

# Datenströme und ihre Visualisierung – KI führt Nutzer durch relevante Analysen

**Die Digitalisierung im Gesundheitswesen führt zu wachsenden medizinischen Datenströmen. Diese bergen werthaltige Informationen für Forschung und Entwicklung. Die heute schon verfügbaren Visualisierungstools ermöglichen beeindruckende Einsichten in den Datenbestand. Ein nächster Schritt ist die statistische Analyse dieser Datenströme, die sich auf Künstliche Intelligenz (KI) stützt und dabei neue Hypothesen generiert für weiterführende wissenschaftliche Untersuchungen, die auf dem klassischen Weg nicht formuliert worden wären. Wie das aussehen könnte, erläutern die Autoren in Folge III ihres Beitrags zur Wertschöpfungskette „Medizinische Daten“.**

| Dr. Joachim Linke, Dr. Frank Schippers, Florian Stöwe

## 1. Ein- und Ausblick

Die zunehmende Digitalisierung im Gesundheitswesen führt zu wachsenden medizinischen Datenströmen. Diese bergen werthaltige Informationen für Forschung und Entwicklung ebenso wie für eine Verbesserung der Diagnose, Therapie und Prävention. Eine strukturierte Aufbereitung, Analyse und Visualisierung des relevanten Inhalts dieser Datenströme erfordert ein interdisziplinäres Vorgehen zwischen Ärzten, Datenmanagern, Statistikern und Softwareentwicklern.

Die zur Verfügung stehenden Visualisierungstools ermöglichen beeindruckende Einsichten in komplex zusammengesetzte medizinische Datenbestände und Datenströme. Und sie erlauben die weltweite Real-time Diskussion der aus der Visualisierung neu gewonnenen Einsichten in einem virtuellen Raum mit anderen Teilnehmern.

Diese Visualisierung ist allerdings nur ein erster Schritt zur Auswertung der Daten. Denn es ergibt sich daraus insbesondere die Frage, ob visuelle Auffälligkeiten (z. B. Outlier oder Trends) auch

tatsächlich einem kausalen Zusammenhang unterliegen und als solche bewiesen werden können oder ob es sich lediglich um Zufallserscheinungen handelt, die keinen weiteren Erkenntnisgewinn erbringen.

Ein möglicher nächster Schritt könnte daher die Verwendung statistischer Analysemethoden sein, die sich auf Künstliche Intelligenz (KI; Artificial Intelligence – AI) stützen. Hierdurch wären fundierte – weil statistisch untermauerte – Aussagen hinsichtlich z. B. Ereignisvorhersagen, Charakterisierungen von Clustern, Zeitreihenanalysen bis hin zu konfirmatorischen statistischen – auch multivariaten – Analysen möglich.

KI-gestützte neue Hypothesengenerierung für weiterführende wissenschaftliche Untersuchungen, die auf dem klassischen Weg nicht formuliert worden wären, könnten dann eine wesentliche Ergänzung des klassisch-wissenschaftlichen Ansatzes darstellen, bei dem zunächst eine wissenschaftliche Frage formuliert und diese dann entsprechend untersucht wird. Doch tatsächlich wird häufig immer noch ein möglicher Erkenntnisgewinn aus einer datengetriebenen Analyse über den gesamten Datenpool für die Hypo-

thesengenerierung außer Acht gelassen.

Es wäre naheliegend, dass die zur Verfügung stehenden Tools zur Datenvisualisierung (3D – dreidimensional/VR – Virtual Reality/AR – Augmented Reality/MR – Mixed Reality) mit den notwendigen, KI-unterstützten statistischen Analysemethoden zusammenwachsen. Die sehr hilfreiche 3D/VR/AR/MR-Betrachtung des Datenpools würde so um eine valide – weil statistisch unterlegte und ggf. bewiesene – Entscheidungsgrundlage ergänzt.

## 2. Einführung

Obwohl Deutschland hinsichtlich der Digitalisierung des Gesundheitssektors im europäischen Vergleich eher schlecht dasteht [1], führt der auch durch verschiedene deutsche und europäische Gesetzesinitiativen geförderte Digitalisierungsprozess im Gesundheitswesen konsequent bereits heute zu exponentiell wachsenden medizinischen Datenströmen. Diese Datenströme können mittels einer sinnvoll etablierten Wertschöpfungskette „Medizinische Daten“ [2] beherrscht und somit genutzt werden. Natur-

lich spielt hier auch das wachsende Bewusstsein der Ärzteschaft für das Thema der Werthaltigkeit und Nutzung medizinischer Datenströme eine große Rolle [3].

Der wesentliche Erfolgsfaktor für die Erstellung und Nutzung einer Wertschöpfungskette „Medizinische Daten“ ist letztlich die Kombination aus high-end IT (Informationstechnik) zusammen mit spezialisierten medizinisch-wissenschaftlich ausgebildeten Ärzten, die neben den medizinischen Kenntnissen auch das nötige technische sowie auch statistische Verständnis mitbringen. Der Endanwender allerdings benötigt diese „Architektur“-Spezialkenntnisse nicht, sondern soll sich auf die Bewertung und Nutzung der verständlich, am besten grafisch visuell dargestellten Informationen konzentrieren können. Die strikte Einhaltung der jeweils gültigen Datenschutzverordnung muss hierbei selbstverständlich eine *conditio sine qua non* sein.

Schnell verfügbare, aber zugleich belastbare – weil validierte – Analyseergebnisse als Grundlage für Managemententscheidungen erfordern im ersten Schritt zunächst eine plausible Datenbasis. Diese validierten Rohdaten müssen auf einer kompatiblen Plattform vereint werden, mit hochmodernen Techniken automatisiert ausgewertet und in

der Folge verständlich dargestellt, also visualisiert werden, so dass auch Nicht-Statistiker schnell die wesentlichen Aussagen der Analysen erfassen können. Dies beschreibt allerdings nur ein sogenanntes hypothesen-getriebenes Vorgehen (die Daten werden mit Fokus auf eine a priori definierte Hypothese ausgewertet) – ein klassisches Vorgehen z.B. in der klinischen Forschung.

Demgegenüber ist eine datengetriebene Datenauswertung dadurch gekennzeichnet, dass verschiedene statistische Verfahren (z.B. Cluster-, Varianz-, Zeitreihenanalysen) ohne vorherige Hypothesenbildung am Datensatz mit x-beliebigen Parameterkombinationen durchgeführt werden („Daten/Informationen schürfen“/„Data Mining“). Daraus können sich wertvolle Analyseergebnisse ergeben, die möglicherweise vorher so nicht bekannt waren und nun auch für eine neue Hypothesenbildung als Grundlage zukünftiger wissenschaftlicher Untersuchungen dienen können.

Letztlich müssen die Analyseergebnisse verständlich und leicht erfassbar sein – und im besten Fall im Team im virtuellen Raum interaktiv visualisiert werden können. Der gemeinsame Besuch sogenannter „virtueller Räume“ (entweder mit VR-Brille oder über AR-Darstellung

auf dem „handheld“ Gerät) ist sicher noch nicht als Standard etabliert, allerdings erschließt sich der Nutzen einer interaktiven Teamdiskussion im dreidimensionalen Raum um eine bewegte 3D-Graphik herum beim ersten Besuch eines virtuellen Raumes sehr schnell.

Bei der Präsentation der Ergebnisse wird in zunehmendem Maße auch künstliche Intelligenz (KI / AI) unterstützen, welche sich an den jeweiligen Daten orientiert und ggf. sinnvolle, zusätzliche Analysen gleich automatisch mit darstellt, mit früheren Erkenntnissen abgleicht und als weiterführende Diskussionsgrundlage anbietet.

Insgesamt verkürzen sich die Entwicklungszyklen moderner Software zur Auswertung und Visualisierung immer schneller, so dass es zunehmend wichtiger wird, die Validität und den Nutzen der zur Verfügung stehenden Programme zu überprüfen. Zusätzlich gilt es auch zu bewerten, ob es sich tatsächlich lohnt, immer sofort das modernste Tool zur Visualisierung zu nutzen, nur um hinterher festzustellen, dass es auch in diesem Bereich immer wieder „Kinderkrankheiten“ der ersten Versionen gibt. Gerade bei Diagnose und Therapieentscheidungen auf der Basis von medizinischen Daten, die mit KI-Hilfe ausgewertet wurden, stehen Verlässlichkeit, Wirksam-

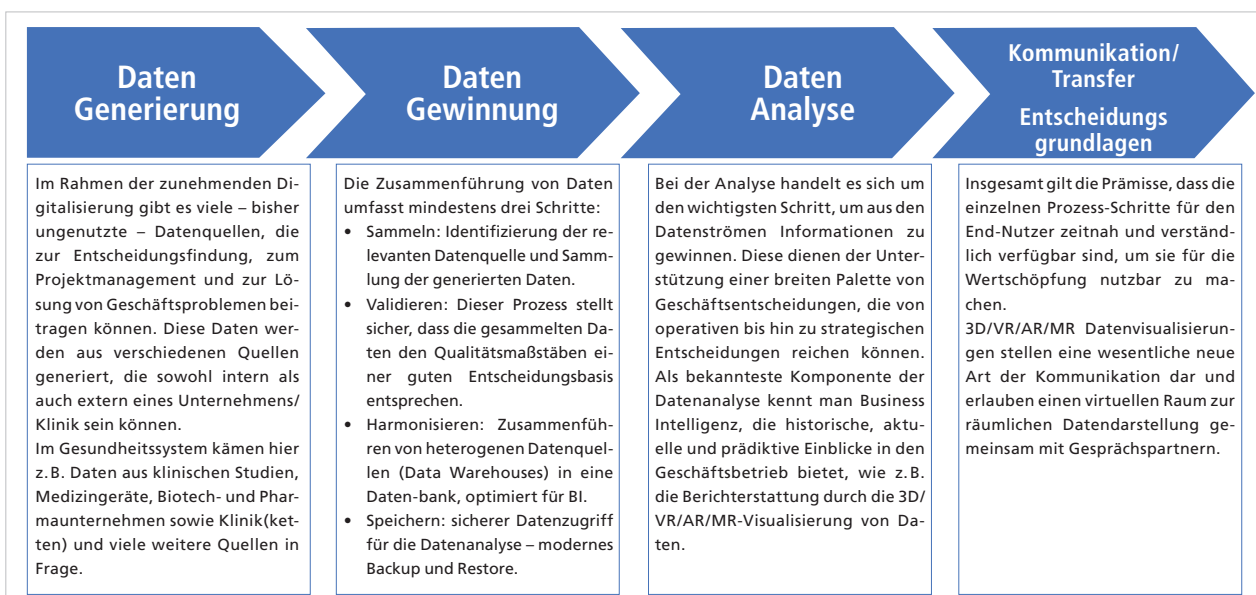


Abbildung 1: Die vier Säulen der Wertschöpfungskette „Medizinische Daten“ (”Medical Data Value Creation Chain“) [CCR GmbH]

keit und Sicherheit an oberster Stelle. Dies muss auch für die eingesetzte Software zu Gewinnung, Analyse und Visualisierung von Daten gelten, wenn diese Software am Ende maßgebliche Entscheidungsgrundlagen erstellt. Beispielsweise fordert das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) ja auch für digitale Gesundheitsanwendungen (DiGAs) den Nachweis von Wirksamkeit und Sicherheit in ähnlicher Form, wie bei der Medikamentenentwicklung [4][5][6].

Die Digitalisierung im Gesundheitswesen ist unumkehrbar und auch von gesundheitspolitischer Seite gewollt – daher wird sie an Geschwindigkeit gewinnen. Mit den daraus zwangsläufig entstehenden medizinischen Datenströmen in den unterschiedlichen medizinischen Disziplinen ergeben sich aufgrund der wachsenden Vielzahl der Möglichkeiten natürlich auch ungeahnte Möglichkeiten und Chancen für alle Teilnehmer im Gesundheitssystem. Bei aller Innovationsbegeisterung sollte dabei bedacht werden, dass es Menschen geben muss, die mit diesen neuen Techniken und Technologien verantwortungsvoll umgehen können und wollen; zudem darf nicht vergessen werden, dass zu diesem Nutzerkreis auch zunehmend ältere Patienten gehören werden.

Vor dem Hintergrund der durch die forcierte Digitalisierung bedingten stetig anwachsenden Datenströme lohnt sich also eine detaillierte Betrachtung der sich schnell verändernden aktuellen Möglichkeiten ebenso, wie die Beteiligung an der Diskussion aktueller und zukünftiger Entwicklungen der Erhebung, Visualisierung (VR/AR/AI) und Auswertung (KI/AI) solcher medizinischer Daten.

Bei der Überlegung, wie man in den wachsenden medizinischen Datenströmen die sogenannten "Hidden Beauties" entdecken und nutzen kann, könnte die bereits genannte Wertschöpfungskette „Medizinische Daten“ [2] eine geeignete Hilfestellung zu sein.

Alle vier Säulen der Wertschöpfungskette „Medizinische Daten“

[2] stellen sich aktuell bereits in technisch sehr weit entwickelter Umsetzung dar. Insbesondere bei der Datenvisualisierung zeigen sich überzeugende, aber auch zugleich intuitiv einfach bedienbare Ansätze, wie z. B. Teamwork in Virtual/Augmented Reality (VR/AR – <https://www.youtube.com/shorts/7FN6cdMZnvM>). Die Darstellung von graphischen Analyse-Ergebnissen in einer virtuellen Umgebung und das gemeinsame Teilen mit anderen Teammitgliedern in dieser VR-Welt ist also nicht etwa die Zukunft, sondern bereits etablierte Realität.

Welche Entwicklungen der nächsten Generation können wir also insbesondere im Bereich der Visualisierung und Analyse der medizinischen Datenströme als sinnvolle Ergänzungen erwarten?

Bevor wir uns dieser Frage zuwenden, lohnt sich ein Blick auf einige aktuell zur Verfügung stehende Möglichkeiten der Datenvisualisierung.

### 3. Beispiele aktueller Datenvisualisierung (3D/VR/AR/MR) [7][8][9][10]

#### 3.1. Tableau®

Tableau ist eine US-amerikanische interaktive Datenvisualisierungssoftware, die sich überwiegend auf die tabellenorientierte und zusätzlich grafische Datendarstellung konzentriert, um die Beziehung zwischen verschiedenen Variablen zu visualisieren. Die Software bietet Desktop-, Online-, Cloud- und weitere Anwenderumgebungen. Tableau ermöglicht den Import verschiedenster Datentypen wie Excel, SQL, Access, JSON und viele mehr – so ist diese Software im Bereich des Datenimports (Tableau Prep Builder) sicher eine der leistungsfähigsten Applikationen.

Bezüglich statistischer Auswertungen kann die Software zusätzlich zu Trendlinien grundlegende statistische Parameter (Mittelwert, Modus, Anzahl und andere) berechnen und sie auf Diagrammen darstellen. Darüber hinaus können

umfangreiche weitere Optionen verwendet werden, um die Daten zu visualisieren (z. B. Streudiagramme, Balken- und Tortendiagramme, Tabellen, Karten, Blasen und viele andere), wobei mehrere Dimensionen gleichzeitig mit verschiedenen Farben, Größen, Achsen, Formen und Etiketten dargestellt werden können.

Diagramme können in einem sogenannten Dashboard verwendet werden oder um eine Präsentation (Story) mit mehreren Dashboards vorzubereiten. Jedes Dashboard hat in der Regel mehrere Plots, die miteinander interagieren können (z. B. gegenseitige interaktive Filterung). Das Filtern der Ergebnisse, um ein bestimmtes Datenelement darzustellen, kann vom Benutzer mühelos bedient werden, um beispielsweise einen bestimmten Trend oder relevante Untergruppen zu untersuchen.

Auch die moderne Serverarchitektur und die sehr einfach bedienbare Lizenzverwaltung sind Gründe für die starke Marktposition von Tableau im Bereich der tabellari-schen und graphischen Datenvisualisierung mit fast 18 Prozent Marktanteil [11].

#### 3.2. Immersion Analytics Visualizer® [12]

Visualizer ist ebenfalls ein US-amerikanisches Produkt für eine neue Art der 3D/VR/AR Datenvisualisierung. Die Software hat verschiedene verfügbare Modi wie die serverbasierte Version sowie Desktop-, Mobil- und VR-Versionen. Der grundsätzliche Ansatz dieser Software besteht darin, die Daten in einer 3D-Ansicht (Cubes) mit 18 Dimensionen darzustellen wie u. a. Größe, Farbe, drei verschiedenen Achsen, Glühen, bewegliche Satelliten, Vibrationen, verschiedene Formen. Die Software ermöglicht es dem Benutzer, mit den Daten zu interagieren, sie nach 18 Dimensionen zu filtern, die Details jedes Datenpunkts zu untersuchen und eine aufschlussreiche Ansicht der möglichen Muster zu erhalten, die von den Daten aus verschiedenen Blickwinkeln generiert werden.

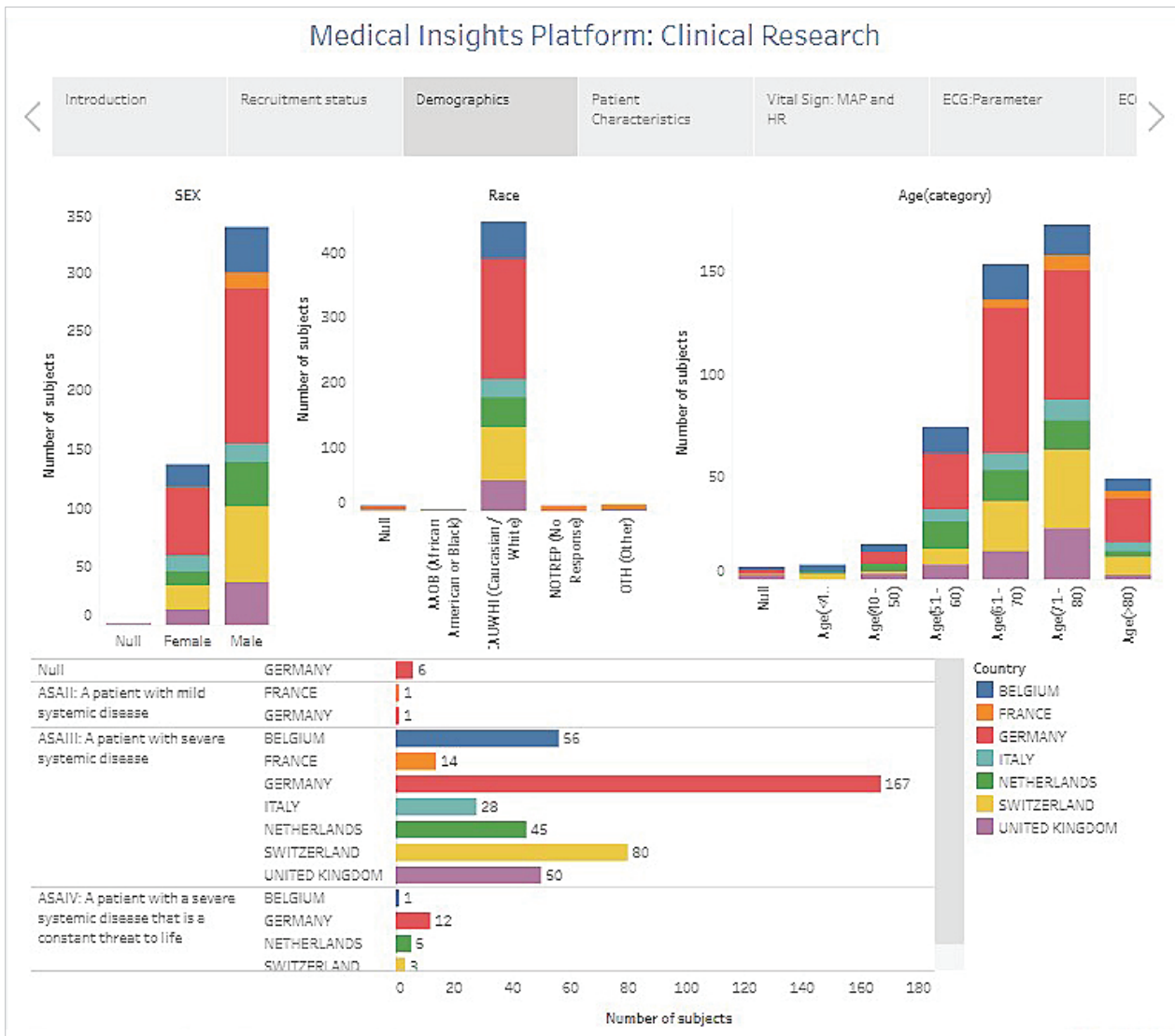


Abbildung 2: Tableau®-Beispiel mit fiktiven Daten – Statistics on the Fly [CCR GmbH]

Mehrere Cubes können mit Hilfe sogenannter Scatterplots oder Bars gleichzeitig dargestellt werden, um mehrere Variablen zu visualisieren und Muster, die sich ergeben können, zu vergleichen. Die Software verfügt über Erweiterungen für Tableau, Qlik Sense, Python und MATLAB, die eine Interaktion zwischen diesen Software Paketen ermöglichen, um die Daten in 3D-Ansicht darzustellen. Allerdings ist die Positionierung der Kamera während der Präsentation bei mehreren Cubes immer noch problematisch und die Anwendung aller 18 Dimensionen macht die Präsentation manchmal komplex und schwer verständlich.

### 3.3. Flow Immersive®

Flow Immersive ist eine server- und webbasierte Software, die sich

mit 3D/VR/MR-Datenvisualisierung befasst. Die Software importiert CSV-Dateien und exportiert Ergebnisse als teilbare Links, Videos oder webbasierte interaktive Präsentationen. Jeder Wert in der Excel-Tabelle wird als Punkt oder Kugel in einer 3D-Ansicht mit mehreren Achsen und Dimensionen dargestellt, einschließlich vieler erweiterter Funktionen wie Kugelgröße, -Farbe, -struktur, Filter, Interaction und Animation.

Die Benutzeroberfläche für die Entwicklung einer neuen Visualisierung ist einfach bedienbar und ermöglicht eine schnelle und einfache Modifizierung der Achsen und Beschriftungen. Texte, Bilder und Bewegungen sowie andere Bedienelemente können leicht hinzugefügt werden. Die genannten Funktionen interagieren miteinander und erge-

ben am Ende eine überzeugende, animierte und farblich untermalte 3D-Präsentation, die aus mehreren Folien bestehen kann. Unter dem Titel "Additional Dimensions" bietet Flow Immersive die Möglichkeit, die Daten mit mehr als nur drei (X, Y, Z) Dimensionen darzustellen. So können auch mehrere Variable gleichzeitig in einem Diagramm dargestellt werden.

Statistische Analysemöglichkeiten sind aktuell noch nicht integriert und die Steuerung der Kamera während einer Präsentation ist gelegentlich eingeschränkt, wenn mehr als ein Diagramm gezeigt wird. Allerdings bietet Flow Immersive dem Benutzer eine 3D/VR/AR Plattform zum Teilen, Präsentieren und Erklären der Daten für das Publikum. Teilnehmer können den Raum mit

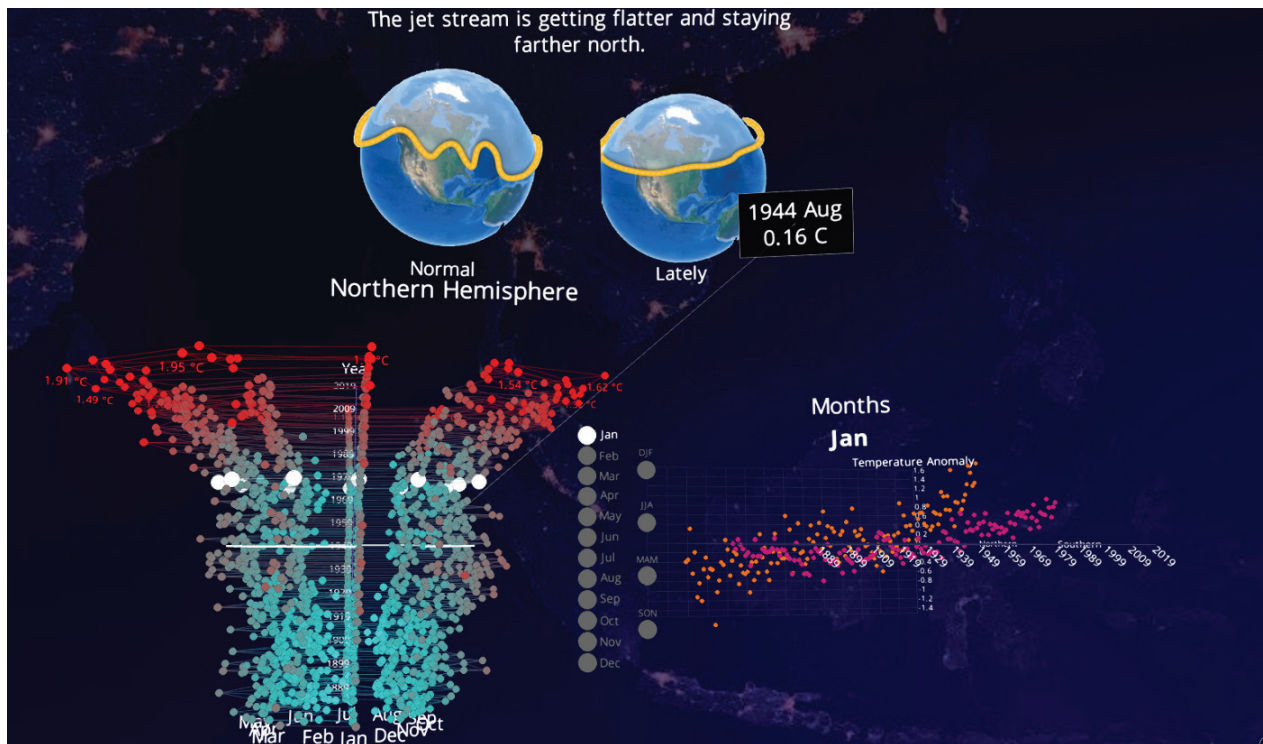


Abbildung 3: Beispiel einer 3D-Visualisierung von Flow Immersive, die als 3D und AR und auch als AR-Darstellung möglich ist [Flow Immersive: Temperature anomalies Story (<https://a.flow.gl/flow/llred3/display>)]

Handy, Laptop oder einem anderen Gerät ohne großen Aufwand betreten – ein üblicher Internetbrowser genügt.

### 3.4. DeltaMaster®

DeltaMaster ist eine Desktop-App mit einer serverbasierten Plattform zur hierarchischen Visualisierung der Daten. Der Hauptvorteil und das Alleinstellungsmerkmal dieser Software sind die integrierten "Data Mining" Möglichkeiten und eine bewegliche, hierarchische Dimensionsdarstellung – der sogenannte "Hyperbrowser". Der Hyperbrowser verwendet eine interaktive Hierarchie, um die Daten mit einem zentralen „Dreh- und Angelpunkt“ darzustellen. Die jeweils multidimensionale Darstellung eines bestimmten Parameters (z. B. Körpergröße) kann durch einfaches Drag & Drop verändert werden. Mit dieser Methode können Millionen von Daten buchstäblich mit der Fingerspitze in Detailbereiche bewegt und analysiert werden. Zusätzlich bietet die Software die Darstellung der grundlegenden statistischen Parameter der deskriptiven Analyse (z. B. Mittelwert, Standardabweichung, etc.).

Grundsätzlich erfolgt die Datenvisualisierung entweder als Tabelle oder mit dem oben erwähnten Hyperbrowser. Bei der farblichen Kennzeichnung von Daten werden ein Blau- und ein Rot-Ton verwendet. Insgesamt ist das Programm im Vergleich zu anderen Präsentationsoptionen effektiv bei der Präsentation von "Big Data" und fasst deren Hauptmerkmale auf einfache und leicht verständliche Weise zusammen – zusätzlich unterstützt durch die Vielzahl der statistischen Analysen.

*ANMERKUNG: Einige der o.g. Visualisierungsbeispiele können auf dem CCR YouTube Kanal betrachtet werden: [https://www.youtube.com/channel/UCK\\_72LwsiuellmnCAEWRhg/videos](https://www.youtube.com/channel/UCK_72LwsiuellmnCAEWRhg/videos) Weitere Beispiele finden sich auf der CCR Insight Plattform <https://mip.ccr-cro.com/#/projects/17> (Benutzername: `ccr-demo-1` Kennwort: `&@4CPyWz%3hb`).*

### 3.5. Vergleich der Visualisierungstools

Eine gelungene Datenvisualisierung erfordert a priori zunächst verschiedene Überlegungen, wie

z. B. technische Faktoren, Datentyp, Datenformat, Datenkomplexität, Verfügbarkeit statistischer Fragestellungen und mögliche Darstellungsvarianten (Linien, Balken, Punktwolken, etc.). Darüber hinaus muss natürlich auch der rote Faden für den Betrachter/Zuhörer erkennbar werden – also die Story als solches mit ihren zentralen Aussagen.

Es gibt eine Vielzahl an Programmen, die sich mit Datenvisualisierung beschäftigen, aber in diesem Beitrag fokussieren die Autoren auf diejenigen Programme, mit denen sie selbst bereits Erfahrungen sammeln konnten, wie z. B. Excel, Tableau, Visualizer, Flow Immersive, DeltaMaster.

Die vorgenannten Applikationen erfüllen einzelne Ansprüche und Anforderungen schon sehr gut, aber jede Software hat sozusagen einen speziellen Fokus und löst damit jeweils fokussiert spezielle Problemstellungen (siehe Tabelle 1).

Visualizer und Flow Immersive ermöglichen eine 3D/AR/VR Datendarstellung; Excel, Tableau und DeltaMaster sind 2D-Programme. Die als Eingabe für die Programme verwendeten Datentypen variieren je nach Software, aber alle unterstützen ei-

Charakteristiken	Excel®	Tableau®	Visualizer®	Flow Immersive®	DeltaMaster®
<b>Fokus</b>	tabellarische Zahlendarstellung (deskriptive und konfirmatorische Statistik eingeschränkt integrierbar)	flexible, interaktive und serverbasierte (multi-user) tabellarische Zahlendarstellung (deskriptive und konfirmatorische Statistik eingeschränkt integrierbar)	3D/VR Darstellung bis zu 18 Dimensionen gleichzeitig  Interaktive Verbindung zu Tableau®	Story Telling, interaktive 3D/VR/AR Darstellung der Daten als Präsentation mit einzelnen Schritten	Übersicht 2D Daten als bewegter, hierarchischer Baum mit „Data Mining“ Fokus (Tabellarische Darstellung, Cluster, Trends, etc.)
<b>Importierbare Datentypen</b>	Exel, CSV	gängige Industriestandards, z. B. auch SAS Dateien importierbar	CSV und interaktive Verbindung zu Tableau®	CSV	Excel, SQL, Access, u. a.
<b>Daten Limitierung</b>	keine	keine	50.000 Rows	10 Megabyte	keine
<b>Räumliche Darstellung</b>	2D	2D	3D/VR	3D/VR/AR	2D
<b>Zahl der Präsentations-einheiten auf einer Folie</b>	viele	viele	limitiert; Einschränkung der Interpretierbarkeit bei max. 18 Dimensionen	viele	viele
<b>Anzahl der Dimensionen (Subgruppen)</b>	2-3	4-6	18	viele	2-3
<b>Verfügbare Grafik-darstellungen (Punkte, Linien, Kugeln, etc.)</b>	limitiert	viele	eine	viele	eine
<b>Verfügbare Farben</b>	limitiert	viele	limitiert	viele	zwei
<b>Statistische Analyse</b>	ja	ja	nein	nein	ja
<b>Cluster Analyse/ "Data Mining"</b>	nein	nein	nein	nein	ja
<b>Interaktion mit den Daten</b>	limitiert	gut	sehr gut	am besten	sehr gut
<b>Animation</b>	nein	sehr limitiert	limitiert	ja	nein
<b>Virtuelle Konferenz (AR und VR)</b>	nein	nein	limitiert	am besten	nein
<b>Teilbarer Link der virtuellen Konferenz</b>	nein	ja	nein	ja	nein
<b>Video-Aufnahmen möglich</b>	nein	nein	nein	ja	nein
<b>Analyse-/ Visualisierungsausgabe</b>	Exel, Doc, PDF, CSV	TWB (proprietär)	VIS (proprietär)	online Videos oder interaktive Shows	DAS (proprietär)
<b>Hardware-Aspekte</b>	Standard Desktop	High End Server/High Resource Use	Standard Desktop + Graphical Processing Units (GPU)	Standard Desktop, Browser Based	High End Server/High Resource Use
<b>Serverarchitektur</b>	On-Premises/Cloud	On-Premises/Cloud/ Linux/Windows	On-Premises/ Windows	Cloud/Linux	On-Premises/ Windows
<b>Lizenzmanagement</b>	User CAL, Device CAL, VL	Floating Licensing, User CAL, Core based	auf Anfrage	User CAL	Floating Licensing

**Abkürzungen:**  
CAL – Client Access License  
2D/3D – zwei- / drei-dimensional  
VR – Virtual Reality  
AR – Augmented Reality  
GPU – Graphical Processing Units  
VL – Volume Licensing  
Access, CSV, DAS, Doc, Exel, PDF, SQL, TWB, VIS – Datei- und Datenbank-Formate

Tabelle 1: Vergleich der untersuchten Programme zur Datenvisualisierung

nen CSV-Datenimport. Tableau bietet darüber hinaus allerdings eine sehr umfangreiche Auswahl an weiteren Datenformaten an, welche importiert werden können.

Visualizer (5.000 Datenzeilen) und Flow Immersive (10 MB Dateigröße) sind bezüglich Datensatzanzahl bzw. Dateigröße limitiert. Ausgabeformate sind zwischen den Applikationen unterschiedlich bzw. proprietär – Flow Immersive ermöglicht aber die Aufzeichnung von Videos.

Von den untersuchten Applikationen verfügt Flow Immersive als 3D-Software über ein erstaunliches Portfolio hinsichtlich Präsentationstypen, Diagrammen, Farben sowie die Anzahl der Dimensionen in einem Diagramm. Darüber hinaus bietet Flow Immersive spezielle Funktionen wie die Interaktion mit Daten, Videos und Animationen, die ein neues virtuelles Erlebnis für den Benutzer schaffen. Interaktion mit Daten bedeutet die Möglichkeit, die Daten in mehreren Ansichten anzuzeigen und mit anderen Sitzungsteilnehmern einen bestimmten Datenbereich auszuwählen und die relevanten Details anzuzeigen oder eben zu „begehen“. Insbesondere die VR- und AR-Datenvisualisierung erlaubt dem Betrachter bei Flow Immersive und Visualizer das besondere Erlebnis „in eine Graphik hinein zu gehen und sich umzusehen“.

Entscheidend bei der Datenvisualisierung ist aus Sicht der Autoren die Möglichkeit, verschiedene Muster gleichzeitig zu erkennen und das Merkmal auszuwählen, auf das sich der Nutzer konzentrieren möchte. Weiterhin stellen einige dieser 3D-Programme bestimmte Datencharakteristika als animierte Merkmale dar, um sie deutlicher hervorzuheben und die auffälligen Muster leichter zu erkennen.

Die Datenvisualisierung dient im ersten Schritt dazu, sich ein Bild der Datenlandschaft und deren Auffälligkeiten zu machen. Konsekutiv ergibt sich aber dann zwangsläufig die Forderung nach deskriptiver und auch konfirmatorischer Statistik, um durch letztere den Beleg zu

erbringen, dass Auffälligkeiten tatsächlich signifikant sind und nicht nur alleine im Auge des Betrachters liegen oder rein zufällig entstanden sind. Der Anspruch an Datenvisualisierung kann daher nicht nur die reine Visualisierung von Datenrohwerten oder Aggregatwerten sein, sondern muss auch deskriptive Statistik (z. B. Mittelwert und Standardabweichungen, etc.) und im besten Fall ebenso konfirmatorische Statistik (mit Signifikanzwerten) und multivariate Analyseansätze wie z. B. MAN(CO)VA sowie Zeitreihen- und Cluster-Analysen einschließen.

Excel, Tableau und DeltaMaster können diese Anforderungen zumindest teilweise erfüllen. Hingegen fehlen bei Flow Immersive und Visualizer einfache statistische Kennzahlen wie z. B. Mittelwerte, Standardabweichung, Median und andere leicht zu berechnende, statistische Parameter.

Aus heutiger Sicht scheint somit die gemeinsame Nutzung aller o.g. Applikationen eine Möglichkeit zu sein, den oben beschriebenen, jeweils unterschiedlichen Mehrwert zu addieren und somit für eigene Aufgabenstellungen zu nutzen.

### 3.6. 3D-Brillen und Hologramme

Nicht nur die digitalen Datenvisualisierungsmöglichkeiten zeigen enorme technische Entwicklungssprünge, sondern auch die entsprechenden sogenannten VR/AR-Brillen [13][14][15][16][17][18][19].

VR-Brillen werden gegenwärtig noch als recht klobige „Kästen“ vor den Augen des Betrachters benutzt. So kann der Betrachter seine reale Umwelt um sich herum nicht mehr wahrnehmen (Ausnahme: manche VR-Brillen erlauben eine „Schwarz/Weiss-Durchsicht“ durch den „Kasten“ und die Darstellung der Umwelt in sehr unscharfen Bildern – Beispiel: Oculus Rift S).

Der Markt der VR-Brillen entwickelt sich allerdings sehr dynamisch, so dass die VR-Brillen der nächsten Generation einen deutlich besseren Tragekomfort erwarten lassen.

Interaktive Hologramme stellen sicherlich eine interessante Mög-

lichkeit der Informationsvermittlung dar. Dies kann einerseits in Form der Einbettung des eigenen Avatars in eine virtuelle Realität geschehen, in der sich der Avatar frei bewegen kann, aber auch als virtuelle Teamsitzung mit mehreren Avataren in derselben virtuellen Umgebung. Die letztgenannten Metaverse-Konzepte stellen heute schon eine klare Idee der zukünftigen Möglichkeiten dar. Es besteht kein Zweifel, dass diese technischen Möglichkeiten virtueller Räume umgesetzt und eine Auswirkung auf die tatsächliche Realität haben werden.

## 4. KI-unterstützte, statistische Datenanalyse

### 4.1. Deskriptive/konfirmatorische Statistik

Die alleinige Visualisierung von Daten führt sehr schnell zur Frage, ob vermeintliche Korrelationen, Outlier oder Trends tatsächlich im Sinne eines objektivierbaren kausalen Zusammenhanges vorhanden oder einfach nur Zufallskorrelationen sind, die keine Ursache-Wirkung-Beziehung aufweisen und damit keinen relevanten Erkenntnisgewinn beinhalten.

Daher liegt die Implementierung statistischer Methoden in die Visualisierungssoftware nahe. Allerdings sollten die statistischen Ergebnisse in einer leicht verständlichen Form und nahezu unbemerkt in die Visualisierung integriert werden, damit das Verständnis und die notwendigen Schlussfolgerungen ebenfalls auch ohne eine statistische Fachberatung erfolgen können.

Auf Grund der relativ stark strukturierten und reglementierten Möglichkeiten für die Auswahl der statistischen Analysemethoden (z. B. t-Test bei normal verteilten metrischen Daten oder Chi-Quadrat-Test bei kategoriellen Daten, etc.) lässt sich in Abhängigkeit vom Datentyp, der Anzahl Gruppen und der Messhäufigkeit sowie der Fragestellung und anderer Parameter ein Entscheidungsbaum aufbauen,



lotet die Chancen einer KI nicht einmal ansatzweise aus.

### 4.2.2. Interaktive, hypothesen-getriebene Formate

Die Präsentation im interaktiven, hypothesen-getriebenen Format ermöglicht es den Teilnehmern, selber mit den Daten zu „spielen“. Das Motto wäre hier: „Sag mir, was Du denkst und ich zeige Dir sofort, ob die Daten Deine Vermutung unterstützen.“

Die Aufgabe der KI besteht dann in der sofortigen (Real-time) Umsetzung der Gedanken und Ideen der Teilnehmer in mathematische Modelle und statistische Verfahren, um die Stichhaltigkeit der Ideen sofort überprüfen und bewerten zu können. Eine visuelle Darstellung, an welcher Stelle eines Entscheidungsbaums man sich gerade befindet, komplettiert dann das Konzept einer hypothesen-getriebenen, jedoch datenbasierten Entscheidungsfindung.

Im Bereich medizinischer Forschung wäre der Mehrwert zum einen in einer freien Zusammenstellung der Daten aus verschiedenen Studien zu sehen und zum anderen in der direkten Überprüfung von verschiedenen Hypothesen. Die Integration multivariater Analysetools in ein lernendes System ermöglicht hierbei die freie Kombination verschiedener Einzelparameter zur Beurteilung von vermuteten Ursache-Wirkungs-Beziehungen [20].

Dies ist insbesondere dann von Vorteil, wenn univariate Analysen keine signifikanten Korrelationen anzeigen. Besondere Aufgabe der KI wäre bei diesem Beispiel der Hinweis, welche univariaten Parameter in Kombination miteinander einen signifikanten Effekt erzeugen. Bei diesem auf Daten gestützten Ansatz käme den menschlichen Experten-Teilnehmern dann die medizinische Einordnung der jeweiligen Kombination zu. Neben der qualitativen Unterstützung einer KI, sprich: welche Parameter sind generell wichtig, kann eine KI auch darüber hinaus eine quantitative Aussage machen, mit welchem Faktor der jeweilige Parameter gewichtet wer-

den muss, um das statistisch beste Korrelationsergebnis zu erzielen.

Das lernende System wäre auch in der Lage, Fehleinschätzungen vorzubeugen, in dem es z. B. darauf hinweist, dass bestimmte Komponenten eines kombinierten Studienendpunkts miteinander in Korrelation stehen bzw. nicht als unabhängig zu betrachten sind. Damit hilft die KI, kostspielige, jedoch wissenschaftlich fehlerbehaftete Entwicklungspläne zu korrigieren. Das Besondere dieser Umsetzung in einem virtuellen Raum besteht dann in der Real-time-Umsetzung von Ideen und Beiträgen der menschlichen Teilnehmer, der Verifizierung bzw. Falsifizierung von vermuteten Korrelationen sowie der Optimierung von Korrelationen.

### 4.2.3. Interaktive, daten-getriebene Formate

Bei der Präsentation im interaktiven, datengetriebenen Format lädt ein „Data Management“ Avatar die Teilnehmer zunächst ein, die „Data Mining“ Ergebnisse gemeinsam anzuschauen. Motto wäre hierbei: „Ich zeige Euch mal, was ich so alles gefunden habe.“

Dieser Ansatz erscheint am ehesten hilfreich zur Lösung von Fragestellungen, bei denen die bisherigen hypothesen-getriebenen Ansätze zu keinem Ergebnis geführt haben oder in der Annäherung zu völlig neuartigen Fragestellungen. Eine KI ermöglicht dabei zunächst die verständliche Darstellung der Ergebnisse, darüber hinaus aber auch in selbiger Sitzung bereits die Exploration möglicher weiterführender bzw. aus den „Data Mining“ Ergebnissen abgeleiteter Ideen. Insgesamt ergibt sich hieraus die Chance auf einen relevanten Innovations- sowie Geschwindigkeitsvorteil in der Umsetzung neuartiger Therapieansätze. Ein patienten-relevantes Beispiel für diesen Ansatz ist die präoperative Optimierung der Ernährungssituation von Patienten vor einer elektiven OP, nachdem die positive Korrelation einer Mangelernährung mit der Länge des postoperativen Intensivaufenthaltes aus

großen Datenanalysen herausgefunden wurde.

## 5. Fazit

In der Zusammenschau ist eine sinnvolle Nutzung der KI in der Lage, eine zielgerichtete Ergebnisfindung zu unterstützen und zu moderieren. Sie zeigt dabei jederzeit transparent und nachvollziehbar den Pfad von Daten über Korrelationen zu möglichen Kausalitäten auf und bietet darüber hinaus als lernendes System die Möglichkeit der Integration des Wissens menschlicher Entscheidungsträger. Aufgrund dieser Interaktivität entsteht ein großes Vertrauen in die eigene Entscheidungsfindung, da KI-unterstützt viele Parameter auf Relevanz überprüft und bewertet werden können. Dadurch sinkt das Risiko kostenintensiver Fehleinschätzungen und damit verbundener Strategiewechsel.

Die Akzeptanz dieser KI-gestützten Systeme und das Vertrauen in sie steigt mit der Einfachheit und dem durchaus wörtlich gemeinten „Begreifen“ der jeweiligen Anwendung. Es ist nicht unwahrscheinlich, dass Anwendungen, die unterschiedliche Sinneseindrücke (visuell, haptisch) ansprechen, hier bevorzugt werden.

Die Verschmelzung der rein mathematisch-statistischen Programme mit einer entsprechenden Visualisierung hat gerade erst begonnen, ist jedoch noch lange nicht an ihrem Ende angelangt.

Man darf also neugierig in die Zukunft sehen, ob und wie die Verbindung von modernen 3D/VR/AR/MR Visualisierungstechniken zusammen mit AI-gesteuerter Statistik neue Erkenntnisse aufzeigen wird. |

### Referenzen

- [1] Linke J, Schipper F: Der medizinische Datenstrom kann zahllose werthaltige Informationen offenbaren. PM QM 2022, 24(1):14-18.
- [2] Linke J, Schippers F: Der medizinische Datenstrom – Hintergrund, praktische Beispiele und Kommentare. PM QM 2022, 24(2):50-60.
- [3] Schmedt M: Gesundheitsdaten – Nutzung pragmatisch regeln. Deutsches Ärzteblatt 2022, 119(39); 30. September 2022.

[4] Imre M, Linke J, Gernert D: Digital Health and Digital Treatments: The New Reality. PM QM 2020, 22(3):134-139.

[5] Linke J, Imre M, Gernert D, Dhabhai K: Building a DiGA – an IT and Medical Multidisciplinary Venture. PM QM 2021, 23(1):4-10.

[6] Imre M, Linke J, Dhabhai K: Planning and Performing Clinical Research for a DiGA – Route to pVE. PM QM 2021, 23(2):76-81.

[7] Novatec Consulting GmbH: Whitepaper – Augmented Reality und Virtual Reality – Vielfalt und Vorteile von Augmented Reality und Virtual Reality. Ohne Publikationsdatum. URL: [www.novatec-gmbh.de/wp-content/uploads/vrarwhitepaper.pdf](http://www.novatec-gmbh.de/wp-content/uploads/vrarwhitepaper.pdf) (letzter Aufruf: 25.10.2022).

[8] Awesome Technologies Innovationslabor GmbH: Whitepaper: Virtual und Augmented Reality. Ohne Publikationsdatum. URL: <https://awesome-technologies.de/whitepaper-virtual-und-augmented-reality> (letzter Aufruf: 25.10.2022).

[9] Vodafone: 5G-Whitepaper – Augmented und Virtual Reality. Ohne Publikationsdatum. URL: [www.vodafone.de/media/downloads/pdf/200103-whitepaper-vodafone-ar-v12-gartne.pdf](http://www.vodafone.de/media/downloads/pdf/200103-whitepaper-vodafone-ar-v12-gartne.pdf) (letzter Aufruf: 25.10.2022).

[10] Bitkom e.V. – Bundesverband Informationswirtschaft, Telekommunikation und neue Medien

e.V.: Augmented und Virtual Reality – Potenziale und praktische Anwendung immersiver Technologien. 2021. URL: [www.bitkom.org/sites/default/files/2021-04/210330\\_if\\_ar\\_vr.pdf](http://www.bitkom.org/sites/default/files/2021-04/210330_if_ar_vr.pdf) (letzter Aufruf: 25.10.2022).

[11] SLINTEL a 6sense company: Tableau Software – Top Competitors and Alternatives of Tableau Software. Letzter Update: 12.10.2022. URL: [www.slintel.com/tech/analytics/tableau-software-market-share/](http://www.slintel.com/tech/analytics/tableau-software-market-share/) (letzter Aufruf: 25.10.2022).

[12] Immersiv Analytics. URL: [www.immersiveanalytics.com](http://www.immersiveanalytics.com) (letzter Aufruf: 25.10.2022).

[13] Ryan Williams: The Best VR Headsets in 2022 (Ranked and Reviewed). ARVRTips – 12.03.2022. URL: <https://arvrtips.com/best-vr-headset/> (letzter Aufruf: 25.10.2022).

[14] Thomas Grau: Apple Glasses – das wird die AR-Brille von Apple können. DeutscheOptiker.de – 23.01.2022. URL: [www.deutscheoptiker.de/ratgeber/apple-glasses/](http://www.deutscheoptiker.de/ratgeber/apple-glasses/) (letzter Aufruf: 25.10.2022).

[15] MacTechNews: Große Fortschritte bei Apples AR-Brille – und eine baldige Markteinführung denkbar. 17.11.2021. URL: [www.mactechnews.de/news/article/Bericht-Grosse-Fortschritte-bei-Apples-AR-Brille-und-eine-](http://www.mactechnews.de/news/article/Bericht-Grosse-Fortschritte-bei-Apples-AR-Brille-und-eine-)

baldige-Markteinführung-denkbar-178974.html (letzter Aufruf: 25.10.2022).

[16] Jonas Bickel: Blick in die Zukunft – AR-Brillen: Erweiterte Realität im Alltag. Pointer.de – 22.01.2019. URL: [www.pointer.de/studium/studi-leben/14130/ar-brillen-im-alltag.htm](http://www.pointer.de/studium/studi-leben/14130/ar-brillen-im-alltag.htm) (letzter Aufruf: 25.10.2022).

[17] [www.google.de](http://www.google.de) Suche nach: „apple ar brille“.

[18] Matthias Bastian: Apple-Brille kommt mit M2 und 16 GB RAM im Januar. mixed.de: – 27.06.2022. URL: <https://mixed.de/apple-brille-kommt-mit-m2-und-16-gb-ram-im-januar-bericht/> (letzter Aufruf: 25.10.2022).

[19] Ben Schwan: Erste Generation von Apples AR-Brille in der Designentwicklung. heise online – 21.06.2022. URL: [www.heise.de/news/Erste-Generation-von-Apples-AR-Brille-in-der-Designentwicklung-7146656.html](http://www.heise.de/news/Erste-Generation-von-Apples-AR-Brille-in-der-Designentwicklung-7146656.html) (letzter Aufruf: 25.10.2022).

[20] Linke J, Topp H, Mueller I, Bark S, Shaffer J, Singer P: Can the use of the multivariate statistical method MANOVA reduce the number of patients needed in clinical trials? Poster 42, 37th ESPEN (European Society for Clinical Nutrition and Metabolism) Congress 2015.

## AUTOREN



**Dr. med. Joachim Linke, MSc** in Applied Statistics, ist als Gründer und Geschäftsführer bei der CCR GmbH für Geschäftsentwicklung und Innovation

sowie insbesondere für die Anwendung spezifischer CCR-Kompetenzen wie Medizin, Statistik und Datenmanagement für die Bereiche Big Data Analysis und Predictive Analytics verantwortlich.

Seiner Meinung nach wird die sich schnell entwickelnde Digitalisierung im Gesundheitswesen zu spannenden neuen Geschäftsfeldern führen, welche insbesondere die kombinierte Nutzung der vorgenannten Kompetenzen erfordern.

Kontakt:  
[jlinke@ccr-cro.com](mailto:jlinke@ccr-cro.com)



**Dr. med. Frank Schippers** ist Facharzt für Anästhesiologie mit Zusatzbezeichnungen in Intensivmedizin und Notfallmedizin. Neben seiner Tätigkeit

als praktischer Intensivmediziner hat er die klinische Forschungsabteilung der PAION Deutschland GmbH geleitet und dort maßgeblich zur Zulassung von Remimazolam in mehreren Regionen beigetragen.

Durch die tägliche Arbeit an der Schnittstelle zwischen klinischer Forschung, Arzneimittelsicherheit und praktischer ärztlicher Tätigkeit sind ihm die Hindernisse, aber auch die Vorteile einer soliden und evidenzbasierten Entscheidungsfindung auf Basis robuster Datenauswertungen bewusst.

Kontakt:  
[fschippers@ccr-cro.com](mailto:fschippers@ccr-cro.com)



**Florian Stöwe** ist als Inhaber eines IT-Systemhauses ein langjährig erfahrener IT-Manager auf höchstem Niveau. Sein beruflicher Fokus und somit seine

Erfahrung liegt auf der Bereitstellung von Serverarchitekturen und Cluster-Systemen für mittelständische Unternehmen, welche er auf dem Weg in eine konsequente Digitalisierung begleitet. Zudem unterstützt er Unternehmen, die sich im Gesundheitsmarkt den komplexen Herausforderungen der Digitalisierung stellen, insbesondere mit den Schwerpunkten belastbares, modernes (Datenbank)-Hosting sowie 3D/VR/AR/MR Visualisierung.

Kontakt:  
[florian@fs-it-service.de](mailto:florian@fs-it-service.de)