

Vorhersage von menschlichen Arbeitsbewegungen zur Verringerung des Kollisionsrisikos bei kollaborativer Arbeit mit mobilen Robotern

Edgar SCHERSTJANOI, Moritz SCHMIDT, Martin SCHMAUDER

*Institut für Arbeitswissenschaft, Technische Universität Dresden
Marschnerstraße 39, D-01307 Dresden*

Kurzfassung: Für die Gestaltung von kollaborativen Mensch-Maschine-Arbeitsplätzen sind mobile, frei fahrende Roboter von immer größerer Bedeutung. Das volle Potential solcher autonomen Systeme kann jedoch nur ausgeschöpft werden, wenn Mensch und Maschine flexibel und sicher in unmittelbarer Nähe miteinander agieren können. Durch das Nutzen von Motion-Capture-Datenströmen für maschinelles Lernen einer künstlichen Intelligenz wurden Vorhersagen der menschlichen Körperbewegung generiert und ausgewertet. Die Einbindung dieser Prognosen in die Routenplanung eines Roboters wurde simuliert und deren potentieller Mehrwert für den Schutz des Menschen dargelegt. Die Ergebnisse liefern somit einen weiteren Beitrag hin zu echter Kollaboration zwischen Mensch und Maschine.

Schlüsselwörter: Motion Capture, Bewegungsvorhersage, Mensch-Maschine-Kollaboration, Maschinelles Lernen, Kollisionsvermeidung

1. Einleitung

Eines der wichtigsten Kriterien für den Betrieb von mobilen Robotern ist das Vermeiden von Kollisionen mit Menschen. Standardmäßig werden diese durch die Erfassung und Überwachung der menschlichen Position seitens des Roboters reduziert. Bei Unterschreiten einer geschwindigkeitsabhängigen Distanz der Akteure wird die Reduzierung der Betriebsgeschwindigkeit bzw. das Auslösen von Notstopps der Maschine eingeleitet (vgl. DIN EN ISO 13857:2019, DIN ISO/TS 15066:2016). Im industriellen Einsatz sind diese Stillstandszeiten jedoch durch einhergehende Einbußen in der Produktivität nachteilhaft.

Ein Ansatz zur Bearbeitung dieser Problematik besteht in der Entwicklung einer künstlichen Intelligenz (KI), die dazu in der Lage ist, eine Prognose zum Pfad einer Arbeitsbewegung des Menschen zu generieren. Es ist anzunehmen, dass bestimmte Abfolgen einer Körperbewegung Indizien zum Weiterverlauf der Bewegung liefern, wie z.B. die Korrelation zwischen dem Pendeln der Arme und der Bewegungsgeschwindigkeit, der Einfluss der Blickrichtung auf die Laufrichtung oder die durch Oberkörperdrehung eingeleitete Kurvenbewegung. Mit solchen Vorhersagen könnten die Routen des mobilen Roboters angepasst werden, um dem Menschen auch prospektiv auszuweichen und somit eine notwendige Bewegungsreduktion zu minimieren.

Anstatt jedes Indiz zur Bewegungsvorhersage explizit in die Programmierung des Roboters zu integrieren, bieten sich KI-Systeme an, die durch maschinelles Lernen (ML) Körperbewegungen einer Datenbasis analysieren, die richtigen Zusammenhänge erkennen und somit Vorhersagen zur menschlichen Körperbewegung selbstständig treffen können.

Zur Aufnahme von Körperbewegung wurde im Rahmen des Vorhabens das inertialbasierte Motion-Capture-System (MoCap) „MVN XSens“ verwendet, welches für 23 Gelenkpunkte jeweils die Orientierung in Form von Quaternionen und die Position im kartesischen Koordinatensystem bei a-priori definierten Segmentlängen und einer Aufnahme Frequenz von 60fps (frames per second; Einzelaufnahmen pro Sekunde) zur Verfügung stellt. Diese Technologie gilt als qualitativ äquivalent zu den Referenztechnologien optischer, Marker-basierter Systeme (Al-Amri et al. 2018; Robert-Lachaine et al. 2017).

2. Vorhersage von Körperbewegung

Mittels MoCap-Technologien ist es möglich, Körperbewegungen aufzunehmen und daraus die Datenbasis zum maschinellen Lernen zu erstellen. Üblicherweise wird eine Bewegungsaufnahme der Länge l als Folge einzelner Körperhaltungen $(K_t)_{t \in \{0..l\}}$ repräsentiert. Jede Körperhaltung K_t zu einem Zeitpunkt t ist wiederum als skelettale Hierarchie von Gelenkpunkten und Körpersegmenten definiert.

Eine vielversprechende Möglichkeit zur Modellerstellung für Bewegungsprognosen mittels ML ist durch sogenannte „rekurrente“ künstliche neuronale Netze (RNN) gegeben. Prinzipiell ermöglichen RNN es, sequentielle Eingaben zu verarbeiten und sequentielle Ausgaben erzeugen. In den Ansätzen von Chiu et al. (2019) oder Martinez, Black & Romero (2017) wurde die Anwendbarkeit von RNN auf Verarbeitung von Körperbewegungen belegt. Als Eingangssequenz wurde dabei die Bewegungsabfolge der Länge l_{in} analysiert, die zur Körperhaltung K_t geführt hat, um daraus eine Vorhersage für eine bestimmte Anzahl (l_{out}) an Körperhaltungen darauf folgender Zeitpunkte zu erzeugen.

Im Gegensatz zu diesen Ansätzen wurde hier das Ziel verfolgt, nicht kommende Körperhaltungen, sondern die zukünftige Positionierung (d.h. der Bewegungspfad) einer Person im Raum zu prognostizieren. Als Eingabesequenz für die KI dienten die Winkelinformationen jedes einzelnen mittels MoCap aufgenommenen Gelenkpunktes. Daraus wurden die Positionsveränderungen des (auf den Boden projizierten) Skelettsprungs (Mitte des Beckens) pro Zeitpunkt als Ausgabesequenz generiert. Aus einer Sammlung von Bewegungsaufnahmen konnten somit Trainingsdaten als Basis für das ML erstellt werden und der Zusammenhang zwischen der Abfolge von Körperhaltungen und dem daran anschließenden Bewegungspfad (d.h. einer Folge von Positionierungen für die kommenden l_{out} Zeitpunkte) „erlernt“ werden.

Eine Aussage zur Validität der generierten Vorhersage konnte getroffen werden, indem für MoCap-Aufnahmen, die kein Bestandteil des Trainingsdatensatzes waren, die erstellte Prognose mit der tatsächlichen Bewegung verglichen wurde. In Abbildung 1 sind dazu Beispiele von K_t visualisiert, jeweils mit der Körperbewegungssequenz, die zu K_t geführt hat: unmittelbar vor dem Stehenbleiben (a), während des Stehens (b), vor dem Beginn des Gehens (c) und während des Gehens (d). Für den Vergleich sind jeweils der prognostizierte (gelb – rot) und der tatsächlich auf K_t folgende (cyan – blau) Bewegungspfad abzulesen.

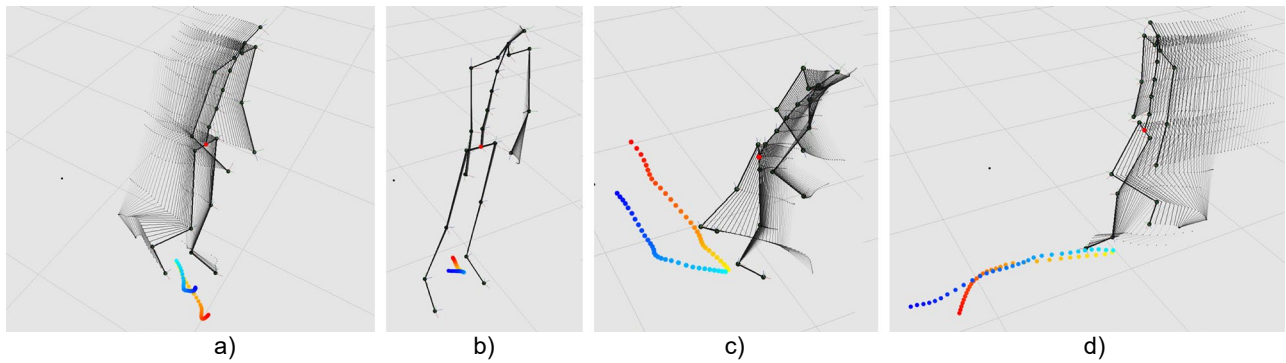


Abbildung 1: Körperbewegungen (grau schattiert), die zu K_t (schwarz) geführt haben und als Grundlage zur Berechnung der Vorhersage verwendet wurden. Der tatsächlich folgende Bewegungsverlauf ist durch Punkte im Farbverlauf von cyan nach dunkelblau, die von der KI berechnete Prognose durch Punkte im Farbverlauf von gelb zu rot dargestellt. a) unmittelbar vor dem Stehenbleiben, b) während des Stehens, c) vor dem Beginn des Gehens, d) während des Gehens.

Sei das Ergebnis der Vorhersage zum Zeitpunkt t ein Bewegungspfad, der durch die Folge der Punkte $P_t^{pred} = (p_i^{pred})$ definiert ist und sei – analog dazu – die Positionsveränderung in der Aufnahme die Punktfolge $P_t^{truth} = (p_i^{truth})$ jeweils für die Indizes $i \in \{1, 2, \dots, l_{out} = 30\}$. Für jeden Zeitpunkt t kann die euklidische Distanz zwischen den 30 Punktpaaren $(p_i^{pred}, p_i^{truth})$ berechnet werden. Dieses Fehlermaß gibt Aufschluss über den Verlauf der Prognose während eines Bewegungsablaufes: je größer die Distanz ausfällt, desto geringer ist die Qualität der Vorhersage zu bewerten. Es ist anzunehmen, dass diese Metrik abhängig von der zukünftig zurückgelegten Strecke der tatsächlichen Bewegung ist. Das bedeutet, dass größere Fehlerwerte entstehen können, je mehr sich die Position der Person ändern wird (vgl. Abbildung 1, in der die Distanz der Punktpaare in der stehenden Situation b) am geringsten und in der gehenden Situation d) am größten ist). Zudem ist zu erwarten, dass mit größerem Index auch der Fehlerwert steigt – das heißt, je weiter eine Prognose in die Zukunft reicht, desto größer ist die Abweichung zur tatsächlichen Bewegung an diesem Zeitpunkt.

In Abbildung 2 sind exemplarisch die Fehlermessungen und die tatsächliche Pfadlänge für eine Bewegungssequenz dargestellt, welche während des „Stehens“ (ca. bis $t = 30$) beginnt und nach dem „Gehen“ wieder in einer stehenden Körperhaltung endet (ca. ab $t = 350$). Der größte Fehlerwert wurde während des Gehens bei $t = 161$ für eine Pfadlänge von $\sim 1,7$ m erzeugt. Für $p_{i=30}^{pred}$ entspricht dort die euklidische Distanz zur tatsächlichen Position $p_{i=30}^{truth}$ 1,305 m, für $p_{i=15}^{pred}$ zu $p_{i=15}^{truth}$ wiederum nur 0,586 m. Obwohl auch bei $t = 225$ eine große Pfadlänge folgen wird, entspricht das Fehlermaß für alle Indizes nur $\sim 0,10$ m, was bedeutet, dass die Vorhersage sehr nahe an der tatsächlichen Bewegung liegt. Die Übergänge von Stehen zu Gehen und vice versa werden weitestgehend richtig vorhergesagt.

Dieses Ergebnis motivierte zum Einbinden des durch die KI generierten menschlichen Bewegungspfades in die Navigation eines Roboters in einer kollaborativen Mensch-Maschine-Umgebung. Es sollte simuliert werden, ob trotz der punktuell fehlerhaften Vorhersage eine Verringerung des Kollisionsrisikos zwischen Roboter und Mensch erreicht werden kann.

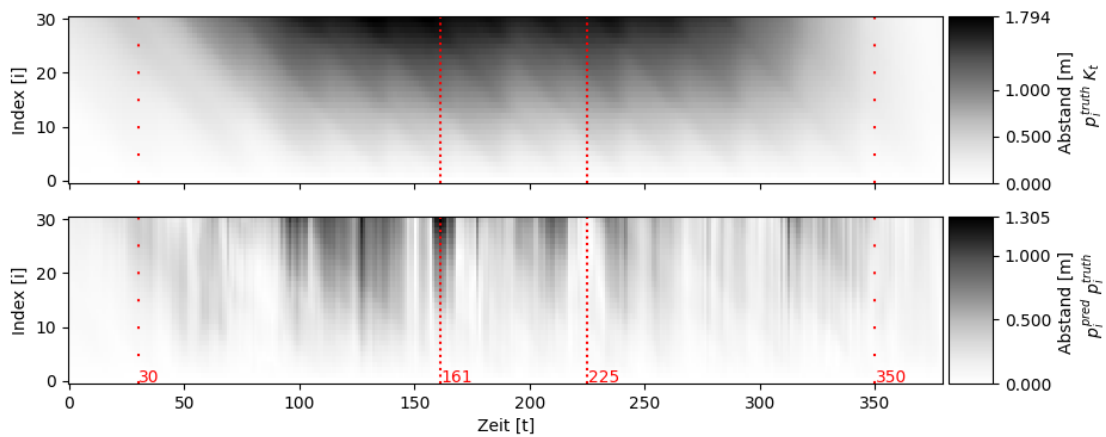


Abbildung 2: Fehlermessung einer Körperbewegung von „Stehen“ (ca. bis $t = 30$) über „Gehen“ (ca. bis $t = 350$) nach „Stehen“. Oben: Tatsächliche Pfadlänge ab dem Zeitpunkt t für die nächsten i Zeitpunkte. Unten: Euklidische Distanz zwischen der Vorhersage am Zeitpunkt t für jedes Punktpaar von Prognose und tatsächlicher Bewegung. Markiert sind zudem die großen Fehlerwerte im Abschnitt „Gehen“ bei $t = 161$ und die geringen Fehlerwerte bei $t = 255$.

3. Simulation in einer virtuellen Umgebung

Das Open-Source-Framework „Robot Operating System“ (ROS - Quigley et al. (2009)) bietet die Möglichkeit, digitale Zwillinge von autonomen freifahrenden Robotern in einer Simulationsumgebung einzusetzen. Für den Betrieb eines Fahrzeuges sind de-facto standardisierte Algorithmen zur Lokalisierung und Navigation sowie die Emulation von industrieüblichen Sensoren implementiert. Mit dem ROS-Framework ist es auch möglich, weitere Informationsquellen wie etwa den Datenstrom eines Motion-Capture-Systems bzw. Ergebnisse der ML-basierten Bewegungsvorhersage des Menschen in die Simulation einfließen zu lassen.

Für die folgenden Versuche wurde eine Simulation des „Robotis Turtlebot3 Waffle Pi“ eingesetzt, welcher zur Kollisionsvermeidung und Navigation den „Dynamic Window Approach“ (Fox, Burgard, & Thrun 1997) verwendet. Wie im Abschnitt 2 beschrieben, generiert die KI für jedes t unter Verwendung der letzten $l_{in} = 30$ Körperhaltungen eine Vorhersage für die kommenden $l_{out} = 30$ Zeitpunkte des Pfades eines in der Simulation agierenden Menschen. Diese Punktfolge wird als dynamisches Hindernis in ROS kommuniziert, welches der Roboter zur Berechnung der kürzesten Route zum Ziel berücksichtigt. In Abbildung 3 ist die Draufsicht der simulierten Umgebung mit der Position des Menschen (roter Punkt) und des Roboters (blau) sowie der berechneten Route zum Ziel (grün) dargestellt. Die grauen Bereiche visualisieren den Einfluss erkannter Hindernisse auf die Routenplanung des Roboters: einerseits, wenn der Mensch nur als dynamisches Hindernis wahrgenommen wird (links), andererseits, wenn auch der prognostizierte Bewegungspfad in die Routenplanung einbezogen wird (rechts). Es ist ersichtlich, dass ohne die Beachtung der Bewegungsprognose eine erhöhte Kollisionsgefahr besteht, da beide Pfade sich kreuzen, insofern die Prognose sich bewahrheitet.

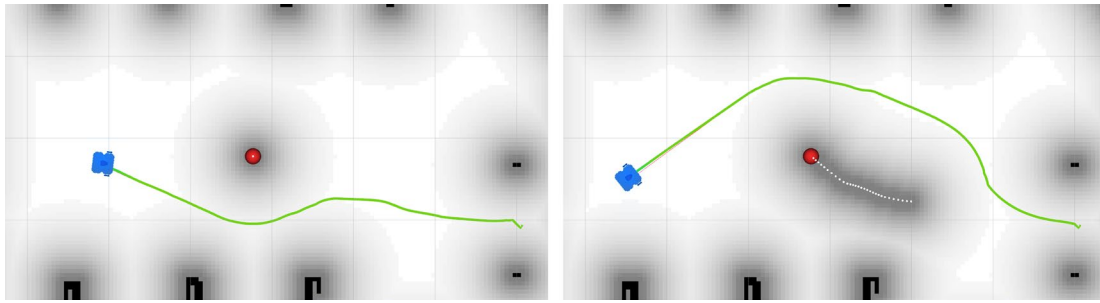


Abbildung 3: Draufsicht der simulierten Versuchsumgebung. Blau: Position des Roboters, Rot: Position des Menschen, Grauwerte: Einfluss der Hindernisse auf die Pfadberechnung des Roboters (grün) bei Betrachtung der Position des Menschen (links) und unter Einfluss der Prognose zum Bewegungspfad des Menschen (rechts).

Mit dem Ziel, eine Überprüfung der Kollisionsrisiken bei Betrachtung eines längeren Zeitraumes durchzuführen, wurden in einem $\sim 50\text{m}^2$ großen Versuchsraum mehrere Stationen festgelegt. In diesem Versuchsraum wurde die Bewegung eines Probanden aufgenommen, welcher sich zwischen zufällig ausgewählten Stationen bewegte. Im virtuellen Äquivalent des Versuchsraumes fand nun die Simulation der Bewegung des Roboters zwischen ebenfalls zufällig ausgewählten Stationen statt. Im Experiment wurde der Abstand zwischen den Positionen von Mensch und Roboter (vgl. Abbildung 4) einerseits mit Einbindung der Bewegungsprognose (grün), andererseits nur unter Berücksichtigung der Position des Menschen als dynamisches Hindernis (grau) in die Navigation des Roboters entlang der Zeit gemessen. Das Ergebnis zeigt, dass die Abstände in beiden Situationen weitestgehend übereinstimmten, allerdings bei den Zeitpunkten $s = 133$ und $s = 160$ die Distanz zwischen Mensch und Roboter ($\sim 0,5\text{ m}$ bzw. $\sim 1,5\text{ m}$) mit der Bewegungsvorhersage deutlich größer war. Zu diesen Zeitpunkten wäre durch den Einfluss der Prognose die Kollisionsgefahr verringert worden.

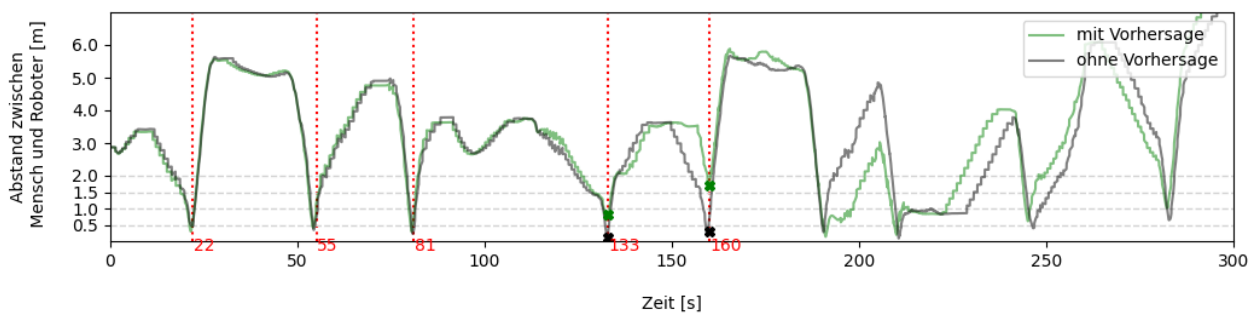


Abbildung 4: Veränderung des Abstandes zwischen Mensch und Roboter während eines Versuchsablaufes. Im Unterschied zur Situation, in der nur die Position des Menschen (grau) in die Routenplanung einfluss, konnten bei Beachtung der Bewegungsprognose (grün) an zwei Zeitpunkten ($s = 133$ und $s = 160$) mögliche Kollisionen prospektiv unterbunden werden.

Im letztgenannten Versuch reagiert der Roboter zwar auf den Menschen, jedoch fanden die Aufnahmen des Menschen ohne den realen Roboter statt. Die simulierte Strategie zur Kollisionsvermeidung wurde somit stets seitens des Roboters initiiert und schlug fehl, wenn der Mensch (unwissentlich) in die Nähe des stehenden Roboters kam bzw. mögliche Ausweichrouten blockierte. Aus diesem Grund konnte an den Zeitpunkten $s = 22, 55$ und 81 die Kollision in der Simulation nicht vermieden werden.

4. Zusammenfassung und Ausblick

Es konnte gezeigt werden, dass durch maschinelles Lernen von Körperbewegung eine Möglichkeit gegeben wird, Vorhersagen für Bewegungspfade zu generieren, welche größtenteils nur geringe Abweichungen von der tatsächlichen Bewegungsabfolge aufweisen. Wird diese Prognose in die Navigationsalgorithmen von mobilen Robotern zur Kollisionsvermeidung einbezogen, kann das Kollisionsrisiko verringert werden und dadurch ein wichtiger Beitrag zur Erhöhung der Personen- und Materialsicherheit geleistet werden.

Dem Untersuchungsergebnis schließen sich weitere Forschungsfragen an:

- Kann eine längere zeitliche Vorhersage erreicht werden (bspw. in Verwendung anderer künstlicher neuronaler Netze)?
- Wie ist das Vorhersagesystem in einer realen Situation unter Berücksichtigung der menschlichen Reaktion auf das Verhalten des Roboters zu bewerten?
- Ist auch ein Mehrwert beim Einsatz mehrerer Akteure zu verzeichnen?
- Eignen sich auch kostengünstigere, wie bspw. Kamera-basierte MoCap-Technologien zur Bewegungsprognose?

Darüber hinaus wurde während der Untersuchungen deutlich, dass beim maschinellen Lernen nur solche Bewegungsmuster einbezogen werden, welche im Trainingsdatensatz enthalten sind. Komplexere Bewegungsabläufe, die in besonderen Arbeitssituationen entstehen, sollten idealerweise auch Bestandteil der dem maschinellen Lernen zugrundeliegenden Bewegungsaufnahmen sein. Es ergibt sich somit der Bedarf nach einer umfangreichen und wohlstrukturierten Datenmenge mit Aufnahmen von Arbeitsbewegungen.

5. Literatur

- Al-Amri M, Nicholas K, Button K, Sparkes V, Sheeran L, Davies JL (2018) Inertial measurement units for clinical movement analysis: reliability and concurrent validity. *Sensors*, 18(3), 719.
- Chiu H, Adeli E, Wang B, Huang DA, Niebles JC (2019) Action-agnostic human pose forecasting. In 2019 IEEE Winter Conference on Applications of Computer Vision (WACV) (S. 1423–1432).
- DIN EN ISO 13857 (2019) Sicherheit von Maschinen - Sicherheitsabstände gegen das Erreichen von Gefährdungsbereichen mit den oberen und unteren Gliedmaßen. Beuth Verlag
- DIN ISO/TS 15066 (2016) Roboter und Robotikgeräte - Kollaborierende Roboter. Beuth Verlag
- Fox D, Burgard W, Thrun S (1997) The dynamic window approach to collision avoidance. *IEEE Robotics & Automation Magazine*, 4(1), 23–33.
- Martinez J, Black MJ, Romero J (2017) On human motion prediction using recurrent neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (S. 2891-2900).
- Quigley M, Conley K, Gerkey B, Faust J, Foote T, Leibs J, Ng AY (2009) ROS: an open-source Robot Operating System. In ICRA workshop on open source software (Vol. 3, S. 5).
- Robert-Lachaine X, Mecheri H, Larue C, Plamondon A. (2017) Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis. *Medical & Biological Engineering & Computing*, 55(4), 609–619.



Gesellschaft für
Arbeitswissenschaft e.V.

Arbeit HUMAINE gestalten

67. Kongress der
Gesellschaft für Arbeitswissenschaft

Lehrstuhl Wirtschaftspsychologie (WiPs)
Ruhr-Universität Bochum

Institut für Arbeitswissenschaft (IAW)
Ruhr-Universität Bochum

3. - 5. März 2021

GfA-Press

Bericht zum 67. Arbeitswissenschaftlichen Kongress vom 3. - 5. März 2021

**Lehrstuhl Wirtschaftspsychologie, Ruhr-Universität Bochum
Institut für Arbeitswissenschaft, Ruhr-Universität Bochum**

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.
Dortmund: GfA-Press, 2021
ISBN 978-3-936804-29-4

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle erhältlich.

Alle Rechte vorbehalten.

© **GfA-Press, Dortmund**

Schriftleitung: Matthias Jäger

im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Ohne ausdrückliche Genehmigung der Gesellschaft für Arbeitswissenschaft e.V. ist es nicht gestattet:

- den Kongressband oder Teile daraus in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) zu vervielfältigen,
- den Kongressband oder Teile daraus in Print- und/oder Nonprint-Medien (Webseiten, Blog, Social Media) zu verbreiten.

Die Verantwortung für die Inhalte der Beiträge tragen alleine die jeweiligen Verfasser; die GfA haftet nicht für die weitere Verwendung der darin enthaltenen Angaben.

Screen design und Umsetzung

© 2021 fröse multimedia, Frank Fröse

office@internetkundenservice.de · www.internetkundenservice.de