

## **Zusammenfassung vom Workshop *Standards für einen aktiven und nachhaltigen Datenaustausch* (14.06.2021)**

Im Zuge einer Analyse vorhandener Ressourcen sowie auch Wünschen hinsichtlich der technischen Infrastruktur und Dienstleistungen für eine kollaborative Datenumgebung, fanden bisher drei Workshops statt. Im Rahmen dieses Workshops zu Standards für einen aktiven und nachhaltigen Datenaustausch trafen sich Teilnehmende aus unterschiedlichen Forschungseinrichtungen, aus dem Qualitätsmanagement und aus der technischen Redaktion, um sich in Diskussionsrunden zu folgenden Themen auszutauschen: (1) *Standards für semantische und technische Interoperabilität*, (2) *Interoperabilität von organisatorischer und rechtlicher Seite*, (3) *Spezialbereiche und fehlende Standards sowie eine mögliche Entwicklung, Etablierung, Anpassung sowie* (4) *Dienstleistungen für Standardisierung, Validierung*.

**Standards erlauben es Ressourcen zu sparen und eine Erleichterung der täglichen Arbeit durch Vereinheitlichung zu erreichen!**

Eine Umfrage zum Einstieg in das Thema erhob die allgemeine Einstellung der Teilnehmenden zum Umgang mit Standards. Zur Frage des Zusammenhangs von Standards und Innovationen herrschte allgemeine Übereinstimmung, dass sich diese gegenseitig begünstigen. Es wurde jedoch angemerkt, dass diese auch Gegenpole darstellen können. Alle Befragten arbeiten zumindest teilweise mit Standards, im Einzelfall bestimmen Spezifikationen sogar den Arbeitsalltag. Die Mehrheit gab an, dass Spezifikationen zwar existieren, aber manche Prozesse erst definiert werden müssen. In einem Fall gibt es im individuellen Arbeitsbereich noch keine definierten Spezifikationen oder es wird bisher erfolgreich ohne etablierte Standards gearbeitet.

Ein Impulsvortrag von Seiten technischer Dokumentation zum Thema „Leben im Standarddschungel für eine verständliche, mediengerechte und gesetzeskonforme Erklärung komplexer Produkte sowie Prozesse im Digitalisierungszeitalter“ diente als Auftakt zur nachfolgenden Diskussion. Es wurde eine Übersicht über rechtliche, normative, technische, wissenschaftliche oder auch vermischte, branchenspezifische bzw. interne Standards im Bereich technischer Dokumentation gegeben.

Projektrelevanz liegt vor aufgrund der Vorgaben zur Mensch-Maschinen-Interaktion gemäß DIN EN ISO 9241 aus dem normativen Bereich, Darwin Information Typing Architecture und DocBook aus dem wissenschaftlichen Bereich sowie der Intelligent Information Request & Delivery Standard aus dem technischen Bereich. Letzteres umfasst ein Metadatenmodell über Informationsart, Produkt-, Funktions-, Verwaltungs- sowie Navigationsmetadaten für den Aufbau von Verzeichnisstrukturen als Resource Description Framework (RDF). Dieses kann Schritt für Schritt über den Standard hinweg branchenspezifisch zu großen Graphenstrukturen mit einer Vernetzung anhand genormter Terminologie führen. Somit würde das Metadatenmodell, als Konstrukt über den eigentlichen Daten liegend, den Austausch dieser erleichtern.

Zum Thema semantischer Interoperabilität wurde anfangs das Thema Software-Maschinen-Kommunikation am Beispiel Computer-aided Modelling angeschnitten, um nachzuvollziehen wozu Standards existieren, aber von Herstellern oftmals abgewandelt werden (müssen). Solche Modellierungssprachen spielen vor allem in der Betrachtung von Teilfragen eine wichtige Rolle. Die Differenzierung zwischen semantischer und

technischer Interoperabilität ergibt sich anhand des Beispiels des RDF, welches als Rahmenwerk dient. Darauf aufbauend gibt es spezifische Modelle in individuellen Domänen. Da Standardisierung einen sehr langsamen Prozess darstellt, wird es immer Heterogenität und sowohl neue Technologien als auch Legacy-Systeme geben. Als realistischer Ansatz wird hierzu empfohlen auf offene Datenformate zurückzugreifen und zusätzlich Metadaten für eine partielle automatisierte Integration unterschiedlicher Domänen zu nutzen. Der Bereich Semantik könnte zusätzlich weiter aufgliedert werden.

Am Beispiel Förderierten Lernens wurde der Punkt Sicherheit zum Bereich technischer Interoperabilität angeschnitten. Dort gibt es ein Spektrum an Verfahren, die unterschiedliche Ausprägung an Performance, Anwendbarkeit oder Sicherheitsgarantien ergeben. Dieses umfasst eine vollständige gemeinsame Nutzung, bilaterale Vereinbarungen über Aggregate bis hin zu Technologien zur Verbesserung der Privatsphäre (u.a. durch homomorphe Verschlüsselung, Differential Privacy oder MPC). Durch die unterschiedlichen Anwendungsfälle ist es schwer eine gesamtheitliche Lösung zu finden, daher sollte das gesamte Spektrum unterstützt werden.

Technische Interoperabilität umfasst auch den Unterpunkt Cloudmanagement. Dieser wird von zwei großen Lagern geprägt, die entweder eine öffentliche oder private Dateninfrastruktur bevorzugen. Hier stehen vor allem die Kostenersparnis gegenüber datenschutzrechtlichen Aspekten im Vordergrund. Eine Tendenz zu hierarchischen Strukturen oder Hybridmodellen wird beobachtet.

Zum Thema organisatorischer und übergreifend auch rechtlicher Interoperabilität wurde angemerkt, dass viele bereichsspezifische Reglements existieren, allerdings ohne Vorgaben für die digitale Welt. Insbesondere bei der Zusammenarbeit von Forschungsorganisationen mit der Industrie kann sich eine Diskrepanz zwischen wirtschaftlichen Interessen, rechtlicher Notwendigkeit gemäß DSGVO und akademischer Veröffentlichungsarbeiten ergeben. Wiederholt

wurde der Handlungsbedarf genannt, einen Mittelweg zu finden, um veröffentlichen und forschen zu dürfen, ohne sich durchgehend mit rechtlichen Grundfragen auseinandersetzen zu müssen. Ein solches bereichsübergreifendes rechtliches Grundgerüst sei bereits mancherorts in Entwicklung. In Bezug auf Validierung und Qualität von Daten oder Applikationen sind Automatismen kaum möglich. Datenqualitätsmodelle seien meist generisch und müssten anwendungsspezifisch gestaltet werden. Forschungsdaten sind auf Basis der Erfahrung oft intuitiv handzuhaben. Eine Unterstützung für Informationsprozesse ist hierbei essentiell.

Fehlende Standards wurden in Spezialbereichen wie Künstlicher Intelligenz oder zu aktuellen Vorhaben der Data Spaces, am Beispiel des Gesundheitsbereichs, genannt. Leitfäden und technische Richtlinien dienen oft als Alternative zu Normen, deren Erstellung, Übersetzung und Anpassung sich als langwierige Prozesse herausstellen und zusätzlich Probleme hinsichtlich des Preises und Lesebeschränkung aufweisen. Im Bereich Datenmanagement fand auch eine Bottom-Up Entwicklung tatsächlicher Standards statt, die auf einer ursprünglichen praktischen Einführung bestimmter Formate, aber auch Anfragesprachen beruht.

Dienstleitungen zur Standardisierung können wie Dienstleistungen im Allgemeinen gehandhabt werden und fehlende Expertise ausgleichen. Im Zuge der aktuellen EU-Regularien um Maschinelles Lernen ist nun ein sehr spezifischer Registrierungsprozess von Hochrisikoapplikationen künstlicher Intelligenz mit einer detaillierten Dokumentation zu Datenvalidierung vorgegeben. Als mögliche Dienstleistung kann eine ähnliche Befolgung von Verhaltenskodexen auch bei allen anderen Applikationen angeboten werden. Weiters spielen nach wie vor eine klare Kommunikation und Marketingaspekte, beispielsweise durch Pilotprojekte, Branchen-Praxisbeispiele als auch Mundpropaganda im Zuge von Digitalisierungsprojekten eine große Rolle, um den Wert von Daten zu vermitteln. Am Beispiel Data Spaces würde eine klarere Kommunikation auch gegen immer

wiederkehrende Fragen nach dem Nutzen helfen. In diesem Zusammenhang wurde wiederum angemerkt, dass Veröffentlichen nicht mit Verschenken gleichzusetzen ist und auch nicht bedeutet, sich Kritik unterziehen zu müssen, sondern primär eigene Werke zu schützen. Weiters muss vermittelt werden, dass Standards letztendlich Ressourcen sparen und eine Vereinheitlichung als Erleichterung der täglichen Arbeit zu sehen ist. Eine entsprechende Datenaufarbeitung kann auch als Schutzfunktion dienen, wenn ein Nachweis zu Spezifikationen,

Konfigurationen oder andere Nachweise zu Fragestellungen verlangt werden und bietet dadurch auch Rechtssicherheit. Wirtschaftliche Aspekte und Statistiken können mitgeführt werden und eine Einschulung von Mitarbeitern verläuft effizienter. Neben der Bereitstellung von Daten, helfen auch die Bewerbung von darauf basierenden neuen Geschäftsmodellen.

An dieser Stelle bedanken wir uns bei allen Workshop-Teilnehmenden für ihre aktive Beteiligung und wertvollen Beiträge!