



Bild: stock.adobe.com\_kossmoss\_339202972

Landwirtschaft, Gesundheit und Klimaschutz: Umweltdaten werden zugänglicher

# Datenräume und ihr Mehrwert für Geodateninfrastrukturen

Der Klimawandel, der Verlust von Ökosystemen und viele andere Veränderungen stellen uns als Gesellschaft vor große Herausforderungen, für die wir Lösungen finden und optimieren müssen. Dabei spielen Daten eine zentrale Rolle. Durch Copernicus, die Inspire- und die Open-Data-Direktive stehen immer mehr offene Geodaten zur Verfügung. Dennoch sind diese nur die Spitze des Eisbergs: Der weit größere Teil der relevanten Daten wird auch zukünftig aufgrund von Sicherheitsanforderungen, wegen gesetzlicher Vorgaben wie der DSGVO oder aufgrund wirtschaftlicher Interessen nicht öffentlich bleiben. Um diese Daten nutzbar zu machen, baut die neue *europäische Datenstrategie* auf einem Konzept auf, das sich unter anderem in der Automobilindustrie bewährt hat: den sogenannten Datenräumen (engl. data spaces). Im Folgenden werden die Unterschiede zwischen Geodateninfrastrukturen (GDI) und Datenräumen erläutert, welche Probleme durch den Aufbau von Datenräumen gelöst werden können und was bei der Umsetzung zu beachten ist.

Autor: Thorsten Reitz

In der Automobilindustrie geht es darum, effektive und effiziente Lösungen für hochgradig komplexe Probleme zu finden. Die Produkte müssen zahlrei-

chen rechtlichen und wirtschaftlichen Anforderungen gerecht werden. Die Entwicklungs- und Produktionskette hierfür ist extrem komplex. Um sie

sicherer und kostengünstiger zu machen, tauschen Automobilhersteller und Zulieferer schon lange in begrenztem Umfang Daten aus.

Ursprünglich hatte aber jedes Unternehmen nur ein Teilbild der gesamten Lieferkette, was von den Bildern der anderen stark abweichen konnte. Weil dies immer wieder zu Problemen führte, wurde der Catena-X-Datenraum umgesetzt. Dieser ermöglicht es allen teilnehmenden Organisationen, ihre Daten in einen gemeinsamen, sicheren Datenraum einzubringen. Wer dann was von diesen Daten sehen und für welchen Zweck verarbeiten darf, entscheiden die Organisationen gemeinsam.

Weil sich dieser Ansatz gut bewährt, und weil sich auch in vielen anderen Bereichen ein enormes Verbesserungspotenzial hinter bisher verschlossenen Daten versteckt, baut die 2021 veröffentlichte europäische Datenstrategie stark auf das Konzept der Datenräume. Diese sollen neun strategische Bereiche unterstützen: Gesundheit, Umwelt, Energie, Landwirtschaft, Mobilität, Finanzen, Fertigung, öffentliche Verwaltung und Kompetenzen. Weiterhin soll ein Datenraum geschaffen werden, der Daten für die Umsetzung des Green Deals enthält. In jedem der Datenräume werden öffentliche und nichtöffentliche Daten von Verwaltungen und Unternehmen zusammengefasst.

**Was ist ein Datenraum?**

Im Mittelpunkt des Datenraum-Konzepts steht die Governance, d.h. die gemeinsamen Regeln und Standards sowie deren technische Umsetzung. Diese definiert, welche Rollen es im Datenraum gibt, und was welche Rolle jeweils mit den Daten

tun darf. Beispielsweise können mehrere Dateneigentümer ihre Daten in einem Pool für das Training von KI-Modellen einander zugänglich machen, aber den Export dieser Daten außerhalb des Datenraums stark einschränken. Neben diesen Regeln müssen auch gemeinsame technische Standards festgelegt werden. Dazu gehören insbesondere Datenmodelle wie zum Beispiel die Inspire-Datenspezifikationen, XPlanung oder 3A/NAS für den Geo- und Umweltbereich.

Wie auch bei einer Geodateninfrastruktur liegen die Quelldaten verteilt vor. Jede Organisation kann ihre eigene Datenkomponente auf einer Infrastruktur nach ihren eigenen Anforderungen betreiben, sei es On-Premise oder in der Cloud. Über einen Adapter, beispielsweise den Eclipse Dataspace Connector, stellt sie den kontrollierten Zugang zu ihren Daten sicher.

Alle Daten in einem Datenraum sind semantisch interoperabel. Das bedeutet nicht zwangsläufig, dass sie alle in genau dem gleichen Schema und Format vorliegen, aber dass sie bei Bedarf vollautomatisiert integriert und harmonisiert werden können. Hierfür kommen Matching- und Mapping-Technologien wie die Annotation zum Einsatz („Dies ist ein Flurstück“). ETL-Werkzeuge wie „hale»studio/hale»connect“ können solche Annotationen verwenden, um Daten aus verschiedenen Teilen des Datenraums automatisiert für verarbeitende Dienste aufzubereiten.

Solche Verarbeitungsdienste sind ebenfalls Teil des Datenraums. Durch die ge-

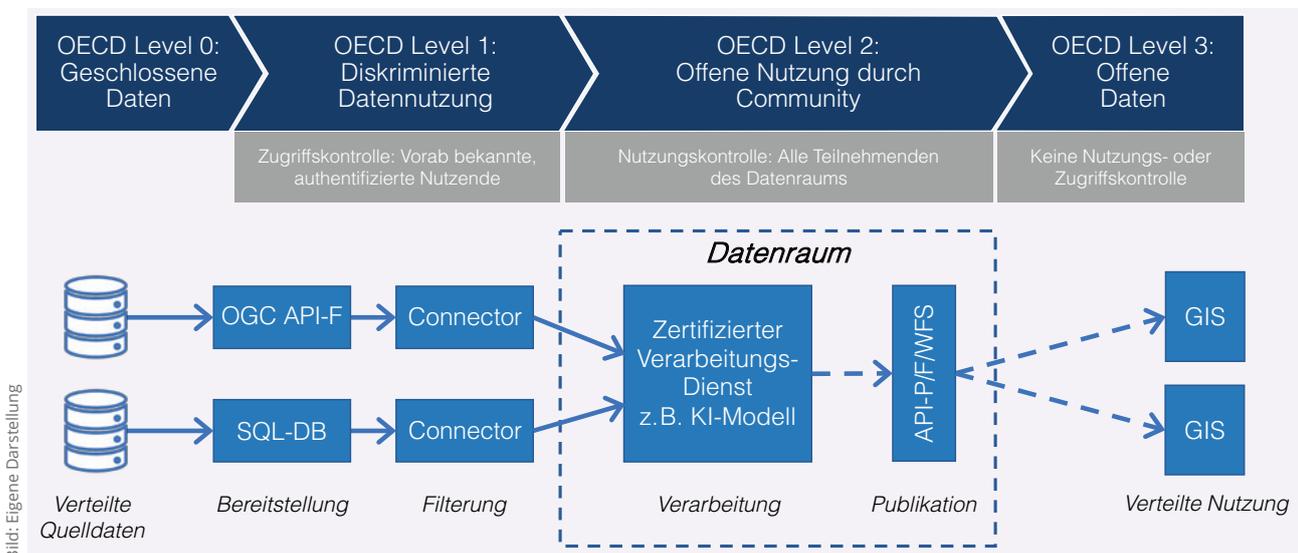
meinsamen Regeln wird festgelegt, was diese Dienste mit den Daten machen dürfen, also ob beispielsweise eine Zwischenspeicherung erfolgen darf. Vertrauen spielt dabei eine wichtige Rolle. Ab 2022 wird es möglich sein, Verarbeitungsdienste zu zertifizieren. Ist ein Dienst zertifiziert, wissen alle am Datenraum Beteiligten, dass dieser Dienst genau das tut, was er gemäß seiner Dienstbeschreibung ankündigt.

**Welches Problem löst ein Datenraum?**

Der Aufbau eines Datenraums ergibt nur dann Sinn, wenn es für konkrete Anwendungsfälle essenzielle Datenlücken gibt, die durch Nutzung bisher nicht zugänglicher Daten geschlossen werden können. Diese Datenlücken müssen zunächst so klar wie möglich beschrieben werden.

Eine solche Datenlücke besteht auch dann, wenn Daten zwar vorhanden sind, aber nicht in ausreichender Menge, um leistungsfähige KI-Modelle zu trainieren. Ein Datenraum kann in einem solchen Fall wesentlich mehr Trainingsdaten zur Verfügung stellen. Da nur das fertige Modell aus dem Datenraum exportiert wird, ist die Vertraulichkeit der Trainingsdaten sichergestellt.

Datenräume lösen noch ein weiteres Problem: Heute schöpfen Plattformen in großem Stil Daten ab und verwerten diese, ohne dass die „Datensubjekte“ – die betroffenen Personen und Unternehmen – dies beeinflussen oder an der Wertschöpfung teilhaben können. Im Datenraum



Grade der Datenoffenheit bei Nutzung eines Datenraums

können Regeln definiert werden, die neben der Souveränität über die eigenen Daten auch eine balancierte Wertschöpfung möglich machen, wie sogenannte „Pay as you go“-Modelle.

Die Datensouveränität kann allerdings nur sichergestellt werden, wenn der Datenraum auch auf Software-, Betriebssystem- und Hardwarekomponenten aufgebaut wird, welche entsprechend entworfen und abgesichert sind. Aus diesem Grund werden die Anforderungen an die Infrastruktur der Datenräume gemeinsam mit GAIA-X, der verteilten Cloud-Plattform für Europa, abgestimmt und umgesetzt.

### Und ganz konkret?

Das alles lässt sich gut veranschaulichen anhand des aktuellen KI-Leuchtturmprojekts „FutureForest.ai“. In diesem schaffen wetransform zusammen mit der TU München und der TU Berlin sowie mit mehreren Staatsforsten und forstlichen Versuchsanstalten einen Datenraum für forstliche Daten. Die Daten aller dieser Organisationen werden zusammengeführt, um Entscheidungen zum klimaangepassten Waldumbau zu unterstützen. Dafür werden öffentliche Daten, wie Höhenmodelle und Landbedeckungskarten, genutzt, sowie private Daten, wie Sensordaten und die detaillierten Informationen aus der Standortkartierung. Auch die Waldeigentümer steuern Daten bei, um von besseren Entscheidungshilfen zu profitieren.

In den letzten zehn Jahren wurden im Rahmen von Geodateninfrastruktur-Initiativen bereits zahlreiche Daten zugänglich gemacht. Leider wurden organisatorische Rahmenbedingungen und die Nutzungsbedingungen für die Daten oft unzureichend betrachtet. Dies wird sich mit der Umsetzung von Datenräumen wie für die Fortwirtschaft ändern.

### Mehr als nur Projekte – die Environmental Data Spaces Community

Die Geo- und GIS-Community steht noch am Anfang, was die Einführung und Nutzung von Datenräumen angeht. Zurzeit werden zahlreiche nationale und internationale Projekte gestartet. Um diese zu vernetzen und eine Kontinuität in der Entwicklung über einzelne Projekte hinaus zu schaffen, hat wetransform die Environmental Data Spaces Community ins Leben gerufen. Gemeinsam mit mehreren Partnern und im Rahmen der International Data Spaces Association, welche die Standards für Datenräume definiert, unterstützt wetransform den Aufbau von vielfältigen Datenökosystemen, um Umweltdaten in einem sicheren Datenraum zugänglich und nutzbar zu machen und gleichzeitig die Souveränität der Datengeber zu schützen. Mehr zur Environmental Data Spaces Community und welches Engagement dort möglich ist, ist unter [1] zu finden.

#### Quelle:

[1] <https://greendataspace.com>

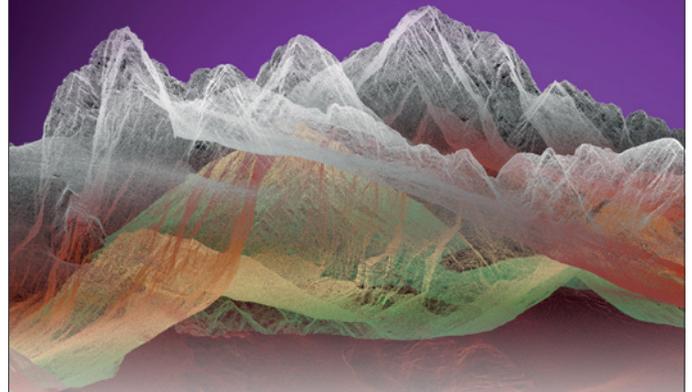
#### Kontakt:

Thorsten Reitz

wetransform GmbH

E: [tr@wetransform.to](mailto:tr@wetransform.to)

## RIEGL AIRBORNE LASER SCANNER & SYSTEME



### RIEGL WAVEFORM LIDAR TECHNOLOGIE für Vermessungseinsätze mittels Flugzeug oder Helikopter

Hier finden Sie den richtigen Scanner für Ihr Vermessungsprojekt!

#### VQ-1560 II / VQ-1560 II-S



Zweikanal-Systeme für die luftgestützte Vermessung von ausgedehnten Arealen und komplexen Umgebungen aus großer Flughöhe

- bis zu 2,66 Mio. Messungen/Sek. am Boden
- Hochleistungs-IMU/GNSS-System und Kamera(s) voll integriert

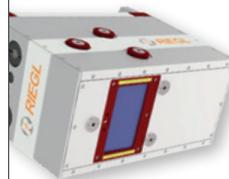


#### VQ-780 II / VQ-780 II-S



Luftgestützte Laserscanner für Kartierungsanwendungen aus großer Flughöhe

- bis zu 1,33 Messungen/Sek. am Boden
- optimiert für die Konfiguration von kundenspezifischen Vermessungssystemen



#### VQ-480 II / VQ-580 II



Luftgestützte Laserscanning-Systeme speziell konzipiert für die Korridor-Vermessung bei Flughöhen von bis zu 4.400 Fuß AGL

- optionale Integration von IMU/GNSS-System und Kamera(s)
- leichtes, kompaktes Design für einen Einsatz auch auf UAVs mit höherer Traglast



Informieren Sie sich auch über unsere Scanner für die UAV-gestützte und bathymetrische Vermessung [www.riegl.com](http://www.riegl.com)

newsroom.riegl.international



Austria | USA | Japan | China | RIEGL WORLDWIDE | Australia | Canada | United Kingdom