

Data Management - Die Durchdringung des Datenschungels

Begriffe wie Big Data, Data Management und Data Science sind im Jahr 2022 längst in aller Munde. Nicht erst seit Corona und der damit einhergehenden forcierten Digitalisierung vieler Unternehmen und Organisationen ist ein exponentieller Zuwachs generierter Daten zu verzeichnen. In den meisten Chefetagen ist man sich heutzutage sicher, dass diese enorme Menge an Daten für Unternehmensentscheidungen sinnvoll zu nutzen ist. Denn was ist die Motivation hinter dieser zunehmenden Datenfaszination? Es ist der Wunsch der Führungsebene nach Entscheidungssicherheit auf Grundlage eindeutiger Daten, nach dem Erkennen wichtiger Trends und dem Aufzeigen klarer Handlungsbedarfe auf einen Blick.

Einige der bekannten Akteure am Markt titulieren sich gar schon als „data-driven Company“, also ein Unternehmen, das überwiegend datengetrieben arbeitet. Um diesem Titel gerecht zu werden, werden Data Analysts, Data Scientists und Data Engineers eingestellt – Jobtitel, die vor etwas mehr als einer Dekade in den meisten Organisationen noch unter der Firmierung Wirtschaftsinformatiker/in kursierten.

Und was macht die Sozialwirtschaft?

Hier scheint die Datenfaszination noch nicht wirklich angekommen. Und dies allein mit der Einschätzung zu begründen, dass der Fokus der Branche eher auf Menschen denn auf Daten liegt, scheint nicht komplett zu greifen. Was ist es also dann, was die Sozialwirtschaft vermeintlich noch so weit von Big Data-Welt entfernt?

Datenanalyse und Datenprognosen sind nur dann sinnvoll und möglich, wenn die richtigen Daten vollumfänglich und zeitnah vorliegen. Doch gerade in der Sozialwirtschaft zeigt sich häufig noch eine andere Realität. Das beste KI-gestützte Analyse- und Vorhersagemodell nutzt vergleichsweise wenig,

wenn in der Organisation Daten aus unterschiedlichen Software-Lösungen nicht sinnvoll zusammengeführt werden, oder die Anfrage nach einem Bericht der historischen Belegungssituation Schweiß auf die Stirn der Verwaltungskraft treibt. Um Daten sinnvoll nutzen zu können, bedarf es im Kern eines zielgerichteten Datenmanagements (**Data Management**).

Was steckt nun aber hinter dem Begriff Data Management?



Man kann Data Management als die Zusammenführung technischer, organisatorischer und konzeptioneller Maßnahmen beschreiben, mit dem Ziel der Bereitstellung und Nutzung von Daten für die optimale Unterstützung von Prozessen und einer datengestützten Entscheidungsfindung in einer Organisation. Die Aufgabe des Data Managements ist es demnach, unstrukturierte Daten aus verschiedenen Quellen zu strukturieren und zusammenzuführen, um sie dann für bestenfalls automatisierte Auswertungen zu verwenden. Die wichtigste Aufgabe im Prozess obliegt dabei immer noch dem Menschen, der auf Basis dieser Vorarbeiten die Organisation beispielsweise besser steuern und lenken kann. Dabei entlastet Data Management gleichzeitig die Verwaltungskräfte von fehleranfälligen Korrekturschritten, bei denen Inhalte aus Software-Exporten und Excel-Tabellen angepasst und von einem Ort an den anderen geschoben werden. Ziel des Data Managements ist also nicht, schlicht möglichst viele Zahlen zu sammeln und wichtige Entscheidungen einem Computer zu überlassen, sondern ganz im Gegenteil, eindeutige und verlässliche Datengrundlagen zu schaffen, die Bereichsleitungen und Geschäftsführung bei wichtigen unternehmerischen Entscheidungen unterstützen.

Die Durchdringung des Datenschungels - Eine Landkarte hilft

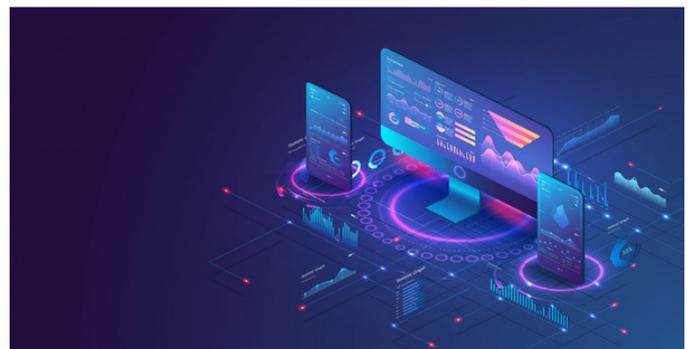
Aber wie kann das gelingen, was über Jahre vermeintlich große Organisationen vor Herausforderungen gestellt hat? Zu Beginn des Prozesses geht es darum, den existierenden „Datenschungel“ zu durchdringen.

Hierzu erstellen wir eine Landkarte, auf der skizziert wird, an welchen Orten in der Organisation und in welchen Systemen eigentlich Daten generiert werden: Belegungsdaten kommen beispielsweise aus der (teilweise je Geschäftsbereich unterschiedlichen) Leistungssoftware, Leistungsabrechnungsdaten ebenfalls, Personaldaten liegen in der Personalabrechnungs- oder der Dienstplanungssoftware vor, während Stundenzettel in Excel-Tabellen und Fahrzeiten womöglich noch händisch in einer Papierliste geführt werden. Am besten eignet sich für diese Sammlung ein großes (ggf. digitales) Whiteboard, auf dem diese Datenquellen eingezeichnet und mit weiteren Informationen zu den dort entstehenden Dateninhalten angereichert werden. Dies ähnelt im ersten Schritt einer Software-Analyse, da der Großteil der generierten Daten aus den großen Software-Lösungen der Organisation stammt. Es lohnt sich jedoch auch, diesen ersten Schritt vom anderen Ende des Prozesses aus zu denken: Welche strategischen und operativen Fragestellungen möchte ich mit Daten beantworten? Die Antwort darauf legt wahrscheinlich auch den Einbezug von Daten nahe, welche aktuell nur analog vorliegen, gar nicht erhoben werden oder nur außerhalb der eigenen Organisation existieren. Hier sind zum Beispiel Daten zu den Preisen und Platzzahlen von Wettbewerbern ebenso spannend wie ortsbezogene demographische Daten und Bevölkerungsprognosemodelle, welche Aufschluss über die Attraktivität eines Standortes geben können.

Die anschließende Visualisierung der Bewegung der Daten in der Organisation in Form von Datenflüssen ist bereits zentrale Grundlage des zukünftigen Datenmanagements. Hierbei werden

langsame Datenflüsse, Sackgassen und Blockaden meist schnell ersichtlich: die Bewerberdaten liegen beispielsweise nur standortbezogen vor und als Austauschmedium wird ein USB-Stick oder gar die Papierakte genutzt. Schlimmstenfalls gibt es innerhalb der Organisationen keinen Austausch über Initiativbewerbungen oder gute Kandidaten. Informationen zu den Fahrzeiten in der ambulanten Pflege werden aus dem System generiert, landen danach aber in einem „Datenfriedhof“ und werden nicht weiter ausgewertet. Arbeitszeit und Face-to-Face-Zeit in der ambulanten Kinder- und Jugend- oder der Eingliederungshilfe werden in unterschiedlichen Systemen erfasst und nicht zur Auswertung der Effizienz herangezogen. Finanz-, Personal- und Leistungsdaten fließen zwar stromlinienförmig und effizient durch die Organisation, kreuzen aber, metaphorisch gesprochen, an keiner Stelle, sodass keine kombinierte Kennzahl für die Steuerung gebildet werden kann.

Nach der initialen Kartografie und dem Offenlegen von Datenquellen, Datenflüssen und Blockaden ist der erste Schritt zwar geschafft, doch die eigentliche Arbeit des Data Managements beginnt genau hier. Denn aus der Karte und den möglichen Problemstellungen leiten sich nun direkt Handlungsmaßnahmen ab. In der eingeführten Analogie bleibend, stellen sich nun konkrete Fragen: Müssen zusätzliche Ströme angelegt werden? Muss das Flussbett verbreitert werden? Wo müssen Dämme gebrochen werden und wo brauchen wir vielleicht einen künstlichen See (Datalake ist nicht ohne Grund ein bekannter Begriff im Data Management), in dem die Datenströme zusammengeführt werden?



Neue Technologie als Ausgangspunkt

Wenn wir die Analogie des Dschungels mit Quellen und Flussläufen einmal verlassen, und uns der Frage zuwenden, was Datenmanagement eigentlich technologisch erfordert, wird klar, die Technologie zur erfolgreichen Umsetzung der oben genannten Schritte ist mittlerweile längst bekannt und vielfach genutzt. Direkte Schnittstellen und sogenannte Data Pipelines ermöglichen den Zugriff auf Daten aus einer Datenquelle, die Transformation dieser Daten und das abschließende Überführen der Daten zu einem Zielort. Solche Zielorte können beispielsweise relationale Datenbanken oder größer angelegte Data Warehouses und Data Lakes sein, die mehrere Datenbanken enthalten. Es kann sich in simplen Fällen aber auch schlichtweg um einen Ordner handeln, in dem Dateien abgespeichert und von dort automatisiert ausgewertet werden können.

Die separaten Datenspeicherorte dienen dem Sammeln und Speichern der transformierten Daten, die dann wiederum analysiert werden können. Moderne Business-Intelligence-Lösungen (BI-Lösungen) ermöglichen heutzutage auch das direkte Einlesen und Auswerten der Daten aus den vorgeschalteten Programmen. Dadurch ist eine separate Speicherung häufig nicht mehr notwendig und die Daten fließen direkt in die BI-Lösung, welche dann anhand vorher definierter Algorithmen entsprechende Dashboards und Auswertungen erstellen kann.

Die vorhin nur beiläufig erwähnte Transformation der Daten ist deswegen notwendig, da insbesondere bei der Verknüpfung von Daten aus mehreren Ursprungssystemen nur selten eine gleiche Nomenklatur und Formatierung existiert. Was in einem System pro Tag kalkuliert wird, berechnet die andere Software vielleicht auf Wochenbasis, sodass ein direkter Vergleich ohne vorherige Transformation der Daten zu nicht aussagekräftigen Ergebnissen führt.

In vielen Fällen sind Daten heutzutage zudem nicht mehr in einer simplen Tabellenform mit Zeilen und Spalten strukturiert, sodass sie in Excel-Tabellen und klassischen relationalen Datenbanken nicht abbildbar sind. Schon 2017 machten solche unstrukturierten Daten sogar etwa 80% der gesamten generierten Datenmenge aus. Bild-, Text-, Video- und Sprachdateien sind jedoch ebenso Daten, mit dem entscheidenden Nachteil, dass sie unstrukturiert sind und keine inhärente Zeilen- und Spaltenposition besitzen. Durch diesen Umstand sind unstrukturierte Daten für Computerprogramme grundsätzlich schwierig zu analysieren. Neueste Entwicklungen, basierend auf der Technologie künstlicher Intelligenz, zeigen jedoch, dass eben auch solche Dateien relevant und sinnvoll auswertbar sind. Algorithmen können im Jahr 2022 beispielsweise nicht nur fehlerfrei Sprache zu Text transformieren, sondern im Anschluss auch den Text auf bestimmte Wortcluster untersuchen. Diesen Prozess nennt man Natural Language Processing (NLP).

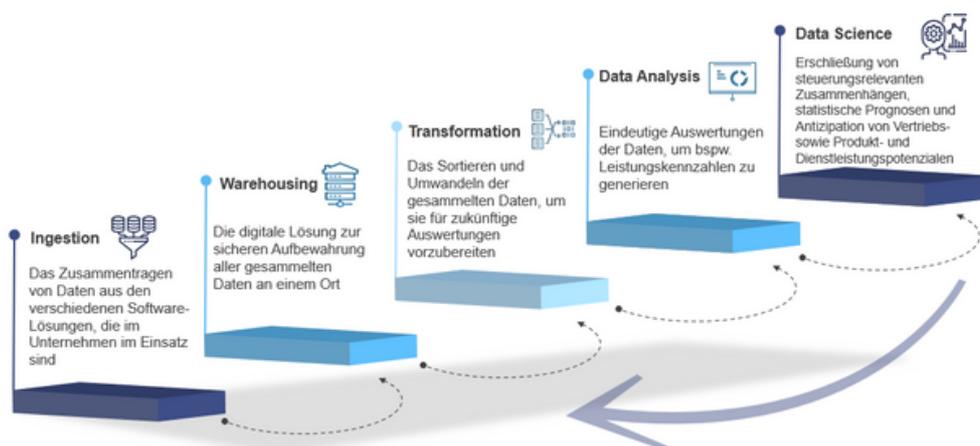


Abb. 1: Die 5 Schritte zur erfolgreichen Nutzung von Daten in Organisationen. Eigene Darstellung

Für die Sozialwirtschaft könnte das heißen: Aus der per Sprache erfassten Pflegedokumentation einer Pflegefachkraft wird ein Textdokument generiert, welches dann automatisch nach bestimmten Schlüsselwörtern durchsucht und ausgewertet wird. Kommen bestimmte Worte und Wortketten im Text vor, erkennt dies die künstliche Intelligenz und erstellt daraus ein strukturiertes Dokument mit allen wichtigen Aspekten. Auf diesem Wege können also auch unstrukturierte Daten in strukturierte Daten umgewandelt werden, welche dann analysierbar sind.

Die rasch wachsende Menge an Daten hat eine weitere revolutionäre Entwicklung forciert: die Cloud-Speicherung. Daten werden immer weniger auf eigenen Computern, Festplatten und lokalen Servern, sondern stattdessen extern bei Drittanbietern gespeichert. Diese Cloud-Anbieter bieten Datenspeicherung als Service, häufig mit monatlichen Gebühren. Neben höherer Datensicherheit und flexibler Skalierbarkeit auf mehr oder weniger Speicherplatz ermöglicht die Cloud Datenzugriff von überall, gemeinsame Bearbeitung von Dateien und einen „Single Point of Truth“, also einen eindeutigen und für alle Mitarbeiterinnen und Mitarbeiter identischen Datenstand statt unterschiedlicher Datenstände auf verschiedenen Endgeräten. Die Cloud-Speicherung ist demnach ein essentieller Bestandteil eines funktionierenden Datenmanagements.

Die Realität der Sozial- und Gesundheitswirtschaft

Auch wenn in der Sozial- und Gesundheitswirtschaft die Kernprozesse der Leistungserbringung naturgemäß nahe am Menschen liegen und Gesetze und Regularien die Etablierung vieler digitaler Lösungen erschweren, so wächst auch hier die Menge an Daten in den letzten Jahren exponentiell. Ob es nun die Stammdaten von Klienten und Kostenträgern (sowie die Hinterlegung, welche Leistungen mit welchem Kostenträger bzw. mit dem Selbstzahler abzurechnen sind), Informationen über gegebenenfalls zusätzlich ab-

rechenbare Leistungsbestandteile, Dokumentationen geplanter Förderziele oder dynamische Einsatzdaten des Personals sind, eines ist klar: heute kursieren in sozialwirtschaftlichen Organisationen mehr Daten denn je, mit steigender Tendenz. Denn neue Regularien, wie beispielsweise das Digitale-Versorgung-und-Pflege-Modernisierungs-Gesetz sehen vor, dass Pflegeeinrichtungen Versichertenstammdaten online abgleichen und Verordnungen elektronisch empfangen müssen. Es entstehen also weitere Datenströme, die effizienter Integration bedürfen.

Umso problematischer wirkt sich aus, dass sozialwirtschaftliche Träger in der aktiven Verwaltung und im Management von Daten anders als viele Unternehmen aus der freien Wirtschaft nicht gut aufgestellt sind. Dies liegt in Teilen sicherlich an den vorherrschenden, abschreckenden Regularien. Insbesondere personenbezogene Daten über Patienten oder Heimbewohner und ihren gesundheitlichen Status genießen in Deutschland aus gutem Grund höchste Datensicherheit. Doch da es im Datenmanagement einer sozialwirtschaftlichen Organisation nur zu einem gewissen Anteil um diese personenbezogenen Daten geht, ist dies nicht der einzige Grund. Zu einem gewissen Grad liegt es auch an einem bis Weilen noch gering ausgeprägtem „Data Mindset“ unter den Mitarbeiterinnen und Mitarbeitern und den Führungskräften. Data Mindset bezeichnet dabei das Denken im Datenkontext und ein grundlegendes Verständnis darüber, wie Daten in einer Organisation fließen und welche Bedeutung ihnen beizumessen ist.

Im Fokus des Tuns steht in der Sozialwirtschaft schließlich immer noch der Mensch und das ist auch gut und richtig so. Und dennoch: Der Druck, sich mit Daten auseinanderzusetzen, wächst. Steuerungsanforderungen und Regularien, marktliche Entwicklungen, neue Player sowie sich wandelnde Anspruchsgruppenbedarfe fordern auch in der Sozial- und Gesundheitswirtschaft eine Professionalisierung des Datenmanagements. Zudem führt eine hohe Digitalisierung sowie ein

funktionierendes Datenmanagement mit automatisierten Workflows unmittelbar zu mehr „Zeit für den Menschen“, da viele manuelle und monotone Schritte wegfallen, die zuvor Arbeitszeit gekostet haben. Zusätzliche Steuerungsinformationen führen dabei zu wichtigen Unternehmensentscheidungen, die die Arbeit am Menschen auch in der Zukunft sichern.

Die Zeit, in der schlechte Datenqualität und unzuverlässige Datengrundlagen zu Entscheidungen „aus dem Bauch heraus“ oder auf Basis unzureichender Informationen zu Fehlallokation führen, muss vorbei sein. Das bedeutet gleichzeitig, dass Organisationen sich viel differenzierter mit der Wahl der richtigen Steuerungskennzahlen beschäftigen können und müssen.

Auch ein Blick auf die großen Software-Unternehmen der Sozialwirtschaft zeigt, dass sich die Zeiten ändern. Zuletzt stark wachsende Anbieter, wie Connext oder CGM, erweitern ihr Angebot an Datenmanagement- und Datenanalysetools für die Sozialwirtschaft stetig, während Software-Giganten wie SAP sich ebenfalls mehr und mehr der Branche zuwenden und inzwischen passgenaue und erschwingliche Lösungen für Datenspeicherung und -auswertung anbieten.

Der wichtigste Faktor sind Ihre Mitarbeiterinnen und Mitarbeiter

Data Management fordert Technologie – keine Frage, aber wie schon im Zusammenhang mit dem angesprochenen Data Mindset festgestellt, steckt ein wichtiger Faktor für erfolgreiches Data Management in den Mitarbeiterinnen und Mitarbeitern der Organisation. Wenn keinerlei Datenverständnis in der Organisation existiert, ist es beinahe unmöglich, Datenmanagement nachhaltig zu implementieren. Hier steckt also auch ein möglicher Fallstrick, der von Beginn an mitgedacht werden muss. Es reicht nicht aus, einzelne Mitarbeiterinnen und Mitarbeiter auf Schulungen zum Thema Datenmanagement zu schicken. Alle in der Organisation müssen verstehen, warum Daten und ihre Pflege für eine spätere Auswertbarkeit so wichtig sind. Zudem ist es wichtig, zu vermitteln, dass die zusätzliche Transparenz nicht schädlich ist oder die Handlungsfreiheit reduziert. Tatsächlich sind Daten immer nur Hinweise auf bestimmte Entwicklungen und Situationen. Entscheidungen können vor dem Hintergrund strategischer Erwägungen auch jeweils entgegen der Datenlage getroffen werden. Um sich diesem Zielbild zu nähern, ist es sinnvoll, den aktuellen Wissensstand zu erheben und das Verständnis in der Organisation

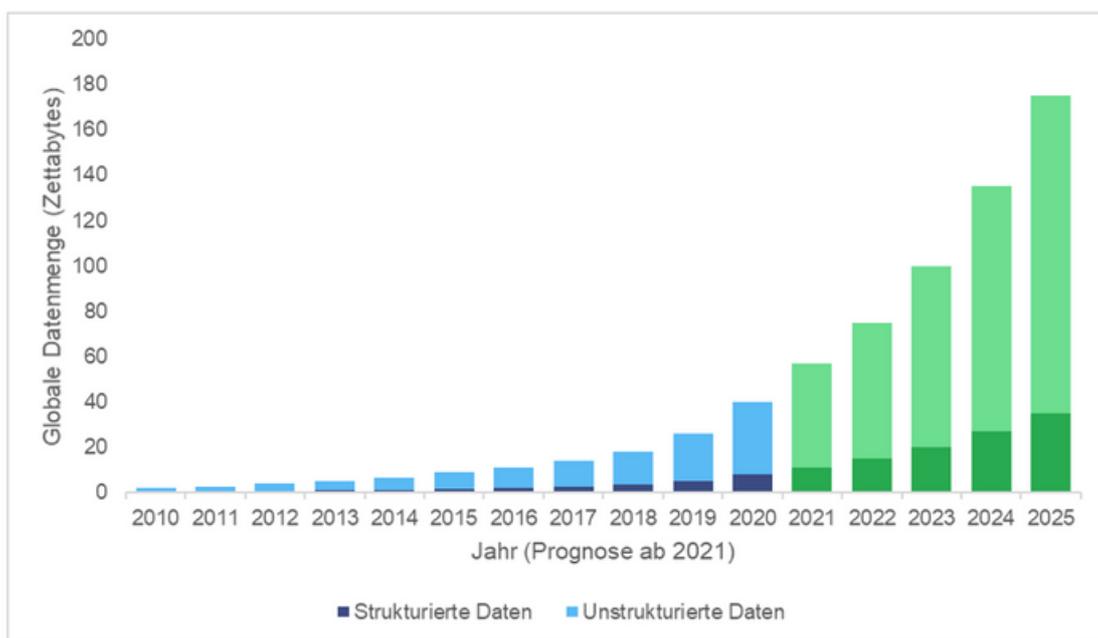


Abb. 2: Exponentielles Wachstum strukturierter und unstrukturierter Daten. Eigene Darstellung

zu ermitteln: Ist die Bedeutung der Daten, die ich pflege, klar? Wann und was entscheide ich auf Basis von Daten? Interpretiere ich Daten richtig?

Je nach Ausgangssituation gilt es dann, genau an den Schwachstellen anzupacken und die Datenkompetenz in der Organisation zu steigern. Unserer Erfahrung nach scheitern deutlich mehr Datenmanagement-Projekte am fehlenden Data Mindset als an technischen Hürden. In manchen Fällen wurden teure Lösungen eingekauft, die in der Lage sind, alle relevanten Daten der Organisation zu aggregieren und sinnvoll auszuwerten, doch die mangelnde Datenkompetenz, das fehlende Verständnis für die Notwendigkeit und Sinnhaftigkeit datengetriebener Auswertungen oder ein ausgeprägtes Ressortdenken ließen das Projekt scheitern.

Die Datenstrategie als strategische Orientierung

Datenmanagement ist ein sehr wichtiger Prozess, der ausgehend von der Datenanalyse über die beschriebenen Schritte die elementaren Voraussetzungen für das Zielbild fundierter, datengestützter Entscheidungen schafft. Daneben braucht es einen übergeordneten Orientierungsrahmen, eine Datenstrategie. Die Wichtigkeit von Daten in der heutigen Zeit wird hierdurch nochmals unterstrichen.

Die Formulierung dieser Teilstrategie fällt vor allem Neulingen im datengetriebenen Denken schwer. Eine Datenstrategie muss, wie die Softwarestrategie, in die Strategie der Organisation eingebettet sein. Neben dem Zielbild und den hierfür notwendigen strategischen Stoßrichtungen wie zum Beispiel Technologie, Kompetenzen und Prozesse, spielen möglicherweise auch rechtliche und ethische Positionen eine wichtige Rolle, die es in der Datenstrategie einer sozialwirtschaftlichen Organisation festzuschreiben gilt.

Die Quintessenz

Welche Erkenntnisse ziehe ich als Sozialwirtschaft-

licher Träger also aus diesen Ausführungen zum Thema Datenmanagement?

1. Die Individualisierung der Leistungsangebote, gesetzliche Änderungen und eine fortschreitende Digitalisierung bringen zwangsläufig eine exponentiell wachsende Datenmenge mit sich.
2. Daten enthalten Informationen, welche für die Steuerung und Lenkung einer sozialwirtschaftlichen Organisation von zunehmender Bedeutung sind.
3. Belastbares und reliables Wissen lässt sich nur aus sorgfältig dokumentierten und gepflegten Daten ziehen, die durch einen strukturierten Datenmanagementprozess zusammengeführt und ausgewertet wurden. Insbesondere der Datenqualität muss dabei von Eingabe bis Auswertung hohe Priorität zugemessen werden.
4. Von zentraler Bedeutung für die nachhaltige und erfolgreiche Umsetzung eines vernünftigen Datenmanagements sind, neben dem Einsatz der richtigen Technologien, die eigenen Mitarbeiterinnen und Mitarbeiter und ihr Verständnis der Bedeutung von Daten.
5. Die Wichtigkeit von Daten in der heutigen Zeit zeigt sich auch in der Festlegung einer ganzheitlichen Datenstrategie für die Organisation.

Der Prozess zur Einführung und konsequenten Umsetzung eines sinnvollen Datenmanagements ist alles andere als trivial und dabei gleichzeitig absolut notwendig. Die exponentiell wachsende Datenwelt bestraft Organisationen, die sich dem Thema erst übermorgen widmen.

Über die oben aufgezeigten Schritte lässt sich für Sie hoffentlich ein niederschwelliger Einstieg in das Thema finden. Die ressortübergreifende Sammlung von Datenquellen und Datenströmen ist dabei schon ein erster, großer Schritt in die richtige Richtung. Natürlich unterstützen wir Sie gerne bei diesem Schritt und begleiten Sie auf dem Weg zu einer datengestützten Steuerung Ihrer Organisation.

Ihr Ansprechpartner:



Sven Buchholz
Fachberater
buchholz@rosenbaum-nagy.de



Besuchen Sie auch unser Data Management Seminar:

06.10.2022: Datenmanagement in der Sozialwirtschaft oder wie man den Datenschungel durchdringt

<https://www.rosenbaum-nagy.de/seminare/datenmanagement-in-der-sozialwirtschaft-oder-wie-man-den-datenschungel-durchdringt/>

und viele weitere Seminare:

<https://www.rosenbaum-nagy.de/seminare/>