LuftDatenNet: Luftqualitäts-Datentreuhandplattform zur Erzielung von Netzwerkeffekten in der Datenproduktion und -verwertung

Schlussbericht zum Vorhaben

Projektlaufzeit:	01. Januar 2022 - 31. Dezember 2024
Projektakronym:	LuftDatenNet
Förderkennzeichen:	16DTM116
Projektleitung:	Robert Heinecke (robert@breeze-technologies.de)
	Breeze Technologies UG (haftungsbeschränkt) Harburger Schloßstraße 6-12 21079 Hamburg

Finanziert von der Europäischen Union
NextGenerationEU

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

1. Einleitung **2**
2. Verwendung der Zuwendung und erzielte Ergebnisse **2**
2.1 Arbeitspaket 1: Analyse des Datenökosystems und potenzieller Datennutzungsszenarien 3
2.2 Arbeitspaket 2: Förderung des Ökosystems von Datengebern und Datennutzern 4
2.3 Arbeitspaket 3: Entwicklung einer Datentreuhandplattform für Luftqualitätsdaten 6
Entwicklung eines einheitlichen Datenaustauschformats für Luftqualitätsdaten. 7
Architektur der Datentreuhandplattform 7
Schaffung notwendiger Prozesse für den Betrieb der Plattform 9
2.4 Arbeitspaket 4: Entwicklung geeigneter Finanzierungs- bzw. Geschäftsmodelle der Plattform 10
Kundensegmente 11
Wertangebot 11
Kundenbeziehungen 11
Kanäle 12
Schlüsselaktivitäten 12
Schlüsselressourcen 12
Schlüsselpartner 13
Einnahmequellen 13
Kostenstruktur 13
SWOT-Analyse 14
2.5 Arbeitspaket 5: Rechte- und Datenmanagement 16
2.6 Arbeitspaket 6: Projektmanagement, -koordinierung und -kommunikation 16
3. Wichtigste Positionen des zahlenmäßigen Nachweises **17**
4. Notwendigkeit und Angemessenheit der geleisteten Projektarbeit **17**
5. Nutzen und Verwertbarkeit der Ergebnisse **18**
6. Anhänge **19**
Anhang 1: Luftdaten Air Quality Measurement Protocol (LAQMP) 19
Beispieldatensatz 19

1. Einleitung

Die Verfügbarkeit und Nutzbarkeit von Daten sind zentrale Treiber für Forschung und Innovation. Innovative Geschäftsmodelle und Dienstleistungen sind zunehmend darauf angewiesen, Daten aus vielen verschiedenen Quellen nutzen und weiterverarbeiten zu können.

Dies betrifft auch die Nutzung und Verwertung von Umweltdaten, wie z.B. Luftqualitätsdaten. Diese werden mittlerweile von unterschiedlichsten Akteuren gesammelt bzw. produziert, aber abseits der klassischen hoheitlichen Luftreinhaltung durch Städte und Länder kaum genutzt. Dabei bieten sie das Potenzial, neue Anwendungen z.B. im Sport-, Smart Home- oder individuellen Gesundheitsbereich zu ermöglichen. Aufgrund der wissenschaftlichen Komplexität von Luftqualitätsdaten ist dies jedoch nicht trivial: Um Daten verschiedener Quellen innerhalb des gleichen Anwendungsfalls nutzbar zu machen, müssen verschiedenste Herausforderungen bei ihrer Zusammenführung, Vereinheitlichung und Harmonisierung überwunden werden.

Das Projekt LuftDatenNet zielte darauf ab, eine Datentreuhandplattform für Luftqualitätsdaten zu entwerfen und aufzubauen. Besonders wichtig war dabei die Schaffung eines einheitlichen Datenformats bzw. -standards auf Basis existierender Luftqualitätsdatenformate und der Beginn des Aufbaus eines Ökosystems um diese Plattform. Daneben sollte ein Ökosystem von Akteuren um die Plattform geschaffen werden, um die initiale Nutzung zu fördern und Pilotprojekte bzw. Demonstratoren zu starten.

Das Projekt knüpfte an den aktuellen wissenschaftlich-technischen Stand an; insbesondere an Aktivitäten von bestehenden Aggregationsplattform für Luftqualitätsdaten und einzelne Datenmonetarisierungsaktivitäten privatwirtschaftlicher Anbieter.

2. Verwendung der Zuwendung und erzielte Ergebnisse

Die Umsetzung des Projektes orientierte sich an definierten Arbeitspaketen. Im Folgenden wird ein Überblick über die Zielsetzungen und die inhaltlichen Schwerpunkte der einzelnen Arbeitspakete und die methodischen Ansätze bei ihrer Bearbeitung, sowie ihrer wichtigsten Ergebnisse gegeben.

2.1 Arbeitspaket 1: Analyse des Datenökosystems und potenzieller Datennutzungsszenarien

Ziel dieses Arbeitspakets war die Erstellung einer Übersicht über Produzenten und Konsumenten von Luftqualitätsdaten, sowie die Identifikation bestehender Standards zum Datenaustausch von Luftqualitätsdaten.

Zur Schaffung der Übersicht von Datenproduzenten wurde sowohl auf das bestehende Know-how des Projektteams in diesem Bereich zurückgegriffen, als auch eine strukturierte Internetrecherche durchgeführt. Weiterhin wurden Fachmessen in diesem Bereich bzgl. ausstellender Organisationen untersucht. Es wurde dabei unterschieden zwischen tatsächlichen Produzenten von Luftqualitätsdaten und Aggregatoren dieser Daten aus verschiedenen Quellen. Teilweise überlappten sich diese Tätigkeiten auch.

Das Projektteam identifizierte mehr als 140 Organisationen, die Luftqualitätsdaten produzieren. Diese ließen sich grundsätzlich in folgende Gruppen einteilen:

- Regierungsorganisationen, die in-situ Messnetze betreiben (z.B. Umweltämter der Bundesländer) (diese Daten stehen meist unter einer Open Source-Lizenz zur Verfügung, allerdings nur aggregiert)
- Weltraumagenturen, die Satellitendaten erfassen (z.B. ESA) (die Daten stehen meist unter einer Open Source-Lizenz zur Verfügung)
- Unternehmen, die Lösungen zur Überwachung der Außenluftqualität entwickeln (die Daten stehen im Regelfall nicht öffentlich zur Verfügung)
- Unternehmen, die Lösungen zur Überwachung der Luftqualität im Inneren von Gebäuden entwickeln (diese lässt ggf. auch Rückschlüsse auf die Luftqualität im Freien zu) (die Daten stehen im Regelfall nicht öffentlich zur Verfügung)
- Aggregatoren, die die Daten verschiedener Quellen zusammenführen und ggf. weiterverwerten oder -veredeln (teilweise werden Schnittstellen zur Datennutzung angeboten)

Zur Identifikation von Datenkonsumenten wurden im Rahmen mehrerer Ideation-Formate mögliche Nutzungsszenarien skizziert. Außerdem wurde eine Internetrecherche hinsichtlich bereits umgesetzter datengestützter Geschäftsansätze durchgeführt.

Es gab in der Vergangenheit bislang nur wenig umgesetzte Projekte, bei denen Luftqualitätsdaten außerhalb hoheitlicher Prozesse genutzt wurden. Mögliche Datenkonsumenten wurden in die folgenden Gruppen kategorisiert:

- Mobilität,
- (Rück-)Versicherungen,
- Medien,

- Real Estate,
- Sport,
- Forschung,
- Tourismus,
- Facility Management (inkl. Smart Home),
- Sonstige.

Mit Vertretern mehrerer dieser Gruppen wurde im Projektverlauf an konkreten Datennutzungsansätzen gearbeitet.

Mit identifizierten Datenproduzenten und -konsumenten wurden Interviews durchgeführt, um Interessen, Wünsche, Anforderungen und Bedürfnisse an eine Datentreuhandplattform zu erfassen und diese dem Anforderungskatalog hinzuzufügen.

Hier ergaben sich insb. abseits der technischen Anforderungen größere Unterschiede hinsichtlich der Wünsche von Datenproduzenten und -konsumenten. Während sich Datenkonsumenten einen möglichst schnellen, unkomplizierten und einfachen Prozess von der Idee der Datennutzung bis hin zum Datenabruf (inklusive harmonisierter und einheitlicher Daten aus mehreren Quellen, die die zentrale Speicherung und vollständig automatische Verarbeitung notwendig machen) und der Datennutzung wünschten, gab es auf Seite der Datenproduzenten größere Vorbehalte und den damit verbundenen Wunsch nach individuellen Freigabeverfahren, sowie der Beschränkung der Funktion der Datentreuhandplattform auf ein reines Durchreichen der Daten aus dem Datenspeicher des Datenproduzenten. Im weiteren Projektverlauf konnten hier aber Bedenken abgebaut werden, sodass mindestens ein Teil der verfügbaren Daten wie oben beschrieben freigegeben und genutzt werden kann. Datenkonsumenten gaben an, dass insbesondere diese Daten für sie den höheren Wert in der Nutzung hätten.

Bestehende Standards waren dem Projektteam größtenteils bereits bekannt. Es wurde dennoch eine Internetrecherche und einzelne Hintergrundgespräche mit Vertretern insb. des öffentlichen Sektors durchgeführt, um mögliche Lücken zu identifizieren. Die meisten Standards wurden in der Vergangenheit von öffentlichen Einrichtungen geschaffen. Oben genannte Datenproduzenten und -aggregatoren verwenden allerdings teilweise auch eigene Standards und Formate.

2.2 Arbeitspaket 2: Förderung des Ökosystems von Datengebern und Datennutzern

Ziel dieses Arbeitspakets war die Schaffung bzw. Förderung eines Ökosystems von Luftqualitäts-Datenproduzenten und -konsumenten. Hierfür wurde zum einen auf die

assoziierten Projektpartner zurückgegriffen und zum anderen wurden Veranstaltungen wie Messen und Konferenzen genutzt, um das Projekt bekannter zu machen und die genannten Vertreter für die Mitarbeit zu gewinnen.

Im Projektzeitraum wurden über 30 solcher Events in Deutschland, Frankreich, Italien, Spanien, den USA und den Vereinigten Arabischen Emiraten besucht. Auf diesen wurde im Rahmen von unterschiedlichen Formaten (Präsentationen, Pitches, Messestände, etc.) auf das Projekt aufmerksam gemacht.

Es wurden zwei Online-Formate und ein physisches Treffen auf einer Veranstaltung organisiert. Außerdem wurde zunächst ein virtuelles Forum aufgesetzt, das den Austausch assoziierter Partner und weiterer möglicher Datenproduzenten und -konsumenten fördern sollte. Hierfür wurden verschiedene Portal- und Foren-Softwares getestet und evaluiert, bis die Lösung “Solved Together” ausgewählt wurde. Da dieses allerdings nicht die gewünschte Nutzung erreichte, wurde der Austausch später auf einen E-Mail-Verteiler verlagert.

Es wurden außerdem mehrere User Stories von möglichen Datennutzungsszenarien entwickelt, u.a.:

- Ein Smart Home-Gerät, was vor hoher Luftbelastung warnt
- Eine Smartphone-App für Sportler, die die gesündesten Laufrouten auf Basis des aktuellen Standorts empfiehlt
- Eine Navigationsanwendung für Asthmatiker, um besonders belastete Gebiete im Stadtgebiet zu vermeiden
- Eine Fachanwendung für Vertreter von Kommunen, um existierende und für ihre Stadt relevante Datensätze im Rahmen der Stadtplanung nutzbar zu machen
- Eine Fachanwendung für Vertreter von Kommunen, um die Luftgüte ihrer Stadt mit der umliegender Städte zu vergleichen

Geplant war auch die Durchführung eines Wettbewerbs, um einen zusätzlichen Use Case zu implementieren. Dieser wurde zunächst aus projektadministrativen Gründen verschoben und musste letztendlich abgesagt werden, da der hierfür ursprünglich ausgewählte Partner aufgrund von Umstrukturierungen bzw. Unternehmensübernahmen nicht mehr zur Verfügung stand. Stattdessen wurde intern ein Prototyp für die Anbindung des LuftDatenNet-Portals entwickelt, um die technische Integration zu testen.

Zunächst war angedacht, auch Verbrauchern die Möglichkeit zu geben, Daten sowohl auf der Datentreuhandplattform einzustellen, als auch solche Daten zu nutzen. Auf Basis von Stakeholderfeedback wurde diese Idee allerdings aufgrund von größeren Herausforderungen bei der Qualitätssicherung sowie beim zusätzlicher rechtlicher Komplexität (Verbraucherrecht vs. B2B) verworfen. Diese Entscheidung

vereinfachte auch die Prüfung der Daten von Datenproduzenten hinsichtlich Qualität und Vertrauenswürdigkeit. Perspektivisch ist die Erweiterung auf Daten von Verbrauchern zukünftig immer noch möglich.

Über die Projektdauer hinweg wurde mit über 20 verschiedenen Datenproduzenten und -konsumenten gearbeitet.

2.3 Arbeitspaket 3: Entwicklung einer Datentreuhandplattform für Luftqualitätsdaten

Das Ziel dieses Arbeitspakets war die technische Spezifikation und Implementierung einer Datentreuhandplattform für Luftqualitätsdaten, inklusive der Konzeption, dem Betrieb und der Wartung der hierfür notwendigen Infrastruktur.

Zunächst erfolgte hierfür die Spezifikation von technischen Anforderungen und darauf basierend die Entwicklung von Skizzen einer möglichen Benutzeroberfläche. Als Tool für letzteres kam die Anwendung *Balsamiq* zum Einsatz.

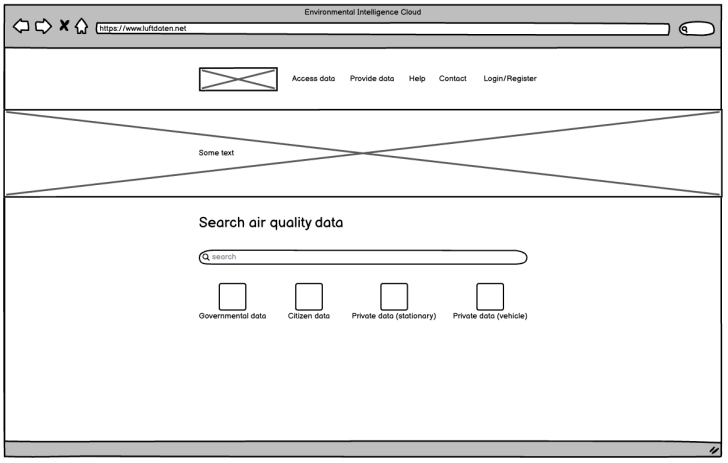

Abbildung: Beispiel-Mockup der Landingpage in Balsamiq

Unter anderem wurden folgende Anforderungen im Anforderungskatalog erfasst:

- Skalierbarkeit der Datenspeicherung und Hochverfügbarkeit von Datentreuhandplattform und ihrer Schnittstellen: Die Plattform muss auch mit großen Datenmengen aus vielen verschiedenen Quellen stets performant arbeiten.
- Interoperabilität mit bestehenden Systemen: Es muss sichergestellt werden, dass die Plattformen und Systeme der Datenproduzenten und -konsumenten leicht zum Zweck des Datenaustauschs angebunden werden können.

- Benutzerfreundlichkeit und Zugänglichkeit: Die Benutzeroberfläche soll von verschiedenen Nutzergruppen verwendet werden, ohne dass umfangreiche technische Kenntnisse erforderlich sind. Als Standardsprache wurde Englisch gewählt, da bereits einige der assoziierten Projektpartner internationale Unternehmen waren.
- Agil änderbare Metadatenverwaltung: Die Speicherung und Verwaltung von Metadaten zu den Luftqualitätsdatensätzen ist wichtig für die Auswahl der Daten durch die Datenproduzenten. Um zukünftige Anforderungen abdecken zu können, sollte dieses System einfach änder- und erweiterbar sein.

Auf Basis dieser Anforderungen erfolgte die agile, iterative Entwicklung der Datentreuhandplattform. Das Entwicklungsvorgehen war angelehnt an Scrum. Ziel war die Beibehaltung hoher Flexibilität bei der (Weiter-)Entwicklung der Plattform. Hohe Feedbackorientierung ermöglichte eine enge Verzahnung von Anforderungen, Konzeptentwicklung und technischer Umsetzung. Hatten die Entwickler Rückfragen zu einzelnen Punkten, konnten die zur Klärung notwendigen Projektbeteiligte schnell eingebunden werden. Im Rahmen von wöchentlichen Statusupdates wurde der Projektfortschritt gesichert und die Erreichung von Entwicklungsmeilensteinen überwacht. Die transparente und strukturierte Entwicklungskoordination sicherte so den erfolgreichen Projektabschluss.

Entwicklung eines einheitlichen Datenaustauschformats für Luftqualitätsdaten

Unter Berücksichtigung der identifizierten Datenformate aus AP1 wurde ein einheitliches Datenaustauschformat entwickelt, um Daten harmonisiert verarbeiten und ggf. speichern zu können. Insbesondere Berücksichtigung fanden dabei Formatbesonderheiten der folgenden Datenmodelle:

- NGSI AirQualityObserved,
- die openAQ-Datenschnittstelle,
- die INSPIRE (EU) Spatial Data Guidelines,
- die (US) EPA AQ Data Exchange (AQDE) Network Standards,
- der Open Geospatial Consortium's (OGC) O&M Data Model Standard,
- NOAAs NCEI NetCDF-CF Templates v2.0 und
- das DEM-AQ: Data Exchange Module for Air Quality data standards.

Die Datenformate wurden u.a. hinsichtlich ihrer Besonderheiten, Vor- und Nachteile und Einfachheit der technischen Integrierbarkeit untersucht. Auf ihrer Basis wurden die relevantesten Parameter und Metadaten identifiziert und darauf basierend das Luftdaten Air Quality Measurement Protocol (LAQMP) entwickelt (siehe Anhang 1: Luftdaten Air Quality Measurement Protocol (LAQMP)). Ziel war, dass Daten jedes potenziellen Datenproduzenten in das LAQMP transformiert werden können und alle für Datenkonsumenten relevanten Informationen erhalten bleiben. Das

LAQMP-Datenmodell ist zukünftig erweiterbar, sodass auch auf Veränderungen bei den Anforderungen der Datenkonsumenten reagiert werden kann und neue technische Entwicklungen im Bereich von Luftqualitätsdaten mit einbezogen werden können.

Architektur der Datentreuhandplattform

Die LuftDatenNet-Datentreuhandplattform basiert auf sorgfältig ausgewählten Technologien, um Zuverlässigkeit, Effizienz und ein nahtloses Benutzererlebnis zu gewährleisten. Die Plattform umfasst ein robustes Backend und ein intuitives Frontend. Die Anwendung wurde serverless entwickelt, um eine zu 100% skalierbare Architektur sicherzustellen. Der Backend-Stack umfasst Hono, Bun, Typescript, Zod, SendGrid, MySQL, Azure Blob Storage und Docker. Zusammen bieten diese Tools eine leistungsstarke, skalierbare und sichere Grundlage für die Datenverarbeitung und -verwaltung. Im Frontend werden React, Typescript, Javascript (HTML5, CSS3), Chakra UI, TanStack Form und TanStack Query, Axios, Zod und Luxon verwendet. Diese Technologiekombination bietet eine reaktionsschnelle, moderne und benutzerfreundliche Oberfläche und gewährleistet den Plattformbenutzern eine reibungslose und effektive Interaktion mit der Plattform.

Eine REST-API übernimmt die Aufgabe, den effizienten Datenaustausch zwischen Frontend und Backend zu erleichtern und sicherzustellen, dass Daten sicher und schnell über die Plattform übertragen werden.

Der Backend-Webserver von LuftDatenNet ist mit modernsten Technologien ausgestattet, um hohe Leistung und Sicherheit zu gewährleisten. Hono fungiert in Kombination mit Bun als Server-Laufzeitumgebung, die für die effiziente Verarbeitung gleichzeitiger Verbindungen optimiert ist und die Leistungsstärke der Plattform sicherstellt. Typescript wurde zum Schreiben typsicheren Codes verwendet, um Fehler zu reduzieren und die Wartbarkeit zu verbessern. Zod wurde zur Schemavalidierung eingesetzt und garantiert damit die Korrektheit und Sicherheit der Dateneingaben vor möglichen Inkonsistenzen. SendGrid wird für die Verwaltung der E-Mail-Kommunikation verwendet, um die zuverlässige Verarbeitung von Warnungen und Benachrichtigungen sicherzustellen. Docker spielt eine entscheidende Rolle bei der Containerisierung der Anwendung, ermöglicht konsistente Bereitstellungen in verschiedenen Umgebungen und macht die Plattform leicht skalierbar.

Die Datenbankkomponenten von LuftDatenNet sind für die Verarbeitung umfangreicher und strukturierter Daten konzipiert, was für die perspektivisch hohe Anzahl an Luftqualitätsdatensätzen von entscheidender Bedeutung ist, da einzelne geografische Messpunkte bereits über eine Million Datensätze pro Jahr erzeugen können. Aufgrund ihrer Robustheit, Zuverlässigkeit und Fähigkeit, strukturierte Datensätze effizient zu verwalten, wurde MySQL als relationale Datenbank

entschieden. MySQL ermöglicht es, Daten schnell und präzise zu speichern, zu organisieren und abzurufen. Zusätzlich wird Azure Blob Storage für die Verwaltung großer Mengen an Luftqualitätsdaten genutzt. Diese Cloud-Speicherlösung bietet die nötige Skalierbarkeit und Zuverlässigkeit, um große Datensätze ohne Leistungseinbußen zu speichern und stellt sicher, dass die Plattform Daten aus zahlreichen Quellen reibungslos verarbeiten kann.

Das Frontend der LuftDatenNet-Plattform muss ein intuitives und reaktionsschnelles Erlebnis bieten. Zur Erreichung dieses Ziels wurden moderne Webentwicklungstechnologien genutzt. React bildet den Kern des Frontend-Frameworks und ermöglichte die Entwicklung dynamischer und effizienter Benutzeroberflächen. Typescript wurde ebenfalls eingesetzt, um die Typsicherheit zu gewährleisten und Entwicklungsfehler zu reduzieren. HTML5 und CSS3 bilden die Grundlage der Webseitenstrukturen und -stile und gewährleisten umfassende Kompatibilität und ein einheitliches Erscheinungsbild auf verschiedenen Geräten. Chakra UI bietet eine Reihe vorgefertigter, zugänglicher und anpassbarer UI-Komponenten, die die Entwicklung beschleunigen und gleichzeitig die Designkonsistenz wahren. TanStack Form und TanStack Query werden für die Statusverwaltung bzw. den Datenabruf genutzt und gewährleisten so eine reibungslose Verwaltung von Benutzereingaben und Backend-Daten. Axios unterstützt HTTP-Anfragen, Zod dient der clientseitigen Schemavalidierung und Luxon verarbeitet komplexe Datums- und Zeitmanipulationen und trägt so zur hochfunktionalen und zuverlässigen Benutzeroberfläche bei. Die Interaktion zwischen Frontend und Backend erfolgt über eine REST-API. Dies ermöglicht einen effizienten und sicheren Datenaustausch, der für die Reaktionsfähigkeit der Plattform und Echtzeit-Updates entscheidend ist.

Im Rahmen des Entwicklungsprojekts wurden regelmäßige Nutzertests durchgeführt, um die korrekte und gewünschte Funktionalität sicherzustellen. Parallel wurde auf eine laufende Dokumentation von Programmcode und Entwicklungsstand geachtet.

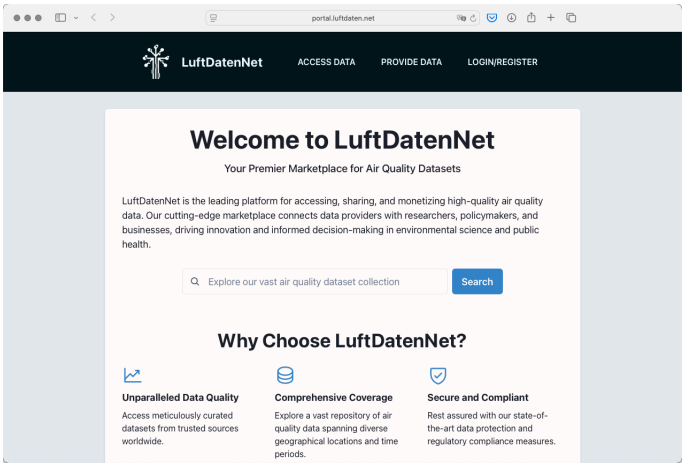

Abbildung: LuftDatenNet-Datentreuhandplattform

Schaffung notwendiger Prozesse für den Betrieb der Plattform

Aus dem Dialog mit Datenproduzenten und -konsumenten ergaben sich unterschiedliche Anforderungen hinsichtlich der prozessualen Begleitung der Datentreuhandplattform. Einige notwendige Prozesse, z.B. im Rahmen des Rechte/Rollen-Managements, konnten in Softwareprozessen umgesetzt werden. Andere Prozesse, z.B. die Vereinheitlichung von Datensätzen, müssen aktuell noch manuell durchgeführt werden. Insbesondere in diesem Bereich ergaben sich Ansätze von zukünftigen Forschungs- und Entwicklungsprojekten. Z.B. im Rahmen von temporaler Interpolation und Aggregation von Luftqualitätsdaten (zur Vereinheitlichung von Zeitstempeln mehrerer Datensätze) oder der Zusammenführung von Datensätzen unterschiedlicher Qualitätsstandards. Relevante Prozesse wurden mit Hilfe von BPMN abgebildet.

2.4 Arbeitspaket 4: Entwicklung geeigneter Finanzierungs- bzw. Geschäftsmodelle der Plattform

Ziel dieses Arbeitspakets war die Identifikation eines Finanzierungs- bzw. Geschäftsmodells für die Datentreuhandplattform, um eine dauerhafte Verwertung sicherzustellen.

Mit mehreren Datenproduzenten und -konsumenten (teilweise auch in einer Doppelrolle) wurden hierfür Case Studies entwickelt. Zur Umsetzung von konkreten “produktiven” Pilotprojekten kam es zwar (noch) nicht, allerdings ist dies im weiteren Betrieb der Datentreuhandplattform noch geplant, sobald sich die Gelegenheit hierfür bietet. Die relevantesten Case Studies kamen aus den Bereichen Smart Home, Versicherungen und Sport (im Rahmen von persönlicher Gesundheit). Mit 3 internationalen Konzernen wurden hierfür mögliche neue Produkte konzipiert. Mit einem weiteren Partner wurde ein erster technischer Prototyp entwickelt.

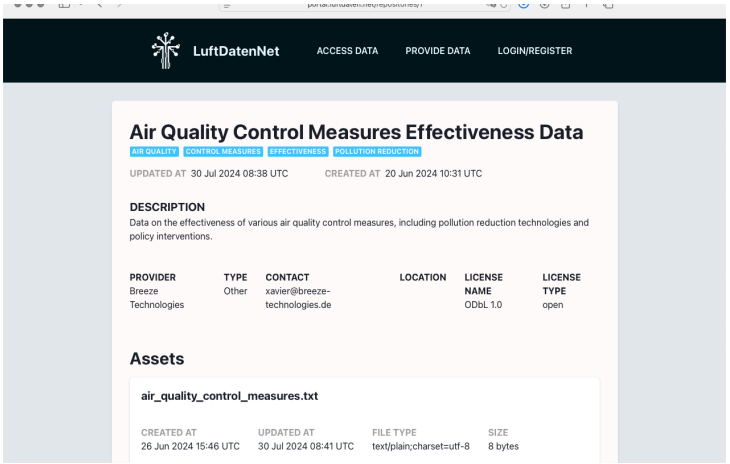

Abbildung: Beispiel für einen Testdatensatz auf der Datentreuhandplattform

Insbesondere von Seiten der Datenkonsumenten wurden Mindestanforderungen definiert, ab welcher Datenmenge und -dichte in welcher geografischen Region für sie ein Pilotprojekt Sinn machen würde. Dieser nun bestehende Anforderungskatalog erleichtert zukünftig auch die Suche nach weiteren Datenproduzenten, da insbesondere solchen, die hier bestehende Datenlücken schließen können, baldige Pilotprojekte in Aussicht gestellt werden können.

Zudem ergab sich aus dem Dialog mit Vertretern der öffentlichen Hand (insbesondere in den USA) noch der Hinweis, dass eine Datentreuhandplattform auch im Rahmen eines (über-)regionalen Datenaustauschs von Luftqualitätsdaten zwischen Städten (und ihrer Bevölkerung) sinnvoll sein könnte. Weitere Schritte werden hier noch evaluiert, wobei die aktuelle politische Situation die Geschwindigkeit weiterer möglicher Schritte stark verlangsamt hat. Keinem der befragten (reinen) Datenproduzenten war ein anderer Datentreuhänder oder eine Datenaustauschplattform für Luftqualitätsdaten bekannt, über die sie ihre Daten wiederverwerten könnten. Dies zeigte nochmals die Notwendigkeit für das LuftDatenNet-Projekt auf.

Auf Basis des Business Model Canvas nach Eric Ries wurden ein Geschäfts- und Finanzmodell für die Fortführung entwickelt:

Kundensegmente

Hier sind Datenproduzenten (inklusive Aggregatoren) und -konsumenten zu unterscheiden:

- Datenproduzenten sind vorwiegend der öffentliche Sektor (Open Source-Daten) und Unternehmen aus dem Bereich Luftqualität (proprietäre bzw. kommerzielle Daten).

- Datenkonsumenten sind vor allem Unternehmen mit digitalen oder teildigitalen Produkten, die durch das Hinzufügen von Luftqualitätsdaten erweitert oder neu entwickelt werden können.

Beide Segmente sind daher also juristische Personen, keine Verbraucher. Dies ist bei der Ausgestaltung der Angebote zu berücksichtigen.

Wertangebot

Für Datenproduzenten liegt der Vorteil der Datentreuhandplattform in der Möglichkeit zur Vermarktung ihrer vorhandenen und neu generierten Daten. Hierüber können ihre Nutzung erhöht, zusätzliche gesellschaftliche Mehrwerte geschaffen bzw. zusätzliche Monetarisierung ermöglicht werden.

Datenkonsumenten profitieren von der Datentreuhandplattform über einen einfacheren und vor allem zentralen Zugang zu für ihren jeweiligen Anwendungsfall notwendigen Luftqualitätsdaten. Ein harmonisiertes Datenschema und die Vermeidung der Notwendigkeit der Anbindung von vielen verschiedenen Schnittstellen und Datenstandards sparen außerdem zusätzliche Zeit und Kosten.

Kundenbeziehungen

In der Anfangsphase der Datentreuhandplattform ist eine individuelle, persönliche Unterstützung der Datenproduzenten und -konsumenten notwendig, um die Plattform zu etablieren und Feedback für die iterative Weiterentwicklung zu sammeln.

Perspektivisch sollen die Plattform und Dienste in ein automatisiertes Selbstbedienungskonzept überführt werden.

Kanäle

Die Gruppe der Datenproduzenten ist klar definiert. Dies ermöglicht die Akquise von neuen Datenproduzenten über die Ansprache auf relevanten Messen und den direkten persönlichen Kontakt mit den oben genannten Organisationstypen. Offene Daten der öffentlichen Einrichtungen können im Rahmen ihrer jeweiligen Lizenz direkt angebunden werden.

Die Ansprache von (möglichen) Datenkonsumenten ist deutlich schwieriger, da diese Gruppe deutlich weniger genau definierbar ist. Es ist deshalb notwendig, dass Unternehmen, die für Projekte nach Luftqualitätsdaten suchen, die Datentreuhandplattform über Internetsuchmaschinen finden. SEO bzw. SEM sind deshalb relevante zukünftige Aktivitäten.

Die Kundenbeziehung wird nach der Akquise vollständig digital geführt und gemanagt (E-Mail, Internetseite, Online-Meetings).

Schlüsselaktivitäten

Die Datentreuhandplattform bietet eine Austauschplattform für Luftqualitätsdaten für Produzenten und mögliche Konsumenten dieser Daten. Da es sowohl wenig bis keine Angebote für die zentrale Bereitstellung solcher Daten gibt, als auch nur wenige öffentlich bekannte Nutzungsszenarien der Daten außerhalb hoheitlicher Aufgaben, müssen solche Anwendungsmöglichkeiten bekannter gemacht werden. Dies erfordert die aktive Teilnahme an Messen, Konferenzen und anderen Veranstaltungen mit Präsentationen und Pitches, sowie weiterführende Publikationen zu diesem Thema, um die Bekanntheit des Themas zu erhöhen.

Eingestellte Daten müssen auf ein einheitliches Datenschema gebracht werden. Je nach dem Format der Ursprungsdaten ist diese Harmonisierung mehr oder weniger aufwändig. Sie stellt in jedem Fall ein wichtiges Kernmerkmal der Plattform dar, da dies die Nutzung von Daten aus mehreren Datenquellen deutlich vereinfacht (z.B. über das Angleichen von Zeitstempeln).

Die bereitgestellten Daten der Datenproduzenten werden auf den Servern der Datentreuhandplattform gespeichert und auf Anfrage von Datenkonsumenten über eine Schnittstelle zur Verfügung gestellt. Wichtig ist die Moderation des ggf. entstehenden Verhältnisses zwischen Datenproduzenten und -konsumenten inklusive vertraglicher Regelungen, z.B. hinsichtlich Nutzungseinschränkungen und Speicherdauer auf Seiten der Datenkonsumenten.

Schlüsselressourcen

Die Schlüsselressourcen der Datentreuhandplattform die die vorhandenen Luftqualitätsdaten. Es handelt sich dabei um extern akquirierte Daten, die der Plattform deshalb nicht selbst gehören. Die Daten stammen von den genannten Datenproduzenten. Es handelt sich dabei um Zeitreihendaten, die die Schadstoffkonzentration und/oder Luftgüte für einen bestimmten geografischen Ort angeben.

Ebenso ist die Expertise der Beschäftigten der Datentreuhandplattform eine Schlüsselressource; auch hinsichtlich der ggf. noch manuell notwendigen Harmonisierung der Daten vor der Nutzung durch einen Datenkonsumenten.

Zudem kann auch die IT-Infrastruktur und der Programmcode für die Datentreuhandplattform als Schlüsselressource verstanden werden, da sie erst das Speichern, Zusammenführen und damit den Austausch der Daten zwischen Datenproduzenten und -konsumenten ermöglichen.

Schlüsselpartner

Schlüsselpartner für die Datentreuhandplattform sind die Datenproduzenten und -konsumenten. Weiterhin können hier auch die Anbieter der technischen Infrastruktur genannt werden, auf denen die Datentreuhandplattform betrieben wird.

Einnahmequellen

Damit die Datentreuhandplattform langfristig betrieben werden kann, muss sie kostendeckend arbeiten. Außerdem stellen viele Datenproduzenten ihre Daten nur zur Verfügung, wenn sie hierdurch monetäre Mehrwerte für sich generieren können. Deshalb soll die Datentreuhandplattform langfristig einem Marktplatz-Modell folgen: Datenproduzenten können den Preis für ihre Daten definieren. Kaufen Datenkonsumenten diese Daten dann ein, erhält die Datentreuhandplattform einen prozentualen Anteil an den Umsatzerlösen. Hierbei wird sich an bestehenden Marktplatz- und App-Store-Modellen orientiert.

Es wird erwartet, dass bei einer Ausgründung mindestens eine Anschubfinanzierung notwendig ist, bis ein kostendeckender Betrieb erreicht werden kann. Deshalb wird die Datentreuhandplattform zunächst von Breeze Technologies innerhalb des Unternehmens fortgeführt. So können Synergien hinsichtlich zukünftiger vertrieblicher Aktivitäten genutzt werden. Außerdem können die von Breeze Technologies selbst generierten Daten über die Datentreuhandplattform vermarktet und ggf. erzielte Umsätze zum Betrieb der Plattform eingesetzt werden. Beim Erreichen der Kostendeckung wird dann evaluiert, ob der Spin-Off in eine eigene Organisation einen zu diesem Zeitpunkt schnelleren Marktzugang gewährleisten könnte.

Kostenstruktur

Die wichtigsten identifizierten Kosten werden im folgenden genannt:

- Fixkosten:
    - Personalkosten (mind. 1 Softwareentwickler für Betrieb, Wartung und Weiterentwicklung, mind. 1 Ökosystem-Manager zur Gewinnung von Datenproduzenten und -konsumenten und den Ausbau der Aktivitäten)
    - Server und Anwendungsinfrastruktur
    - Allgemein notwendige Kosten für den Betrieb eines Unternehmens
- Variable Kosten:
    - Reisekosten
    - Kosten für die Speicherung großer Datenmengen der Datenproduzenten

Hierbei ist wichtig klarzustellen, dass die Kosten für den reinen Betrieb der Datentreuhandplattform bei einer Grenzkostenbetrachtung gegen Null gehen, die Kosten für die Speicherung von Luftqualitätsdaten aber mit der Datenmenge bzw.

-größe nahezu linear skalieren. Deshalb muss sichergestellt werden, dass alle gespeicherten Daten auch die notwendige Relevanz haben, um gegenüber möglichen Datenkonsumenten vermarktbar zu sein. Es ist jedoch wichtig, dass bei einer einzelnen Stelle die Geschäftskontinuität nicht ausreichend gesichert sein könnte, z.B. bei Personalwechseln oder längerer Krankheit. Perspektivisch sollten die o.g. Rollen deshalb idealerweise jeweils doppelt besetzt sein.

SWOT-Analyse

Neben dem Business Model Canvas wurde auch eine SWOT-Analyse für das LuftDatenNet-Projekt durchgeführt:

Stärken	**Schwächen**
- Bereits vorliegende Daten aus dem Kerngeschäft von Breeze Technologies für Kick Off der Datentreuhandplattform - Blue Ocean	- Wirtschaftliche Lage insb. hinsichtlich Innovationsprojekten - Langfristige Finanzierung (insb. für aggressive Skalierung)
Chancen	**Risiken**
- Finanzierung & Projektförderung - Zusammenarbeit mit Partnern - Schaffung neuer Incentives für die Generierung von Luftqualitätsdaten (“Data as a service”)	- Konkurrenzdenken von Datenproduzenten - Henne-Ei-Problem - Möglicherweise Churn, sobald sich Datenproduzent und -konsument gefunden haben - Datenqualität bei Datenproduzenten

Die größte Stärke des Vorhabens liegt darin, dass es noch keine andere spezialisierte Datentreuhandplattform im Bereich der Luftqualitätsdaten gibt (“Blue Ocean”) und dass bereits vorliegende Daten aus dem Kerngeschäft von Breeze Technologies (neben Daten anderer Datenproduzenten) genutzt werden können, um die Datentreuhandplattform initial mit Daten zu befüllen.

Auf der anderen Seite wurde erwartet, dass insbesondere Datenkonsumenten aus dem Bereich der Digitalwirtschaft und aus Innovationsteams von Unternehmen zur Schaffung neuer digitaler Produkte kommen. Aufgrund des wirtschaftlichen Abschwungs wurden aber in genau diesen Bereichen in den letzten Jahren Einsparungen vorgenommen, was die Umsetzung von neuen Projekten erschwert. Eine aggressive Skalierung der Plattform würde hohe Kosten bei Personal und Marketing mit sich bringen und ist zum aktuellen Zeitpunkt nicht realistisch. Stattdessen muss hier auf organisches Wachstum gesetzt werden.

Dies könnte aber z.B. durch weitere Projektförderung oder andere Finanzierungsquellen auf Seiten der Chancen abgemildert und der weitere Projektfortschritt so deutlich beschleunigt werden. Auch die Zusammenarbeit mit Partnern insbesondere aus dem Kontext von breiteren Datenmarktplätzen bietet die Möglichkeit, die Anzahl möglicher Datenkonsumenten drastisch zu erhöhen. Hierfür muss die Datentreuhandplattform aber eine ausreichende technische und wirtschaftliche Reife mitbringen. Letztendlich birgt auch die Schaffung der Plattform als solche die Chance, das Luftqualitäts-Datenökosystem voranzubringen, da zukünftig die Generierung von Luftqualitätsdaten (ähnlich wie heute Wetterdaten) ein gewinnbringendes Geschäftsmodell sein könnte, wenn ausreichend Vermarktungskanäle zur Verfügung stehen. Dies könnte mittelfristig ein relevantes Feld für Venture Capital-Investitionen und langfristig ein Feld für Private Equity darstellen.

Auf Seiten der Risiken ist unter anderem zu nennen, dass mögliche Datenproduzenten aufgrund der Doppelrolle von Breeze Technologies nicht gewillt sein könnten, ihre Daten auf der Plattform einzustellen. Sollte dies zukünftig einen signifikanten Anteil möglicher Datenproduzenten betreffen, ist über die Auslagerung der Aktivitäten in ein eigenes Unternehmen nachzudenken. Im Kontext des Förderprojekts war dies jedoch nicht der Fall. Es zeigte sich allerdings, dass auch für die Umsetzung von konkreten geschäftsrelevanten Projekten für die Datenkonsumenten größtenteils noch nicht genügend Daten zur Verfügung stehen ("Henne-Ei-Problem"). Deshalb wurden Datenanforderungen solcher Projekte gesammelt. Sie werden in der weiteren Vermarktung berücksichtigt (z.B. "flächendeckende Daten aus den drei größten Städten des Landes"). Ein weiteres Risiko ist möglicher Abwanderung (Churn) von Datenproduzenten und -konsumenten von der Plattform, sobald diese sich dort für ein potenziell längerfristiges Projekt zusammengefunden haben. Dieses Risiko muss zukünftig vertraglich minimiert werden. Das Risiko bei Projekten, die die zusammengeführten Daten unterschiedlicher Datenproduzenten nutzen, wurde geringer eingeschätzt. Eine weitere Herausforderung besteht in der Sicherstellung der Datenqualität der Datenproduzenten. Diese muss durch den jeweiligen Datenproduzenten selbst angegeben werden, da eine unabhängige Prüfung im Einzelfall sehr aufwändig und kostenintensiv wäre.

Insgesamt überwiegen die Stärken des Vorhabens die Schwächen, da letztere durch die bereits genannten Maßnahmen abgefedert werden können.

2.5 Arbeitspaket 5: Rechte- und Datenmanagement

Ziel des Arbeitspakets war die eingehende Behandlung der Themen Rechtssicherheit und Datenschutz. Hierfür wurden mehrere Kanzleien hinsichtlich ihrer Einschätzung zu Luftqualitätsdaten befragt. Insgesamt kamen hierbei alle befragten Kanzleien zu dem Schluss, dass es sich bei Luftqualitätsdaten weder um

personenbezogene, noch besonders schützenswerte Daten handelt und ihre Speicherung und Verarbeitung deshalb ohne weitere Vorkehrungen möglich ist.

Mögliche zukünftig relevanter Themen im Rechte- und Datenmanagement, wie z.B. Zertifizierungen, die für die Nutzung der Daten durch bestimmte Organisationen möglich sein könnten (z.B. ISO 27001), wurden bei dem Entwurf der Softwarearchitektur (AP 3) berücksichtigt. So ist bereits eine detaillierte Protokollierung von Datenzugriffen in das System integriert.

Sonstige technisch-organisatorisch notwendige Maßnahmen zum Datenschutz leiten sich sonst nur von individuellen Vereinbarungen mit Datenproduzenten (und ggf. -konsumenten) auf Basis ihrer individuellen Anforderungen zur Nutzung der Plattform ab und waren zum Zeitpunkt der Projektdurchführung nicht notwendig. Eine Pseudonymisierung der Daten ist ebenso nicht notwendig und wäre kontraproduktiv, da die einzigen identifizierenden Kriterien Standort und Zeitpunkt sind, die jedoch zwingend für die Weiterverarbeitung der Daten notwendig sind. Die Datenquelle und Informationen z.B. über die Erfassungsmethode sind wichtig, um die Nutzbarkeit der Daten zu bewerten.

Für die zukünftige Nutzung der Datentreuhandplattform wurden Vertragsmuster entworfen, die die relevantesten Regelungen wie Dateneigentum, Speicher- und Verarbeitungserlaubnissen und Datenhaftung enthalten.

2.6 Arbeitspaket 6: Projektmanagement, -koordinierung und -kommunikation

Ziel des Arbeitspaketes war die effektive und effiziente Koordination des Projektes und die Sicherstellung eines effizienten Berichtswesens.

Es wurde hierfür unter anderem auch Kontakt zu bestehenden Datenaggregatoren und zu weiteren datentreuhändischen Anbietern aufgenommen, deren Erfahrungen mit ihren Geschäftsmodellen für die Entwicklung von LuftDatenNet interessant waren.

Zur internen und externen Projektkommunikation und zur Vernetzung wurde eine Website (www.luftdaten.net) eingerichtet, die zunächst das Projekt beschrieb. Später wurde diese durch den Softwareprototypen ersetzt. Zum Zweck der Projektkommunikation wurden außerdem Projekt-E-Mail-Postfächer angelegt. Außerdem wurde via Pressemitteilungen über das Projekt informiert.

Ursprünglich war vorgesehen, für das Projekt größtenteils neues Personal zu akquirieren. Da die ausgeschriebenen Stellen bis zum Projektstart allerdings unbesetzt blieben, wurde, um den Projekterfolg nicht zu gefährden, bestehendes

Personal umbesetzt. Dabei wurde darauf geachtet, dass die ursprünglich definierten Anforderungen an die einzelnen Rollen eingehalten wurden. Hierdurch kam es aber zu leichten Verschiebungen bei der Umsetzung der geplanten Personenmonate.

Im Rahmen von wöchentlichen Projektstatusmeetings wurde der Projektfortschritt und -erfolg überwacht, um frühzeitig nachsteuern zu können.

Die Begleitforschung wurde während dem Projekt durch die Teilnahme an Interviews und Vernetzung bei Vernetzungsveranstaltungen unterstützt.

3. Wichtigste Positionen des zahlenmäßigen Nachweises

Die Mittel für Personalkosten von ca. 673 T€ wurden wie beantragt im Projektzeitraum für Personal zur Bearbeitung der oben genannten Arbeiten innerhalb des Förderprojektes ausgegeben.

4. Notwendigkeit und Angemessenheit der geleisteten Projektarbeit

Die Durchführung des LuftDatenNet-Projekts war notwendig, um die erste Datentreuhandplattform für Luftqualitätsdaten zu entwickeln und zu implementieren. Sie wird zukünftig den Austausch von Luftqualitätsdaten zwischen Datenproduzenten und -konsumenten ermöglichen und damit das Datenökosystem in diesem Bereich fördern.

Die geleisteten Arbeiten waren zur Erreichung der Projektziele, d.h. zur Konzeption und zum initialen Aufbau einer Datentreuhandplattform für Luftqualitätsdaten, notwendig. Der Bedarf für eine solche Datentreuhandplattform wurde während der Projektlaufzeit im Rahmen von Marktrecherchen und Stakeholderinterviews bestätigt (siehe u.a. AP4). Die Datenproduzenten und -konsumenten, die die Datentreuhandplattform zukünftig nutzen werden, kommen vor allem aus den Bereichen Mobilität, (Rück-)Versicherungen, Medien, Real Estate, Sport, Forschung, Tourismus und Facility Management. Die Datentreuhandplattform entlastet dabei sowohl Datenproduzenten als auch -konsumenten bei der Nutzbarmachung und konkreten Nutzung von Luftqualitätsdaten über hoheitliche Aufgaben hinaus.

Insbesondere aufgrund des “Henne-Ei-Problems”, dass eine Datentreuhandplattform sowohl genügend Datenproduzenten, als auch -konsumenten benötigt, um kostendeckend arbeiten zu können, war die Durchführung des Förderprojektes notwendig. Die Projektergebnisse zeigen, dass hier immer noch eine Lücke besteht.

Mit der Entwicklung von mehreren Case Studies und technischen Prototypen konnten die umgesetzten Anforderungen und die Funktionalität der Plattform getestet und evaluiert werden. Dank der Projektdurchführung mit einem interdisziplinären Team konnten alle hierfür notwendigen Fachkompetenzen und Netzwerkzugänge gebündelt und in das Projekt eingebracht werden.

Die gewählte Methodik und der Umfang der Projektarbeiten waren angemessen, um die genannte Zielsetzung zu erreichen. Das Projekt konnte innerhalb des gesteckten Zeitraums durchgeführt und abgeschlossen werden.

Insgesamt konnte damit ein signifikanter Beitrag zur Weiterentwicklung von Datentreuhandmodellen und der Datenökonomie allgemein erzielt werden. Die Datentreuhandplattform wurde mit einem Technologiereifegrad (TRL) von 6 erprobt.

5. Nutzen und Verwertbarkeit der Ergebnisse

Die Projektergebnisse sind unmittelbar anschlussfähig. Seit Projektende betreibt Breeze Technologies die Plattform weiter. Für die dauerhafte Etablierung der Datentreuhandplattform am Markt müssen jedoch noch weitere Partner gewonnen und zukünftig Umsätze über konkrete Projektimplementierungen erzielt werden. Hierfür wird weiter für die Plattform und das Projekt geworben. Dank der thematischen Überschneidungen hinsichtlich des aktuellen Produktportfolios des Unternehmens ist dies einfach möglich.

Im Verlauf des Projekts kristallisierte sich noch die weitere Nutzungsmöglichkeit als Austauschplattform zwischen Städten heraus, um Vergleiche und Zusammenarbeit zwischen Verwaltungen zu ermöglichen. Dieses Nutzungsszenario soll zukünftig auch weiter untersucht werden. Hierfür hat sich bereits ein potenzieller Partner in den USA angeboten.

Entwickelte Programmteile und Microservices lassen sich zum Teil auch auf andere Forschungskontexte übertragen. Sie bieten eine flexible Grundlage, den Austausch von Daten zwischen Stakeholdern auch in anderen Kontexten sicherzustellen.

Insgesamt sieht das Projektteam damit ein hohes Verwertungspotenzial für die entwickelten Projektergebnisse.

Darüber hinaus haben sich mit den Ergebnissen insbesondere aus AP 3 weitere Forschungsbedarfe ergeben, die im Rahmen von zukünftigen Forschungsprojekten untersucht werden sollten:

- Temporale Interpolation und Aggregation von Luftqualitätsdaten zur Vereinheitlichung von Zeitstempeln mehrerer Datensätze in paralleler Verarbeitung (z.B. Daten von einem Messpunkt um 16:00:01 und Daten eines anderen Messpunkts um 16:00:27 Uhr)
- Zusammenführung von Luftqualitätsdaten unterschiedlicher Qualitätsstandards, z.B. hinsichtlich der Qualität des harmonisierten Datenproduktes

6. Anhänge

Anhang 1: Luftdaten Air Quality Measurement Protocol (LAQMP)

```
{
   "measurementID":"string",
   "location":{
      "country":"ISO-country-code",
      "city":"city-name",
      "address":"specific-location-address",
      "coordinates":{
         "latitude":"number",
         "longitude":"number"
      }
   },
   "measuredAt":"ISO-time",
   "data":[
      {
         "param":"param",
         "units":"unit",
         "value":"value"
      }
   ],
   "source":{
      "organizationName":"name",
      "url":"url",
      "device":{
         "deviceType":"string"
      }
   }
}
```

Beispieldatensatz

```
{
   "measurementID":"AQ239482034",
   "location":{
      "country":"USA",
      "city":"San Francisco",
      "address":"742 Evergreen Terrace",
      "coordinates":{
         "latitude":37.7749,
         "longitude":-122.4194
```

```
        }
    },
    "measuredAt":"2023-11-29T14:30:00Z",
    "data":[
        {
            "param":"PM2.5",
            "units":"µg/m³",
            "value":12.5
        },
        {
            "param":"PM10",
            "units":"µg/m³",
            "value":20.3
        },
        {
            "param":"CO",
            "units":"ppm",
            "value":0.7
        },
        {
            "param":"SO2",
            "units":"ppb",
            "value":5.6
        },
        {
            "param":"O3",
            "units":"ppb",
            "value":30.2
        }
    ],
    "source":{
        "organizationName":"City of San Francisco",
        "url":"sf.com/airquality",
        "device":{
            "deviceType":"Breeze Sensor"
        }
    }
}
```