

FINSOZ e. V.

Datenstrategie für die Sozialwirtschaft

Praxisleitfaden für effiziente Geschäftsprozesse und rechtskonforme Datennutzung

Mai 2026

www.finsoz.de



Autor:innen und Mitwirkende

FINSOZ-Fachgruppe „Digitales Arbeiten in der Sozialwirtschaft“, 2025

Carolyn Schneider	d.velop AG (Autorin)
Thomas Schönweitz	Whitespring (Autor)
Henning Brandes	Konica Minolta Business Solutions Deutschland GmbH
Thordis Eckhardt	FINSOZ e. V.
Dana Hieronymus	Diakonie Nord Ord Ost in Holstein gGmbH
Anne Lena Knall	Die Zieglerschen e. V.
Melanie Miermeister	Stadt Nürnberg NürnbergStift
Jürgen Wintermantel-Menzel	LVR-Verbund für WohnenPlusLeben

Inhalt

Überblick	2
1. Grundlagen der Datenstrategie	3
Zwei mögliche Herangehensweisen zur Erstellung einer Datenstrategie	5
Anforderungsaufnahme.....	9
2. Datenqualität	10
3. Strategische Entscheidungen zur Speicherung der Datenstrukturen	12
4. Schnittstellen	13
5. Zentrale Auswertung	14
6. Entwicklung von Szenarien	15
Szenario 1: Kundenzentrierung, Angebotserweiterung, Stabilisierung am Markt.....	15
Szenario 2: Mitarbeitermangel: Effizienzen durch Daten erhöhen.....	16
Szenario 3: Abrechnung.....	17
Szenario 4: CSRD – Nachhaltigkeitsberichterstattung durch effizientes Datenmanagement.....	17
Szenario 5: Innovation – Datenstrukturen als Katalysator für soziale Innovation	19
Zusammenfassung und Schlusswort.....	20
Glossar	21

Überblick

Der vorliegende Praxisleitfaden gibt Informationen und Hinweise zu einer systematischen Datenstrategie für Organisationen der Sozialwirtschaft, die vier Kernbereiche abdeckt:

1. **Grundlagen** (Datenmodellierung, Geschäftsprozessanalyse, rechtliche Compliance)
2. **technische Architektur** (Datenqualitätsstufen, API-Integration, dezentrale vs. zentrale Speicherung)
3. **operative Umsetzung** (Schnittstellen, zentrale Auswertung) und
4. **strategische Anwendungsszenarien**

Darüber hinaus zeigen fünf konkrete Szenarien praktische Anwendung auf:

1. Kundenzentrierung durch 360°-Kundendaten und DSGVO-konforme Prozesse
2. Effizienzsteigerung bei Personalmangel durch KI-gestützte Automatisierung
3. optimierte Abrechnungsprozesse mit digitalen Beweisketten
4. CSRD-konforme Nachhaltigkeitsberichterstattung und
5. innovationsfördernde Datenstrukturen

Kernbotschaft des Praxisleitfadens: Eine wirksame Datenstrategie beginnt nicht mit der Technik, sondern mit einer klaren Zieldefinition und einer strukturierten Anforderungsaufnahme. Erst wenn klar ist, welche Daten für welche Zwecke benötigt werden, lassen sich tragfähige Lösungen entwickeln. Gute Schnittstellen und offene Datenarchitekturen spielen dabei eine zentrale Rolle, denn sie ermöglichen es, bestehende Systeme intelligent zu verknüpfen, Effizienzpotenziale zu heben und Innovationen gezielt zu fördern.

Zielgruppe des Praxisleitfadens: Das Dokument ist nicht für Einsteiger:innen gedacht, sondern für Fach- und Führungskräfte wie z. B. Strategische Entscheider:innen, IT-Verantwortliche, Digitalisierungsbeauftragte, Fachbereichsleitungen, Projekt- und Change-Manager:innen sowie Multiplikator:innen im FINSOZ Verband. Das Dokument enthält kurze Exkurse zu Fachthemen mit Managementbezug sowie technische Details. Diese werden für den interessierten Lesenden, der mehr Kontext möchte, kurz erläutert, sind aber für das Verständnis dieses Dokuments nicht zwingend nötig.

1. Grundlagen der Datenstrategie

Eine Datenstrategie ist ein strukturierter Plan, der beschreibt, wie ein Unternehmen mit Daten umgeht, um seine spezifischen Geschäftsziele zu erreichen. Da jedes Unternehmen unterschiedliche Ziele, Prozesse und Rahmenbedingungen hat, ist auch jede Datenstrategie individuell. Dennoch lassen sich zentrale Bestandteile identifizieren, die in nahezu jeder Strategie berücksichtigt werden sollten. Dazu zählen die Zielsetzung der Datennutzung, Fragen zur Datenqualität, klare Verantwortlichkeiten und Rollen, sowie die Prozesse zur Erhebung, Verarbeitung und Nutzung von Daten.

Eine Datenstrategie existiert nie zum Selbstzweck. Sie soll konkrete Verbesserungen in der Organisation ermöglichen, insbesondere in der Altenhilfe und Eingliederungshilfe, wo Fachkräftemangel, Dokumentationspflichten und individuelle Teilhabeziele besondere Anforderungen an Daten und Prozesse stellen. Um dies zu erreichen, sollten von Beginn an folgende Fragen beantwortet werden:

- a. Was genau soll mit der Datenstrategie erreicht werden? (z. B. Reduktion von Bearbeitungszeiten in der Pflegedokumentation)
- b. Bis wann sollen erste Ergebnisse sichtbar sein? (Dies hängt vom Projektumfang und den Zielen der Unternehmung ab)
- c. Woran erkennen wir, dass die Strategie erfolgreich ist? (z. B. anhand messbarer KPIs wie:
 - Anteil strukturierter Klient:innendaten in zentralen Systemen
 - Anzahl automatisierter Übergabeprozesse zwischen Diensten
 - Zufriedenheit der Mitarbeitenden mit der Datenverfügbarkeit
 - Reduktion von Doppelpflege in Dokumentation und Abrechnung)
- d. Wie und wann wird überprüft, ob die Ziele erreicht wurden? (z. B. durch quartalsweise Auswertung definierter KPIs, Feedbackrunden mit Fach- und Pflegekräften oder durch Reviews in Projekt- oder Leitungsgremien)

Diese Fragen sollten nicht nur zu Beginn gestellt, sondern regelmäßig überprüft und angepasst werden. Dabei empfiehlt es sich, Ziele stets nach dem SMART-Prinzip zu formulieren (also spezifisch, messbar, erreichbar, realistisch und terminiert). Nur so wird die Datenstrategie zu einem wirksamen Instrument der Organisationsentwicklung.

Praxisbeispiel zur Zieldefinition:

Eine stationäre Pflegeeinrichtung verfolgt das Ziel, die Pflegedokumentation vollständig, nachvollziehbar und rechtssicher zu gestalten und definiert dies wie folgt:

- Spezifisch (S): Ziel ist die lückenlose Erfassung aller pflegerischen Maßnahmen, Übergaben und Medikationen.
- Messbar (M): Als KPI wird der Anteil vollständig dokumentierter Pflegeprozesse pro Bewohner:in erfasst (z. B. 90 % innerhalb von 24 Stunden).
- Erreichbar (A): Die Einrichtung verfügt über die notwendige technische Infrastruktur, personelle Ressourcen zur Einführung sowie die finanziellen Mittel.
- Realistisch (R): Die Maßnahme wird durch Schulungen und externes Consulting unterstützt.
- Terminiert (T): Zielerreichung innerhalb von 12 Monaten nach Projektstart.

Eine Datenstrategie existiert nicht nur nicht zum Selbstzweck, sondern ist eng mit anderen strategischen Managementansätzen verknüpft und sollte stets im Zusammenspiel mit diesen gedacht und entwickelt werden. Daher wollen wir vorab auf einige **Managementansätze** eingehen, die hier relevant sein können:

Exkurs zu relevanten Managementansätzen:

1. Geschäftsprozessmanagement (GPM):

Daten bilden die Grundlage für die Analyse, Optimierung und Automatisierung von so gut wie allen Geschäftsprozessen und insbesondere in der Altenhilfe und Eingliederungshilfe sind Prozesse oft stark an gesetzliche Vorgaben (z. B. SGB IX, SGB XI) gebunden. Ein durchdachtes Datenmodell unterstützt die Modellierung von Prozessen und ermöglicht eine präzise Abbildung der realen Abläufe. Gleichzeitig liefert das GPM wichtige Impulse für die Datenstrategie, insbesondere die Frage, welche Prozesse welche Daten benötigen und von hier abgeleitet: Welche Softwareprodukte und Schnittstellen hier pro Prozess relevant sind.

2. IT-Management:

Die Umsetzung einer Datenstrategie ist ohne eine enge Verzahnung mit der IT-Strategie nicht denkbar. Aspekte wie Systemarchitektur, Cloud-Nutzung oder Datenbanktechnologien müssen abgestimmt werden. Die IT stellt die technische Infrastruktur bereit, auf der datengetriebene Anwendungen aufbauen.

3. Balanced Scorecard (BSC):

Für die Steuerung eines Unternehmens über KPIs ist eine verlässliche Datenbasis unerlässlich. Die Datenstrategie gewährleistet, dass relevante und qualitativ hochwertige Daten für alle vier Perspektiven der BSC (Finanzen, Kunden, interne Prozesse sowie Lernen und Entwicklung) verfügbar sind. Nur so lassen sich strategische Ziele messbar und steuerbar machen.

Die BSC kann zudem auch für den Einsatz im Sozialwesen angepasst werden, z. B. durch die Integration von Perspektiven wie „Klientenwohl“ oder „Wirkung sozialer Arbeit“. Daten sind hier essenziell, um z. B. Zufriedenheit, Teilhabe oder Selbstständigkeit zu messen.

4. Unternehmensstrategie:

Letztlich ist die Datenstrategie ein Instrument zur Umsetzung der übergeordneten Unternehmensstrategie. Sie leitet sich aus dieser ab und unterstützt deren Realisierung. Ein Unternehmen, das beispielsweise Klienten- bzw. Bewohner-Zentrierung setzt, muss in seiner Datenstrategie den Fokus auf Klienten- bzw. Bewohner-Daten legen und dies von der Erhebung über die Analyse bis hin zur Nutzung für personalisierte Angebote auch in der Datenstrategie mitdenken.

In der Altenhilfe kann dies z. B. bedeuten, Bewohner-Daten zu Lebensqualität, Pflegebedarfen oder Angehörigenzufriedenheit zu priorisieren. In der Eingliederungshilfe stehen z. B. Teilhabeziele oder Förderverläufe der Klienten im Fokus.

Aber auch die relevanten Managementansätze sind für jede Einrichtung individuell. Neben den obenstehenden vier klassischen Ansätzen gibt es hier noch weitere Bereiche, die mitgedacht werden sollten. Zum Beispiel im Bereich der Sozialen Arbeit noch das Thema Sozialraumorientierung, welche die Erhebung und Auswertung raumbezogener Daten (z. B. Demografie, Infrastruktur, Bedarfslagen) nötig macht, um passgenaue Angebote zu entwickeln. Daneben kann allgemein im Sozialwesen noch das Thema Qualitätsmanagement genannt werden, denn Daten dienen der Dokumentation, Evaluation und Steuerung von Qualität (z. B. Klienten oder Bewohnerzufriedenheit, Prozesskennzahlen etc.). Auch das Wissensmanagement kann profitieren, denn gerade in der Eingliederungshilfe und Altenhilfe ist viel Wissen implizit vorhanden. Eine gute Datenstrategie kann helfen, dieses Wissen zu strukturieren, zu dokumentieren und zugänglich zu machen. Dies nur als Beispiele, denn diese Aufzählung kann u.a. über Risikomanagement, Personalmanagement, Change-Management und viele andere Themen weitergeführt werden.

Die vorgestellten Managementansätze sollen verdeutlichen, wie vielfältig die Eingangsquellen und Anforderungen an eine Datenstrategie sein können. Im nächsten Kapitel betrachten wir unterschiedliche Herangehensweisen zur Entwicklung einer solchen Strategie, jeweils mit den Stärken und Schwächen sowie anhand praktischer Beispiele.

Zwei mögliche Herangehensweisen zur Erstellung einer Datenstrategie

Bei der Entwicklung einer Datenstrategie lassen sich grundsätzlich zwei Herangehensweisen unterscheiden, die sich auch sinnvoll ergänzen können:

Einerseits der **top-down-Ansatz**, der von den strategischen Zielen des Unternehmens und der Geschäftslogik ausgeht. Hier steht die Frage im Vordergrund, welche realen Entitäten, Prozesse und Beziehungen innerhalb der Organisation abgebildet werden sollen, um mit IT langfristige Ziele wie Effizienz, Qualität, Analysefähigkeit oder KI-Nutzung zu unterstützen.

Andererseits gibt es den **bottom-up-Ansatz**, der pragmatisch ansetzt und zunächst eine Bestandsaufnahme der vorhandenen Daten, Systeme und Schwachstellen vornimmt. Dieser Ansatz ermöglicht es, schnell greifbare Probleme zu identifizieren und erste Verbesserungen umzusetzen.

	Top-down-Ansatz	Bottom-up-Ansatz
Vorteile	+ Klare Ausrichtung an strategischen Zielen und Vision	+ Schnelle Identifikation von Schwachstellen
	+ Systematische Identifikation relevanter Entitäten und Prozesse	+ Quick Wins möglich
	+ Fördert interdisziplinäre Abstimmung (IT, Fachbereiche, Management)	+ Geringer Initialaufwand
	+ Technologieneutrale Analyse im ersten Schritt	+ Gut geeignet für Pilotprojekte oder erste Umsetzungen
Nachteile	- Hoher Initialaufwand für Analyse und Modellierung	- Gefahr der Orientierung an suboptimalen Strukturen
	- Erfordert methodisches Vorgehen und enge Abstimmung	- Möglicherweise fehlende strategische Ausrichtung
		- Risiko von Insellösungen
		- Langfristige Ziele (z. B. KI, Interoperabilität) können vernachlässigt werden
		- Wichtige Entitäten oder Prozesse könnten übersehen werden

Tabelle 1: Vergleich zwischen Top-Down und Bottom-Up Ansatz

Im Folgenden wird zunächst der Top-down-Ansatz näher betrachtet, da er die Grundlage für eine strategisch fundierte Datenarchitektur bildet.

Ausgangspunkt bei diesem Vorgehen ist die Frage, welche übergeordneten Unternehmensziele verfolgt werden.

Nehmen wir als praktisches Beispiel eine Altenhilfeeinrichtung mit folgendem übergeordneten Unternehmensziel:

„Sicherstellung einer qualitativ hochwertigen, menschenwürdigen und wirtschaftlich tragfähigen Pflege und Betreuung älterer Menschen“

Aus diesem zentralen Unternehmensziel und Leitbild lassen sich verschiedene untergeordnete Ziele ableiten, die bereits Anforderungen an Daten und Informationsflüsse stellen könnten und damit relevant für die Datenstrategie sind. Dazu zählen beispielsweise:

- Sicherstellung einer lückenlosen Pflegedokumentation
- Erfüllung gesetzlicher Qualitätsanforderungen (z. B. MDK-Prüfungen)
- Optimierung der Auslastung und effiziente Ressourcensteuerung
- Nachweisfähigkeit gegenüber Aufsichtsbehörden und Kostenträgern
- Entlastung des Pflegepersonals
- Transparente Kommunikation und Wissensmanagement

Eine top-down-basierte Datenstrategie muss somit nicht bei der Technologie beginnen, sondern bei der Frage, welche Informationen für die Zielerreichung benötigt werden.

Basis für jede Datenstrategie ist somit die Überlegung: „Was soll abgebildet werden?“ Es geht darum, ein strukturiertes Datenmodell zu entwickeln, das die realen Entitäten, Prozesse und Beziehungen innerhalb der Organisation widerspiegelt. Diese Modellierung sollte erfolgen, bevor man in die technische Konzeption einsteigt, denn nur so lässt sich sicherstellen, dass die Datenarchitektur langfristig strategisch ausgerichtet ist.

Relevant sind für den Top-down-Ansatz folgende Schritte, welche wir auch hier mithilfe eines Beispiels veranschaulichen wollen. Als untergeordnetes Ziel nehmen wir „Sicherstellung einer lückenlosen Pflegedokumentation“, denn es ist zum einen in der Altenhilfe sehr zentral und trägt zum anderen sowohl rechtliche, qualitative als auch organisatorische Anforderungen mit sich:

- Unternehmensziele und Geschäftsprozesse:** Wie bereits erwähnt, stellt sich zuerst die Frage nach den Unternehmenszielen und den konkret vorhandenen Prozessen, die zur Erreichung dieser Ziele nötig sind. Dabei ist zu klären, welche Dateneinheiten, Attribute und Beziehungen zwischen diesen Einheiten existieren – dies sowohl innerhalb der Organisation als auch im Austausch mit externen Akteuren (z. B. Leistungserbringende, Leistungsträger, Kund:innen/Klient:innen).
 - **Praktisches Beispiel:** Aus dem Ziel, eine lückenlose Pflegedokumentation sicherzustellen, ergeben sich Prozesse wie die Pflegeplanung und -Dokumentation, die Übergaben zwischen Schichten, die Medikamentengabe sowie die Kommunikation mit Ärzten und Angehörigen. Zudem sind Entitäten wie Pflegekraft, Bewohner:innen, Ärzt:innen, Angehörige, Pflegemaßnahmen und Medikamentengaben zu berücksichtigen, welche in Beziehungen zueinander stehen und die Grundlage für ein strukturiertes Datenmodell bilden.
- Datenanforderungen:** Im nächsten Schritt werden aus den identifizierten Zielen und Prozessen konkrete Anforderungen an die Daten abgeleitet. Dazu gehört, sich zu fragen, welche Daten benötigt werden, welche Qualitätsanforderungen bestehen, welche Anforderungen sich aus regulatorischen Vorgaben ergeben (bei Themen wie Datenschutz und Datensicherheit kann und sollte auf ein Informations- oder IT-Sicherheitskonzept verwiesen werden) und welche analytischen Anforderungen es gibt, denn wichtig für das Controlling von Geschäftszielen ist immer die Auswertbarkeit der Daten. Dies nicht nur in Bezug auf Kennzahlen, sondern auch im Hinblick auf die Möglichkeiten, welche KI bietet. Eine ganzheitliche Datenstrategie berücksichtigt daher immer auch **analytische Anforderungen**.
 - **Praktisches Beispiel:** Aus dem Ziel der lückenlosen Pflegedokumentation ergeben sich dann zum einen Qualitätsanforderungen an die Vollständigkeit, Aktualität und Nachvollziehbarkeit (wer hat was wann dokumentiert?) und zum anderen regulatorische Anforderungen wie z. B. die

Einhaltung der Dokumentationspflichten nach SGB XI und der Datenschutz gemäß DSGVO. Darüber hinaus bestehen analytische Anforderungen an die Auswertbarkeit.

- c. **Ableitung konkreter Maßnahmen und Umsetzungsschritte:** Auf Grundlage der zuvor identifizierten Geschäftsprozesse und -ziele werden nun konkrete Anforderungen an Daten, Systeme und Zuständigkeiten abgeleitet. Am Ende dieses Vorgehens steht dann die Auswahl geeigneter Technologien und Werkzeuge für die technologische Umsetzung, um möglichst lange technologieoffen zu bleiben. Projekte, in welche die definierten Maßnahmen umgesetzt werden, sind dann oft (aber nicht immer) klassische Projekte nach dem „Wasserfallprinzip“, da die Anforderungen durch die ausführliche Analyse (upfront design) gut definierbar sind.
 - o **Praktisches Beispiel:** Basierend auf den strategisch abgeleiteten Anforderungen an eine lückenlose Pflegedokumentation wird ein zentrales digitales Dokumentationssystem konzipiert, das technologieoffen und skalierbar ist. Durch die hohen Kosten und Schulungsaufwände erfolgt die Einführung aber deutlich später als eigentlich in der Einrichtung benötigt. Die mobile Erfassung wird von Beginn an organisationsweit mitgedacht, inklusive Spracheingabe und barrierearmer Bedienung. Validierungsregeln und Pflichtfelder werden einheitlich definiert, um eine konsistente Datenqualität sicherzustellen. Die Speicherung erfolgt in einem standardisierten, maschinenlesbaren Format, das sowohl interne Auswertungen als auch zukünftige KI-gestützte Analysen ermöglicht.

Ein weiteres mögliches Vorgehen ist der Bottom-up-Ansatz, der bei der Praxis beginnt und daraus eine Strategie ableitet.

Im Gegensatz zum Top-down-Ansatz, der von strategischen Zielen ausgeht, setzt der Bottom-up-Ansatz an der bestehenden Realität im Unternehmen an. Er beginnt mit einer detaillierten Analyse der vorhandenen Datenbestände, Systeme und Prozesse und der Probleme mit diesen. Ziel ist es, aus der Praxis heraus konkrete Schwachstellen, Verbesserungspotenziale und bereits funktionierende Strukturen zu identifizieren und daraus schrittweise eine tragfähige Datenstrategie zu entwickeln.

Auch hier lässt sich das Vorgehen an einem Beispiel pro Schritt besser greifbar machen. Hier haben wir aber, im Gegensatz zum Top-down-Konzept, kein zentrales Ziel, sondern folgende Ausgangssituation:

In einer Altenpflegeeinrichtung fällt auf, dass es regelmäßig zu Doppel- und Fehleinträgen in der Pflegedokumentation kommt, insbesondere bei der Medikamentengabe. Dies führt zu Unsicherheiten im Pflegealltag und erhöhtem Aufwand bei der Nachdokumentation.

Relevant sind bei diesem Vorgehen folgende Schritte:

- a. **Bestandsaufnahme und Analyse der Ist-Situation:** Im ersten Schritt erfolgt eine strukturierte Erhebung der vorhandenen Daten, Schnittstellen und Systeme. Dabei werden auch bestehende Probleme wie Dateninkonsistenzen, Medienbrüche oder redundante Datenhaltung sichtbar.
 - o **Praktisches Beispiel:** Bei der Analyse wird festgestellt das es mehrere parallele Dokumentationssysteme (Papier, Excel, Pflegesoftware) gibt und keinen Standard. Dort werden Daten mehrfach erfasst, aber nicht synchronisiert. Eine Auswertbarkeit ist nicht gegeben. Die Pflegekräfte dokumentieren zudem nicht teilweise zeitverzögert oder doppelt.
- b. **Ableitung von Anforderungen aus der Praxis:** Die im Bottom-up-Ansatz gewonnenen Erkenntnisse dienen als Grundlage, um Anforderungen realitätsnah zu formulieren. In welchen Geschäftsprozessen fließen die Daten von wo nach wo? So entsteht eine Datenstrategie, die auf tatsächlichen Bedarfen

basiert und vorhandene Strukturen sinnvoll weiterentwickelt. Im Gegensatz zum Top-down-Ansatz ist in diesem Schritt dafür aber oft keine Technologieneutralität mehr gegeben.

- **Praktisches Beispiel:** Aus diesen Beobachtungen ergeben sich direkt konkrete Anforderungen wie z. B. die Einführung eines einheitlichen digitalen Dokumentationssystems mit Pflichtfeldern und Validierungsregeln sowie mobile Erfassungsmöglichkeiten für Pflegekräfte.
- a. **Umsetzung der Anforderungen in konkreten Projekten:** Die identifizierten Maßnahmen werden in Form von Projekten umgesetzt. Dies kann bei diesem Ansatz teilweise sehr schnell erfolgen. Eine Möglichkeit, wenn Zeit ein ausschlaggebender Faktor ist, ist hier dann auch der Start von Pilotprojekten, die iterativ weiterentwickelt werden.
 - **Praktisches Beispiel:** Basierend auf den Anforderungen, die sich aus der Analyse ergeben haben, wird kurzfristig ein zentrales digitales Dokumentationssystem eingeführt, das auf die bestehende Infrastruktur abgestimmt ist und daher schnell live gehen kann. Hierfür wird eine bereits vorhandene Lösung kostengünstig erweitert, obwohl sie langfristig nicht skalierbar ist. Die mobile Erfassung wird nur auf bestimmten Stationen pilotiert, da nicht alle Endgeräte kompatibel sind. Die Validierungsregeln werden individuell konfiguriert, was zu unterschiedlichen Datenqualitätsstandards zwischen den Bereichen führt.

Beide Ansätze (Top-down und Bottom-up) führen im Idealfall zum Ziel. Doch sie unterscheiden sich grundlegend in ihrer Herangehensweise, ihren Stärken und ihren potenziellen Risiken. Aufmerksame Leser:innen werden bereits die Unterschiede des jeweils letzten Praxisbeispiels in den beiden Ansätzen c) erkannt haben. Das Ergebnis beim Bottom-Up-Ansatz ist hier bewusst negativ, um die potenziellen Risiken klarer zu machen. Ist man sich dieser Risiken nicht bewusst, besteht beim Bottom-up-Ansatz die Gefahr, dass strategische Zielsetzungen aus dem Blick geraten und sich Insellösungen verfestigen, die langfristig schwer integrierbar sind.

Eine Datenstrategie entwickelt man in der Regel nicht „auf der grünen Wiese“, sondern auf Basis einer bestehenden Applikationslandschaft mit zugehörigen Datenbeständen und technischer Infrastruktur. Dies macht es schwer, möglichst oft im Top-down-Verfahren alles möglichst optimal neu zu denken.

In der Praxis hat sich daher gezeigt, dass eine Kombination aus Top-down- und Bottom-up-Ansatz sinnvoll sein kann: Während der Top-down-Ansatz für strategische Orientierung und langfristige Zieldefinition sorgt, liefert der Bottom-up-Ansatz wertvolle Einblicke in bestehende Strukturen und ermöglicht schnelle, praxisnahe Verbesserungen.

Damit diese Anforderungen jedoch erfolgreich umgesetzt werden können, ist eine enge Verzahnung mit der IT-Infrastruktur unerlässlich. Denn technologische Maßnahmen wie etwa Datenmigrationen, die Einführung neuer Plattformen oder die Umsetzung von Analyse- und KI-Funktionalitäten sind nur dann realisierbar, wenn die zugrunde liegende Infrastruktur dies auch unterstützt. Aspekte wie Datensicherheit, Verfügbarkeit, Skalierbarkeit und Performance müssen daher von Anfang an mitgedacht werden. Eine ganzheitliche Datenstrategie betrachtet deshalb Applikationen, Daten und Infrastruktur stets im Zusammenspiel, wenngleich die Infrastrukturthemen nicht in der Datenstrategie selbst definiert werden.

Anforderungsaufnahme

Unabhängig davon, ob eine Datenstrategie eher aus der Praxis heraus entwickelt (Bottom-up) oder strategisch geplant (Top-down) wird, ist die strukturierte Aufnahme von Anforderungen der zentrale Erfolgsfaktor. Nur wenn klar ist, welche Daten für welche Zwecke benötigt werden, welche Prozesse unterstützt werden sollen, und welche Rahmenbedingungen gelten, kann eine Datenstrategie wirksam und nachhaltig gestaltet werden.

Dabei geht es nicht nur um technische Spezifikationen, sondern vor allem um das gemeinsame Verständnis zwischen Fach- und IT-Abteilungen. Ein weiterer wichtiger Aspekt bei der Anforderungsaufnahme ist daher die Übersetzung der Fachsprache in IT-Sprache. Dies stellt sicher, dass die technischen Anforderungen klar und verständlich sind und die IT-Abteilung die Geschäftsprozesse und Ziele effektiv unterstützen kann.

Dabei ist es essenziell, die Rollen und Verantwortlichkeiten zwischen IT- und Fachabteilungen im Kontext der Datenstrategie klar zu definieren. In vielen Organisationen werden Fachapplikationen direkt in den Fachabteilungen administriert. Es sollte daher bei Bedarf auch die Einrichtung eines zentralen Daten- und Anwendungsmanagements geprüft werden.

Vor dem Hintergrund ergeben sich verschiedene Möglichkeiten zur Umsetzung:

Workshops und Schulungen: Gemeinsame Veranstaltungen, bei denen Fachabteilungen und IT-Abteilung zusammenkommen, um gegenseitiges Verständnis zu fördern. In sozialen Einrichtungen treffen dabei häufig Mitarbeitende mit unterschiedlichen kulturellen Prägungen, beruflichen Hintergründen und digitalen Erfahrungswerten aufeinander. Um eine wirksame Anforderungsaufnahme zu ermöglichen, sollten Workshops so gestaltet sein, dass alle Teilnehmenden unabhängig von Vorwissen oder Herkunft aktiv eingebunden werden können. Dies erfordert eine passende Moderation, klare Sprache sowie praxisnahe Methoden, die unterschiedliche Perspektiven wertschätzen und sichtbar machen.

1. **Interdisziplinäre Teams:** Bildung von Teams, die sowohl Fach- als auch IT-Expertise vereinen, um einen kontinuierlichen Austausch sicherzustellen. Diese Teams können projektbezogen organisiert sein, zum Beispiel bei der Einführung eines neuen Pflegemanagementsystems, oder dauerhaft im Rahmen eines zentralen Daten- und Anwendungsmanagements bestehen. Denn auch ein zentrales Daten- und Anwendungsmanagement kann wertvolle Impulse zur Anpassung und Weiterentwicklung der Datenstrategie liefern, da es bereichsübergreifende Anforderungen, Schnittstellen und Nutzungsmuster systematisch erfasst und bewertet. Entscheidend ist, dass alle Teams, die zur Umsetzung und Weiterentwicklung der Datenstrategie beitragen, interdisziplinär aufgestellt sind, unabhängig von ihrer organisatorischen Verankerung oder Laufzeit.
2. **Anforderungsmanagement:** Der Einsatz eines geeigneten Projektmanagement-Frameworks bzw. Vorgehensmodells mit starkem Fokus auf Anforderungen, z. B. das „V-Modell XT“ oder „PRINCE 2“, hilft, Anforderungen strukturiert zu erfassen. Ein agiles Framework, welches ein leichtgewichtiges Vorgehensmodell zum Anforderungsmanagement mitbringt, wäre z. B. „Scrum“. Es beinhaltet die kontinuierliche Erfassung und Priorisierung von Anforderungen in Form von User Stories. Darüber hinaus bietet „Design Thinking“ ebenfalls iterative Methoden, die sich sehr gut zur nutzerorientierten Anforderungserhebung einsetzen lassen. Strukturierte und agile Modelle können sich, klug eingesetzt, dabei ergänzen.

Zudem sollte das Thema Barrierefreiheit bereits bei der Anforderungsaufnahme berücksichtigt werden inkl. der Beteiligung von Menschen mit den jeweiligen Beeinträchtigungen. Hierfür gibt es mittlerweile auch mit dem Barrierefreiheitsstärkungsgesetz (BFSG) und der EU-Richtlinie 2019/882 zunehmend verbindliche Anforderungen. Daten müssen also so bereitgestellt werden, dass sie für alle zugänglich sind. In Bezug auf die Datenstrategie könnte dies die Verwendung strukturierter, maschinenlesbarer Formate sowie die Zugänglichkeit der Systeme durch Schnittstellen, insbesondere zur Datenerfassung und -auswertung mit assistiven Technologien sein.

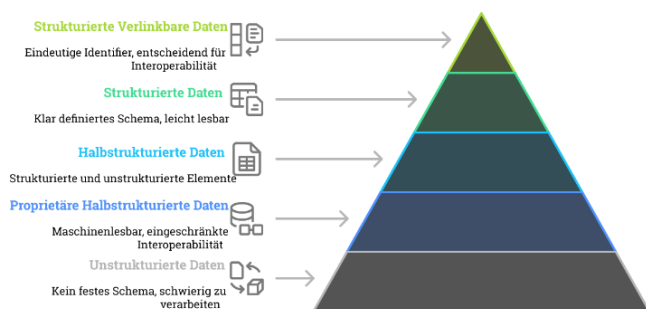
Auf dieser Basis kann ein **Bild der „Daten-Zukunft“** definiert werden. Vorab sind jedoch einige Rahmenbedingungen zu beachten. Im Bereich der Pflege gelten neben der Datenschutz-Grundverordnung (DSGVO) der EU auch spezifische nationale Regelungen. Die DSGVO legt grundlegende Prinzipien wie Datenminimierung, Zweckbindung und Transparenz fest und schützt insbesondere Gesundheitsdaten als besonders sensible personenbezogene Daten. Ergänzend dazu regeln das Sozialgesetzbuch (SGB) I und X den Umgang mit Sozialdaten, insbesondere im Rahmen der sozialen Pflegeversicherung. Das SGB XI enthält darüber hinaus spezielle Datenschutzvorgaben für Pflegeeinrichtungen und Pflegekassen. Diese Vorschriften konkretisieren das Sozialgeheimnis und stellen sicher, dass personenbezogene Daten nur unter engen gesetzlichen Voraussetzungen verarbeitet werden dürfen. Ein zentrales Instrument zur Umsetzung dieser Anforderungen ist aber dann ein umfassendes Informations- bzw. IT-Sicherheitskonzept auf, welches die Datenstrategie verweisen sollte. Dieses getrennte Dokument beschreibt dann die technischen und organisatorischen Maßnahmen, mit denen Pflegeeinrichtungen die Sicherheit der Datenverarbeitung gewährleisten, etwa durch Zugriffskontrollen, Verschlüsselung, regelmäßige Schulungen und Notfallmanagement. Zusätzlich ist der geplante AI-Act der EU relevant, der den Einsatz von Künstlicher Intelligenz auch im Pflegebereich regulieren und ethisch absichern soll.

Wichtig für ein umfassendes Verständnis sind zudem Grundlagen im Bereich der Datenarchitektur. Dazu mehr in den folgenden Kapiteln.

2. Datenqualität

Im Zentrum jeder erfolgreichen Datenstrategie steht die Qualität der zugrunde liegenden Daten. Denn nur wenn Daten konsistent und maschinenlesbar sind, können sie problemlos genutzt werden. Datenqualität beschreibt dabei nicht nur die Korrektheit oder Vollständigkeit einzelner Datensätze, sondern umfasst auch deren Strukturierungsgrad und Interoperabilität. Gerade im Sozial- und Gesundheitswesen, wo zahlreiche Akteure und Systeme miteinander kommunizieren müssen, ist dies von zentraler Bedeutung.

Datenqualitätshierarchie



Übergreifend lässt sich aber definieren: Daten von hoher Qualität sind genau, vollständig, konsistent, aktuell und haben eine einheitliche Struktur¹.

Betrachtet man die Qualität aus dem Blickwinkel der Interoperabilität, ist primär die Verfügbarkeit der enthaltenen Informationen wichtig. Diese lässt sich in folgende Stufen einteilen:

1. **Unstrukturierte Daten** sind Formate wie Textdateien, Bilder (JPEG, PNG), PDFs. Diese Daten haben kein festes Schema und sind daher am schwierigsten maschinell zu verarbeiten. Sie erfordern oft zusätzliche Verarbeitungsschritte wie OCR-Analyse. Bisher mussten beispielsweise Rechnungen gescannt und per OCR ausgelesen werden. Ein Beispiel wäre eine eingescannte Bilddatei eines Schreibens, auf welchem eine Klienten-Nummer mit angegeben ist.

¹ Weitere Infos zur Datenqualitätshierarchie finden Sie zum Beispiel unter: <https://5stardata.info/>

2. **Proprietäre Halbstrukturierte Daten** sind Formate wie XLSX oder DOCX. Diese Formate können maschinenlesbar sein, aber ihre Interoperabilität ist oft eingeschränkt, da sie keine selbsterklärende Struktur aufweisen oder spezifische Software oder Lizenzvereinbarungen erfordern. Ein Beispiel wäre eine Word Datei mit einem Schreiben, in welchem eine Klienten-Nummer mit angegeben ist.
3. **Halbstrukturierte Daten** sind Formate wie HTML und E-Mails. Diese Daten enthalten sowohl strukturierte als auch unstrukturierte Elemente. Sie sind teilweise maschinenlesbar und erfordern oft spezielle Parser, um die relevanten Informationen zu extrahieren. Ein Beispiel wäre eine E-Mail im HTML-Format in welcher eine Klienten-Nummer angegeben ist.
4. **Strukturierte Daten** sind Formate wie JSON, CSV, XML und praktisch sämtliche in relationalen Datenbanken wie MS-SQL oder MySQL gespeicherte Daten. Diese Daten sind in einem klar definierten Schema organisiert, was ihre Auslesbarkeit und Maschinenlesbarkeit erleichtert. Als Beispiel wäre z. B. in FHIR (HL7) wäre die Klientennummer z. B. im Patient-Ressourcentyp maschinenlesbar enthalten.
5. **Strukturierte verknüpfbare Daten** liegen in Datenbanken und darauf basierenden Anwendungen wie Klientenmanagement- oder DMS-Systemen vor. Für eine systemübergreifende Interoperabilität sind eindeutige Identifier entscheidend um Datensätze in unterschiedlichen Anwendungen zuordnen zu können. Ein Beispiel wäre, dass eine eindeutige Klienten-Nummer in allen Systemen verfügbar ist in welchen Informationen zu Klienten erfasst und strukturiert gespeichert werden.

Praxisbeispiel zur Erfassung von strukturierten Daten:

In einer Pflegeeinrichtung werden Medikamentengaben bislang handschriftlich dokumentiert und abgelegt (Stufe 1). Diese unstrukturierten Daten werden durch die Einführung eines digitalen Medikationsmoduls ersetzt, das strukturierte Eingabefelder nutzt (Stufe 5). Dadurch können Fehler reduziert und Auswertungen erstellt werden.

Doch Struktur allein reicht nicht aus: Damit Daten auch systemübergreifend genutzt werden können, braucht es gemeinsame Austauschstandards. Beispiele für bereits im Sozialwesen existierende Austauschstandards, welche interoperablen Datenaustausch zwischen Akteuren ermöglichen, sind:

1. eGK, e-PA, E-Rezept, Gesundheits-ID
2. Einheitlicher eMedikationsplan (eMP)
3. KIM (Kommunikation im Medizinwesen)
4. HL7-basierte Formate (z. B. eArztbrief, eWundbericht, Reha-Entlassbericht, Krankenförderung)
5. PIO (Überleitungsbogen)
6. MIO (digitaler Impfpass)

Diese Standards nutzen in der Regel strukturierte technische Austauschformate wie XML, JSON oder CSV. Dabei ist zu beachten, dass einige Dienste, wie z. B. KIM, auch unstrukturierte Inhalte (z. B. Freitext in E-Mails) enthalten können und daher nur teilweise strukturiert sind.

Bei der Rechnungsverarbeitung sind ab 2027 folgende elektronische Formate vorgeschrieben:

1. ZUGFeRD
2. XRechnung

Datenqualität ist entscheidend für die Nutzung und den Austausch von Informationen im Sozial- und Gesundheitswesen. Je strukturierter und interoperabler die Daten, desto besser sind sie maschinell nutzbar.

3. Strategische Entscheidungen zur Speicherung der Datenstrukturen

Nach der Datenqualität stellt sich auch die Frage, wie und wo gespeichert werden soll. Bei der Wahl zwischen zentraler und verteilter Datenstruktur sind folgende Aspekte zu berücksichtigen:

1. **Zentral oder dezentral:** Zentrale Datenstrukturen ermöglichen eine einheitliche Datenhaltung und einfachere Governance. Dezentrale Strukturen bieten mehr Flexibilität und eine bessere Anpassung an lokale Anforderungen. Sie können (bei entsprechender technischer Umsetzung) auch eine höhere Ausfallsicherheit ermöglichen. In der Praxis hängt dies jedoch stark von der jeweiligen Infrastruktur, Wartung und Backup-Strategie ab.
2. **Cloud oder On-Premise:** Cloud-basierte Datenlösungen bieten in der Regel bessere Skalierbarkeit, geringere Einstiegskosten und moderne Sicherheitsmechanismen. On-Premise-Ansätze ermöglichen hingegen mehr Kontrolle über sensible Daten und können regulatorische Anforderungen besser erfüllen.
3. **USA oder Europa:** Die Wahl der Plattform oder des Anbieters hat auch geopolitische Relevanz. Zwar betreiben US-Anbieter wie Amazon oder Microsoft Rechenzentren in Europa, die der DSGVO unterliegen, dennoch bestehen rechtliche Unsicherheiten hinsichtlich des Zugriffs durch ausländische Behörden. Europäische Initiativen wie das EU-Projekt 8ra setzen daher auf digitale Souveränität und europäische Kontrolle. Dies ist ein zunehmend wichtiger Aspekt angesichts globaler Spannungen und Datenschutzbedenken.

Hinweis zum Zusammenspiel zwischen Datenstruktur und IT-Infrastruktur:

Auch wenn hier der Fokus auf Datenstrukturen liegt, ist die zugrunde liegende IT-Infrastruktur stets ein entscheidender Faktor. Sie bildet die technische Basis, auf der Daten gespeichert, verarbeitet und ausgetauscht werden. Entscheidungen zur Datenstruktur sollten daher immer im Zusammenspiel mit der IT-Infrastruktur getroffen werden.

Zentrale / verteilte Datenstrukturen beeinflussen, wie Daten gespeichert und verarbeitet werden. Zentrale Systeme bieten Einheitlichkeit, dezentrale mehr Flexibilität. Die Wahl zwischen Cloud und On-Premise sowie zwischen US- und EU-Anbietern betrifft Datenschutz, Kontrolle und digitale Souveränität.

Insbesondere angesichts wachsender Datenmengen und unterschiedlicher Strukturierungsgrade ist aber auch die konkrete Speicherung der Daten, also verteilt oder zentral, ein wichtiger Faktor.

Verteilte vs. zentrale Informationsspeicherung

Bei der Erstellung einer Datenstrategie stehen Organisationen der Sozialwirtschaft vor der Herausforderung, große Mengen an Informationen effizient zu speichern, zu verarbeiten und nutzbar zu machen. Die Entscheidung für eine zentrale oder verteilte Speicherung ist dabei eng mit der Art der vorhandenen Daten verknüpft. Denn ein erheblicher Teil der Daten in Organisationen der Sozialwirtschaft ist un- oder halbstrukturiert (siehe Kapitel 2). Gleichzeitig besteht ein wachsender Bedarf an Echtzeiteinblicken bei begrenzten Ressourcen. In diesem Kontext ist der Einsatz klassischer, strukturierter Speicherlösungen wie **Data Warehouses** nicht immer wirtschaftlich oder praktikabel. Zwar ermöglichen sie schnelle und komplexe Auswertungen und entlasten andere Systeme, erfordern jedoch eine aufwendige Datenmodellierung, Datenhaltung und Pflege.

Data Lakes hingegen speichern Rohdaten unabhängig vom Verwendungszweck und eignen sich daher besser für explorative Analysen und Big-Data-Szenarien. Insbesondere in heterogenen, dynamischen Datenlandschaften.

Exkurs: Data Warehouse vs. Data Lake

Ein **Data Warehouse** ist eine zentralisierte Datenbank, die speziell für strukturierte Daten und analytische Auswertungen konzipiert ist. Bevor Daten dort gespeichert werden, müssen sie aber in ein festgelegtes Format gebracht werden. Das sorgt für eine hohe Datenqualität und ermöglicht schnelle, komplexe Abfragen. Diese Struktur eignet sich besonders für standardisierte Berichte und Business-Intelligence (BI)-Anwendungen, erfordert jedoch hierzu einen höheren Aufwand.

Ein **Data Lake** hingegen verfolgt einen flexibleren Ansatz: Hier werden Daten im Rohformat gespeichert. Die Daten werden erst bei der Auswertung in ein passendes Format gebracht. Dadurch eignet sich ein Data Lake besonders für explorative Analysen, Machine Learning und Big-Data-Szenarien.

Ein zentrales Problem vieler Organisationen in der Sozialwirtschaft ist eine **historisch gewachsene, fragmentierte Anwendungslandschaft** mit zahlreichen, teils isolierten Fachanwendungen. Eine Konsolidierung dieser Systeme ist daher ein notwendiger erster Schritt. Gleichzeitig stellt auch die spätere Abhängigkeit von einem einzelnen **führenden System** eine Herausforderung dar: Dieses muss nicht nur Datenqualität, Konsistenz und Aktualität gewährleisten, sondern kann (je nach technischer Offenheit) die Integration neuer Technologien erschweren. Dies kann die Innovationsfähigkeit einschränken und dazu führen, dass auf sich schnell ändernde Anforderungen nicht flexibel reagiert werden kann.

4. Schnittstellen

Aufbauend auf den Überlegungen zur zentralen oder verteilten Speicherung von Datenstrukturen (Kapitel 2) wird deutlich, dass die technische Umsetzung einer vernetzten Datenlandschaft maßgeblich von der Offenheit der Systeme abhängt. Insbesondere in dezentralen Systemarchitekturen sind **APIs** (Application Programming Interfaces) ein wichtiger Schlüssel zur nahtlosen Integration verschiedener heterogener Systeme.

Idealerweise handelt es sich hierbei um eine **Public API**. Eine Public API ist offen und kann von jedem externen Entwickler oder Unternehmen genutzt werden. Bei APIs gibt es wiederum unterschiedliche Standards. Die wichtigsten sind **REST** und **SOAP**. REST bietet gegenüber SOAP mehrere Vorteile: Es ist weniger komplex und leichter zu implementieren, da es auf HTTP-Standards basiert und in der Regel JSON verwendet, was weniger Overhead hat. REST ist schneller, effizienter, und weniger fehleranfällig, da es zustandslos ist.

Exkurs: REST-Schnittstellen ermöglichen den standardisierten Datenaustausch zwischen Systemen. Sie nutzen sogenannte „HTTP-Methoden“ wie GET (Daten abrufen), POST (Daten senden), PUT (Daten aktualisieren) und DELETE (Daten löschen). Jede Ressource, z. B. ein Klient, ist über eine eindeutige URL erreichbar.

Ein Beispiel: In einem Klientenverwaltungssystem könnte eine REST-Anfrage zum Abruf der Stammdaten eines bestimmten Klienten so aussehen:

```
GET https://api.klientenverwaltung.de/klienten/12345
```

Diese Anfrage ruft die Stammdaten des Klienten mit der Nummer 12345 ab. Zum Beispiel Name, Geburtsdatum, Adresse oder Betreuungsstatus. Die Antwort erfolgt in der Regel im JSON-Format und kann von so von anderen Systemen leicht weiterverarbeitet werden.

APIs ermöglichen es also, dass verschiedene Softwareanwendungen, auch ohne permanente manuelle Eingriffe, miteinander kommunizieren und Daten auszutauschen. Dies ermöglicht Interoperabilität und gewährleistet so ein nahtloses Zusammenspiel. Das bedeutet: Gute Schnittstellen sind also in vielen Fällen wichtiger als „ein“ zentrales System. Gleichzeitig gilt aber auch: Schnittstellen sind kein Allheilmittel. Gerade bei komplexen, stark integrierten Prozessen (etwa in der Personalwirtschaft oder im Finanzwesen) stoßen API-basierte Lösungen an Grenzen. Hier kann ein zentrales System Vorteile bieten, etwa durch konsistente Datenmodelle und durchgängige Prozesslogik.

Die Entscheidung für Schnittstellen oder zentrale Systeme sollte daher immer kontextabhängig und mit Blick auf die konkreten Anforderungen getroffen werden.

Ein Beispiel für den Einsatz von APIs zur Erreichung von Interoperabilität findet sich im Dokumenten- und Klientenmanagement. Diese Systeme enthalten strukturierte Metadaten, etwa zur Klassifizierung von Klienten und / oder Dokumenten.

Praxisbeispiel zum Zugriff auf ein DMS aus dem Klientenverwaltungssystem:

Eine Einrichtung nutzt ein Klientenverwaltungssystem und ein separates DMS. Über eine REST-API werden automatisch alle Dokumente, die einem Klienten zugeordnet sind, im DMS verschlagwortet und mit der Klientennummer verknüpft. So können Mitarbeitende direkt aus der Klientenakte auf relevante Dokumente zugreifen.

Über offene APIs können diese Informationen so abgerufen und weiterverarbeitet werden. Dies ermöglicht nicht nur systemübergreifenden Zugriff, sondern (wie im nächsten Abschnitt beschrieben) auch zentrale Auswertungen.

5. Zentrale Auswertung

Wie vorab erwähnt, ist eine zentrale Datenspeicherung in den meisten Fällen keine ausreichende Lösung: Es werden nie alle Daten gebündelt werden können. Um übergreifende Auswertungen zu ermöglichen, bieten sich folgende Optionen an:

- Eine Lösung kann beispielsweise ein Enterprise-Search-Tool sein, welches anwendungsübergreifend Informationen aus verschiedenen Systemen bündelt und über einen zentralen Index in Form einer „Google-Suche“ erreichbar macht.
- Ein anderer Ansatz sind Business-Intelligence-Tools, die Daten aus verschiedenen Quellen visuell aufbereiten und analysierbar machen. Dies ermöglicht interaktive Dashboards, Berichte und Datenvisualisierungen zur Unterstützung datenbasierter Entscheidungen.
- Ein weiterer Ansatz kann sein, Dokumente und Informationen zu diesen Dokumenten aus verschiedenen Anwendungen (Pflegemanagement, Finanzen, Personalmanagement etc.) in einem DMS zu bündeln. Dort besteht im ersten Schritt die Möglichkeit, einen Volltextindex über diesen Bestand zu erzeugen und zu durchsuchen. Damit hat man zudem im DMS einen abgegrenzten Bereich mit hochwertigen, strukturierten Dokumenten. Mithilfe von souverän gehosteten KI-Modellen besteht nun die Möglichkeit, in diesem Dokumentenbestand einen „Datenschatz“ zu heben und in Folge die enthaltene Information durch einfache Anfragen in natürlicher Sprache („Prompts“) den Mitarbeitenden zur Verfügung zu stellen.
- Auch bei großen, strukturierten Datenbeständen, wie etwa einer Klientendatenbank mit 10.000 Klienten und Millionen von Einzeleinträgen, ist eine Auswertung möglich, wenn bestimmte Voraussetzungen erfüllt sind. Grundlegend stellt sich die Frage, wie diese große Menge an Informationen effizient und intelligent zugänglich gemacht werden kann. Zwei grundsätzliche Ansätze stehen dabei zur Verfügung: der Einsatz vorberechneter Indizes (Prebuild-Index) und der dynamische Zugriff über Schnittstellen.
 - Ein vorab erstellter Prebuild-Index ist besonders dann sinnvoll, wenn häufig wiederkehrende Suchanfragen durchgeführt werden sollen. Hierbei werden die Daten vorab analysiert, strukturiert und in einem durchsuchbaren Format abgelegt. Der Vorteil liegt in der hohen Geschwindigkeit und Effizienz: Die KI kann direkt auf vorbereitete Strukturen zugreifen, ohne jedes Mal auf die Originaldaten zugreifen zu müssen. Das ist besonders bei Dokumenten oder statischen Daten sinnvoll, die sich selten ändern.
 - Demgegenüber steht der dynamische Zugriff über Schnittstellen, bei dem die KI bei jeder Anfrage in Echtzeit auf die zugrunde liegenden Systeme zugreift. Etwa auf eine operative Klientendatenbank. Dieser Ansatz bietet den Vorteil, dass stets mit den aktuellen Daten gearbeitet wird. Allerdings ist er

abhängig von der Leistungsfähigkeit der Datenbank und der Qualität der Schnittstellen. Zudem kann die Verarbeitung komplexer Abfragen bei sehr großen Datenmengen zu Performanceproblemen führen.

Anmerkung: Aufmerksamen Leser:innen wird hier das gleiche Dilemma (bzw. die damit verbundenen Vor- und Nachteile) aufgefallen sein wie beim Kapitel „Verteilte vs. zentrale Informationsspeicherung“. Insgesamt zeigt sich aber: Je nach Art der Daten und Fragestellung ist ein hybrider Ansatz am sinnvollsten. Statische oder häufig genutzte Informationen können in einem Index vorgehalten werden, während dynamische, aktuelle Daten über Schnittstellen oder ein Data Warehouse zugänglich gemacht werden. So entsteht ein leistungsfähiges Gesamtsystem, das sowohl Geschwindigkeit als auch Aktualität und Flexibilität vereint.

Praxisbeispiel zur KI-gestützten Analyse:

Die Einrichtung führt ein KI-Modell ein, um aus archivierten Pflegeberichten häufige Risikofaktoren für Stürze zu identifizieren. Die Ergebnisse fließen dann in die Pflegeplanung ein und helfen, präventive Maßnahmen zu gestalten.

Zentrale Auswertungen gelingen somit nicht allein durch zentrale Speicherung, sondern durch intelligente Verknüpfung verteilter Daten über geeignete Schnittstellen und Analysewerkzeuge; unter anderem mit KI-Unterstützung.

6. Entwicklung von Szenarien

Die Definition und Analyse verschiedener Szenarien ermöglicht fundierte strategische Entscheidungen und hilft, Maßnahmen unter verschiedenen Bedingungen besser einzuschätzen. Zudem fördert die Einbindung dieser Szenarien die Zusammenarbeit und das Engagement der Stakeholder, was zu einer umfassenderen und robusteren Datenstrategie führt.

Nachfolgend werden fünf exemplarische Szenarien skizziert.

Szenario 1: Kundenzentrierung, Angebotserweiterung, Stabilisierung am Markt

Ausgangslage und Potenziale:

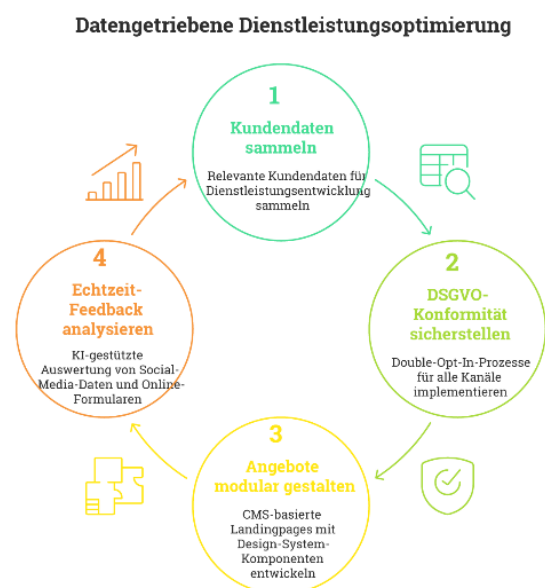
Sozialwirtschaftliche Einrichtungen stehen vor der Herausforderung, bedarfsgerechte Dienstleistungen bei gleichzeitigem Kostendruck anzubieten.

Eine integrierte Datenstrategie ermöglicht hier:

- **Präzise Zielgruppenansprache** durch analysierte Interessent:innen- und Kund:innenbedürfnisse
- **Rechtskonforme Kommunikation** mittels Double-Opt-In-Verfahren
- **Schnelle Marktanpassung** durch modular erweiterbare Angebotsportale

Spezifische Anforderungen:

- **360°-Kunden- und Interessentendaten:** Integration von Kontaktdaten, Nutzungsverhalten und Feedback
- **DSGVO-konforme Erfassung:** Implementierung von Double-Opt-In-Prozessen für alle Kommunikationskanäle
- **Modulare Angebotsarchitektur:** CMS-basierte Landingpages mit Design-System-Komponenten

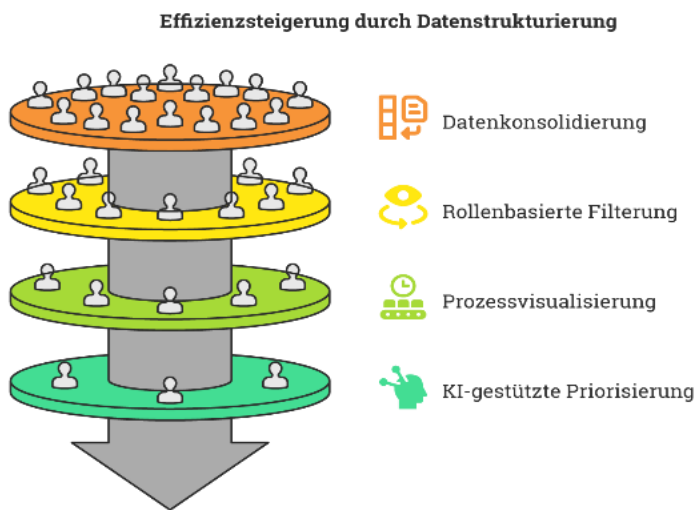


- **Echtzeit-Marktfedback:** KI-gestützte Auswertung von Social-Media-Daten und Online-Formularen

Datenbezogene Fragestellungen:

- Welche Kunden- und Interessentendaten sind für unsere Dienstleistungsentwicklung wirklich relevant?
- Wie implementieren wir einen DSGVO-konformen Double-Opt-In-Prozess über alle Kanäle?
- Welche CMS-Funktionen benötigen wir für schnell änderbare Landingpages?
- Wie können wir externe Tools (z. B. HubSpot, Mailchimp, ...) nahtlos integrieren?

Szenario 2: Mitarbeitermangel: Effizienzen durch Daten erhöhen



Ausgangslage und Potenziale:

Personalmangel führt zu Überlastung bei Dokumentation, Koordination und Klientenbetreuung. Gleichzeitig liegen wertvolle Daten in Excel, Word und E-Mails verstreut, die durch intelligente Strukturierung Arbeitszeit einsparen könnten.

Potenziale:

- Reduktion redundanter Dateneingaben durch Konsolidierung, Struktur und Automatisierung
- Steigerung der Prozessqualität durch klare Handlungsprotokolle
- Freisetzung von Kapazitäten für Kernaufgaben durch KI-Assistenz

Spezifische Anforderungen

- **Rollenbasierte Filterung:** Jeder Mitarbeitende sieht nur relevante Klientendaten und Aufgaben
- **Automatisierte Datenkonsolidierung:** Zusammenführung von Excel-Tabellen, Word-Berichten und E-Mail-Inhalten sowie Datenbestände z. B. über Retrieval Augmented Generation, Knowledge Graphs und ChatBots.
- **Prozessvisualisierung:** Klare Definition von „Wer macht was bis wann“ pro Klient:in
- **KI-gestützte Priorisierung:** Algorithmische Empfehlungen für Tagesabläufe basierend auf Dringlichkeit

Datenbezogene Fragen:

1. Welche manuellen Prozesse verursachen aktuell den höchsten Zeitaufwand?
2. Wie können wir Excel/Word-Dokumente z. B. durch Formulare in strukturierte Daten überführen?
3. Welche KI-Funktionen würden die tägliche Arbeit spürbar entlasten?
4. Wie stellen wir sicher, dass Mitarbeitende nur für sie relevante Daten sehen?

Szenario 3: Abrechnung

Ausgangslage und Potenziale:

Manuelle Abgleichprozesse zwischen Dokumentation, Controlling, FiBu und Kostenträgern führen zu Fehlern und Verzögerungen.

Potenziale:

- Echtzeit-Controlling der Liquidität
- Automatische Plausibilitätsprüfung von Leistungsnachweisen
- Dynamische Anpassung von Zusatzleistungskatalogen

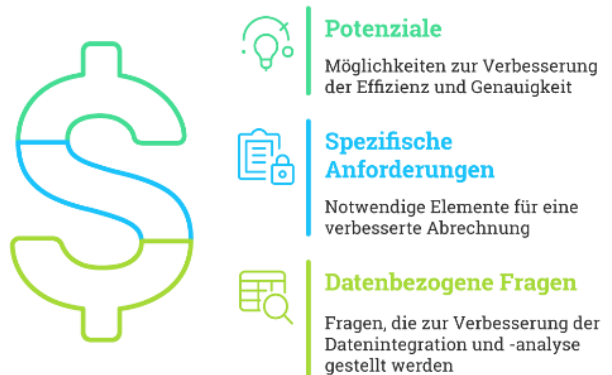
Spezifische Anforderungen:

- **Datenketten:** Lückenlose Verbindung zwischen Leistungserfassung
→ Abrechnungsposition → FiBu-Buchung
- **Smart Contracts:** Automatisierte Honorarvereinbarungen mit Kostenträgern
- **Digitale Beweiskette:** Integrierte Signaturen mit klarer Nachvollziehbarkeit und Revisionsicherheit.

Datenbezogene Fragen:

1. Welche Daten benötigen wir für vorausschauende Liquiditätsplanung?
2. Wie integrieren wir digitale Signaturen rechtskonform in bestehende Prozesse?
3. Welche Schnittstellen zu Banken/Kostenträgern sind prioritär?

Verbesserung der Abrechnungsprozesse



Szenario 4: CSRD – Nachhaltigkeitsberichterstattung durch effizientes Datenmanagement

Herausforderungen und Ansatzpunkte:

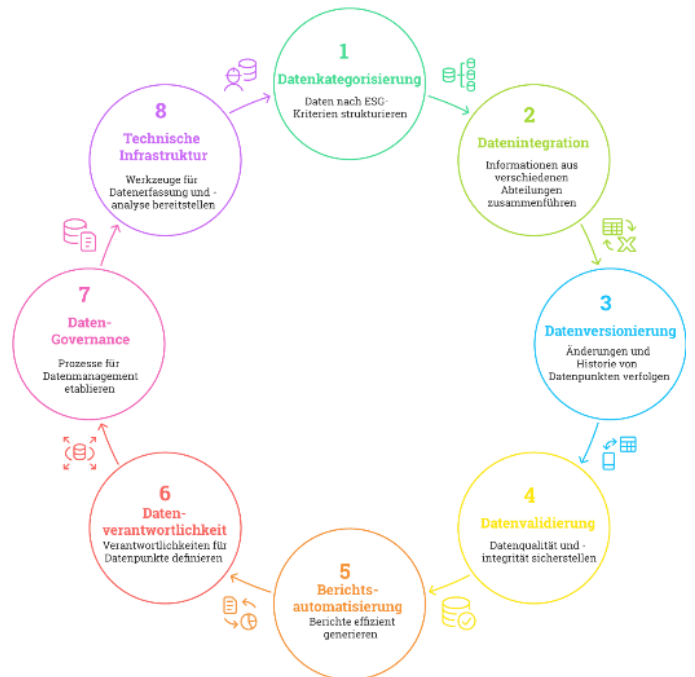
Die Corporate Sustainability Reporting Directive (CSRD) verstärkt den Fokus auf die Offenlegung nichtfinanzieller Informationen durch Unternehmen, um die Transparenz und Vergleichbarkeit ihrer ESG-Performance zu verbessern.

Für einen aussagekräftigen Nachhaltigkeitsbericht benötigen Unternehmen Zugang zu umfangreichen Daten aus verschiedenen Quellen.

Spezifische Datenmanagement-Anforderungen

- **Datenkategorisierung nach ESG-Kriterien:** Strukturierte Erfassung von Umwelt-, Sozial- und Governance-Daten
- **Datenintegration aus Abteilungen:** Zusammenführung von Informationen aus HR, Facility Management, Finanzen, und Fachbereichen
- **Versionierung und Datenpunkthistorie:** Nachvollziehbarkeit der Entwicklung von Nachhaltigkeitskennzahlen
- **Validierungsprozesse:** Mechanismen zur Sicherstellung der Datenqualität und -integrität
- **Automatisierte Berichtsgenerierung:** Technische Lösungen für die standardisierte Berichterstellung

CSRD-Datenmanagement-Zyklus



Datenstruktur-Empfehlungen für CSRD

- 1. Zentrales ESG-Datenmanagementsystem einrichten mit:**
 - a. Definierter Datenverantwortlichkeit für jede ESG-Kennzahl
 - b. Standardisiertem Erfassungsprozess mit Qualitätssicherung
 - c. Zeitgesteuerten Automatismen für regelmäßige Datenerhebungen
- 2. Daten-Governance etablieren**
 - a. Wer ist verantwortlich für welche Datenpunkte?
 - b. Wie werden Daten validiert und freigegeben?
 - c. Welche Daten müssen zwingend berichtet werden (Wesentlichkeitsanalyse)?
- 3. Technische Infrastruktur**
 - a. Software zur ESG-Datenerfassung und -konsolidierung
 - b. Dashboard-Lösung für kontinuierliches Monitoring
 - c. XBRL-fähige Systeme für die maschinelle Lesbarkeit der Berichte

Datenbezogene Fragen für CSRD-Vorbereitung

- Welche CSRD-relevanten Datenpunkte sind für unsere Organisation wesentlich?
- Werden Nachhaltigkeitsdaten bereits systematisch erfasst, und wenn ja, wo?
- Wie können wir Datensilos zwischen verschiedenen Unternehmensbereichen überwinden?
- Welche Stakeholder müssen in den Datenerfassungsprozess eingebunden werden?
- Wie kann die Datenerhebung so gestaltet werden, dass sie neben der Berichtspflicht auch strategischen Mehrwert liefert?

Szenario 5: Innovation – Datenstrukturen als Katalysator für soziale Innovation

Ausgangslage und Potenziale:

Innovationen sind auch in der Sozialwirtschaft aktiv durch Managementprozesse aufzuspüren, zu unterstützen und weiterzuentwickeln. Neue Lösungen, wie soziale Dienstleistungen organisiert, finanziert und erbracht werden können, sind gefragt. Erfolgreiche Innovationen in der Sozialwirtschaft zeichnen sich dadurch aus, dass sich das „Neue“ als sinnvoll erweist und Einzug in die Praxis nimmt.

Datenmanagement zur Innovationsförderung

- **Bedarfserhebungsdaten:** Strukturierte Erfassung von Nutzer-Feedback und gesellschaftlichen Herausforderungen
- **Trendmonitoring:** Systematische Beobachtung von Entwicklungen im Sozialsektor
- **Kollaborative Plattformen:** Digitale Räume für Co-Creation und Ideenentwicklung
- **Wirkungsmessung:** Datenstrukturen zur Bewertung von Innovationserfolgen
- **Wissensmanagementsysteme:** Dokumentation von Lernprozessen und Best Practices



Empfehlungen für innovationsfördernde Datenstrukturen

Offene Datenarchitektur schaffen:

- APIs für einfache Integration neuer Tools und Methoden
- Modulare Systeme, die flexibel anpassbar sind
- 5-Star-Open-Data-Prinzipien als Leitlinie für Datenzugang

Kollaborative Datennutzung ermöglichen:

- Plattformen für interdisziplinären Austausch
- Datenaustausch mit Partnern und externen Innovatoren
- Echtzeit-Feedback-Systeme für kontinuierliche Verbesserung

Experimentierräume für Datennutzung:

- Sandboxed Environments für innovative Datennutzung
- Prototyping-Tools für datengestützte Dienstleistungen
- A/B-Testing-Infrastrukturen für neue Angebote

Datenbezogene Fragen für Innovationsvorhaben

1. Welche Daten können uns helfen, Bedarfe unserer Zielgruppen besser zu verstehen?
2. Wie können wir Daten nutzen, um unsere Dienstleistungen kontinuierlich zu verbessern?
3. Welche Datenquellen könnten uns neue Perspektiven auf soziale Herausforderungen eröffnen?
4. Wie können wir sicherstellen, dass innovative Ideen systematisch erfasst und bewertet werden?
5. Welche Dateninfrastruktur benötigen wir, um schneller von der Idee zur Umsetzung zu kommen?

Zusammenfassung und Schlusswort

Eine nachhaltige Datenstrategie entsteht durch das Zusammenspiel von strategischer Zielorientierung und praktischer Realität. Daraus ergibt sich eine ganzheitliche Sicht auf Daten und Prozesse. Besonders wichtig ist dabei die strukturierte Erfassung von Anforderungen: Nur wenn klar ist, welche Daten wofür benötigt werden, lassen sich Lücken erkennen, Prioritäten setzen und langfristig tragfähige Lösungen entwickeln, die sich am Unternehmensziel ausrichten. So wird die Datenstrategie zum verbindenden Element zwischen Fachlichkeit, Technik und Organisation.

An dieser Stelle lohnt sich ein Blick zurück an den Anfang dieses Dokuments, wo wir uns gefragt haben: Was soll erreicht werden und damit verbunden: Was sind Messkriterien und wann und wie wird gemessen? Hierzu ein Zitat welches dem „Management-Guru“ Peter Drucker zugeschrieben wird:

„What gets measured gets managed.“

Daraus lässt sich ableiten, dass man nichts verbessern kann, was man nicht gemessen hat. Daher sollte man sich immer die Frage stellen: Wurden die eingangs formulierten Ziele erreicht? Sind die angestrebten Verbesserungen (etwa in der Reduktion von Bearbeitungszeiten, der Verfügbarkeit strukturierter Daten oder der Zufriedenheit der Mitarbeitenden) messbar eingetreten? Denn die regelmäßige Überprüfung anhand definierter KPIs und Feedbackrunden ist essenziell, um die Wirksamkeit der Datenstrategie zu bewerten und gezielt nachzusteuern.

Ohne messbare Ziele lässt sich zum einen der Erfolg von Veränderungsprozessen nicht bewerten, und ohne aktives Change-Management zum anderen werden Veränderungen oft nicht nachhaltig umgesetzt. Denn die Einführung einer Datenstrategie ist nicht nur ein technisches oder organisatorisches Projekt, sondern bestenfalls eingegliedert in einen umfassenden Veränderungsprozess. Damit dieser Wandel gelingt, braucht es gezielte Maßnahmen zur Begleitung und Steuerung, etwa durch offene Kommunikation, die Beteiligung relevanter Stakeholder, regelmäßige Abstimmungen sowie klar definierte Rollen und Verantwortlichkeiten. Die Messbarkeit der Ziele unterstützt dabei nicht nur die Erfolgskontrolle, sondern schafft Transparenz und Vertrauen, was wiederum die co-kreative Zusammenarbeit fördert. So entsteht eine lernende Organisation, die ihre Datenstrategie kontinuierlich weiterentwickelt und strategische Initiativen wirksam umsetzt.

Glossar

Begriff

Datenstrategie

Datenmodellierung

Interoperabilität

API (Application Programming Interface)
REST / SOAP

Datenqualität

Strukturierte Daten

Unstrukturierte Daten
Data Warehouse
Data Lake

Bottom-up-Ansatz

Top-down-Ansatz

SMART-Ziele

KPI (Key Performance Indicator)
Digitale Souveränität

5-Star-Open-Data-Modell

Barrierefreiheit

Informationssicherheitskonzept

Definition

Strukturierter Plan zur Erhebung, Verarbeitung, Nutzung und Auswertung von Daten zur Erreichung organisatorischer Ziele.

Strukturierung von Datenobjekten, Attributen und Beziehungen zur Abbildung realer Prozesse.

Fähigkeit verschiedener Systeme, Daten nahtlos auszutauschen und zu verarbeiten.

Programmierschnittstelle zur Anbindung und Kommunikation zwischen Softwareanwendungen. Zwei gängige API-Standards. REST ist leichtgewichtig und JSON-basiert, SOAP XML-basiert und komplexer.

Maß für Genauigkeit, Vollständigkeit, Konsistenz, Aktualität und Struktur von Daten.

Daten mit festem Schema, z. B. in Datenbanken (CSV, JSON, XML).

Daten ohne festes Schema, z. B. PDFs, Bilder, Freitext.

Zentrale Datenbank für strukturierte, analysierbare Daten.

Speicher für Rohdaten in beliebigem Format, oft für explorative Analysen.

Entwicklung der Datenstrategie ausgehend von der Praxis und bestehenden Systemen.

Entwicklung der Datenstrategie ausgehend von strategischen Zielen und Geschäftslogik.

Zieldefinition: spezifisch, messbar, erreichbar, realistisch, terminiert.

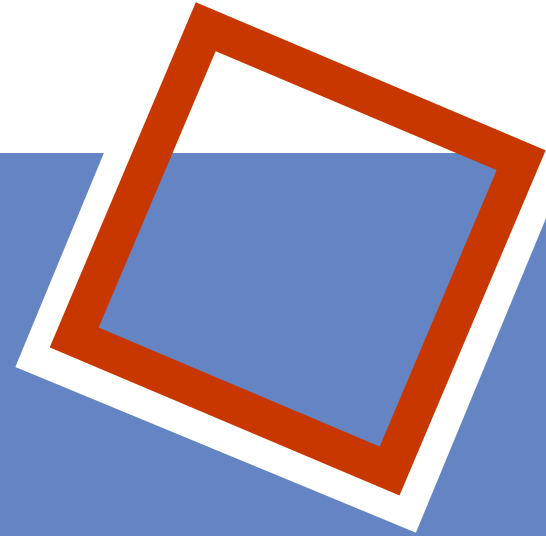
Kennzahlen zur Messung des Erfolgs von Maßnahmen.

Fähigkeit, digitale Systeme und Daten unabhängig und selbstbestimmt zu betreiben.

Modell zur Bewertung der Offenheit und Verknüpfbarkeit von Daten.

Gestaltung von Daten und Systemen, sodass sie für alle Menschen zugänglich sind.

Dokumentation technischer und organisatorischer Maßnahmen zum Schutz von Daten.



Kontakt:
FINSOZ e. V.
**Fachverband Informationstechnologie
in Sozialwirtschaft und Sozialverwaltung**
Ostseestraße 109
10409 Berlin
Tel.: 030 42084-512
E-Mail: info@finsoz.de
www.finsoz.de

V.i.S.d.P.: Helmut Ristok, Vorsitzender des Vorstandes

Ansprechpartner für den Praxisleitfaden:
FINSOZ-Fachgruppe „Digitales Arbeiten in der Sozialwirtschaft“
Lena Knall
Lena.Knall@finsoz.de