

Eine modulare Software-Pipeline zur Synthese funktioneller menschlicher Bewegungen

Dennis WITTCHEN^{1,2}, Markus WACKER¹

¹ Hochschule für Technik und Wirtschaft Dresden
Friedrich-List-Platz 1, D-01069 Dresden

² Institut für Technische Logistik und Arbeitssysteme
Professur für Arbeitswissenschaft, Technische Universität Dresden
Marschnerstraße 39, D-01307 Dresden

Kurzfassung: Im Rahmen des Promotionsprojekt „MotionSynth“ wird eine Software-Pipeline entwickelt, die zur Synthese menschlicher Bewegungen eingesetzt werden soll. Damit sollen funktionelle Bewegungen, mit Hilfe datengetriebener Verfahren, erzeugt werden. So sollen bereits aufgenommene Bewegungen aus Datenbanken rekombiniert, oder gänzlich neue Bewegungen synthetisiert werden. Die Entwicklung der Pipeline erfolgt zunächst mit Fokus auf die Anwendungsgebiete Rehabilitationsmedizin und virtuelle Ergonomie und wird entsprechend der anwendungsspezifischen Anforderungen validiert. Methodische Ziele der Promotion liegen in der Ableitung generischer Schnittstellen und im modularen Aufbau der Pipeline, um sie für andere Anwendungsgebiete adaptieren zu können.

Schlüsselwörter: Bewegungssynthese, Bewegungsanalyse, Animation, Computer Simulation, künstliche Intelligenz

1. Einleitung und Motivation

Das hier vorgestellte Promotionsprojekt „MotionSynth“ entstand aus einem Forschungsprojekt der HTW Dresden (Kinetek 2019), in dem Bewegungen von Parkinson-Patient*innen analysiert wurden. Ziel war es Krankheitssymptome in Aufnahmen von ausgewählten, funktionellen Bewegungen, z.B. Sit-to-Stand (vgl. Millington et al. 1992) im Rahmen von Ergotherapien, zu detektieren und den Verlauf der Krankheit zu dokumentieren. Dazu wurden die Patient*innen mit einem Mehrkameranystem und einer Microsoft Kinect V2 aufgezeichnet. Zur Bewegungsanalyse kamen datengetriebene Verfahren bzw. Verfahren basierend auf künstlicher Intelligenz (KI) zum Einsatz. KI-basierte Verfahren benötigen für adäquate Resultate oft große Datensätze mit einer anwendungsbezogenen Diversität und Validität der Daten. Andernfalls besteht die Gefahr von Diskriminierung oder Marginalisierung bestimmter Personengruppen (Shankar et al. 2017). Für die Analyse von Bewegungen in der Rehabilitationsmedizin bedeutet dies, dass möglichst viele Proband*innen z.B. unterschiedlichen Alters, Geschlechts und Schweregrad der Erkrankung die gleiche Bewegung unter den gleichen Bedingungen mehrmals ausführen müssen. Wir haben erkannt, dass diese Anforderungen in der Praxis nahezu nicht realisierbar sind, was letztendlich dazu führte, dass die angewendeten KI-basierten Analysen keine verlässlichen Ergebnisse lieferten.

Daraus entwickelte sich der hier vorgestellte Ansatz, fehlende Diversität in den Datensätzen durch synthetisierte Daten zu kompensieren. Ein technisches System, wel-

ches diese Varianz anthropometrischer Variablen und die Reproduzierbarkeit von Bewegungsabläufen erzeugen kann, könnte die wissenschaftliche Community mit diskriminierungsarmen Datensätzen bei der Weiterentwicklung von Bewegungsanalysen unterstützen. Die Dokumentation der Datensätze nach dem Beispiel von Gebru et al. (2018) könnte aus solch einem System ebenfalls automatisiert abgeleitet werden und somit das Forschungsdatenmanagement vereinfachen sowie erzielte Ergebnisse nachvollziehbar machen.

Die Bewegungssynthese soll auf Basis von KI-Verfahren entwickelt werden, da sich damit solche komplexen Anwendungen flexibler modellieren lassen als mit regelbasierten Systemen. Methodische Herausforderungen bestehen diesbezüglich darin, natürlich wirkende und biomechanisch plausible Bewegungen anhand von Bewegungsprimitiven (Krüger et al. 2007) zu synthetisieren, oder den Stil einer Bewegung (z. B. altersbedingte Verlangsamung) auf eine andere Bewegung zu übertragen. Charakteristische Merkmale der Zielbewegung sollen dabei durch die Anwender*innen der Software parametrisierbar sein (z. B. Auswahl des Alters).

Zur Sicherstellung valider Ergebnisse wird eine anwendungsnahe Entwicklung angestrebt. Neben der Rehabilitationsmedizin soll dazu die virtuelle Ergonomie als geeignetes Anwendungsgebiet näher betrachtet werden. Auf dem Gebiet der virtuellen Ergonomie gibt es laut Mühlstedt (2016) Entwicklungsbedarf für modulare Softwarelösungen, die über komplexe Funktionalitäten verfügen. Dies betrifft u. a. die (Weiter-)Entwicklung digitaler Menschmodelle (DMM) und deren Manipulationsfunktionen. Zudem soll deren Entwicklung in Form von Open-Source-Projekten forciert werden. Auch bei bestehenden Datenbanken sieht Mühlstedt Erweiterungsbedarf – insbesondere für dedizierte Posen und Bewegungen. Wir sehen hier Schnittmengen zu Anforderungen, die sich aus dem Projekt Kinetek heraus entwickelt haben.

Im Weiteren ist dieser Artikel wie folgt strukturiert. Abschnitt 2 gibt einen Überblick zum Stand der Forschung in ausgewählten Fachgebieten. Darauf aufbauend stellt Abschnitt 3 die Zielstellung des Promotionsvorhabens heraus. In Abschnitt 4 werden die methodischen Ansätze zur Entwicklung der Pipeline vorgestellt.

2. Stand der Forschung

In den folgenden Abschnitten werden einige der für dieses Promotionsprojekt relevanten Forschungsfelder betrachtet.

2.1 Integration digitaler Menschmodelle

Um technische Systeme zur digitalen Synthese menschlicher Bewegungen entwickeln zu können, muss zunächst eine digitale Repräsentation des Menschen verfügbar sein. Zahlreiche solcher DMM wurden bereits u. a. im Bereich der Arbeitswissenschaft entwickelt (Bullinger-Hoffmann & Mühlstedt 2016: S. 73-182). Diese besitzen neben allgemeinen Funktionen (z.B. Laden gespeicherter Modelle) jeweils spezialisierte Funktionen zur Manipulation (z.B. Inverse-Kinematics), Analyse (z. B. Haltungsanalysen) und Ausgabe (z.B. Visualisierung) (vgl. Bullinger-Hoffmann & Mühlstedt 2016: S. 81-91). In den Projekten Kinetek (2019) und TISRA (2019) der HTW Dresden wurden bereits einfache DMM zur Erstellung von Avataren anhand anthropometrischer Variablen mit der freien Software MakeHuman-Community (MakeHuman 2019) oder zur Analyse und Visualisierung von Oberflächendeformationen (Loper et al. 2015) einge-

setzt. Problematisch bei der Verwendung verschiedener DMM ist u.a. die unterschiedliche Anzahl und Hierarchie von Gelenkpunkten, die zur Übertragung von Bewegungen genutzt werden können. Hierfür wurde im Projekt Kinetek eine verallgemeinerte Skelettdefinition anhand mehrerer Referenzsysteme erarbeitet. Terlemez et al. (2014) haben mit einer Master Motor Map (MMM) ebenfalls eine Vereinheitlichung verschiedener Systeme erreicht und stellen für diese jeweils Converter bereit.

2.2 Nutzung bestehender Bewegungsdatenbanken

Als Basis für die spätere Bewegungssynthese werden in diesem Promotionsprojekt bestehende Bewegungsdatenbanken genutzt. In diesen sind mannigfaltige Bewegungen enthalten, entsprechend klassifiziert und teilweise annotiert (Mandery et al. 2015). Diese umfassen funktionelle Bewegungen (z.B. Greifbewegung), sportliche Aktivitäten, Interaktionen zwischen mehreren Menschen oder Interaktionen mit Objekten der Umgebung (CMU 2008) (Ionescu et al. 2013). Erfasst werden die Bewegungen mit unterschiedlichen Sensoren (z.B. Motion-Capture, Monochrom-/RGB-Kamera, Tiefenkamera, Inertial Measurement Unit), wobei nicht in allen Datenbanken die gleichen Verfahren zur Datenerhebung zum Einsatz kommen.

In den einzelnen Datenbanken werden die Daten zudem oftmals in unterschiedlichen Dateiformaten und Datenstrukturen bereitgestellt – darunter Videos, 3D-Skelettdaten (Gelenkposition, -rotation, -beschleunigung), Oberflächendaten (Mesh) und weitere Metainformationen. Entsprechend der verwendeten Sensorik unterscheiden sich auch die zugrundeliegenden Topologien der 3D-Skelettdaten (z. B. Anzahl und Hierarchie der Gelenke). Mandery et al. (2015) zeigen hierzu einen schnittstellenbasierten Ansatz zur Vereinheitlichung und Konvertierung zwischen einzelnen Datenbanken.

2.3 Bewegungssynthese

Schwerpunkt des Promotionsprojektes ist die computergestützte Bewegungssynthese. Hierbei werden u. a. zuvor aufgezeichnete Daten automatisiert neu arrangiert oder neuartige Bewegungen erzeugt, die in dieser Form nicht im Datensatz enthalten sind. Methoden zur Synthetisierung menschlicher Bewegungen werden bereits seit den 1990er Jahren entwickelt und erforscht. Zu Beginn konnten Bewegungen mit nur sehr wenigen Freiheitsgraden (DoF) erzeugt werden (Lamouret & van de Panne 1996). Einen Überblick zur geschichtlichen Entwicklung dieses Forschungszweiges sowie eine Kategorisierung der entwickelten Methoden geben Wang et al. (2014) und Gu et al. (2016). Wang et al. (2014) unterscheiden grundsätzlich vier Gruppen methodischer Ansätze: manuelle Methoden, Methoden basierend auf physikalischen Modellen, videobasierte Methoden und datengetriebene Ansätze unter Nutzung von Motion-Capture-Daten. Jede der genannten Gruppen weist dabei spezifische Vor- und Nachteile auf, die bei der Integration solcher Verfahren in technische Systeme zu berücksichtigen sind.

Durch die fortwährend nahezu konstante Leistungssteigerung moderner CPUs und GPUs werden seit einigen Jahren auch KI-basierte Verfahren für die Bewegungssynthese interessant. Hierbei können große Datensätze und komplexe Bewegungsabfolgen von neuronalen Netzen automatisiert analysiert, segmentiert und zu neuen Bewegungen synthetisiert werden. Die Synthese kann dabei aus den zuvor trainierten Netzen in Echtzeit erfolgen, um einen virtuellen Avatar in einer Szene mit komplexen Untergründen zu bewegen (vgl. Holden et al. 2017). Auch Interaktionen mit Objekten in der Umgebung des Avatars wurden bereits in Echtzeit umgesetzt (Starke et al. 2019).

Mit der Zeit wurden sehr viele Architekturen neuronaler Netze für diverse Anwendungsgebiete entwickelt und optimiert (vgl. van Veen & Leijnen 2019). Für die Bewegungssynthese kommen aktuell unter anderem Recurrent Variational Autoencoder (Habibie et al. 2017), Reinforcement Learning (Peng et al. 2018) oder Convolutional Neural Networks (Zhou 2019) zum Einsatz.

3. Konkrete Zielstellung

Primär verfolgt dieses Promotionsvorhaben das Ziel, KI-gestützt menschliche Bewegungsdaten auf Basis realer Datensätze (z.B. aus Motion-Capture-Datenbanken) zu synthetisieren. Dabei sollen Charakteristiken, u.a. der Bewegungsstil, realer Bewegungen erhalten bleiben. Für Beispielanwendungen der Rehabilitationsmedizin und virtuellen Ergonomie sollen funktionelle Bewegungen synthetisiert werden, ohne reale Personen tatsächlich in einer zeit- und kostenaufwändigen Prozedur in großem Maßstab aufzunehmen.

Die Synthese erfolgt dabei in einer zu entwickelnden Verarbeitungspipeline, in der bestehende digitale Menschmodelle integriert und Bewegungsdatenbanken angebunden werden können. Unabhängig vom konkreten Anwendungsbereich sollen parametrisierbare Datensätze mit möglichst großer Varianz (z.B. anthropometrische Variablen) entstehen, um den Bias KI-gestützter Bewegungsanalysen zu reduzieren. Die durch KI-Verfahren generierten Bewegungen sollen zudem für Anwender*innen der Software bzgl. der zugrundeliegenden Parameter jederzeit nachvollziehbar und reproduzierbar sein.

Aus den konkreten Anwendungsfällen soll eine Abstraktion zu generischen Schnittstellen für die Eingabe (z.B. Datenbanken und Sensorik), Analyse und Synthese sowie zur Ausgabe (z.B. 2D-/3D-Visualisierung) menschlicher Bewegungen abgeleitet werden. Damit soll die einfache Adaption der Pipeline für weitere Anwendungsbereiche gewährleistet werden.

4. Entwicklung der Pipeline

Konzeptionell wird die Software-Pipeline in vier Abstraktionsebenen entworfen.

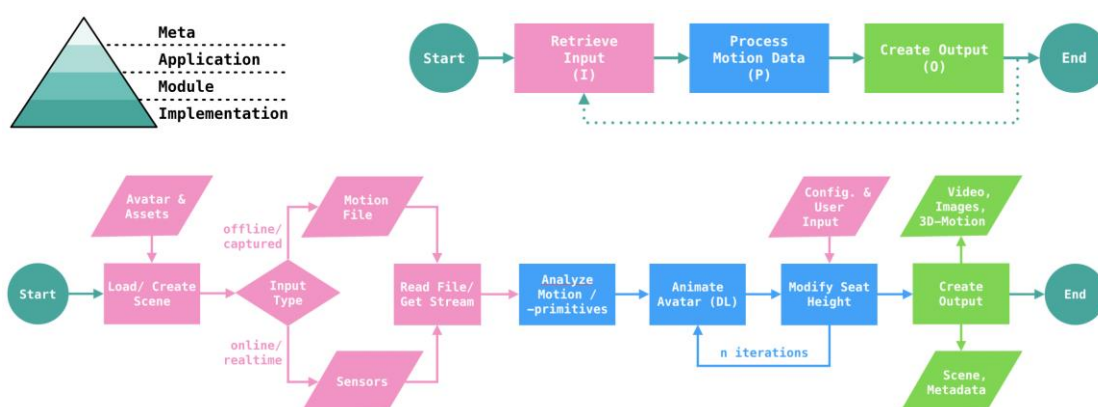


Abbildung 1: Die Abstraktionsebenen (oben links). Die gesamte Pipeline auf der (obersten) Metaebene (oben rechts). Ein Beispiel aus der Rehabilitationsmedizin (Synthese der Sit-to-Stand Übung) auf der Anwendungsebene (unten).

Auf der Ebene der höchsten Abstraktion (Metaebene) definiert sich die Pipeline durch die Verschaltung dreier Komponentenklassen – Eingabe, Verarbeitung und Ausgabe. Auf der darunterliegenden Ebene (Anwendungsebene) erfolgt die Beschreibung anwendungsspezifischer Anforderungen und der dazu benötigten Module je Komponentenklasse. Einzelne Module werden hinsichtlich ihrer Funktionsweise auf der Modulebene beschrieben. Die Ebene mit der geringsten Abstraktion ist die Implementierungsebene. Hier werden die algorithmischen Details einzelner Module entwickelt und dokumentiert. Abbildung 1 stellt die Abstraktionsebenen exemplarisch in der Notation von Flussdiagrammen dar.

Die Implementierung der Pipeline folgt einem iterativen Schema und knüpft an die vorgestellten Abstraktionsebenen an (Abb. 2). In der ersten Iteration werden die Basiskomponenten der softwareinternen Prozesse für die Datenerfassung (I), Verarbeitung und Synthese (P), Ausgabe (O) und Validierung bzw. Analyse (V) definiert. In der zweiten Iteration wird das System um Module erweitert, die anwendungsspezifische Anforderungen umsetzen. Anhand der spezifischen Implementierungen werden in der darauffolgenden Iteration generische Schnittstellen abgeleitet. Nachdem prototypische Pipelines für zwei Anwendungsgebiete umgesetzt wurden und aus diesen generische Schnittstellen abgeleitet werden konnten, werden die einzelnen Module knotenbasiert implementiert. Hierbei hat jedes Modul definierte Eingangsdaten und -parameter, eine Verarbeitungsroutine und definierte Ausgangsdaten. Somit können Module von Anwender*innen einfach und sicher zu komplexeren Einheiten verknüpft werden. Bei der Entwicklung soll zudem darauf geachtet werden, bereits etablierte Standards einzuhalten und bevorzugt freie Software (z. B. Bibliotheken) einzubinden. Die zu entwickelnde Software-Pipeline soll unter einer Open-Source Lizenz frei verfügbar sein und somit die Adaption durch Dritte erleichtern.

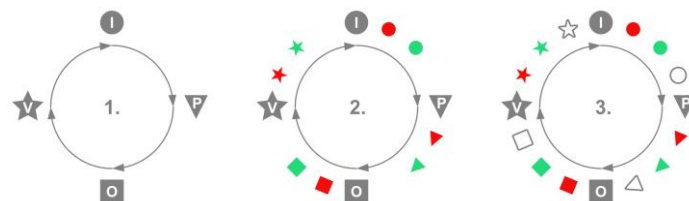


Abbildung 2: Entwicklung der Software-Pipeline in drei Iterationen (1. Basiskomponenten, 2. anwendungsspezifische Module, 3. generische Schnittstellen).

Mit erfolgreicher Umsetzung der Software-Pipeline wird Anwender*innen ein Werkzeug zur Verfügung stehen, mit dem sie diverse Bewegungsdatensätze mit Hilfe datengetriebener Verfahren parametrisiert synthetisieren können. Dazu können sie die Pipeline ihren Anforderungen entsprechend aus Modulen zusammensetzen, oder die generischen Schnittstellen nutzen, um der Pipeline weitere Module hinzuzufügen.

5. Literatur

- Bullinger-Hoffmann, A. C., & Mühlstedt, J. (Eds.). (2016). Homo Sapiens Digitalis - Virtuelle Ergonomie und digitale Menschmodelle. Springer Berlin Heidelberg.
- CMU Carnegie Mellon University (2008). [Online]. CMU Graphics Lab Motion Capture Database. <http://mocap.cs.cmu.edu/>. letzter Zugriff: 07.01.2020
- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumeé III, H., & Crawford, K. (2018). Datasheets for datasets. arXiv preprint arXiv:1803.09010.

- Gu, P., Zheng, Y., Wang, D., & Wu, Q. (2016, August). A review of human motion synthesis and motion retrieval technology. In 2016 11th International Conference on Computer Science & Education (IC-CSE) (pp. 976-980). IEEE.
- Habibie, I., Holden, D., Schwarz, J., Yearsley, J., & Komura, T. (2017). A Recurrent Variational Autoencoder for Human Motion Synthesis. In BMVC.
- Holden, D., Komura, T., & Saito, J. (2017). Phase-functioned neural networks for character control. ACM Transactions on Graphics (TOG), 36(4), 42.
- Ionescu, C., Papava, D., Olaru, V., & Sminchisescu, C. (2013). Human3.6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. IEEE transactions on pattern analysis and machine intelligence, 36(7), 1325-1339.
- Kinetek – Programm und Messsystem zur Bewegungserfassung und Trainingstherapie von Parkinson Patienten (2019). [Online]. <http://www.drematrix.de/projects/software-and-sensor-system-for-the-motion-detectio/>. letzter Zugriff: 07.01.2020
- Krüger, V., Kragic, D., Ude, A., & Geib, C. (2007). The meaning of action: A review on action recognition and mapping. Advanced robotics, 21(13), 1473-1501.
- Lamouret, A., & van de Panne, M. (1996). Motion synthesis by example. In Computer Animation and Simulation'96 (pp. 199-212). Springer, Vienna.
- Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., & Black, M. J. (2015). SMPL: A skinned multi-person linear model. ACM transactions on graphics (TOG), 34(6), 248.
- MakeHuman (2019). [Online]. MakeHuman Community: Open Source tool for making 3D characters. <http://makehumancommunity.org>. letzter Zugriff: 07.01.2020
- Mandery, C., Terlemez, Ö., Do, M., Vahrenkamp, N., & Asfour, T. (2015, July). The KIT whole-body human motion database. In 2015 International Conference on Advanced Robotics (ICAR) (pp. 329-336). IEEE.
- Millington, P. J., Myklebust, B. M., & Shambes, G. M. (1992). Biomechanical analysis of the sit-to-stand motion in elderly persons. Archives of Physical Medicine and Rehabilitation, 73(7), 609-617.
- Mühlstedt, J. (2016). Virtuelle Ergonomie. In Homo Sapiens Digitalis-Virtuelle Ergonomie und digitale Menschmodelle (pp. 183-225). Springer Vieweg, Berlin, Heidelberg.
- Peng, X. B., Kanazawa, A., Malik, J., Abbeel, P., & Levine, S. (2018, December). Sfv: Reinforcement learning of physical skills from videos. In SIGGRAPH Asia 2018 Technical Papers (p. 178). ACM.
- Shankar, S., Halpern, Y., Breck, E., Atwood, J., Wilson, J., & Sculley, D. (2017). No classification without representation: Assessing geodiversity issues in open data sets for the developing world. arXiv preprint arXiv:1711.08536.
- Starke, S., Zhang, H., Komura, T., & Saito, J. (2019). Neural state machine for character-scene interactions. ACM Transactions on Graphics (TOG), 38(6), 1-14.
- Terlemez, Ö., Ulbrich, S., Mandery, C., Do, M., Vahrenkamp, N., & Asfour, T. (2014, November). Master Motor Map (MMM)—Framework and toolkit for capturing, representing, and reproducing human motion on humanoid robots. In 2014 IEEE-RAS International Conference on Humanoid Robots (pp. 894-901). IEEE.
- TISRA – Nachwuchsforschergruppe (2019). [Online]. <http://www.drematrix.de/projects/tisra-junior-researcher-group/>. letzter Zugriff: 07.01.2020
- Van Veen, F., & Leijnen, S. (2019). [Online]. The Neural Network Zoo. <https://www.asimovinstitute.org/neural-network-zoo/>. letzter Zugriff: 07.01.2020
- Wang, X., Chen, Q., & Wang, W. (2014). 3D human motion editing and synthesis: A survey. Computational and Mathematical methods in medicine, 2014.
- Zhou, D., Feng, X., Yi, P., Yang, X., Zhang, Q., Wei, X., & Yang, D. (2019). 3D Human Motion Synthesis Based on Convolutional Neural Network. IEEE Access.

Danksagung: Dieser Beitrag wäre ohne die Unterstützung durch die Europäische Union und das Land Sachsen (FKZ 100380879) nicht möglich gewesen (Abb. 3). Ein ganz besonderer Dank gilt Loreen Pogrzeba für die konstruktive Unterstützung bei der Konkretisierung der Projektschwerpunkte.



Europäische Union

Europa fördert Sachsen.



Europäischer Sozialfonds



Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des vom Sächsischen Landtag beschlossenen Haushaltes.

Abbildung 3: Förderung (Stipendium) durch ESF/SAB für eine Landesinnovationspromotion.



Gesellschaft für
Arbeitswissenschaft e.V.

Digitale Arbeit, digitaler Wandel, digitaler Mensch?

66. Kongress der
Gesellschaft für Arbeitswissenschaft

TU Berlin
Fachgebiet Mensch-Maschine-Systeme

HU Berlin
Professur Ingenieurpsychologie

16. – 18. März 2020, Berlin

GfA-Press

Bericht zum 66. Arbeitswissenschaftlichen Kongress vom 16. – 18. März 2020

**TU Berlin, Fachgebiet Mensch-Maschine-Systeme
HU Berlin, Professur Ingenieurpsychologie**

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.
Dortmund: GfA-Press, 2020
ISBN 978-3-936804-27-0

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

© 2020 fröse multimedia, Frank Fröse

office@internetkundenservice.de · www.internetkundenservice.de