

Die „Data Intelligence Offensive“ (DIO) zielt auf die Forcierung und Förderung der Datenwirtschaft und des optimierten Einsatzes von Technologien ab, vorrangig der Künstlichen Intelligenz (KI). DIO fördert Perspektivenwechsel von einem Austauschkonzept zu einem Handelskonzept, von Data Exchange zu Data Sharing und Trading. DIO bezweckt die Implementierung von europäischen Werten wie Datenschutz und Datensouveränität im technischen Grunddesign und in den regulatorischen Vorschriften.

Stellungnahme

zum Austrian Micro Data Center

Erweiterung für ein Österreichisches Datenökosystem für Forschungsdaten

29. Juli 2021. Wien, Data Intelligence Offensive.

Das Bundesministerium für Bildung, Wissenschaft und Forschung (BMBWF) plant ein zentrales Datenspeicherungszentrum für Grundlagenforschungsdaten unter dem Schirm der Statistik Austria, das sogenannte „Austrian Micro Data Center“ (AMDC).

Die Data Intelligence Offensive (DIO), Österreichs Vernetzungs- und Austauschplattform für die Data Community, setzt sich für das Teilen von relevanten Daten ein und befürwortet daher auch die rechtliche Schaffung des AMDC.

Die DIO begrüßt die Zielsetzung und Grundintention, öffentliche Registerdaten mit der Wissenschaft teilbar zu machen.

Allerdings entspricht ein zentrales Datenspeicherungszentrum nicht mehr dem Stand des Wissens und der Technik.

Die DIO fordert deshalb gleichzeitig eine Erweiterung vom AMDC für ein GAIA-X-konformes dezentrales Daten-Ökosystem, das Kooperation, Dezentralität und Datensouveränität als zentrale Maximen verfolgt.

Hintergrund: Wozu ein Ökosystem?

Ökosysteme sind nichts Neues. In einem der ältesten Ökosysteme – dem Ökosystem Wald – ermöglicht das Zusammenspiel von vielfältigen Akteur*innen aus der Pflanzen- und Tierwelt ein gesundes Miteinander. Sind solche Beispiele auf die Datenwelt anwendbar? Ist dies ein Vorbild für komplexe Sicherheits- und Wertschöpfungsstrukturen in Datenräumen? Ja! So wie es in der Natur gelingen muss, einen Ausgleich zwischen den Interessen aller Teilnehmer*innen des Ökosystems zu schaffen, gilt es auch in einer Datenwelt, die Interessen aller Akteur*innen bestmöglich zu wahren. Ein Datenwert generiert sich erst durch die Verbindung mit weiteren Daten. Die Zusammenarbeit mit Dritten erfordert oft zusätzliche Ressourcen, wie zum Beispiel von Rechtsexpert*innen unterstützte Verhandlungsteams, um die Zusammenarbeit mit potenziellen Data Providern und Data Usern auszuhandeln und zu strukturieren. Hierbei können Technologien wie Smart Contracting oder API Management entscheidend von Vorteil sein.

Was ist ein Datenökosystem?

Einfach ausgedrückt stellt ein Datenökosystem eine dezentrale Vernetzung von Daten dar. Die Daten von zahlreichen Anbieter*innen werden kombiniert und durch die Nutzung dieser verarbeiteten Informationen wird ein Mehrwert geschaffen. Ein erfolgreiches Ökosystem balanciert drei Prioritäten:

- 1) **Data Sovereignty as Guiding Light:** Datensouveränität beziehungsweise Datenhoheit bezeichnet die größtmögliche Kontrolle und Herrschaft über eigene Daten. Diese muss im Kontext von Datennutzung als

oberste Prämisse gesehen werden und kann durch zentrale Speicherung und Verarbeitung gefährdet werden.

- 2) **Economy of Scale:** Durch die Skalierung von Ausgangsdaten und deren Verarbeitung können neue Produkte geschaffen werden. Zudem muss das Ökosystem über das Kernprodukt hinaus klare Kund*innennutzen und Abhängigkeiten generieren, um langfristig einen Nutzen zu etablieren. Dadurch steigert sich der Innovationsgehalt der Datennutzung rasant.
- 3) **Using the same productivity:** Der Aufbau eines Kooperationsnetzwerks motiviert eine Vielzahl von Parteien mit ähnlichen Interessen (am konkreten Beispiel Forschungseinrichtungen), ihre Kräfte zu bündeln und ähnliche Ziele zu verfolgen. Einer der Hauptvorteile eines Datenökosystems ergibt sich aus der Beteiligung mehrerer Kategorien von Akteuren (wie App-Entwicklern und App-Benutzern). Dadurch werden Parallelstrukturen vermieden, wodurch sich wiederum der Innovationsgehalt erhöht und der Wissenschafts- und Entwicklungsstandort stärker befeuert wird.

Was macht ein dezentrales Datenökosystem sinnvoller als ein zentrales Datenspeicherungszentrum?

Sicherheit: Zentrale und daher auch Verdoppelung von Datenspeicherung hat sich oft als problematisch erwiesen, als leichter manipulierbar (defacing). Diese mangelnden Sicherheitsaspekte können mit einem dezentralen Ansatz mitigiert werden. Gleichzeitig bringt die Verknüpfung die gleiche Verfügbarkeit wie eine zentrale Datenbank, nur sicherer. Denn auch wenn „die Daten das Haus nicht verlassen“, heißt das nicht, dass niemand die Daten aus dem Haus entwendet oder manipuliert. Durch die zentrale Speicherung und Pseudonymisierung sind alle Prozesse potentiellen Eingriffen von Fremden ausgesetzt. Zusätzlich zum technologischen Aspekt, werden rechtliche und ethische Rahmenbedingungen benötigt, die erst adäquat ausgearbeitet werden müssen (wie z.B. klar definierte Zugangs- und Nutzungsregelungen). In einem dezentralen Datenmarkt (wie es nationale und internationale Initiativen anstreben, z.B. Data Market Austria, GAIA-X, IDSA, BDVA, etwaige EU-Projekte, etc.), wären diese Regelungen bereits definiert, und wären zusätzlich innovationsfördernd (direkte und indirekte Wertschöpfung entstehen).

Europäischer Weg: Daten machen also nicht an Gebäudegrenzen halt, so auch nicht an nationalen Grenzen. Daher ist es wichtig, in pan-europäische Datenökosysteme eingebunden zu sein und so als starker Akteur wichtige Werte (z.B. Datensouveränität) zu vertreten. Die Argumentation, der Wissenschaftsstandort würde durch das Austrian Micro Data Center gestärkt werden, verlangt aus Sicht der DIO, dass eine solche Stärkung im Kontext unserer globalen Gesellschaft nur stattfinden kann, wenn die Innovationsaktion mit einem richtigen Technologieansatz und unter der Wahrung der Privatsphäre (privacy preservation) innovative Anwendungsforschung auf nationalem und pan-europäischem Level berücksichtigt.

Wir fassen daher zusammen: Der Aufbau eines für alle Stakeholder*innengruppen gewinnbringenden Systems, kann nur über die Säulen Kooperation, Souveränität und vor allem Dezentralität gewährleistet werden.

Daten bleiben an keinen physischen Grenzen stehen, daher ist es wichtig, internationale (pan-europäische) Lösungen zu suchen, die mehr Innovation, mehr direkte und indirekte Wertschöpfung fördern und gleichzeitig die Souveränität des Individuums höchstens achten (nach dem Schema von Datenräumen/Data Spaces).

Das AMDC ist ein guter erster Schritt – konkrete Schritte für die Erweiterung in Richtung dezentralisiertes Ökosystem sollen aber gleichzeitig unternommen werden, um den Zugang zu Micro- und Registerdaten für die Wissenschaft und hier gerade auch für die vorbeugende Gesundheitsforschung sowie die Wirtschafts- und Sozialforschung nachhaltig und zukunftsorientiert zu verbessern.