

**we**  
transform



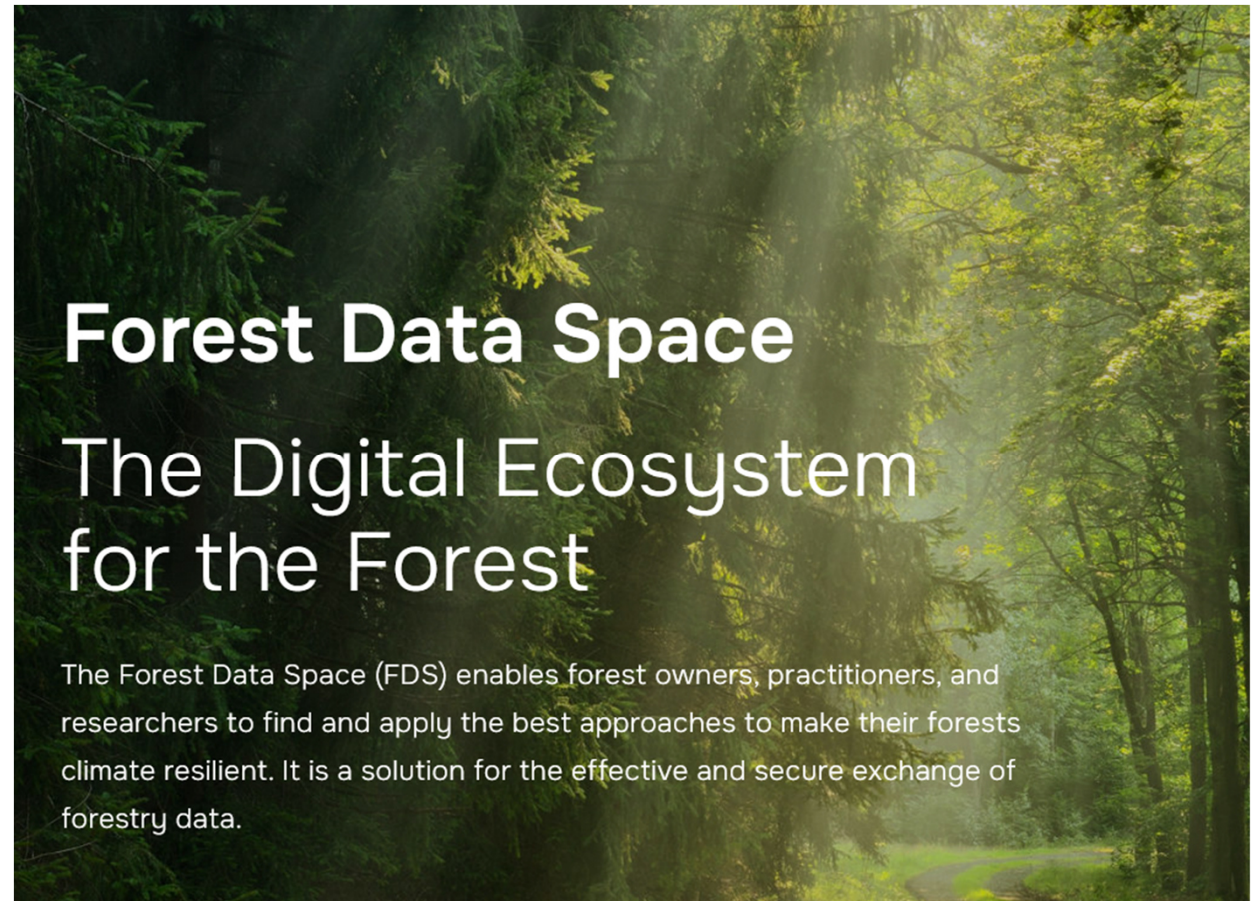
# INSPIRE-Workshop

## Datenräume – praktische Anwendungen

Thorsten Reitz, Founder/CEO wetransform GmbH  
Linz, 05.12.2024

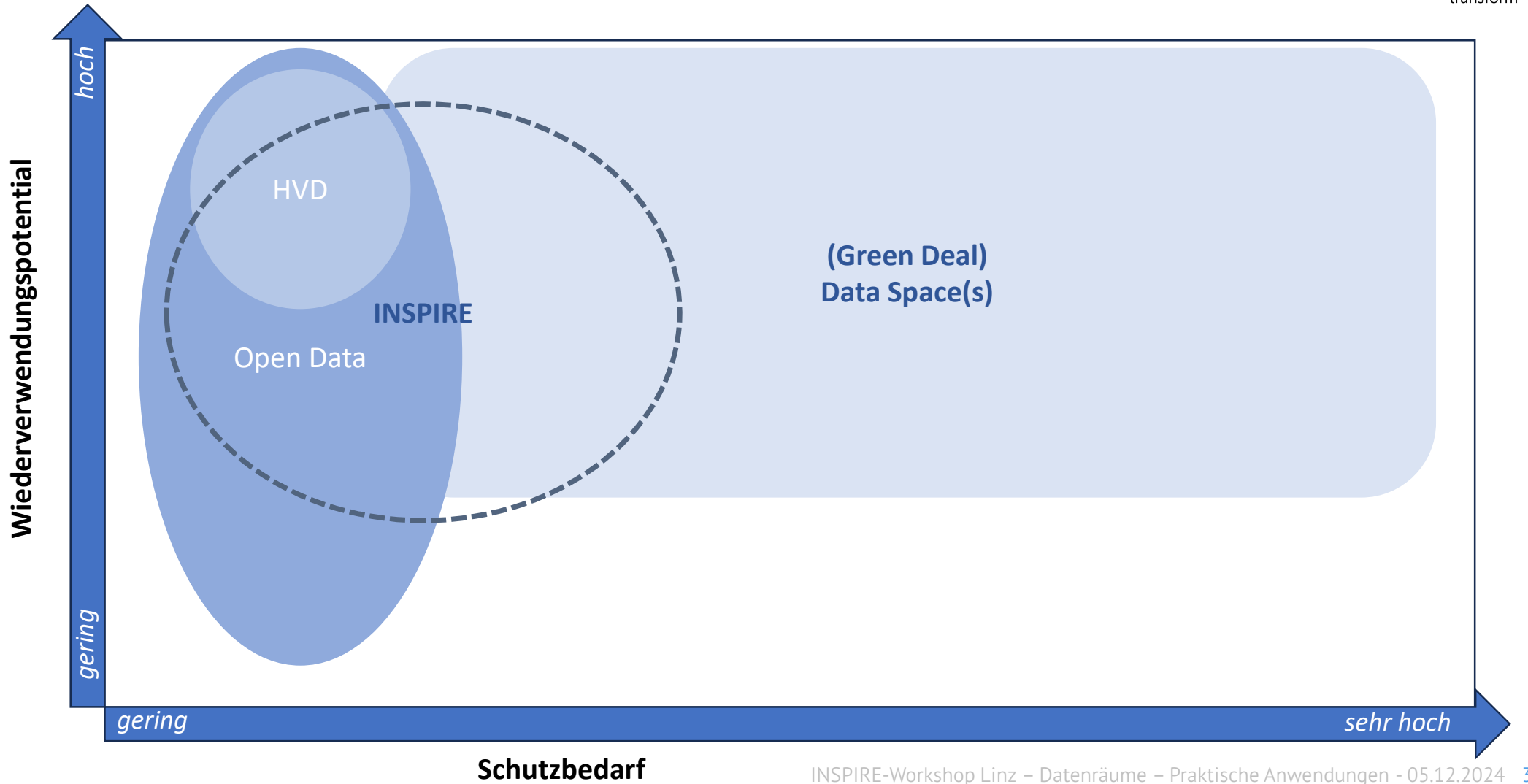
# Agenda

- 🌲 Welches Problem lösen wir jetzt eigentlich, wofür braucht man diese Datenräume wirklich?
- 🌲 Wie sehen Datenräume konkret aus?
  - 🌲 Forest Data Space
  - 🌲 InGeoDTM
  - 🌲 SAGE
- 🌲 Konkrete Architekturen
- 🌲 Interoperabilität in Datenräumen
- 🌲 Austausch



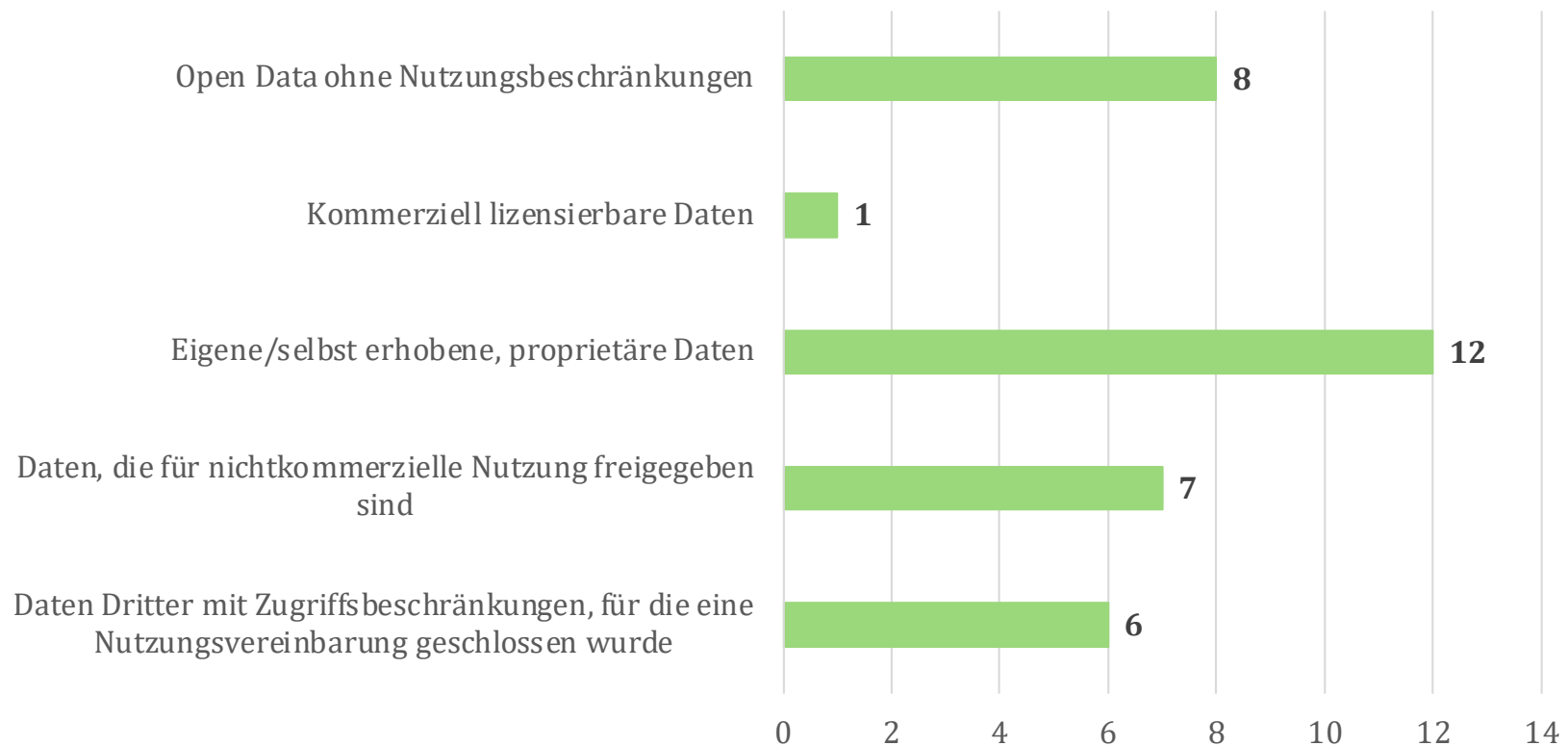
<https://forestdata.space>

# Mehr Daten wiederverwendbar machen



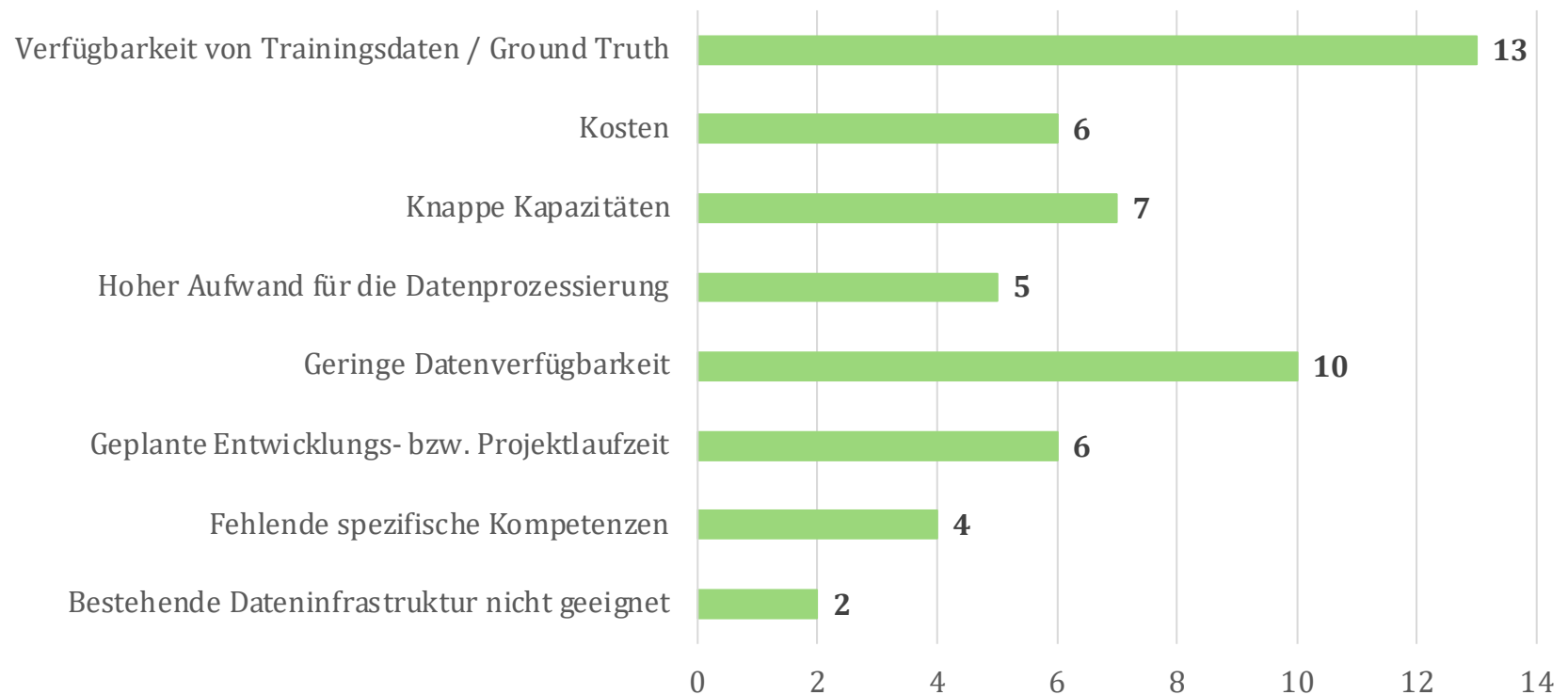
# Datenbedarfe

„Unter welche Datenlizenzen fallen die Daten, die Sie genutzt haben?“



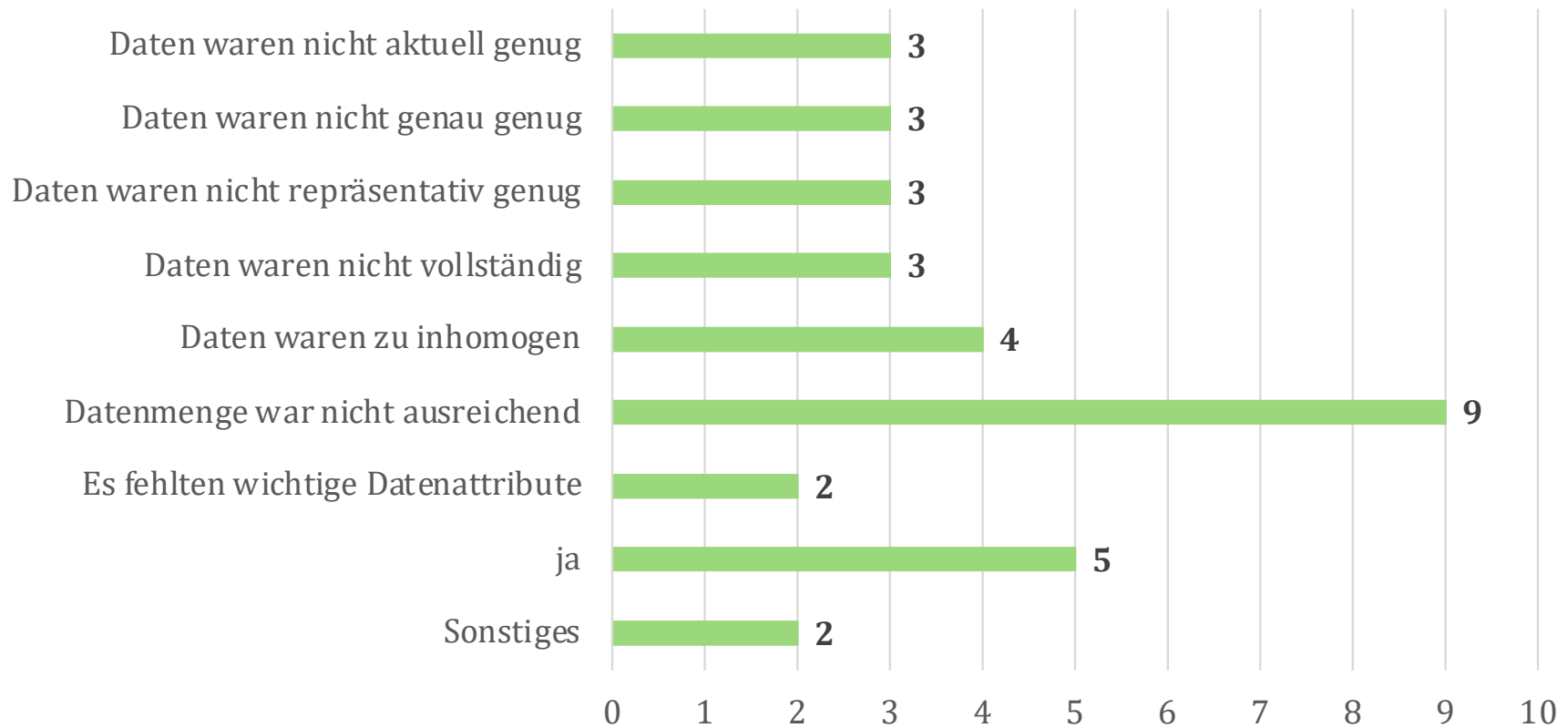
# Datenbedarfe

„Was sind ihre begrenzenden Faktoren bei der KI-Modellentwicklung?“



# Datenbedarfe

„War die Datenqualität für ihren Anwendungsfall ausreichend?“



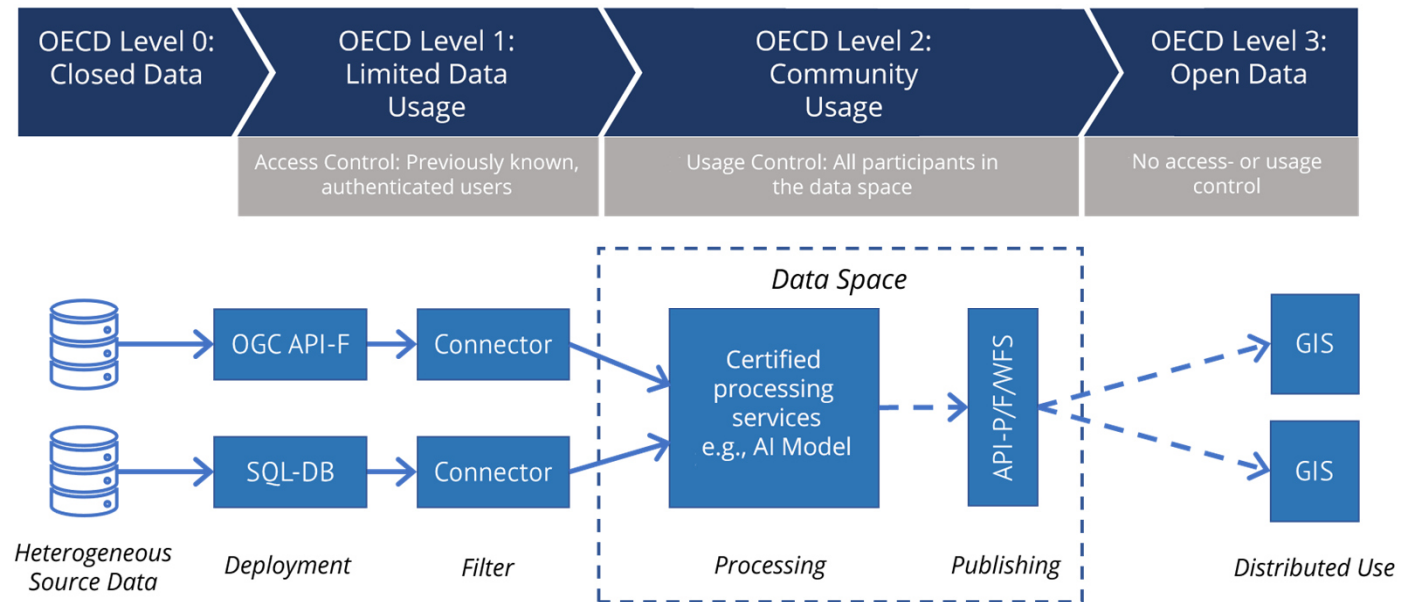
# Chancen von Datenräumen

- 🌳 Personenbezogene, proprietäre und sicherheitskritische Daten können unter kontrollierten Bedingungen genutzt werden
- 🌳 Pooling kleiner Datensätze führt zu besseren KI-Modellen
- 🌳 Standardisierte Anwendungen
- 🌳 Datenräume als perfektes Gegenmodell zu zentralisierenden, allmächtigen Plattformen
- 🌳 Souveränität auf Daten- und Verarbeitungsebene
- 🌳 Adaption allgemeiner, standardisierter Lösungen fördert Zugänglichkeit, reduziert Lock-In



# Datenräume: Schließen der Datenlücke

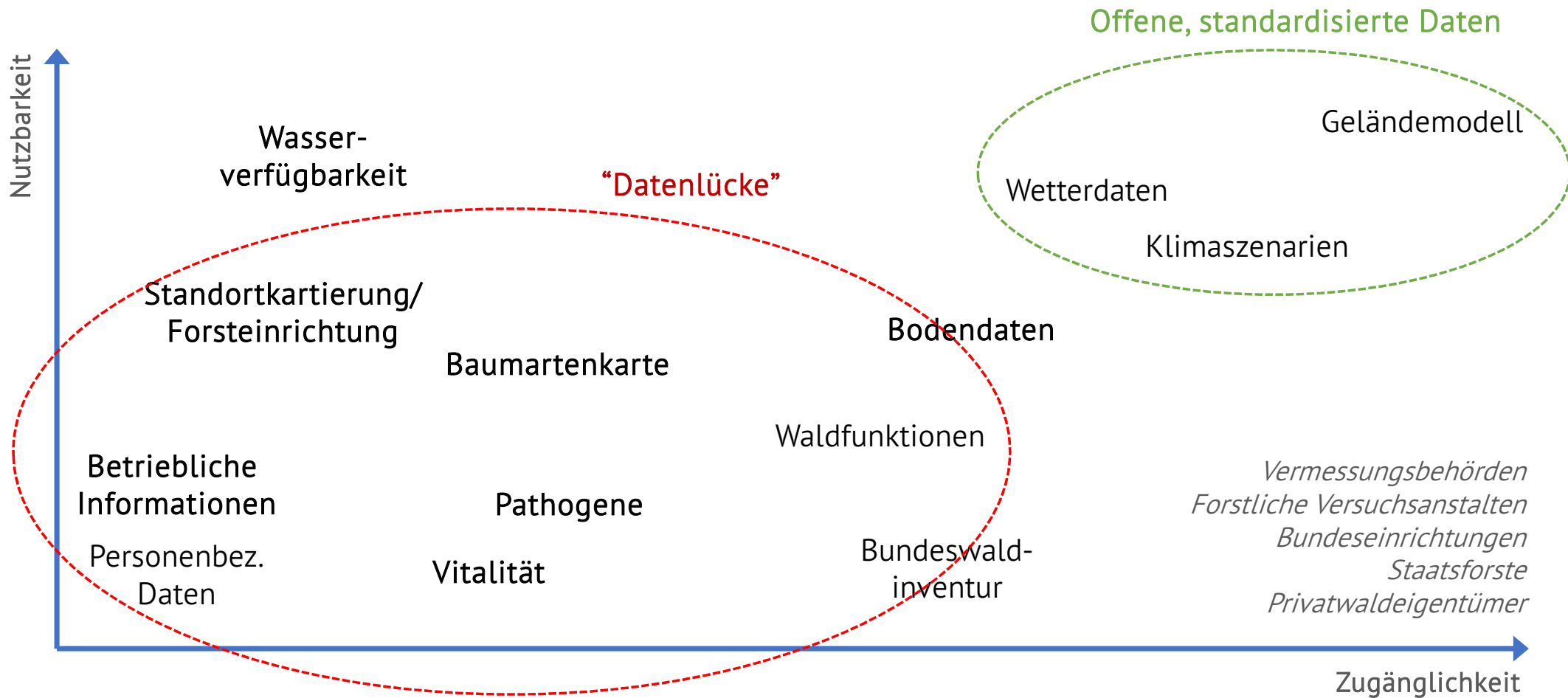
- ☛ Kontrolle über Datenzugang und Datennutzung für Datensubjekte und Datenanbieter
- ☛ Zugang zu wertvollen Daten zu transparenten Bedingungen
- ☛ Offenes mehrseitiges Ökosystem, es gibt keine Monopole





# Beispiel 1: Forest Data Space

# Waldumbau: Datenzugänglichkeit und Nutzbarkeit



# Für wen ist der Forest Data Space?

Waldumbau: Mit dem Klimawandel umgehen



| <b>Datenhaltende Stellen</b>  | <b>Multiplikatoren</b>  |
|---|---|
| Vermessungsverwaltungen<br>Forstverwaltungen<br>Forschungseinrichtungen<br>Fernerkundungsunternehmen    | Forstliche Versuchsanstalten<br>und Forschungs-<br>einrichtungen<br>Umweltbehörden    |
| <b>Nutzende</b>   | <b>Mehrwertdienste</b>  |
| Private Waldbesitzende<br>Staatsforste<br>Körperschaften,<br>Genossenschaften, ...<br>Forstverwaltungen | Saatgutanbieter<br>Baumschulen<br>Inventurapps<br>UAV-Betreiber<br>Maschinenverleiher |

## ... und welche Vorteile bringt der FDS?

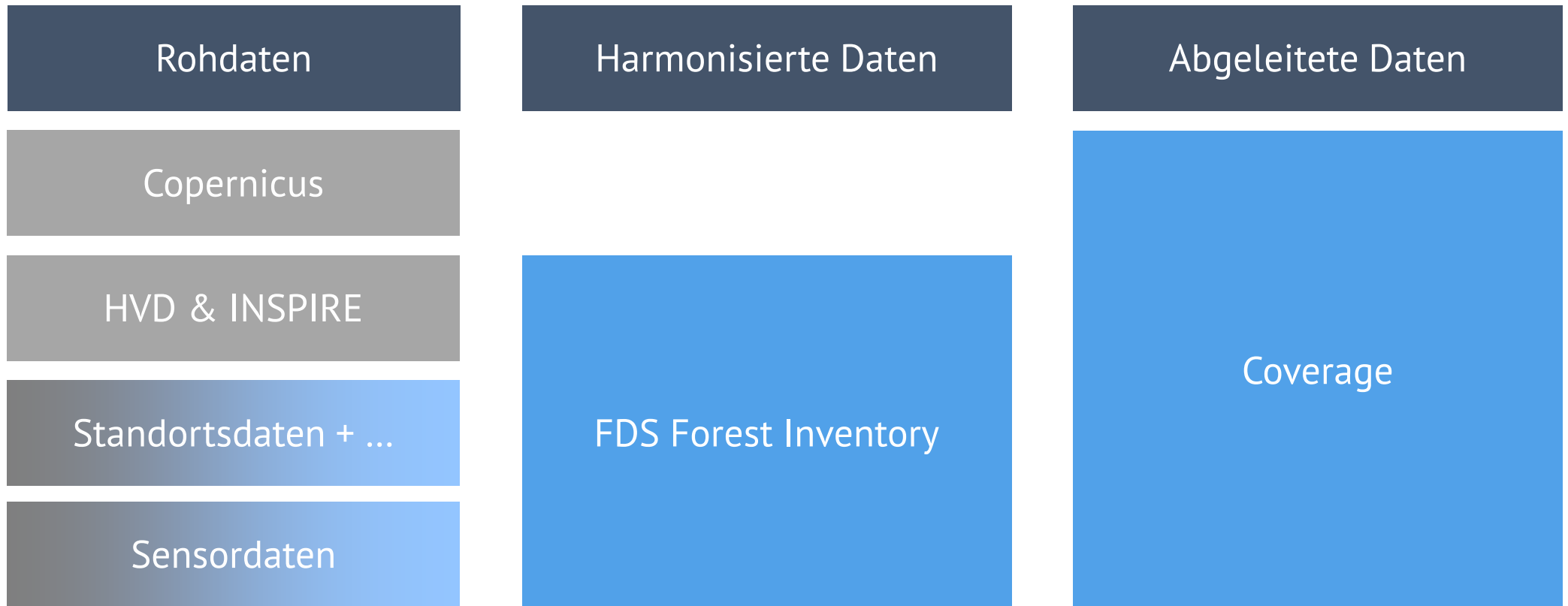
In offenen Ökosystemen wird Mehrwert geschaffen und fair verteilt



| Datenhaltende Stellen  | Multiplikatoren  |
|--|--|
| <p>Harmonisierung<br/>Zugriff auf Daten anderer datenhaltender Stellen<br/>Erfüllung rechtlicher Vorgaben</p> <p><i>Werteinheit: Datensatz</i></p> | <p>Maximieren der Nutzbarkeit und Wirksamkeit von Daten und Forschung<br/>Nutzende aktivieren</p> <p><i>Werteinheit: Datenpool</i></p> |
| Nutzende   | Mehrwertdienste  |
| <p>Einheitliche, vorbereitete Daten zur Entscheidungsunterstützung</p> <p><i>Werteinheit: Bestandsfläche</i></p>                                   | <p>Innovationen können schneller umgesetzt werden<br/>Kundengewinnung</p> <p><i>Werteinheit: API-Aufrufe</i></p>                       |

# Daten im Datenraum

Welche Daten können im Datenraum verwaltet und genutzt werden?

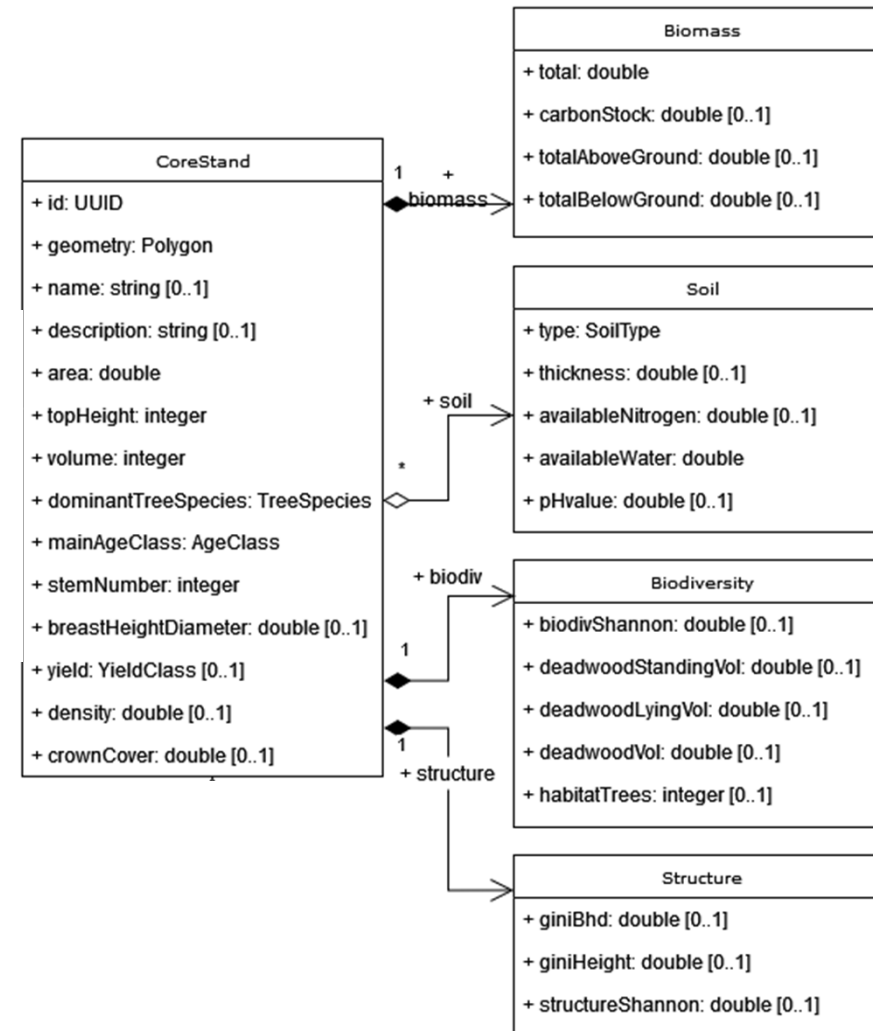




# Formate und Datenmodelle

## Standardisierung

- 🌳 Standardmodell für Inventurdaten
- 🌳 GeoJSON/JSON
- 🌳 GML/XML
- 🌳 GeoPackage
- 🌳 GeoTIFF
- 🌳 Typed Arrays



# Nutzungsbedingungen

## Wer darf wo was mit welchen Daten wofür machen?



*Waldbesitzende und – Praktiker, die im Namen von Waldbesitzern handeln*

und ihre eigenen Daten in den FDS hochladen. Sie werden gemäß der EU-Datenverordnung als „Datensubjekte“ bezeichnet.



*Software-Anbieter und -integratoren,*

Die auf Daten im FDS-Coverage Service und im ProMCDA-basierten Ranking-Dienst zugreifen möchten, um Szenarien in ihrer Software anzubieten



*Der Datenintermediär (Data Intermediary)*

betreibt die zentrale Infrastruktur des Datenraums und ist für die Entwicklung der technischen und organisatorischen Infrastruktur verantwortlich. Diese Rolle übernimmt derzeit wetransform GmbH.



*Forstdatenverwalter,*

die Daten hochladen, die sie erstellt oder gespeichert haben, um Bestände im Auftrag von Waldeigentümern und Praktikern zu verwalten. Sie werden gemäß der EU-Datenverordnung als „Dateninhaber“ bezeichnet und umfassen ebenfalls den Staatsforst, Forstverbände und Verwaltungsgesellschaften.



*Forschende,*

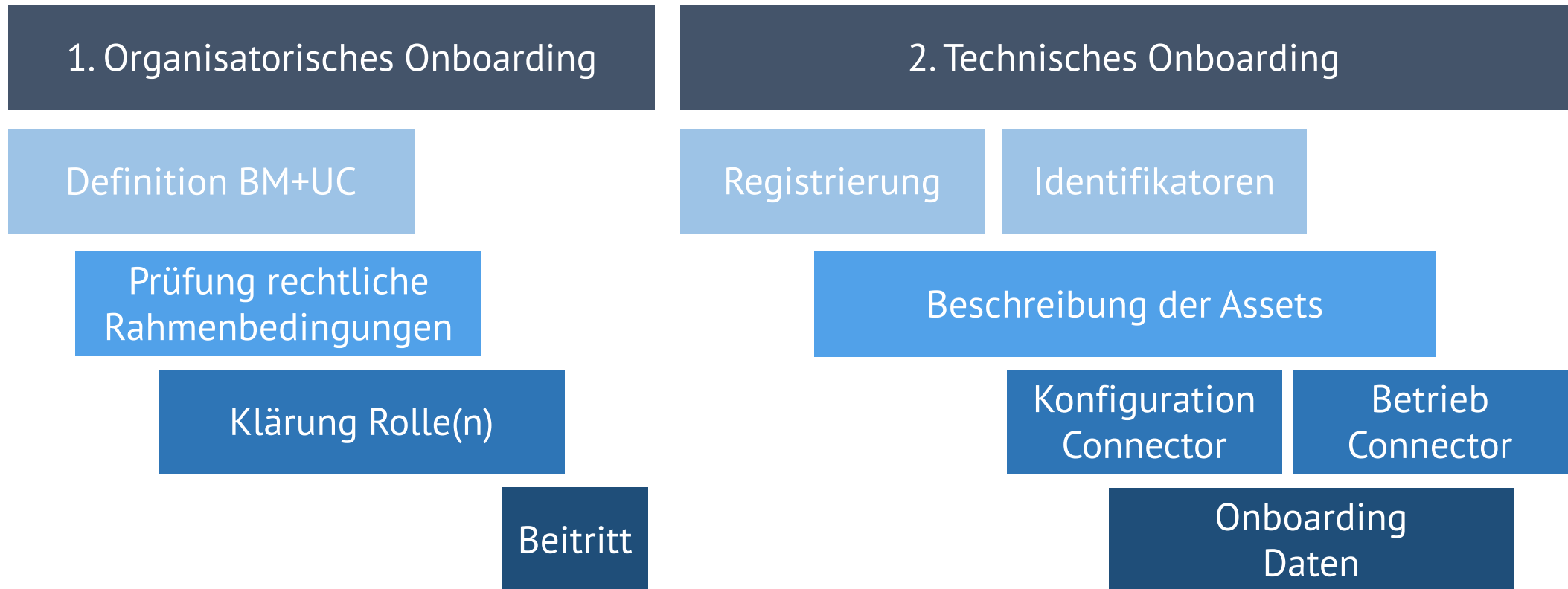
die auf Basis eines großen Datenpools neue Applikationen und Modelle entwickeln möchten, z.B. Vertreter Forstlicher Versuchsanstalten und Universitäten

<https://forestdataspace.com/de/datennutzungsrichtlinien/>



# Wie kann man Teil des FDS werden?

Beispiel: Als datenhaltende Stelle



# Beispiel 2: InGeoDTM

# Datentreuhänder: Der einfachere Weg?

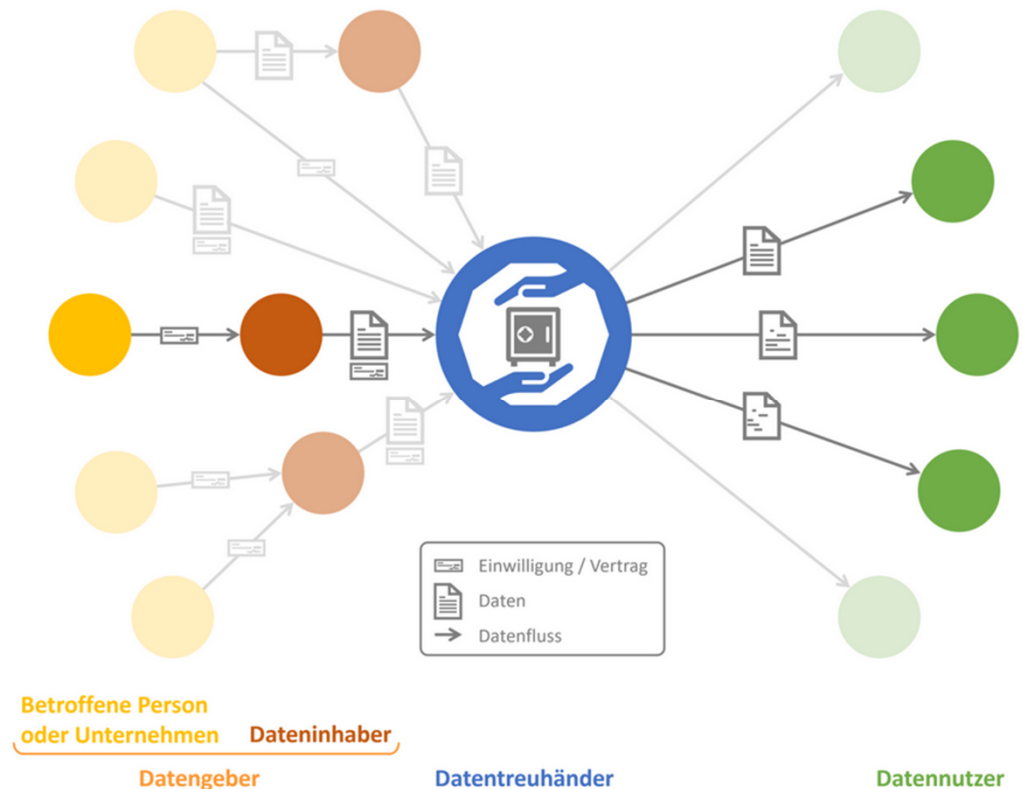
... vor allem für kleinere Organisationen?

## ■ Datentreuhänder

- „Ein Datentreuhänder ist eine Vertrauensinstanz, die schützenswerte Daten zwischen Datengebern und Datennutzern unter Wahrung der Interessen beider Seiten digital vermittelt.“
- Soll neutral und unabhängig sein
- Standardregeln zur Verwertung
- Übernimmt das technische Onboarding weitestgehend

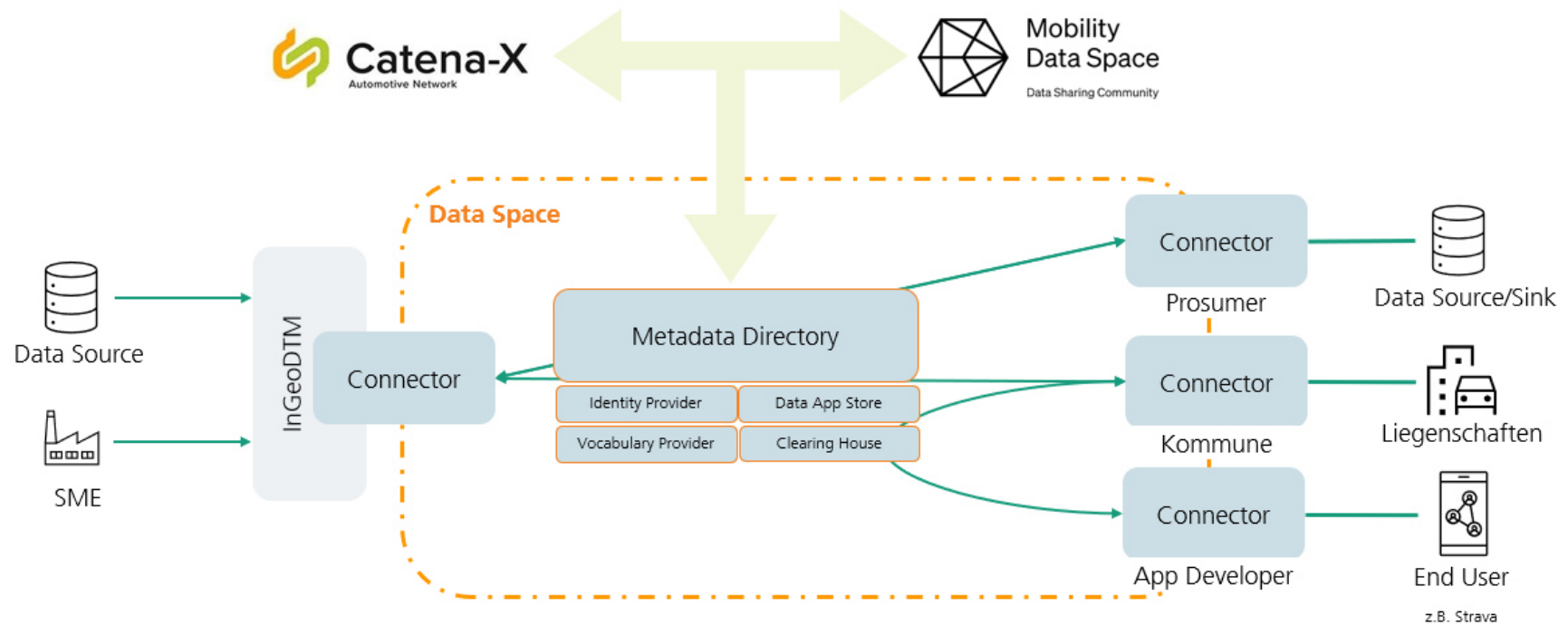
## ■ Datenintermediäre

- Im Data Governance Act definiert
- Muss Datenanbieter-neutral sein, darf also nicht bestimmte Daten vorziehen
- Muss den Datenzugang neutral gestalten, darf also nicht einzelne Nutzende bevorzugen
- **Darf Daten nicht selbst verwerten**



# InGeoDTM und InGeo-X

Datentreuhänder und Datenraum für schützenswerte Geodaten

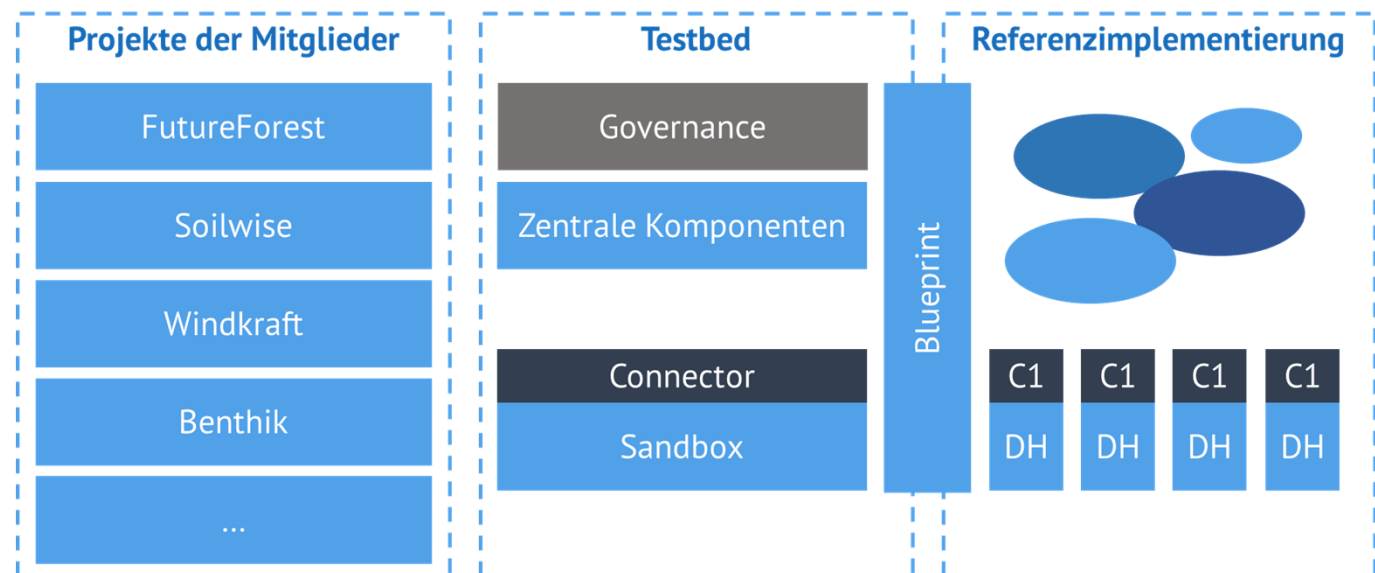


# InGeoDTM Blueprint

## Spezifische Lösungen für Geodaten in Datenräumen

Beitragende:

- Datenhaltende Stellen
- Daten- und dienstennutzende Organisationen
- Software-Entwickler (Dienste, Apps, Infrastruktur)
- Föderatoren und Anbieter zentraler Infrastrukturen (z.B. Identität, Clearing House + Broker, Vokabulare, Appstore...)
- Governance-Gremien (z.B. Interoperability)

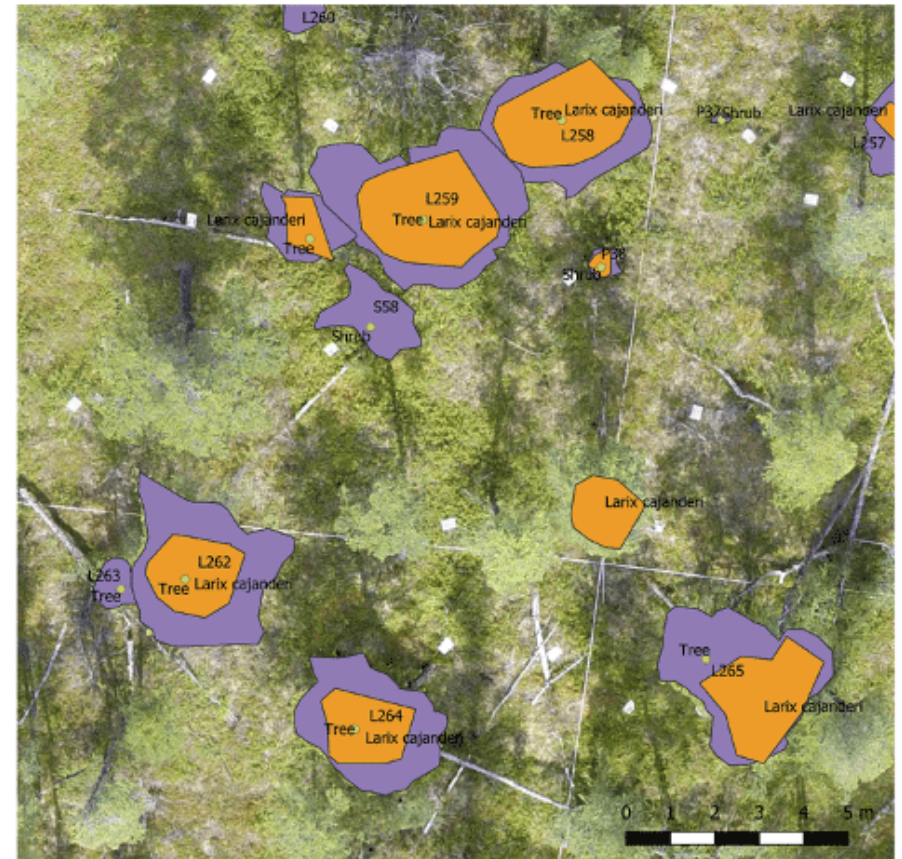


# Location Data Policies

## Spezifische Lösungen für Geodaten in Datenräumen

Ortsbezug nutzen können, trotz Schutzbedarf

- 🌲 Masking
- 🌲 Verrauschen
- 🌲 Aggregation, Shuffling
- 🌲 Anonymisierung, z.B. durch Nutzung synthetischer Daten
- 🌲 Geohashing (Pseudonymisierung)
- 🌲 Verteiltes Training und Processing



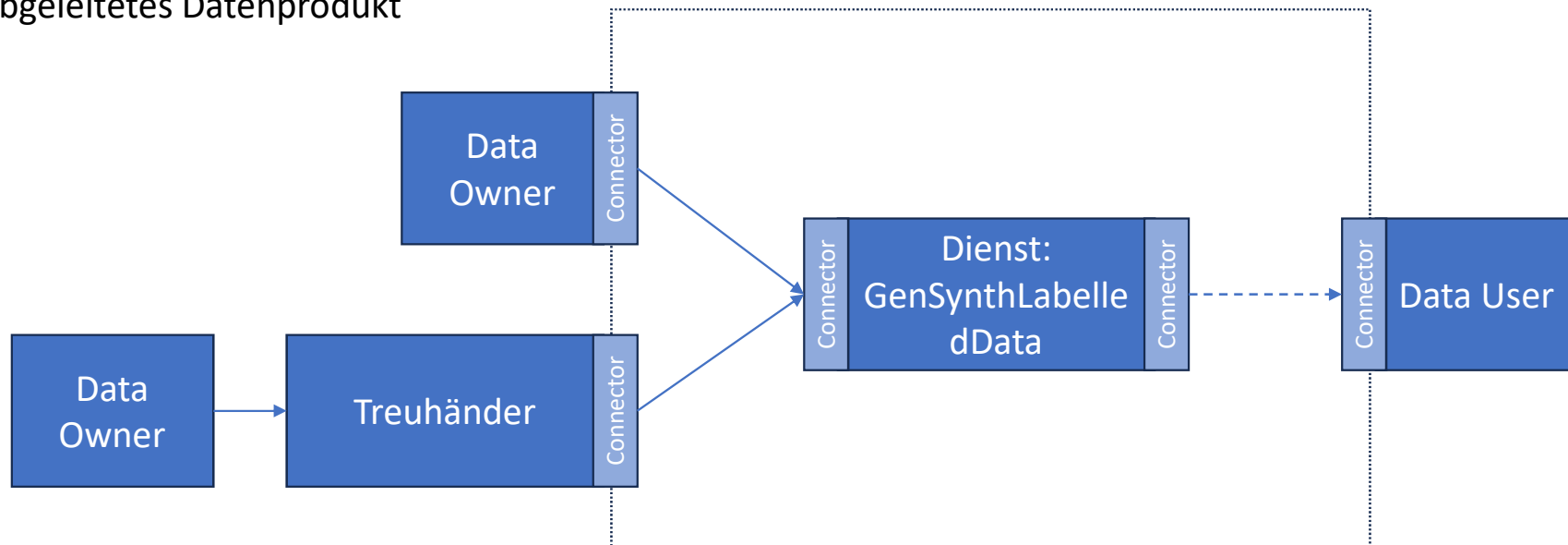
Quelle: <https://essd.copernicus.org/articles/14/4967/2022/#&gid=1&pid=1>

# Anonymisierte Trainingsdaten

Spezifische Lösungen für Geodaten in Datenräumen

➤ Originäre Daten

➤ Abgeleitetes Datenprodukt



# Beispiel 3: SAGE

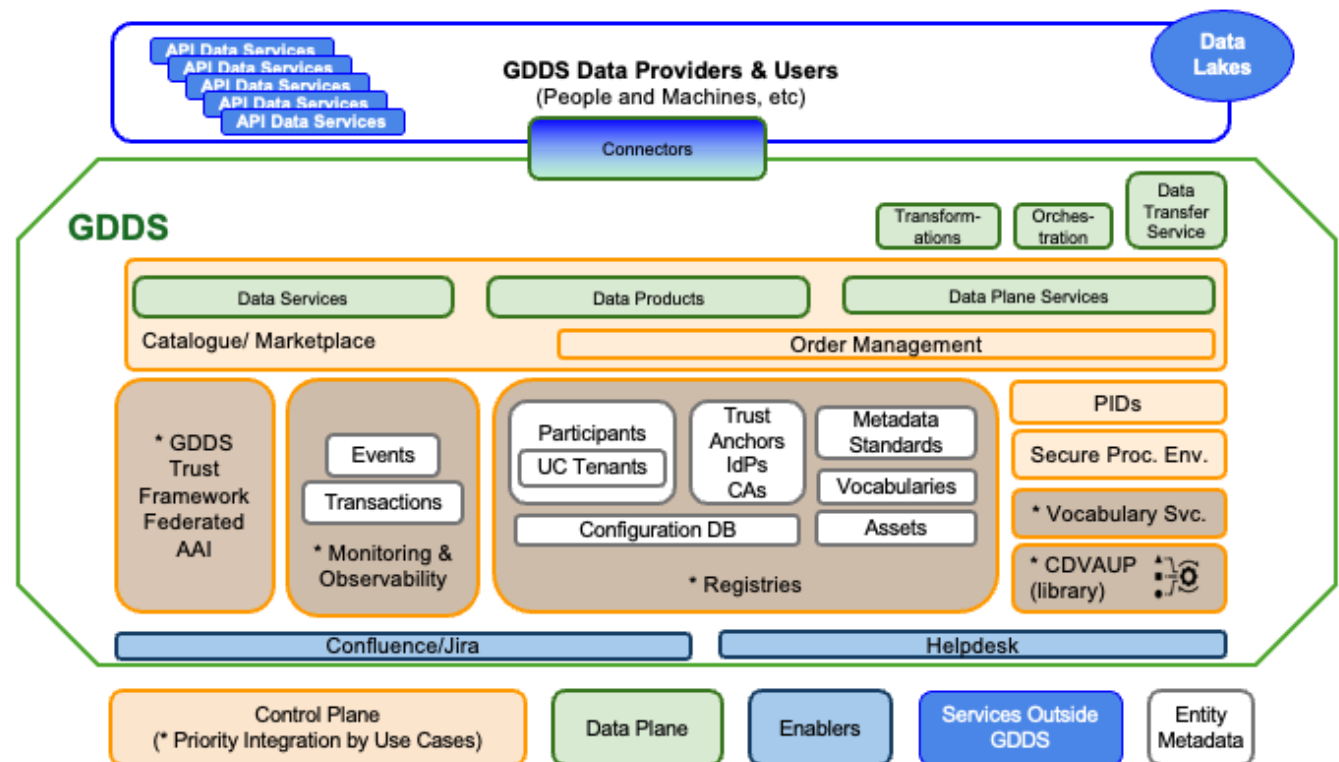
Deployment Project for the Green Deal Data Space



# SAGE

*! Unter Vorbehalt - Projektvertrag noch nicht unterzeichnet!*

- 🌳 40+ Partner
- 🌳 16M€ Budget, 8M€ Förderung
- 🌳 10 Use Case Cluster:
  - 🌳 Forestry
  - 🌳 Soil Circularity
  - 🌳 Environmental Hazards
  - 🌳 Circular Textiles
  - 🌳 ...
- 🌳 Ziele:
  - 🌳 Governance aufbauen
  - 🌳 GDDS Kerndienste entwickeln, deployen und betreiben
  - 🌳 Zentrale Datensätze integrieren
  - 🌳 Nachhaltiges Geschäftsmodell entwickeln und validieren



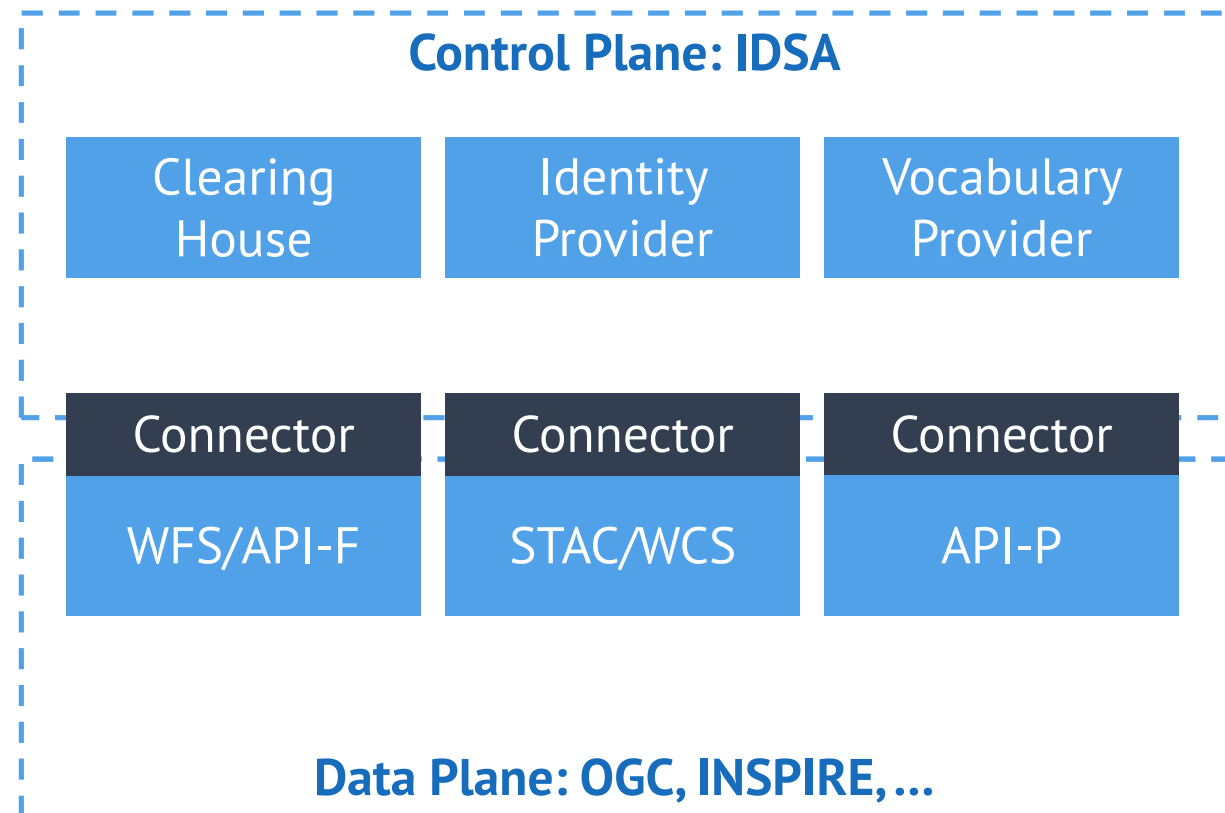
# Architektur und Standards

Deployment Project for the Green Deal Data Space

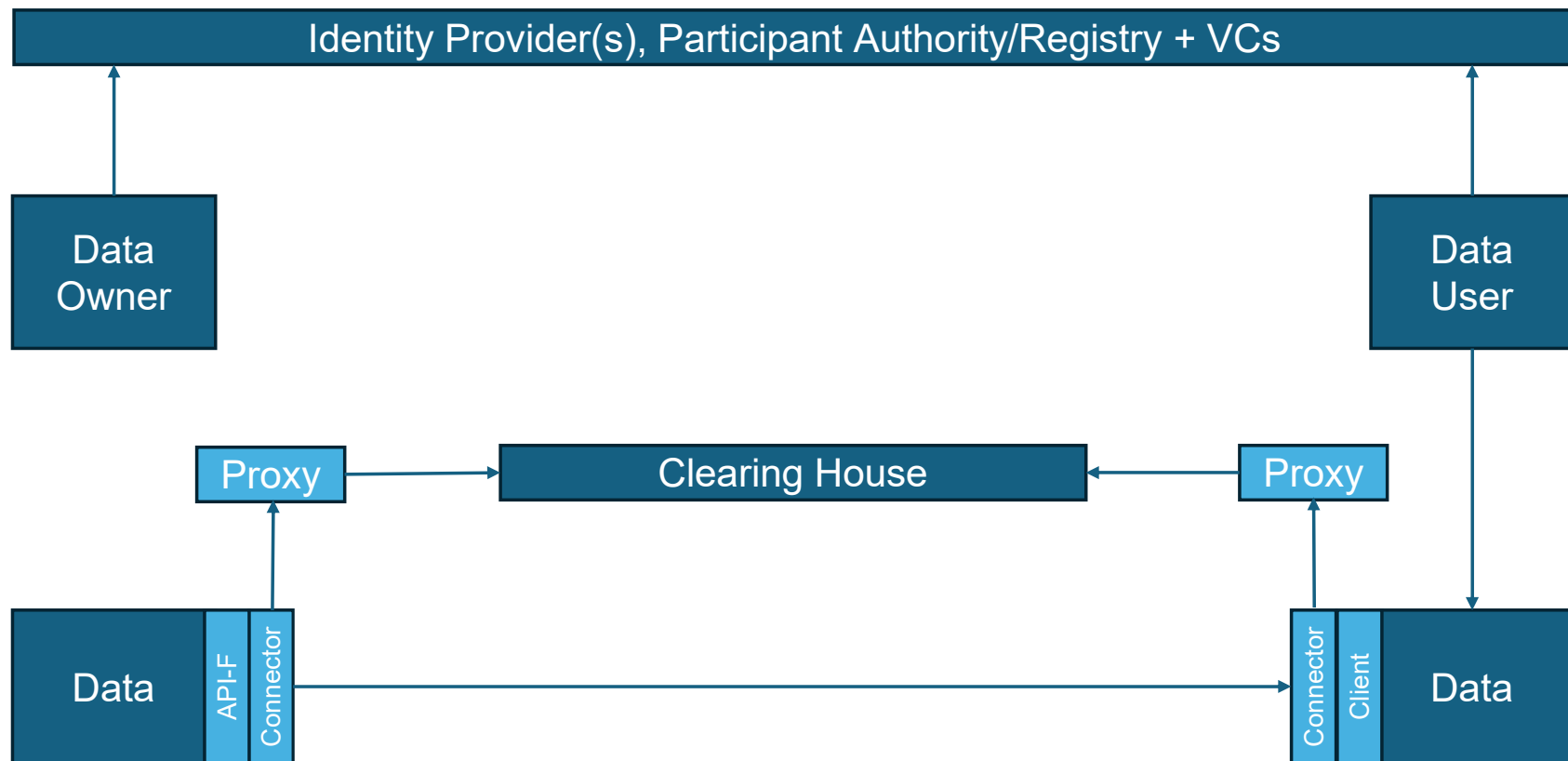
# Architektur und Interoperabilität

Ohne geht es nicht 😊

- Jede datenhaltende Stelle kann einen eigenen Connector betreiben
  - On top von bestehenden Strukturen und APIs möglich
- Zentrale Infrastruktur, z.B.:
  - Identity Provider (kann auch verteilt sein)
  - Vocabulary Provider
  - Clearing House
  - Broker/Hub
  - App Store
  - ....

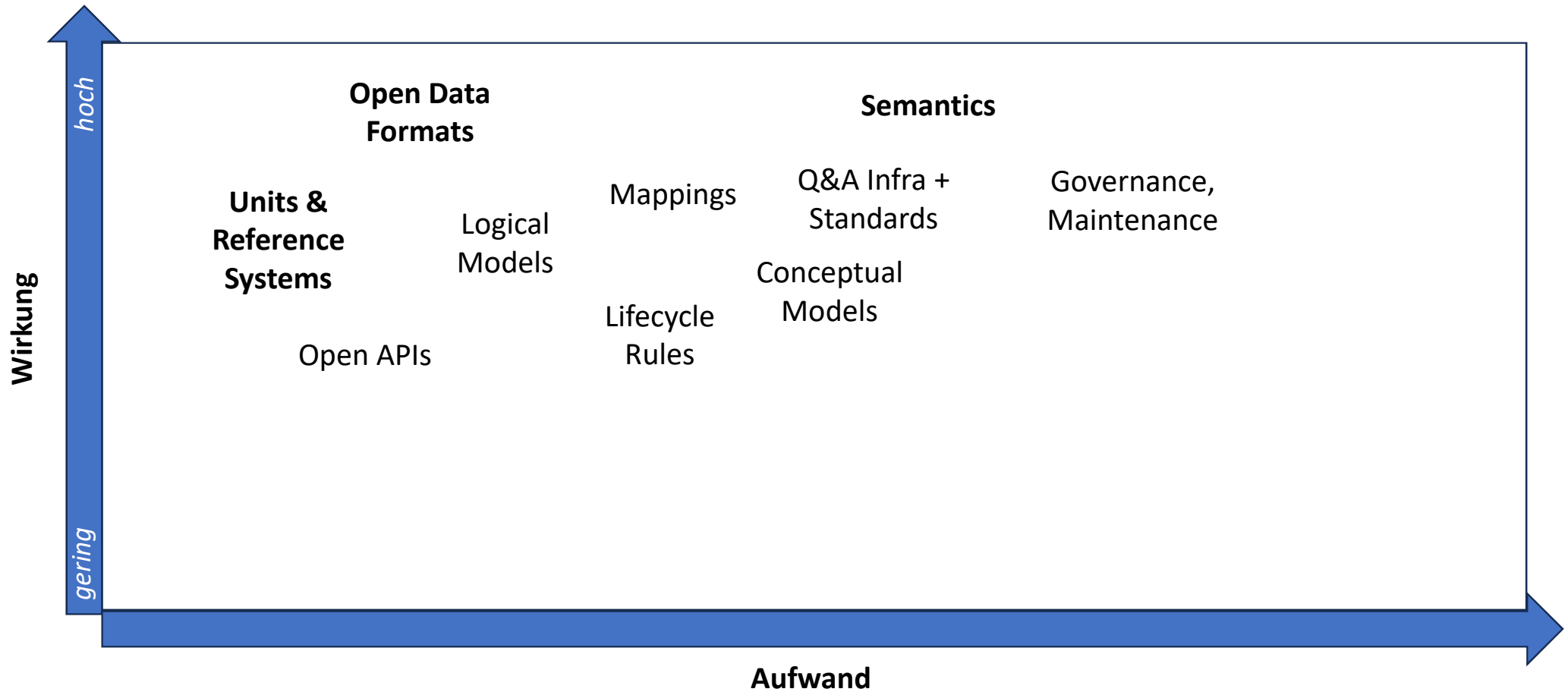


## ... im Detail



# Iterativ zur Interoperabilität

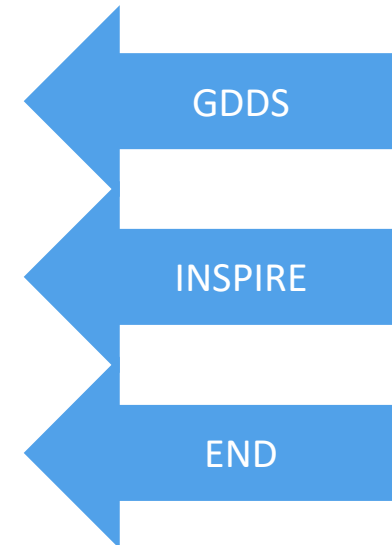
Was trägt am Meisten zu Interoperabilität bei?



# Stufen der Harmonisierung

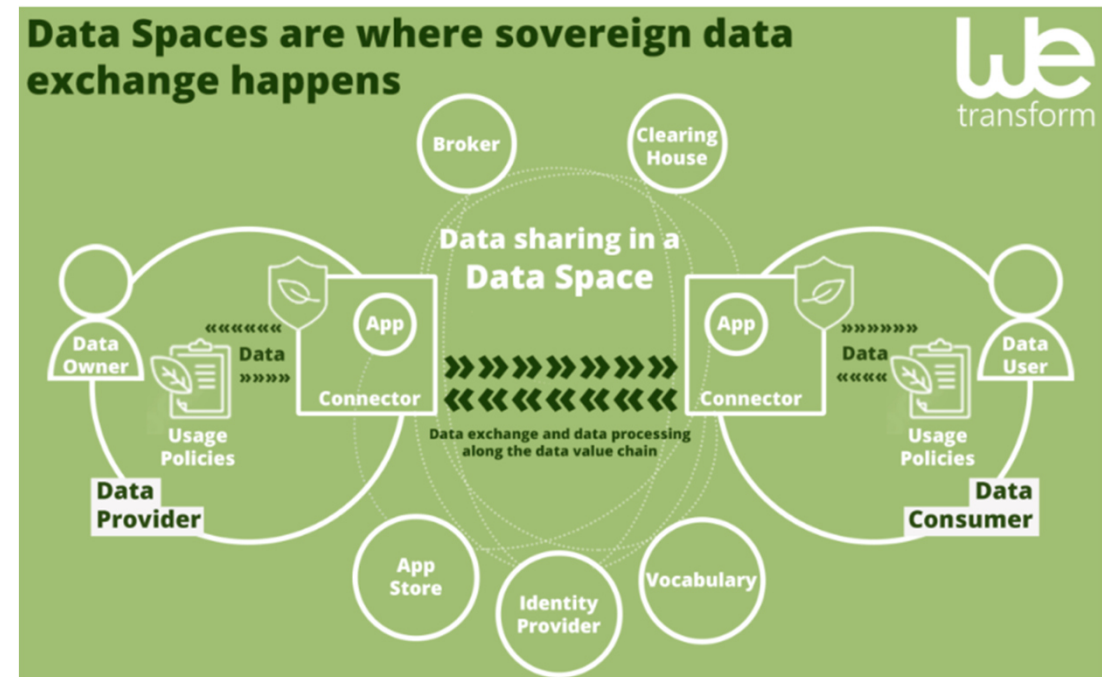
... und eine faire Aufgabenverteilung

- 🌳 All data partials are Findable and Accessible in a homogeneous way (common metadata, API, file format).
- 🌳 The structure, units and semantics of all data partials are fully documented, ideally in a machine-readable form.
- 🌳 All data partials are harmonised to a common vocabulary on a semantic level, and remaining mismatches to partials are well documented.
- 🌳 The structure of all data partials is fully harmonised.
- 🌳 The data quality (thematic accuracy, completeness, positional accuracy, logical/topological inconsistency) of all data partials is homogeneous.
- 🌳 There is full harmonisation along the process from data acquisition to provision.



# Zusammenfassung

- Datenräume fördern den Datenaustausch, Innovationen und die rasche Anwendung guter Lösungen
- Die Hauptherausforderungen liegen weniger in der Technologie, sondern in der Governance und in der Regulierung
- Gemeinsame Schnittstellen und Formate sind zentral; ohne Interoperabilität wird der Mehrwert der Datenräume nicht realisiert
- Wir freuen uns auf den Austausch mit Ihnen!



# KONTAKT

**Thorsten Reitz** Co-Founder/CEO



tr@wetransform.to



+49 179 59 08 203



Fraunhoferstr. 5,  
64283 Darmstadt (Germany)

