
Farbliches Trace Highlighten	Als Betrachter möchte ich Abweichungen von Traces untersuchen können. Deshalb soll es möglich sein, Traces farblich hervorzuheben	-

Danksagung Dieser Beitrag und die dazugehörige Forschung wäre ohne die Unterstützung von Capgemini in Deutschland nicht möglich gewesen. Besonderer Dank gilt insbesondere Andreas Eßbaumer, Andreas Brillung und Lisa Zimmermann.

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

- Betz S, Eichhorn D, Hickl S, Klink S, Koschmider A, Li Y, Oberweis A, Trunko R (2008) 3D representation of business process models. In: Loos P, Nüttgens M, Turowski K, Werth D (Hrsg) Modellierung betrieblicher Informationssysteme (MobIS 2008). Gesellschaft für Informatik e.V., Bonn, S 73–87
- Billinghurst M, Cordeil M, Bezerianos A, Margolis T (2018) Collaborative immersive analytics. In: Marriott K, Schreiber F, Dwyer T, Klein K, Riche NH, Itoh T et al (Hrsg) Immersive analytics. Springer, Cham, S 221–257
- Chandler T, Cordeil M, Czaderna T, Dwyer T, Glowacki J, Goncu C et al (2015) Immersive analytics. In: 2015 Big data visual analytics (BDVA). IEEE Computer Society, Washington, D.C., S 1–8
- Cordeil M, Cunningham A, Bach B, Hurter C, Thomas BH, Marriott K et al (2019) IATK: an immersive analytics toolkit. In: 2019 IEEE conference on virtual reality and 3D user interfaces (VR 2019) Osaka, Japan, 23-27 March 2019 IEEE Computer Society, Piscataway, S 200–209
- Drescher A, Koschmider A, Oberweis A (2017) Modellierung und Analyse von Geschäftsprozessen. De Gruyter Oldenbourg, Berlin
- Eichhorn D, Koschmider A, Li Y, Stürzel P, Oberweis A, Trunko R (2009) 3D support for business process simulation. In: Proceedings of the 33rd annual IEEE international computer software and applications conference, COMPSAC. IEEE Computer Society, Washington, D.C., S 73–80
- Gall M, Rinderle-Ma S (2019) Process attribute visualization in 3D and virtual reality. In: Depaire B, de Smedt J, Dumas M, Fahland D, Kumar A, Leopold H et al (Hrsg) Proceedings of the dissertation award, doctoral consortium, and demonstration track at BPM 2019. CEUR workshop proceedings, Bd. 2420, S 184–188 (Verfügbar unter: <http://ceur-ws.org/Vol-2420/papeDT14.pdf>)
- Garcia-Hernandez RJ, Anthes C, Wiedemann M, Kranzlmüller D (2016) Perspectives for using virtual reality to extend visual data mining in information visualization. In: 2016 IEEE aerospace conference (AERO 2016). IEEE, New York, S 1–11
- Laue R, Koschmider A, Fahland D (Eds.) (2020) Prozessmanagement und Process-Mining, De Gruyter Studium, De Gruyter 2020, ISBN 978-3-11-050015-8
- Lee B, Cordeil M, Prouzeau A, Dwyer T (2019) FIESTA: a free roaming collaborative Immersive Analytics system. In: Lee B, Lee G, Scott S, Tory M, Kim J (Hrsg) Proceedings of the 2019 ACM international conference on interactive surfaces and spaces. ACM, New York, S 335–338
- Marriott K, Chen J, Hlawatsch M, Itoh T, Nacenta MA, Reina G et al (2018) Just 5 questions: toward a design framework for immersive analytics. In: Marriott K, Schreiber F, Dwyer T, Klein K, Riche NH, Itoh T et al (Hrsg) Immersive analytics. Springer, Cham, S 259–288 https://doi.org/10.1007/978-3-030-01388-2_9
- Moran A, Gadepally V, Hubbell M, Kepner J (2015) Improving big data visual analytics with interactive virtual reality. In: 2015 IEEE high performance extreme computing conference (HPEC). IEEE, New York, S 1–6
- Su S, Perry V, Bravo L, Kase S, Roy H, Cox K et al (2020) Virtual and augmented reality applications to support data analysis and assessment of science and engineering. *Comput Sci Eng* 22(3):27–39. <https://doi.org/10.1109/MCSE.2020.2971188>
- Ware C, Mitchell P (2008) Visualizing graphs in three dimensions. *ACM Trans Appl Percept* 5(1):1–15. <https://doi.org/10.1145/1279640.1279642>