Heino Rudolf

Unsere Daten neu denken

„Aktuelle, zugängliche und gut strukturierte oder gar verknüpfte und weiterverwendbare maschinenlesbare Daten sind eine Grundvoraussetzung für wegweisende Datennutzungen, aber auch eine wichtige Grundlage für gemeinwohlorientierte Datenarbeit, für die Information der Öffentlichkeit und die Entwicklung innovativer digitaler Lösungen in Wirtschaft und Zivilgesellschaft.“ - Digitalstrategie Deutschland 2022 [1, S.46].

Wir erheben jetzt schon unfassbar viele Daten, und es werden ständig mehr. Sinnvoller Weise sollten die einmal erhobenen Daten als Open Data angeboten werden, um diese für digitale Wertschöpfungen verfügbar zu haben. Aber genau das ist derzeit äußerst aufwendig: Wir transferieren die Daten in eine themenspezifische Zielstruktur über ETL-Prozesse (ETL = Extract, Transfom, Load) und setzen dann die Dienste auf. Die Dienste stellen die Daten in sehr phänomenalen Strukturen zur Verfügung. Wer sie jetzt weiterverarbeiten will, muss sich mit diesen Strukturen auseinandersetzen und konkrete Programme entsprechend schreiben. So entsteht eine riesige Menge Programmcode. Wo sollen so viele Programmierer/-innen sowohl für die Datenbereitstellungen als auch für deren Nutzungen in Zeiten des Fachkräftemangels eigentlich herkommen? Also dauert das alles sehr lange, und viele durchaus wichtige Daten können in absehbarer Zeit nicht bereitgestellt werden. Und wenn sie dann verfügbar sind, müssen auch erst die Programme für deren Nutzung geschrieben werden. Das kostet alles sehr viel Geld!

Mit diesen heute praktizierten Verfahren werden wir die Digitalisierung nur sehr schwerfällig bewältigen. Es müssen dringend verbesserte und effiziente Methoden der Datenverarbeitung eingeführt werden, um die notwendigen Entwicklungsschübe zu erreichen.

1. Daten und Digitalisierung

„Daten sind das Öl des 21. Jahrhunderts.“

Dieser Vergleich wird heute von Politikern/-innen, Wissenschaftlern/-innen, Ökonomen/-innen, Wirtschaftsmanagern/-innen, Journalisten/-innen u. a. sehr häufig bemüht. Erstmalig hat ihn wohl der britische Mathematiker Clive Humby 2006 [5] verwendet.

Wir können diese Metapher durchaus wirtschaftspolitisch deuten, denn die großen Ölkonzerne wurden von den IT-Riesen wie Google, Amazon, Apple, Facebook, Microsoft als reichste Unternehmen der Welt abgelöst. Und wieder einmal konzentrieren sich die Gewinne in den Händen einiger Weniger. Malte Spitz stellt in seinem Buch [4] Ansätze zusammen, was wir aus der Öl-Vergangenheit für eine demokratische, gleichberechtigte und v. a. nachhaltige Daten-Zukunft ableiten können.

Angela Merkel griff diesen Vergleich seit 2015 mehrfach auf (z. B. [6]) und wandelte ihn ein wenig ab:

„Daten sind Rohstoffe des 21. Jahrhunderts.“

Damit wird deutlicher auf die Wirkungen der Daten als notwendige Basis für die Wirtschaft und Produktion fokussiert. Diese Feststellung dürfte mittlerweile zum Allgemeingut avanciert sein. Wir sprechen in diesem Zusammenhang in Deutschland von „Digitalisierung“, die

gewaltige gesellschaftliche, soziale, wissenschaftliche und wirtschaftliche Umwälzungen nach sich ziehen wird.

Digitalisierung

Bereits in den 90er Jahren gab es die Digitalisierung – als Prozess der Schaffung digitaler Daten aus analogen Vorgaben. Heute steht der Begriff für eine digitale Transformation der Gesellschaft, die alle Bereiche des Lebens erfasst.

Oftmals wird in diesem Kontext über die vierte industrielle Revolution (bekannt auch als „Industrie 4.0“) diskutiert. Es ist sehr auffällig, dass immer technische Innovationen die Triebkräfte der vorangegangenen drei industriellen Revolutionen waren: (1) Maschinen, Wasser und Dampfkraft; (2) Fließbandproduktion, Elektrifizierung, Motoren, Telekommunikation; (3) Elektronik, Computer, Informationstechnik. Diese Kräfte waren für jedermann sicht- und anfassbar.

Das ist jetzt ganz anders, denn Digitalisierung findet im Virtuellen statt. „Die wahren technischen Hintergründe erschließen sich nicht auf den ersten Blick, also werden die technischen und infrastrukturellen Voraussetzungen, die Gestaltung und Bedienung der allgemein genutzten Applikationen, Visualisierungen, die spürbaren rechtlichen, sozialen, gesellschaftlichen u. a. Effekte diskutiert“ [2, Kapitel 9]. Wir reden von Digitalisierung, aber meinen wir alle dasselbe? Wenn wir uns die Potenziale der Digitalisierung erschließen und die Auswirkungen beeinflussen wollen, müssen wir die (technischen) Ursachen und Entwicklungen verstehen. Ich greife im Folgenden auf meine Ausführungen in [2] zurück, um „Digitalisierung“ zu definieren:

(1) Umwandeln von analogen Werten in digitale Formate; Erstellen digitaler Repräsentationen von physischen Objekten, Ereignissen, analogen Informationen

Die ersten Digitalisierungen bezogen sich auf einfache Umwandlungen von analogen Werten in digitale Formate; zunächst einfache Zeichen und Zahlen in binären Code, später Texte, Musikstücke, Fotos, Karten, Filme … Es entsteht ein binäres Abbild des analogen Gegenstands. Wesentlich ist die Definition des Formats, wie dieser Gegenstand codiert wird. Damit ist ein entscheidender Aspekt der Digitalisierung gegeben: ein verständliches, leicht und effizient weiterverarbeitbares und auf zukünftige Anforderungen überführbares binäres Format.

Für die Verwaltung der digitalen Daten werden Speichermedien benötigt. Für die Erzeugung, der digitalen Daten, vor allem aber auch ihre Rückübersetzung in die ursprünglichen (analogen) Formen sind DV-Programme notwendig. Es entsteht eine zweite, digitale (virtuelle) Welt mit gespeicherten digitalen (binären) Daten und zugehörigen Erzeugungs- und Verarbeitungsprogrammen.

→ Letztendlich kann alles digital erfasst werden, was der Mensch sich ausdenkt, was er analysiert und dokumentiert – natürlich immer als Modell mit bestimmten Formaten und damit in einer binären Beschreibung wesentlicher Eigenschaften des Originals.

(2) Wandel in Wirtschaft, Wissenschaft, Gesellschaft und Politik zu digitalen Prozessen

Zunächst waren Digitalisierungen auf Rechentechnik fokussiert. Durch die immer komplexeren Formate, die Bereitstellung großer Speicherkapazitäten und die gewaltige Erhöhung der digitalen Verarbeitungsgeschwindigkeiten wurden immer neue Bereiche digitalisiert und

digitale Prozesse implementiert: Geodatendarstellung und -bereitstellung, Funk und Fernsehen, Musik-, Foto-, Kommunikationsindustrie, Aspekte des Autofahrens, Verkehrssteuerungen, Prozesssteuerungen in der Industrie…

Die Digitalisierung bietet im Vergleich zur analogen Informationsverarbeitung wesentliche Vorteile: einfache Speicherung, Verteilung, Wiedergabe der Informationen, maschinelle Lesbarkeit und damit schnelle und einfache Verarbeitung der Daten, Langzeitspeicherung ohne Qualitätsverluste, Komprimierbarkeit der Informationen. Damit führt der Umstieg auf digitale Systeme und Steuerungen zu einer völlig neuen Qualität in Wirtschaft, Wissenschaft, Gesellschaft und Politik.

(3) digitale Transformation: durch die Digitalisierung ausgelöster Umbruch, der einen Wandel der Technik sowie (fast) aller Lebensbereiche bewirkt

Technisch eingeleitet wurde die digitale Transformation durch die Vernetzung von Geräten, Maschinen, Unternehmen usw. Werden die Daten in standardisierten digitalen Formaten bereitgestellt, können sie jederzeit ausgewertet und weiterverarbeitet werden. So werden Prozesse angestoßen und gesteuert; es werden auch beschreibende Zustandsdaten zur Optimierung der Prozesse herangezogen u. v. m. Nach (2) sind immer mehr Maschinen und Geräte mit DV-Programmen ausgestattet, sodass sie digitale Daten auswerten, über Daten gelenkt werden und auch wieder Daten zur Weiterverarbeitung erzeugen. Der Trend ist heute längst erkennbar: Wenn so viele digitale Daten wie möglich gespeichert und digitalen Prozessen zur Verfügung gestellt werden, kann die Realität virtuell umfassend erklärt; können immer mehr real ablaufende Prozesse vorhergesagt, gemanagt und beschrieben werden. Und das betrifft (noch mit unterschiedlicher Intensität) alle Bereiche der Gesellschaft: beginnend mit Speicherungen menschlicher Verhaltensweisen, über Gesundheitsdaten, Verkehrs- u. a. Prozessabläufe usw. - aber auch Erfassung der Umweltvorgänge und –zustände.

2. Wie sollten Datenverarbeitungen ablaufen?

Es dürfte aus allem bisher Ausgeführten völlig klar sein, dass wir für unsere digitale Gestaltung der Zukunft dringend Daten benötigen – so viele wie möglich und in einer Form, dass wir sie einfach, schnell und effizient weiterverarbeiten können.

In diesem Kapitel möchte ich beschreiben, wie ich mir einzelne Verarbeitungsprozesse vorstelle. Im nachfolgenden Kapitel schildere ich dann zum Vergleich die heutigen Abläufe. Dafür wähle ich drei Persona aus:

- Datenbereitsteller/-innen in der datenhaltenden Stelle
- Programmierer/-innen, die aus den bereitgestellten Daten neue Software erstellen
- Planer/-innen, die mit den Daten räumliche Planungsaufgaben bearbeiten.

Datenbereitsteller/-innen

Aus den Fachsystemen erhalten wir die Daten entweder per Dienst oder per Datenexport. Moderne Systeme stellen Dienste zur Verfügung, meist in konkreten, fachspezifisch geprägten Strukturen. Sollten diese Möglichkeiten nicht gegeben sein, können wir aber davon ausgehen, dass in den Fachsystemen Exportfunktionalitäten implementiert sind.

Es gibt eine (!) klar definierte Zielstruktur. Diese ist nicht mehr proprietär auf spezielle Anwendungsfälle zugeschnitten. – In [2] und [3] habe ich diese allgemeingültigen Strukturen

ausgearbeitet. Als Datenbereitsteller/-in kenne ich diese Struktur. Ich muss jetzt „mappen“, d. h., die Strukturen der Fachsysteme den Zielstrukturen zuordnen. Dieser Schritt ist der aufwendigste und natürlich immer durchzuführen. Aber gerade hier besteht ein großes Einsparpotenzial: einerseits weil die Zielstrukturen immer gleich und sehr überschaubar sind und sich damit das „Mapping“ wiederholt und professionalisiert; andererseits weil durch diese Vereinheitlichung die nachfolgenden Prozesse beim Konvertieren automatisiert werden können.

Nach dem „Mapping“ werden automatisch die neuen Klassen für die Datenbereitstellung und die durchzuführenden Konvertierungen angelegt. Danach muss ich mir als Datenbereitsteller/-in noch einmal die konkreten Klasseninhalte und Konvertierungsregeln anschauen und ggf. anpassen. Komplizierte Konvertierungsregeln sind manuell anzulegen.

Dann führen wir den Datenimport durch. Im Ergebnis liegen die Daten in den Zielklassen vor. Alle notwendigen Metadaten werden automatisch angelegt. Sobald die Daten importiert sind, stehen sie per Dienst zur Verfügung.

Abschließend kontrolliere ich als Datenbereitsteller/-in die Qualität der importierten Daten. Dazu werden im Ergebnis des Imports automatisch die Daten in die Quellstrukturen rückkonvertiert (soweit das möglich ist). Diese Daten können wir mit den Ausgangsdaten vergleichen.

Programmierer/-innen

Die Daten werden immer in einheitlichen Strukturen bereitgestellt. (Vgl. mit [2] und [3].) Das trifft sowohl auf die Objektklassen als auch deren Verknüpfungen zu. Der programmtechnische Zugriff auf die Daten muss demnach nur einmal allgemeingültig programmiert werden. Ich als Programmierer/-in konzentriere mich jetzt voll und ganz auf die fachlichen Inhalte, die ich implementieren will.

Planer/-innen

Immer wieder benötige ich als Planer/-in raumbezogene Daten. Ich konvertiere die notwendigen Daten in die oben beschriebene einheitliche Struktur, womit ich mir einen immer wieder gleichen Zugang zu den Daten und deren Verknüpfungen schaffe. Analog zu den Datenbereitstellern/-innen bereite ich mir die Daten auf. Einmal erstellte Konvertierungen kann ich immer wieder benutzen, sodass auch ein Update der Daten jederzeit möglich ist.

Aus den einheitlichen Strukturen kann ich mir die Daten auf ganz einfache und intuitive Weise zu Szenarien zusammensetzen. Szenarien sind spezielle Sichten auf die Daten, die ich aus meinen fachlichen Anforderungen heraus definieren kann.

Sowohl die eingespielten Quelldaten als auch die Daten der Szenarien stehen als Dienste zur Verfügung und können in mein GIS oder Fachsystem eingebunden werden.

Für mich als Planer/-in wird die Datenverarbeitung deutlich einfacher, und ich kann mich auf meine fachlichen Aufgaben konzentrieren.

3. Wie werden die Daten heute verarbeitet?

Häufig hören wir, dass heute bereits Daten zur Weiterverarbeitung angeboten werden, dass es die notwendige interoperable Datenbereitstellung bereits gibt – z. B. wurden Geodateninfrastrukturen geschaffen, die auch den Download von (Geo-) Daten ermöglichen. - Aber:

Für komplexe Projekte, Simulationen, für den Einsatz künstlicher Intelligenz u. ä. reicht es nicht aus, einfach Layer (Feature) mit einer Geometrie, angereichert um spezielle Sachdaten, herunterzuladen. Viel wichtiger sind meist die Verknüpfungen zwischen den Datensätzen, wie das im obigen Zitat aus der Digitalstrategie auch festgestellt wird. Diese werden dann mit z. T. komplizierten Algorithmen erzeugt oder nachträglich übernommen.

Und zum weiteren Verkomplizieren haben alle bereitgestellten Datensätze auch noch eigene, ganz speziell auf die Featureklasse zugeschnittene Datenstrukturen.

Häufig wird dann auch behauptet, dass es gar nicht anders gemacht werden könne. Wir brauchen heute einerseits Interoperabilität, andererseits werden aus Ermangelung adäquater Methoden aktuelle Standards (z. B. XÖV [C]), die etwa 20 Jahre alte Technologien verwenden, zu Gesetzen für die Datenbereitstellung erhoben.

„Die Methoden sind nach meinem Dafürhalten für eine interoperable Datenbereitstellung von Informationen zu unserer Umwelt unbefriedigend. Sie sind schwer verständlich, kompliziert, auf bestimmte Phänomene zugeschnitten und äußerst schwerfällig (wenn überhaupt) an erweiterte Anforderungen anpassbar. Wir dürfen auch nicht vergessen, dass für die gewünschte Weiterverarbeitung der Daten zu neuen Wertschöpfungen komplizierte Programme für das Entschlüsseln, insbesondere bei einem Zusammenführen von Daten verschiedener Themen, notwendig werden.“ [2, Kapitel 8]

Datenbereitsteller/-innen

In der folgenden Abbildung symbolisiert ein Zylinder der oberen Reihe die Datenhaltung für ein Fachsystem. Die unteren Fünfecke stehen für Datenbereitstellungen (es sind Auszüge aus INSPIRE-Datenmodellen – die Inhalte der Kästchen und Zylinder sind nicht wichtig, die Methode steht im Fokus).

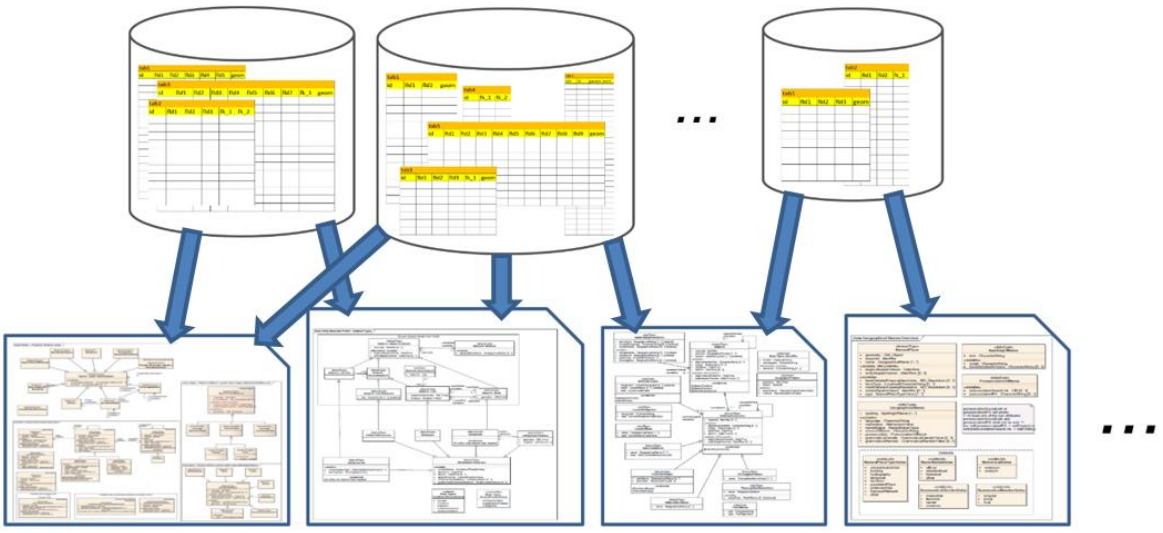

Abbildung 2: Datenbereitstellungen heute

Es wird Code geschrieben, um die Daten von einem Fachsystem in die Strukturen für eine Datenbereitstellung zu überführen (in der Abbildung durch dicke blaue Pfeile skizziert, die oftmals recht aufwendig zu programmieren sind). Diese Verfahrensweise ist heute allgemein verbreitet, kreiert wurde sie i. W. durch die Vermessungsverwaltungen. Sie erheben ihre Daten im Fachsystem fast ausschließlich zum Zwecke der Bereitstellung von geografischen Informationen, denn das ist ihre ureigene Aufgabe. So müssen sie diese bereits zum Veröffentlichen gedachten Daten einfach in das konkrete Datenformat umwandeln. Das wird dann „Datenharmonisierung“ genannt.

Unglücklicher Weise sind die Datenstrukturen für die Bereitstellung (in der Abbildung die unteren Fünfecke) immer auf spezielle inhaltliche Phänomene fokussiert. Allein bei INSPIRE haben wir 34 Datenmodelle zu den einzelnen Anhängen, es gibt beispielsweise mehr als 60 Umweltberichterstattungen an die Europäische Kommission. Ich hatte in meinem Buch [2] schon prophezeit, dass es ein „Wuchern“ von Datenharmonisierungen geben wird. Und jetzt können wir tatsächlich auch noch ein „Wuchern“ an Standards beobachten, denn neue fachlich-inhaltliche Aufgaben benötigen ein neues Modell zur Datenbereitstellung.

Alles, was an Inhalten und Verknüpfungen offeriert werden soll, muss bei dieser Methode natürlich in den Quellsystemen vorgehalten werden. Deshalb werden Interoperabilitäts-kriterien definiert, um bereits in den Fachsystemen notwendige Verlinkungen aufzunehmen.

Genau diese Methode der „Datenharmonisierung“ wird jetzt auch auf die Bereitstellung von Fachdaten angewandt. Doch das funktioniert bei den Fachexperten/-innen so einfach nicht. Sie nutzen die IT-Systeme zur Erfüllung ihrer konkreten Aufgaben, z. B. das Betreiben und Auswerten von Messsystemen, Bearbeiten von Genehmigungsverfahren, Umweltanalysen und –prognosen. Darauf sind die Fachsysteme zugeschnitten, um beispielsweise riesige Messzeitreihen oder komplexe Auswertealgorithmen performant zu verarbeiten. Die Daten-bereitstellung spielt nur eine untergeordnete Rolle in ihren Fachsystemen. Die „Datenhar-monisierungen“ können jetzt noch aufwendiger werden. Und vieles, was an bereitzustellen-den Verknüpfungen gewünscht wird, muss irgendwie nachträglich im Fachsystem einge-arbeitet werden, also wären ggf. diese Systeme sogar anzupassen, was dann noch kost-spieliger wird.

Im Ergebnis der „Datenharmonisierungen“ liegen die Daten in Datenbanken bzw. Dateien. Es muss noch ein weiterer Schritt zum Aufsetzen der konkreten Dienste vorgenommen werden. Und dieses Aufsetzen ist für jeden Dienst auf ein Neues notwendig.

Programmierer/-innen

Als Programmierer/-in muss ich mich erstmal mit der konkreten phänomenalen Struktur der bereitgestellten Daten auseinandersetzen. Wie oben bereits beschrieben, gibt es beispiels-weise allein in INSPIRE 34 Datenmodelle. Mit Kenntnis des Datenmodells kann ich dann meine Software schreiben, wobei alles proprietär auf die konkrete Struktur zugeschnitten ist. Durch die Standardisierung der Datenmodelle konzentriert sich dieser Prozess, was aber durch das oben erläuterte „Wuchern“ an neuen Bereitstellungsstrukturen wieder aufgehoben wird.

Besonders kompliziert wird es, wenn ich Daten aus verschiedenen Bereitstellungsstrukturen gemeinsam verarbeiten muss. Verknüpfungen über die Strukturen hinweg sind fast immer nicht gegeben – jetzt ist also meine persönliche Kreativität gefragt…

Planer/-innen

Auf unserer Sitzung des Fachausschusses „Digitale Wasserwirtschaft“ der Deutschen Vereinigung für Wasserwirtschaft, Abwasser und Abfall e.V. (DWA) im Juni 2022 berichteten die Kollegen/-innen aus Nordrhein-Westfalen über ihre Arbeiten zur Aufarbeitung des Hochwassers der Ahr 2021: ca. 20% der Arbeitszeit benötigten sie zum Auffinden der Daten, ca. 60% zum Aufbereiten der Daten, und dann verblieben nur etwa 20% für die eigentlichen fachlichen Auswertungen.

In der folgenden Abbildung haben wir mal skizziert, wie es einem Akustiker/einer Akustikerin bei der Bereitstellung der Daten für die Lärmkartierung ergehen könnte (und das alle fünf Jahre wieder).

Grafische Gestaltung: Sirid Fillinger Visuelle Kommunikation

4. Interoperabilität

Interoperabilität heißt ganz einfach: die Daten sind so bereitzustellen, dass sie miteinander operieren, also verarbeitet werden können. Die Definition von Interoperabilität (des gemeinsamen Miteinanders) erschließt sich aus unseren Erfahrungen mit gesundem Menschenverstand: Wir müssen uns unterhalten wollen (psychologische Voraussetzungen). Wir müssen miteinander reden dürfen (rechtliche Voraussetzungen). Wir müssen uns treffen oder zumindest miteinander kommunizieren können (technische Voraussetzungen). Wir brauchen Regeln für dieses Miteinander, denn sonst agieren wir nicht gemeinsam (organisatorische Standards). Wir brauchen für alle verständliche Kommunikationsformen, über die wir uns austauschen (Standards der Datenbereitstellung). Wir brauchen eine gemeinsame Sprache, die wir alle verstehen (datenstrukturelle Standards).

Im Grundsatz müsste uns das bereits seit INSPIRE [A] bekannt sein, wo konkrete organisatorische, technische und sogar datenstrukturelle Festlegungen zur Datenbereitstellung in einem Gesetz festgeschrieben wurden. Warum haben wir dann aber diese Herausforderung bis heute nur teilweise gemeistert und sind vom eigentlichen Ziel recht weit entfernt?

Im Folgenden möchte ich einige Gedanken zu den genannten sechs Aspekten der Interoperabilität andiskutieren und damit zu lösungsorientierten Debatten anregen:

Voraussetzungen für ein gemeinsames Datenmanagement

Psychologische Voraussetzungen

Immer wieder wird behauptet, Corona habe der Digitalisierung einen gewaltigen Impuls gegeben. Wenn wir Digitalisierung auf die Bereitstellung und breite Nutzung von Softwarelösungen reduzieren, dann mag das stimmen. Aber eigentlich hat doch Corona viel deutlicher gezeigt, welch riesigen Nachholbedarf wir haben: Wissenschaftler/-innen beklagen immer wieder eine nicht ausreichende Datenbasis, die Bearbeitungsprozesse und der Datenaustausch waren oftmals analog und führten zu Engpässen in der Berichterstattung, bei Nachverfolgungen und Statistiken, viele Einrichtungen waren gar nicht am Internet angeschlossen.

Die Bereitschaft der Menschen, die digitalen Veränderungen anzunehmen oder gar aktiv mitzugestalten, ist nach meinem Dafürhalten das größte Problem bei der weiteren digitalen Transformation der Gesellschaft. Das betrifft vor allem die Umsetzung von digitalen Prozessen, insbesondere auch beim Zusammenspiel von Softwarelösungen (sodass z. B. nicht der Nächste in der Prozesskette die Daten seiner Vorgänger erstmal konvertieren oder gar wieder neu digitalisieren muss). Wie oft höre ich als Erwiderung auf konzeptionelle Vorschläge für den Einsatz der Informationstechnologien: „Das haben wir aber immer schon so gemacht."

Datenverarbeitung wird oft als „Spielplatz" der Experten/-innen gesehen, viele Menschen schieben sie von sich weg. Dahinter verbergen sich auch Ängste vor den Veränderungen. Geschürt wird das Ganze auch durch die Experten/-innen selber, indem immer neue „Trends" deklariert werden: z. B. Augmented Reality, Big Data, Internet of Things, Künstliche Intelligenz, Social Networks (entnommen aus [9] zu den aktuellen Digitalisierungstrends mit Auswirkungen auf den Umweltschutz). Wir brauchen dringend eine Entmystifizierung der Wirkmechanismen der digitalen Welt, sodass alle Menschen die Digitalisierung als ihre eigene große Chance sehen: Digitalisierung ist etwas Großartiges – so erreichen wir Nachhaltigkeit bei der Gestaltung der Prozesse und werden dann tatsächlich etwas für den Umweltschutz bewirken.

Rechtliche Voraussetzungen

Der Trend ist ganz klar: INSPIRE [A], GeoZG [D], UIG [E] u. a. Gesetze und Verordnungen legen fest, dass Behörden die von ihnen erhobenen Informationen offen zur Verfügung zu stellen haben. Im GeoZG des Bundes [D] sowie in den meisten Geodatenzugangs- bzw. Geodateninfrastrukturgesetzen der Länder (je nachdem, wie das Gesetz im Bundesland bezeichnet wurde) ist darüber hinaus festgelegt, dass alle Geodaten aktiv (!) bereitzustellen sind:

Geodaten sind im GeoZG und in Ländergesetzen folgendermaßen definiert: „Geodaten sind alle Daten mit direktem oder indirektem Bezug zu einem bestimmten Standort oder geografischen Gebiet." (Zitat aus [D §3(1)]). Danach sind alle Umweltdaten Geodaten – auch Messwerte, Bewertungen, Berechnungsergebnisse u. ä. Die betroffenen Geodaten werden im GeoZG [D §4(1)4.] und in Ländergesetzen auf die Themen der INSPIRE-Anhänge fokussiert, was aber bei der breiten inhaltlichen Aufstellung von INSPIRE keine Einschränkungen bzgl. der betroffenen Umweltdaten darstellt.

„Die geodatenhaltenden Stellen stellen sicher, dass für die von ihnen erhobenen, geführten oder bereitgestellten Geodaten und Metadaten mindestens die nachfolgenden Dienste bereitstehen […]“ [D §6(1)]. D. h. doch zusammengefasst, dass alle Umweltdaten als Geodaten für die die Dienste eigenständig von der datenhaltenden Stelle bereitzustellen sind. Die Gesetzgeber haben also klar erkannt und manifestiert, welche Bedeutung die Geodaten bei der weiteren Digitalisierung einnehmen.

Wie am Anfang des Artikels bereits ausgeführt, setzt der Bund dabei in der Digitalstrategie auf Open Data. Auch in den Ländern wurden Open Data Initiativen gestartet, z. B. ist in Nordrhein-Westfalen 2022 die Open-Data-Verordnung in Kraft getreten, die die Anforderungen an das Bereitstellen von Daten regelt.

Auch die Fachbereiche werden folgen: z. B. hat die Geologie mit GeoIDG [F] gesetzlich definiert, dass die Behörden die Daten zu den geologischen Aufschlüssen digital zu veröffentlichen haben, und mehr noch: alle Bohrungen sind den Behörden anzuzeigen und diese haben die Daten in kurzer Zeit bereitzustellen.

Technische Voraussetzungen

Mit INSPIRE ist in der Verwaltung viel geschehen. Es wurden Geodateninfrastrukturen aufgebaut, die Daten können über Dienste bereitgestellt werden. In der Wissenschaft und Wirtschaft ist die Rechentechnik nicht mehr wegzudenken. Bis zu jedem Mitarbeiter in den Behörden, an unseren Schulen… ist das noch nicht 100%ig durchgedrungen. Richtig schlimm ist der Anschlussgrad mit dem (schnellen) Internet. (Könnten wir uns eine Ortschaft vorstellen, zu der keine Straße führt?)

Standards für ein gemeinsames Datenmanagement

Organisatorische Standards

INSPIRE hat gerade diesbezüglich klare Rahmenbedingungen für die Datenbereitstellungen gesetzlich festgelegt. Darauf Bezug nehmend, hat Stephan Klemm in seiner Masterarbeit [7] Qualitätskriterien für die Datenbereitstellung zusammengestellt: Datenstruktur, Verfügbarkeit, Lizenzmodell/rechtlicher Rahmen, Vollständigkeit, Aktualität, Genauigkeit, Inhalt, Interoperabilität. Er hat untersucht, wie die Daten der Mobilithek (ehemals mCLOUD) des Ministeriums für Digitales und Verkehr diesen Kriterien entsprechen und inwieweit sie für Weiterverarbeitungen überhaupt geeignet sind (konkret analysiert am Beispiel des automatisierten Fahrens). Fazit: Einfach hoch geladene Daten, die nicht die oben genannten Qualitätskriterien erfüllen, sind für neue Wertschöpfungen und Verarbeitungen ungeeignet.

Damit haben wir das große Dilemma vieler Datenbereitstellungen der Open-Data-Initiativen: Die Daten können vielleicht für anschließende Projekte hilfreich sein, aber für die digitale Umgestaltung unserer Gesellschaft sind wertlos. Halten wir uns doch auch in Deutschland einfach an die Ansätze von INSPIRE.

Standards der Datenbereitstellung

Hier geht es um die „Verpackung“: wie werden die Daten angefragt, wie wird geantwortet. Das ist natürlich zu formalisieren (standardisieren), damit die Maschinen kommunizieren können. Weit verbreitet sind die Standards des Open Geospatial Consortiums (OGC): z. B. Web Map Service (WMS zur Übertragung von Bildern), Web Feature Service (WFS zur

Übertragung von Objekten/Features). Neue Standards sind die API Services (Application Programming Interface), z. B. der OGC API-Features, der den WFS ablöst. Die API Services sind einfacher zu interpretieren und haben noch den Vorteil, dass die Daten selbst in beliebigen, allgemein bekannten Formaten übertragen werden können, z. B. JSON, CSV, XML.

Das ist alles sehr ausgereift. Wir können also Daten anfragen und gelieferte Datenpakete entgegennehmen. Damit sind aber die eigentlichen Dateninhalte (welche Attribute gehören zum Datenpaket und wie sind die Attribute codiert) noch nicht erklärt.

Datenstrukturelle Standards

INSPIRE [A] hat Datenstrukturen gesetzlich festgeschrieben. Für die 34 Annexthemen wurden Datenmodelle definiert, die für die zu übertragenden Objektklassen die Attribute mit ihren Inhalten und Codierungen und die Verknüpfungen festlegen. Eine ganz wesentliche Erkenntnis, denn nur so können wir die offerierten Daten auch inhaltlich weiterverarbeiten und damit wirkliche Interoperabilität erreichen.

Wie oben im Text bereits ausgeführt, sind alle Datenmodelle auf fachliche Phänomene zugeschnitten. Die eigentliche Ursache dafür liegt in der genutzten Modellierungsbasis. INSPIRE setzt auf die Datenstrukturen der ISO 191xx-Serie [G] auf. In [3] habe ich begründet, dass bereits durch die Verwendung dieser Normen fachlich zugeschnittene Datenmodelle vorherbestimmt sind. Wenn wir so modellieren, geht es tatsächlich nicht anders als heute üblich.

5. Datenverarbeitungen neu denken

Kommen wir wieder zurück zur Interoperabilität und den Methoden, sie herzustellen:

(1) Wir fassen die Fachsysteme nicht an. Wir lassen sie so, wie sie sind. Einerseits erfüllen die Systeme die Aufgaben der Fachkollegen/-innen; und diese werden es uns danken, wenn wir sie weiter so arbeiten lassen, wie sie es möchten. Andererseits können wir diese Systeme doch nicht mit Verknüpfungen u. ä. überlasten, denn so entstehen die oft zitierten „Spaghettistrukturen“.

(2) Es muss eine andere Modellierungsmethode ausgearbeitet werden. Diese habe ich in meinem Buch [2] kreiert, die konkrete Umsetzung habe ich in [3] erläutert. Viele denken nun, ich habe mir ein großes neues „Weltdatenmodell“ überlegt. Natürlich nicht, ich habe mich bei der Modellierung von der Graphentheorie leiten lassen und diese ganz konsequent angewandt. Ein Graph besteht aus genau zwei Elementen: Knoten und Verbindungen. Wenn wir die Datenhaltung als Graphen begreifen, dann brauchen wir eine Datenbanktabelle für die Verwaltung der Knoten (Entities) und eine Datenbanktabelle für die Verwaltung der Verbindungen (Relationships). Und tatsächlich können alle Objektklassen mit ihren Verbindungen in genau zwei Datenbanktabellen verwaltet werden! Hinzu kommen noch ergänzende Tabellen zu den sachlichen und grafischen Ausprägungen. Heute sprechen wir von einem Wissensgraphen.

Die Datenstruktur ist also völlig unabhängig vom fachlichen Gegenstand. Wir können uns das Modell wie ein Spinnennetz vorstellen, es kann jederzeit verdichtet oder vergrößert werden. Welche Inhalte sich hinter den einzelnen Netzknoten bzw. Netzverbindungen verbergen, wird über beschreibende Metadaten hinterlegt.

(3) Wir erstellen ein eigenständiges Tool für interoperable Datenzusammenstellungen. Das nennen wir **Simplex**4TwIS. TwIS steht für Twin Information System, wobei einerseits die Datenstruktur als Zwilling die Realität mit all ihren Ursache-Wirk-Beziehungen, Lagebeziehungen und Wechselwirkungen abbildet und andererseits ein (!) Zwilling aus den verschiedenen Fachinformationssystemen entsteht. Die Datenstruktur ist als Wissensgraph angelegt, wir können beliebige Objektklassen in dieses Tool einspielen. Je mehr Objekte wir übernehmen, umso dichter und vermaschter wird unser „**Simplex**4TwIS-Spinnennetz“.

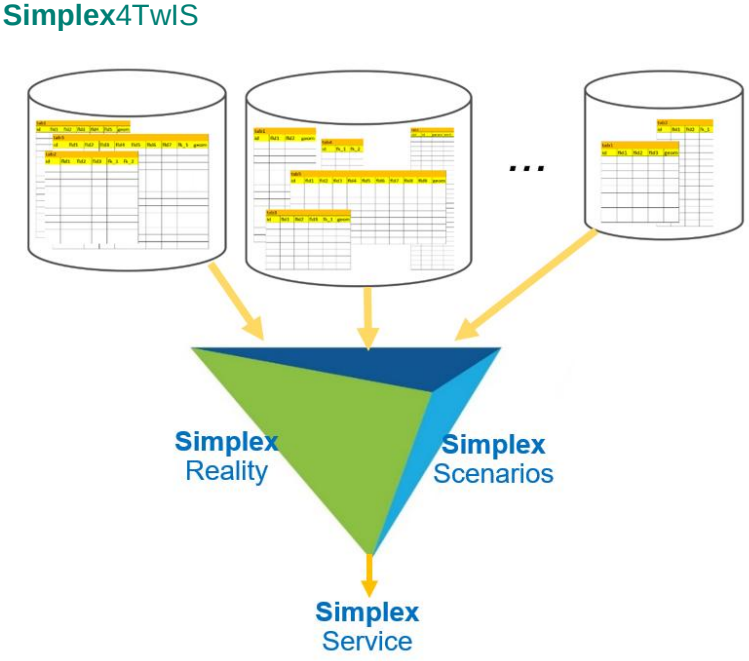

Abbildung 3: Datenmanagement neu denken

Die Datenhaltung gestalten wir zweistufig: Hinter **Simplex**Reality verbirgt sich das oben beschriebene graphentheoretische Datenmodell. **Simplex**Scenarios sind Sichten auf die Daten. So können wir schnell und einfach Daten für Auswertungen, Präsentationen, aber auch für fachliche Projekte zusammenstellen.

Da die **Simplex**Reality-Struktur sehr einfach ist, sind jetzt auch die Datenübernahmen aus den Fachsystemen weniger kompliziert. (Ich habe das in Abbildung 3 mit dünnen gelben Pfeilen symbolisiert, in Abbildung 2 waren sie dick blau.)

Natürlich können wir als **Simplex**Scenarios auch alle Daten entsprechend der heute gängigen Standards zusammenstellen. Wir haben das für mehrere INSPIRE-Anhänge auch umgesetzt. – Aber, warum eigentlich? Müssen wir weiter die komplizierten phänomenalen Datenstrukturen liefern oder können wir nicht einfach die Daten allgemeingültig in der **Simplex**Reality-Struktur veröffentlichen? Jedenfalls wären sie dann sehr einfach und auf immer gleiche Art und Weise verarbeitbar!

Janik Großmann hat in [8] den **Simplex**Service am Beispiel des implementierten OGC API-Features Service beschrieben. Die folgende Abbildung zeigt die Nutzung des **Simplex**Service an einem Beispiel der Stadt Dresden.

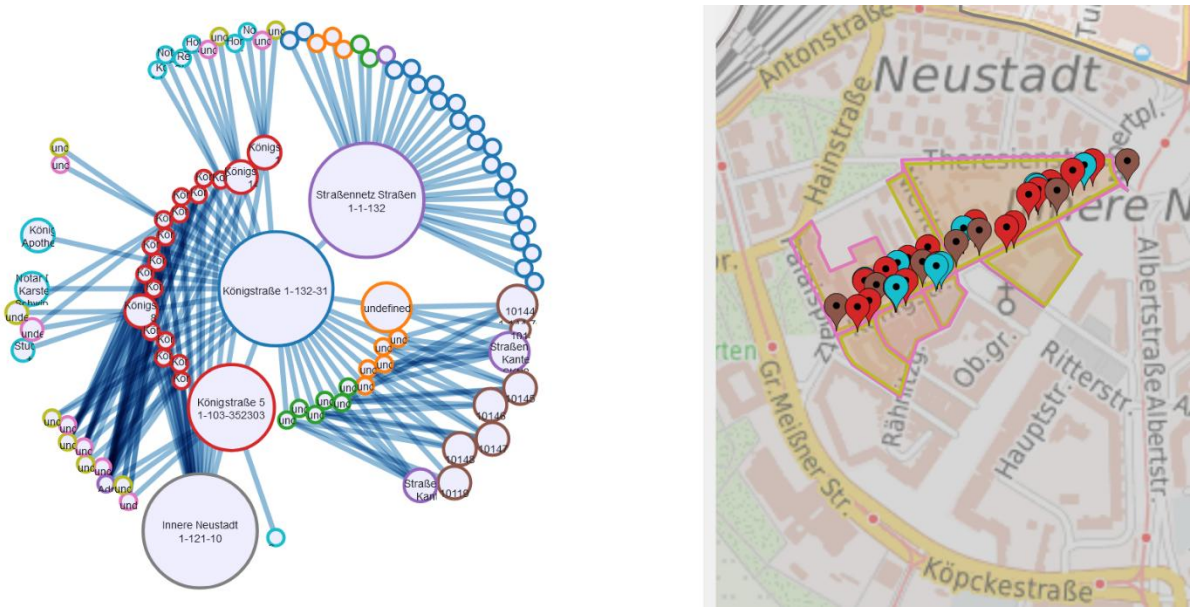

Abbildung 4: Beispiel für die Datenhaltung in einem Wissensgraphen mit **Simplex**4TwIS

Mit einer einzigen Abfrage in **Simplex**4TwIS können z. B. alle Objekte entlang einer Straße (hier die Königstraße in Dresden) selektiert werden. Links sehen wir in einem Graphen die Königstraße mit allen ihr zugeordneten Objekten: Adressen (rot), Straßenknoten (braun), Nachbarstraßen (als Linien), kulturelle, soziale und technische Einrichtungen (blau), dem zugehörigen Stadtteil (als Fläche) ... Rechts werden die Objekte grafisch dargestellt. Wir können uns immer weiter im Graphen vom Objekt zu einem benachbarten Objekt usw. bewegen. Die Daten wurden in Zusammenarbeit mit dem Vermessungsamt Dresden aus über 100 Quellen aus dem Open-Data-Portal in **Simplex**4TwIS übernommen und zusammengespielt. Die Abfrage erfolgte über den Simplex API-Features Service.

Wie revolutionieren wir die Verarbeitungsprozesse mit Simplex4TwIS?

In der Digitalstrategie der Bundesregierung [1] sind die Zielsetzungen für die digitale Transformation definiert und auch konkret bis 2025 benannt. Die interoperable Bereitstellung von weiterverarbeitbaren Daten ist dabei eine Kernaufgabe. Das „Wie“ kann aus der Digitalstrategie indirekt abgeleitet werden:

„Der Zugang zu Daten und Diensten über offene und einheitliche Schnittstellen ist deutlich erweitert und verbessert. Die Interoperabilität von Daten, insbesondere Mess- und Forschungsdaten, wird basierend auf internationalen Standards sichergestellt. Damit ist eine wichtige Grundlage gelegt für die Entwicklung innovativer Geschäftsmodelle, für Wertschöpfungen durch digitale und soziale Innovation durch zivilgesellschaftliche Akteure, Start-ups, Mittelstand und Industrie.“ [1, S.8]

Der Weg geht also über offene und einheitliche internationale Standards. Das sehe ich heute als eine Selbstverständlichkeit an. Die Annahme im obigen Zitat, dass sich der Zugang zu den Diensten deutlich erweitert und verbessert hat, kann m. E. einerseits mit der größeren Mengen an verfügbaren Open Data und andererseits mit der Ergänzung neuer fachlicher Datenmodelle begründet werden. Vergleichen wir die bereitstehenden Open Data aber mit den Anforderungen aus den Gesetzen (s. o. unter „Rechtliche Voraussetzungen“), so steht uns heute nur ein ganz kleiner Teil der verpflichtend bereitzustellenden Geodaten über Services zu Weiterverarbeitungen zur Verfügung. Beachten wir (wie oben erläutert), dass die

neuen fachlich-phänomenalen Datenmodelle zu einem „Wuchern“ an Datenstrukturen führen, so wird die programmtechnische Weiterverarbeitung für anspruchsvolle Fragestellungen (z. B. komplexe Projekte, Simulationen, Einsatz künstlicher Intelligenz) immer aufwendiger und komplizierter.

Das ist auch nicht verwunderlich, wenn wir bedenken, dass wir heute noch die Datenbereitstellungen mit über 20 Jahre alten Methoden umsetzen und viele auch an diesen Methoden festhalten wollen (zumeist mit der Frage „Wie sollt es denn auch anders gehen?“ begründet). Auch durch die Digitalstrategie [1] wird suggeriert, dass wir bei der Datenbereitstellung wichtige Grundlagen für Innovationen und Wertschöpfungen bereits verfügbar haben und einfach so weiter machen können. Diese Einschätzung empfinde ich als sehr oberflächlich, und dieses „Weiter so“ der Bedeutung der Digitalisierung nicht angemessen. So werden wir niemals die notwendige Verpflichtung erfüllen können, alle Geodaten aktiv für neue Wertschöpfungen zur Verfügung zu stellen. Wir laufen den Daten immer weit hinterher, beschäftigen dabei sehr viele Fachkräfte und verbrauchen große Ressourcen, und alles kostet uns riesige Summen.

Dieses „Weiter so“ ist so lange vertretbar, so lange es keine Alternativen dazu gibt. Aber mittlerweile geht es mit **Simplex**4TwIS erwiesenermaßen auch anders. Und in den Praxisprojekten zeigen sich allein für die Datenbereitstellung Kosteneinsparungen von bis zu 80%; es ist noch gar nicht absehbar, was bei der Datennutzung alles möglich und bezahlbar wird.

Wie verändern sich die Prozesse bei Nutzung von **Simplex**4TwIS?

- Wir fassen die Fachsysteme nicht an, sondern nutzen deren Möglichkeiten zur Datenbereitstellung. Es entfallen Programmanpassungen in den Fachsystemen.
- Es gibt nur noch das eine (!), bereits komplett ausgearbeitete Datenmodell. Sollten neue Fachinhalte bereitgestellt werden, entfällt die doch sehr aufwendige und von Experten vorzunehmende Modellierung.
- Das „Mapping“ (die Zuordnung der fachlichen Quellstrukturen zu den Zielstrukturen) ist weiterhin der aufwendigste Prozessschritt. Er vereinfacht sich aber enorm, indem in **Simplex**4TwIS zu den Importklassen und -feldern direkt die Zielobjekte und deren Attribute eingetragen werden. Mit dem sehr überschaubaren Zielmodell wird diese Zuordnung deutlich einfacher als bisher.
- Die eigentlichen Konverierungsprozesse können automatisch aus den in **Simplex**4TwIS eingetragenen „Mappings“ gebildet werden. Es entfällt die Erstellung der zum Teil recht komplizierten ETL-Prozesse.
- Die automatisch erzeugten Konvertierungsprozesse sind qualitativ zu kontrollieren, ggf. sind komplizierte Konvertierungsalgorithmen manuell zu ergänzen. Bei diesem Schritt ist auch zu empfehlen, weitere beschreibende Metadaten in **Simplex**4TwIS einzutragen.
- Durch eine automatische Rückkonvertierung der Importergebnisse in Quellstrukturen ist eine sehr effiziente Qualitätssicherung gegeben.
- Sind die Daten importiert, so stehen sie sofort als API-Features Service in einem Simplex-Standardszenario zur Verfügung. Das Aufsetzen von Diensten entfällt komplett.
- Die Metadaten sind nichts anderes als eine spezielle Fachapplikation in **Simplex**4TwIS, die nach genau denselben Strukturen aller importierten Fachdaten aufgebaut ist. Damit gibt es automatisch ein Simplex-Metadatenszenario, sodass auch die Metadaten über den API-Features Service bereitgestellt werden. Ein gesondertes Aufsetzen von Metadatendiensten entfällt.
- Beliebige Simplex-Fachszenarien werden auf ganz einfache Art und Weise mit dem ScenarioEditor kreiert (siehe Abbildung 5). Die Simplex-Fachszenarien stehen natürlich auch automatisch als API-Features Service zur Verfügung. So können in ganz kurzer Zeit neue Sichten auf die Daten definiert und als Dienste zur Weiterverarbeitung angeboten werden.

- Die Daten aller Szenarien stehen via Dienst zur Verfügung. Und sie haben immer die gleichen Datenstrukturen, was ihre Weiterverarbeitung massiv erleichtert! Die Erstellung neuer Wertschöpfungen wird wesentlich vereinfacht. Die **Simplex**Services können auch in GIS u. a. Fachsysteme eingebunden werden.

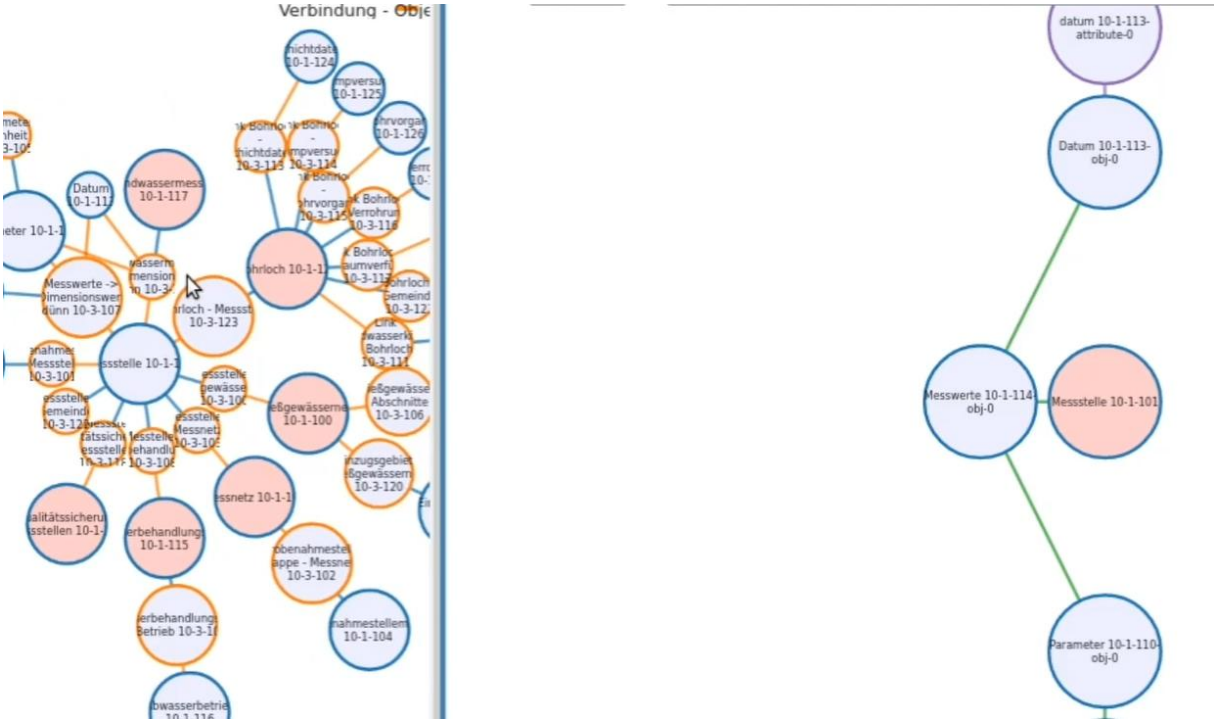

Abbildung 5: Beispiel für die Erstellung eines Simplex-Fachszenario mit dem **Simplex**ScenarioEditor

Die Darstellung auf der linken Seite erfolgt mit **Simplex**Visual (vgl. mit Abbildung 4). In diesem Fall werden lediglich die Metadaten selektiert, was durch dieselben Datenstrukturen bei Fach- und Metadaten möglich wird. Wir sehen die Objektklassen (blau umrandet) und deren Verbindungen (orange umrandet). Im ScenarioEditor können wir jetzt einzelne Objektklassen in das rechte Fenster ziehen. Wenn wir dort eine Objektklasse anklicken (im rechten Fenster rot markiert) werden im linken Fenster die über Verbindungen zugeordneten Objektklassen ebenfalls rot markiert. Eine dieser Klassen kann jetzt ausgewählt und ins rechte Fenster gezogen werden. Der ScenarioEditor erstellt automatisch die Verbindungsabfragen. Ergänzend können im rechten Fenster spezielle Suchfelder (violett umrandet) und auch Filter gesetzt werden. Das Ergebnis wird eine konkrete Sicht, zu der dann auch der API Service automatisch aufgesetzt ist.

Auch das Management mit Metadaten ist komplett modernisiert und vereinfacht worden. Dadurch entstehen Synergieeffekte, die wir im ScenarioEditor ausnutzen. Das Erstellen beliebiger Dienste über die im ScenarioEditor kreierten Sichten eröffnet ungeahnte Möglichkeiten der Datenbereitstellung.

Ausblick, Möglichkeiten und Chancen

Haben wir uns eigentlich schon gefragt, warum die INSPIRE-Datenbereitstellungen denn so wenig genutzt werden und ob es überhaupt neue programmtechnische Wertschöpfungen unter Nutzung der INSPIRE-Services gibt? Ähnlich sieht es doch auch bei anderen Datenbereitstellungen nach der heute üblichen Methode entsprechend der ISO 191xx-Normen [G] aus. Sie werden doch fast ausschließlich für Downloads oder als Schnittstellen verwendet, nicht aber in innovative Neuentwicklungen integriert.

Es gibt sehr schöne Lösungen für Datenbereitstellungen, die von den Mitarbeitern/-innen der datenhaltenden Stellen mit viel Aufwand und Engagement erarbeitet wurden. Die aufbereiteten Daten werden in sehr ansprechenden Portalen präsentiert und meist zum Download angeboten. – Es ist einfach schade: ein Riesenaufwand und hohe Kosten mit relativ wenigen Ergebnissen, gemessen an den Herausforderungen bei der Bereitstellung der Geodaten!

In unserer Facharbeitsgruppe INSPIRE-Umweltüberwachung haben wir mehrfach diskutiert, dass die für INSPIRE gewählte Methode nach der ISO 191xx-Serie für die Bereitstellung von Umweltdaten nur bedingt geeignet ist. Mit einigen Kollegen/-innen kamen wir zu der Erkenntnis, dass es anders gehen muss! Das inspirierte mich zu meinem Buch [2], wo ich diese Methoden analysierte und einen alternativen Modellansatz ausarbeitete. In [3] habe ich dann begründet, warum die gewählte Art der Modellierung nach den ISO 191xx-Normen [G] zu diesen phänomenalen Datenmodellen und damit zu den heute noch anzutreffenden komplizierten Datenbereitstellungen führen muss.

Die Umsetzung der von mir neu ausgearbeiteten Methode wurde vom Bundesministerium für Digitales und Verkehr im mFUND-Projekt envVisio-GI [B] gefördert.

Wir sind uns sicher, dass es einen Paradigmenwechsel bei der Datenbereitstellung geben wird und damit auf ganz einfache Art und Weise die Daten interoperabel und verarbeitbar angeboten werden. Die Datenbereitstellungen werden nachhaltig, die Digitalisierung wird durch die Verfügbarkeit der Daten in einfach umsetzbaren Strukturen einen riesigen Schub erfahren.

Simplex4TwIS ist offen zugänglich und entsprechend dokumentiert. **Simplex**4TwIS soll allen die Möglichkeit eröffnen, ihre Daten auf einfache Art und Weise interoperabel zu nutzen und auch anderen zur Verfügung zu stellen.

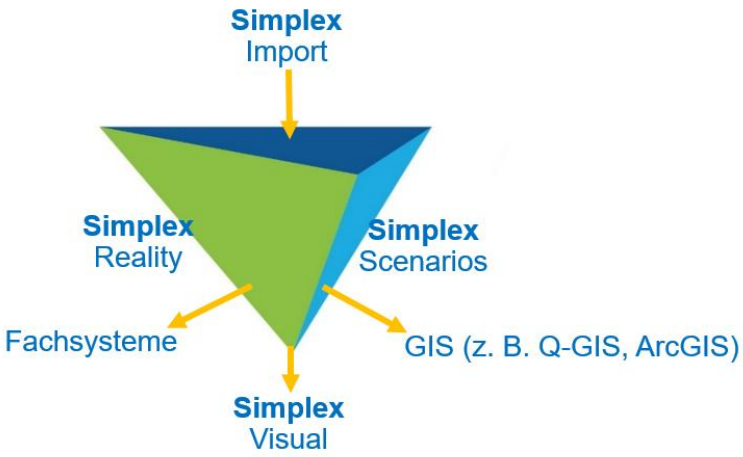

Abbildung 5: Unsere Daten neu denken

Für Nutzer/-innen des Systems ergibt sich die Möglichkeit, Daten aus den Fachsystemen herauszuheben und mit Daten anderer Fachsysteme zu verknüpfen. Sie erhalten einen Überblick über alle Daten und Verbindungen.

Die **Datenbereitsteller/-innen** müssen nicht mehr programmieren bzw. ETL-Prozesse erstellen; sie konfigurieren einfach. Das Aufsetzen von Diensten für die Geo- und Metadaten entfällt komplett, da in **Simplex**4TwIS die API-Features Services automatisch generiert werden. Es wird alles deutlich überschau- und beherrschbarer, mit weitaus effektiveren Möglichkeiten der Qualitätssicherung.

Programmierer/-innen müssen nicht erst die Bereitstellungsstrukturen studieren, auf immer gleiche Art und Weise werden die Daten geliefert. Damit wird die Einbindung der Dienste sehr einfach, und sie können sich voll auf die Programmieraufgabe konzentrieren. Das wird dann ganz sicher zu neuen Wertschöpfungen (z. B. neue Berechnungsalgorithmen, Prozessunterstützungen oder Anwendungen mit künstlicher Intelligenz) führen.

Planer/-innen können die aktuell noch heterogen angebotenen Daten in **Simplex**4TwIS importieren. Die Prozesse der Datenaufbereitung werden sich deutlich vereinfachen. Sie gewinnen wieder Zeit für ihre eigentlichen Fachaufgaben.

Und für alle Anwender/-innen gibt es noch den ScenarioEditor, mit dem schnell und einfach gewünschte Sichten auf die Daten zusammengestellt oder (besser gesagt) grafisch zusammengeklickt werden können - ohne dafür irgendetwas programmieren zu müssen. So können in wenigen Minuten neue Datenzusammenstellung via API-Features Service konfiguriert werden, die dann jeder wieder für seine Aufgaben nutzen kann.

Das Produkt ist fertig. Da es themenunabhängig die Daten generalisiert, via Metadaten beliebige fachliche Strukturen bearbeiten kann und die Metadatenerfassung im System integriert ist, sind für den Einsatz von **Simplex**4TwIS keine Konzeptionen oder Spezifikationen notwendig. **Simplex**4TwIS kann (so wie es ist) einfach ausprobiert und eigentlich auch sofort eingesetzt werden. Aktuell entschließen sich immer mehr Fachbehörden und Kommunen, ihr Datenmanagement mit **Simplex**4TwIS zu komplettieren bzw. auf innovative moderne Methoden umzustellen.

Wesentliche Informationen stehen unter www.simplex4data.de.

Dresden, September 2023

Veröffentlicht in:

- Bundesingenieurkammer (Hrsg.): Deutsches Ingenieurblatt 9-2022, Schiele & Schön GmbH, Berlin 2022
- Harzer, C. (Hrsg.): GIS-Report 2022/23, Bernhard Harzer Verlag GmbH, Karlsruhe 2022

Überarbeitete Version:

- Harzer, C. (Hrsg.): GIS-Report 2023/24, Bernhard Harzer Verlag GmbH, Karlsruhe 2023

Quellenverzeichnis

[1] Digitalstrategie der Bundesregierung, 2022: https://digitalstrategie-deutschland.de/static/fcf23bbf9736d543d02b79ccad34b729/Digitalstrategie_Aktualisierung_25.04.2023.pdf (abgerufen am 29.08.2023)

[2] Rudolf, H.: Umweltdatenmanagement: Eine Geo-Inspiration. Bernhard Harzer Verlag GmbH, Karlsruhe 2018

[3] Rudolf, H.: Umweltdaten-Intelligenz. In: Freitag, U., Fuchs-Kittowski, F., Abecker, A., Hosenfeld, F. (Hrsg.): Umweltinformationssysteme – Wie verändert die4Digitalisierung unsere Gesellschaft? Springer Vieweg, Wiesbaden 2020

[4] Spitz, M.: Daten – das Öl des 21. Jahrhunderts?: Nachhaltigkeit im digitalen Zeitalter. Hoffman und Campe, Hamburg 2017

[5] Humby, C.: Talk at ANA Senior marketer's summit, Kellogg School, 04.11.2006

[6] Merkel, A.: Video-Podcast zur Eröffnung der CeBIT, Presse- und Informationsamt der Bundesregierung (BPA), 12.03.2016

[7] Klemm, S.: INSPIRE und Open Data für Dienste zum automatisierten Fahren. – Eine Untersuchung der Verwertbarkeit dieser Daten im Rahmen des Forschungsprojektes Cartox², Masterarbeit Hochschule Anhalt, Dessau-Roßlau 2019

[8] Großmann, J.: envVisio Service – ein universeller Dienst für Umweltdaten. In: Fuchs-Kittowski, F., Abecker, A., Hosenfeld, F. (Hrsg.): Umweltinformationssysteme – Wie trägt die Digitalisierung zur Nachhaltigkeit bei? Springer Vieweg, Wiesbaden 2022

[9] Umweltbundesamt: Umweltmanagement und Digitalisierung - Praktische Ansätze zur Verbesserung der Umweltleistung), Dessau-Roßlau 2019

[A] Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE). In: ABl. der EU, L 108, S.1-14, 2007

[B] envVisio-GI - GeodatenIntelligenz für interoperables Datenmanagement mit der envVisio-Methode – envVisio-GI, mFUND-Projekt des Bundesministeriums für Digitales und Verkehr, https://www.bmvi.de/SharedDocs/DE/Artikel/DG/mfund-projekte/envvisio-gi.html (aufgerufen am 29.08.2023)

[C] XML in der öffentlichen Verwaltung (XÖV). Offizielle Website zum Standard: https://www.xoev.de/ (aufgerufen am 29.08.2023)

[D] Gesetz über den Zugang zu digitalen Geodaten (Geodatenzugangsgesetz GeoZG) vom 10.02.2009. BGBl. I S. 278

[E] Umweltinformationsgesetz (UIG) vom 08.07.1994. BGBl. I S. 1490; Neubekanntmachung vom 27.10.2014. BGBl. I S. 1643

[F] Geologiedatengesetz (GeolDG) vom 19.07.2020. BGBl. I S. 1387

[G] ISO 191xx-Serie – Standardnormen der International Organization for Standardization zur Bereitstellung von Geodaten