



Datengovernance Handlungsfeld Datenintegration

CUT | T1 | 01.10.2023

Partnerstädte:



Gefördert durch:



Disclaimer:

Dieses Dokument bietet einen Überblick über die aktuellen Datenintegrationsprozesse im Kontext der Datengovernance in den drei beteiligten Städten. Die präsentierte Information basiert auf der Perspektive und Erfahrung in der Verwaltung der Geodateninfrastruktur (GDI) und der Urbanen Datenplattform (UDP).

1. **Zweck und Perspektive:** Dieses Dokument bietet einen Überblick über die Datenintegrationsprozesse, insbesondere in Bezug auf die GDI. Es zielt darauf ab, die Strategien, Herausforderungen und Erfolge der Datenintegration in den drei CUT-Städten darzustellen.
2. **Begrenzte Abdeckung:** Bitte beachten Sie, dass sich dieses Dokument auf die Datenintegration in die GDI und die UDP konzentriert. Die bereitgestellten Informationen decken nicht alle Aspekte der Datenintegration ab, insbesondere nicht weitergehende Themen wie Emerging Technologies wie KI oder Business Intelligence.
3. **Keine professionelle Beratung:** Die bereitgestellte Information dient nur Informationszwecken. Dieses Dokument ersetzt keine professionelle Beratung. Für umfassende Einblicke oder Empfehlungen zu neuen Technologien wird empfohlen, qualifizierte Fachleute zu konsultieren.
4. **Änderungen und Aktualisierungen:** Bitte beachten Sie, dass sich der Inhalt dieses Dokuments ändern kann. Dies ist auf die Weiterentwicklung der Datenintegration und der Technologie zurückzuführen.

Durch die Nutzung dieses Dokuments erkennen Sie diesen Haftungsausschluss an.

Inhaltsverzeichnis

1	Einführung.....	4
2	Festlegung von Rollen und Zuständigkeiten	4
3	Strukturierung (Formalisierung) der Datenintegration	7
3.1	Datenintegrationsprozess.....	7
3.1.1	Datenbedarf melden.....	7
3.1.2	Datenquellen identifizieren.....	7
3.1.3	Anforderungen erfassen.....	8
3.1.4	Visualisierung	8
3.1.5	Metadaten erfassen.....	8
3.1.6	Daten freigeben.....	9
3.1.7	Datenübernahme	9
3.1.8	Erstellung von Webdiensten.....	9
3.1.9	Test & Veröffentlichung	10
3.2	Weitere Aspekte	10
3.2.1	Zielgruppen, Zugangswege und Nutzerberechtigungen	10
3.2.2	Daten- Aktualisierungszyklen.....	11
3.2.3	Turnus Datenintegration	11
4	Dokumentation & Kommunikation des Datenintegrationsprozesses.....	11
5	Ausblick	13
6	Anhang	15
6.1	Anhang 1: Data Governance-Rollen im Kontext GDI/UDP-Datenmanagement in drei CUT-Städten	15
6.2	Anhang 2: Freigabeerklärung der FHH.....	17
6.3	Anhang 3: Formular Metadatenerfassung.....	17
6.3.1	UDP FHH - Handzettel Metadatenatz erfassen	17
6.3.2	GDI München - Metadatenkatalog Formular	17

6.3.3 Geoportal München – Metadatenkatalog Formular.....	19
6.4 Anhang 4: Dokumentation Datenintegrationsprozess.....	20
6.4.1 Beispiel Leipzig: Grafik zum Datenintegrationsprozess im Intranet.....	20
6.4.2 Beispiel Hamburg: Handzettel Datenintegration.....	20

Abbildungsverzeichnis

Abbildung 1: DG Rollen im Datenintegrationsprozess	6
Abbildung 2: Generalisierter Datenintegrationsprozess in den CUT-Städten (Quelle: LGV)	7
Abbildung 3: Darstellung des Prozesses der Datenveröffentlichung im ODP-L im Intranet (Quelle: Leitfaden Open Data Portal Leipzig, abgerufen am 2.12.2022).....	20

1 Einführung

Das vorliegende Dokument baut auf der ersten Definition der Handlungsfelder der Datengovernance (DG) für Städte und Kommunen auf, die im Rahmen des CUT-Projektes 2022 zur Schaffung eines gemeinsamen Verständnisses der Begriffe erarbeitet wurde.

Ziel des Dokumentes ist eine nähere Betrachtung des Handlungsfeldes „Datenintegration“ unter Berücksichtigung der Erfahrungen der drei CUT-Städte Hamburg, München und Leipzig und der Dokumentation von Best Practice Beispielen.

Unter Datenintegration (DI) wird ein Prozess verstanden, der Daten aus - ggfs. unterschiedlichen - Datenquellen in eine andere Dateninfrastruktur (z.B. in die Geodateninfrastruktur oder urbane Datenplattform einer Kommune) einbindet. In Hamburg, Leipzig und München erfolgt die Dateneinbindung durch ETL-Prozesse (Extract Transform Load), wobei die Daten konsolidiert in einem Data Warehouse gespeichert und über Kataloge und Dienste verfügbar gemacht werden. Andere Methoden der Datenverwaltung wie Data Lakes, die zum Speichern von unverarbeiteten Rohdaten genutzt werden, spielen in den drei Städten (noch) keine größere Rolle. Ziel der Datenintegration ist die Verknüpfung zuvor getrennt vorgehaltener Daten und die Vermeidung von Datenredundanz. Ziel ist es zudem, bei der Datenintegration offene Standards zu nutzen und bedarfsbezogen Datensätze und Dienste bereitzustellen, die

- maschinenlesbar sind;
- in internen und externen Anwendungen wie komplexen Simulationswerkzeugen analysierbar und auswertbar sind;
- in internen und externen Anwendungen wie Kartenclients oder Dashboards visualisierbar sind;
- über Metadaten auffindbar und gut beschrieben sind.

Innerhalb des Prozesses sind verschiedene Akteure mit verschiedenen Rollen und Zuständigkeiten beteiligt.

Im Rahmen einer Datengovernance werden zunächst Rollen und Zuständigkeiten festgelegt (Kapitel 2). Der Prozess der Datenintegration innerhalb einer Kommune wird strukturiert und z.B. mittels Formularen standardisiert (Kapitel 3). Auch sollte der Prozess möglichst transparent und nachvollziehbar beschrieben und kommuniziert werden (Kapitel 4).

2 Festlegung von Rollen und Zuständigkeiten

Die Rollen der Datengovernance lassen sich auf drei Ebenen gliedern: Strategie, Datengovernance und Daten-Management (vgl. Handlungsfeld Rollen). Die Mehrzahl der Rollen, die im Kontext des Datenintegrationsprozesses eine bedeutende Rolle spielen, fallen in die Kategorie

der Daten-Management-Ebene. Diese Rollen sind in Abbildung 1 dargestellt und wie folgt definiert¹:

- **Koordinierungsstelle UDP/UDZ:** Die (zentrale) Rolle **Koordinierungsstelle UDP/UDZ** ist verantwortlich für den fachlichen Betrieb der zentralen Datenplattform (UDP/UDZ) und trägt für diese Budgetverantwortung [...].
- **Dateneigentümer/-in:** Die dezentrale Rolle des Dateneigentümers / der Dateneigentüme-
rin setzt die stadtweite Strategie und Datengovernance auf Organisationseinheit (OE) oder
Datendomänenebene (DDO) um. Sie gewährleistet Datenkontinuität, -qualität und -bereit-
stellung für Nutzer/-innen sowie Überwachung der Daten- und Prozessqualität [...].
- **Datenlotse/-in:** Die (dezentrale) Rolle Datenlotse/-in setzt die Datengovernance in das
operative Datenmanagement einer OE bzw. DDO um. Sie trägt die fachliche Datenverant-
wortung, löst Datenverwendungs- und Integrationsprobleme (mit Datenpfleger/-innen)
und agiert als Bindeglied zur Koordinierungsstelle UDP/UDZ, Datengovernance Manager/-
in und IT-Betreuung [...].
- **Datenpfleger/-in:** Die (dezentrale) Rolle **Datenpfleger/-in** ist fachlich zuständig für spe-
zifische Datensätze bzw. Fachanwendungen. [...] Sie ist im Datenintegrationsprozess fach-
lich und ggf. technisch gesehen Ansprechpartner/-in für die Koordinierungsstelle
UDP/UDZ. Konkrete Aufgaben der Rolle umfassen u.a.: Die Erhebung und Pflege von Da-
ten/Metadaten (z.B. Exceltabelle, Fachsystem, Data Warehouse) in einer OE bzw. DDO; die
Freigabe von Daten (Zugriffserteilung) für Weitergabe (fachliche Richtigkeit) [...].

Zusätzlich zu den oben genannten Rollen, die auf gesamtstädtischer Ebene definiert sind, kann es aus der operativen Perspektive des Datenintegrationsprozesses weitere Rollen in den Berei-
chen Beratung, Entwicklung von Datenintegrations-Workflows und Bereitstellung von
APIs/Diensten geben. Vor allem die folgenden Rollen sind wichtig:

- **Dateningenieur/-in:** Die Verantwortlichkeiten eines Dateningenieurs erstrecken sich
auf die präzise Umsetzung von Vorgaben und die Schaffung von datenbasierten Diens-
ten. Dies schließt die Konzipierung, Entwicklung und Implementierung von ETL-Prozes-
sen (Extract, Transform, Load) ein, die essenziell für die effiziente Datenaufbereitung
und -integration sind. Des Weiteren obliegt es dem Dateningenieur, diese Dienste zu
generieren, was bedeutet, dass sie sowohl die technische Umsetzung als auch die be-
reitgestellten Funktionen und Datenleistungen orchestrieren. In dieser Funktion sind sie

¹ Die in diesem Dokument genannten Rollen und ihre Beschreibungen sind dem Dokument "T1 DG Field of Action Roles" entnommen. Die vollständige Beschreibung der Rollen finden Sie im Originaldokument.

maßgeblich daran beteiligt, sicherzustellen, dass die Daten in den gewünschten Formaten und Strukturen vorliegen, um den Anforderungen der Dateneigentümer und -bedarfsträger gerecht zu werden.

- **IT-Betreuer/-in:** Ein IT-Dienstleister ist maßgeblich an der Umsetzung, Verwaltung und Sicherstellung des reibungslosen Betriebs der IT-Infrastruktur, Daten und -Dienste für die GDI/UDP beteiligt. Der IT-Dienstleister arbeitet eng mit GDI/UDP zusammen. Komponentenverantwortliche innerhalb des IT-Dienstleisters übernehmen die Konfiguration und Administration der Server sowie teilweise die Verwaltung der Zugriffsrechte.

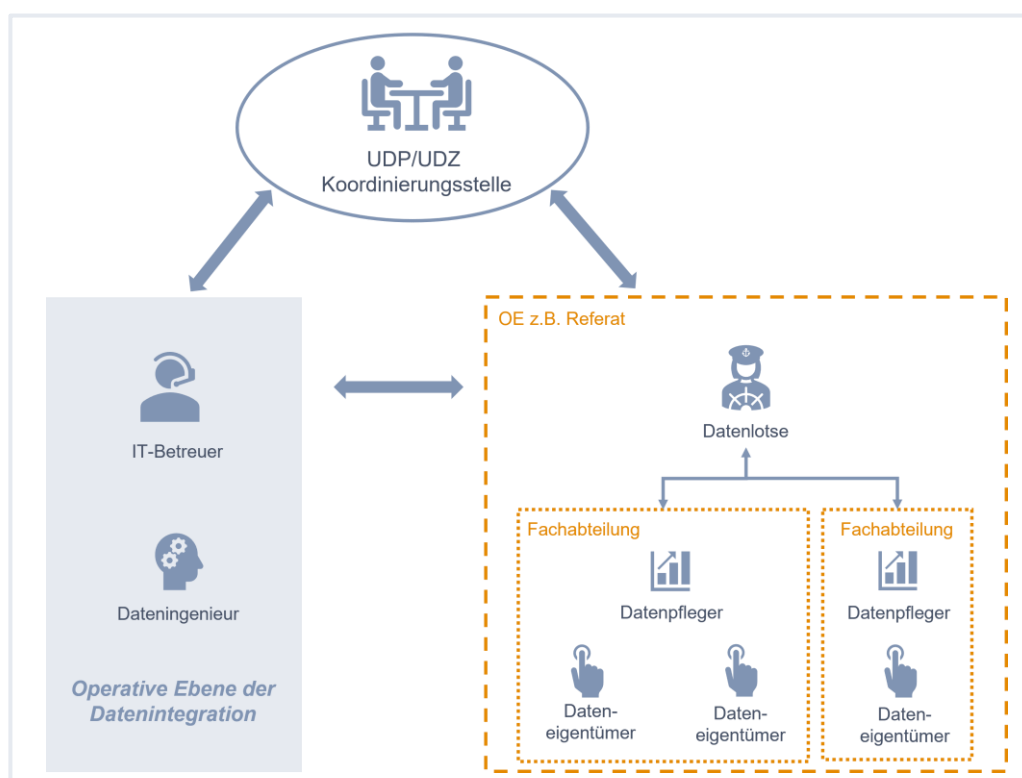


Abbildung 1: DG Rollen im Datenintegrationsprozess

Parallel zur Rolle des Datenlotsen kann es in der GDI/UDP Koordinierungsstelle die Rolle des „Datenmanagers“ (siehe Hamburg in Anhang 6.1) bzw. des „Anwendungsfallmanagers“ (wie in Leipzig) geben. Beide übernehmen dann ähnliche Aufgaben wie der Datenlotse in der OE nur auf Ebene der Koordinierungsstelle GDI/UDP. Sie ermitteln konkrete Anforderungen für die Dateningenieure im Dialog mit den Datenlotsen der OEs oder den jeweiligen Datenpflegern.

Hamburg, Leipzig und München haben im Rahmen ihres jeweiligen GDI/UDP Datenintegrationsprozesses Datenmanagement-Rollen definiert, die sich inhaltlich in den oberen Rollen wiederfinden, jedoch je nach Stadt z.T. andere Bezeichnungen tragen (siehe Anhang 6.1). Die

momentan vorhandenen Rollen werden im Rahmen der aktuell zu erarbeitenden Datengovernance-Strategie in jeder Stadt betrachtet und je nach Bedarf entsprechend der in diesem Kapitel definierten Rollen erweitert oder angepasst.

3 Strukturierung (Formalisierung) der Datenintegration

3.1 Datenintegrationsprozess

Der Datenintegrationsprozess in den drei CUT-Städten Hamburg, Leipzig und München ist in nachfolgender Abbildung generalisiert dargestellt und umfasst neun Schritte, die je nach Stadt in ihrer Reihenfolge etwas variieren. Die dargestellte Reihenfolge ist daher beispielhaft und sinnvoll, aber nicht festgelegt. Jedoch sind die einzelnen Schritte klar definiert und werden im Folgenden kurz beschrieben.

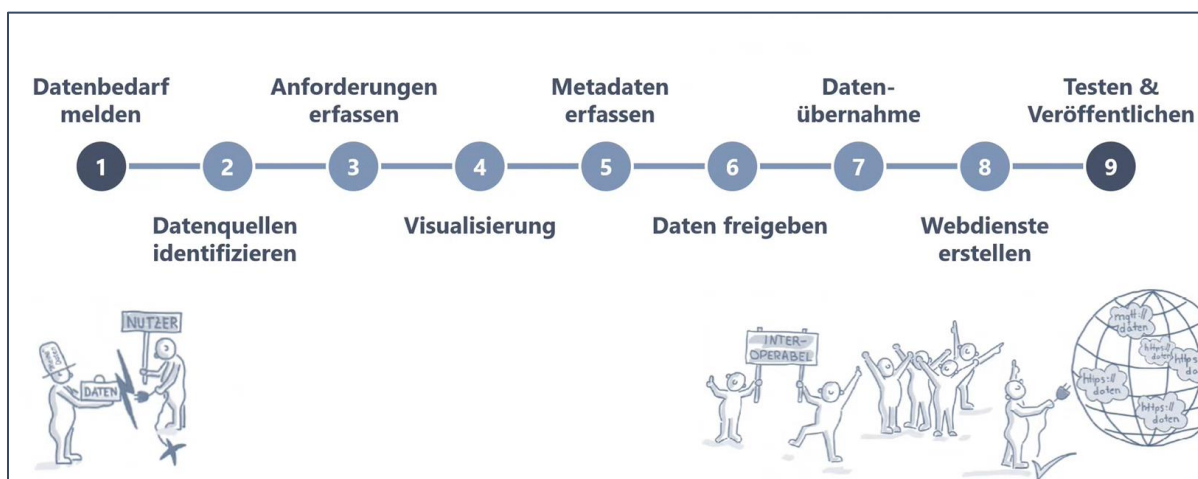


Abbildung 2: Generalisierter Datenintegrationsprozess in den CUT-Städten (Quelle: LGV)

3.1.1 Datenbedarf melden

Der erste Schritt bei der Datenintegration ist immer die formale (z.B. Online-Formular, Ticket-system) oder formlose Meldung eines Datenbedarfs bei der GDI oder UDP Koordinierungsstelle durch einen Dateneigentümer, der seine Daten veröffentlichen möchte, oder einen Kunden (oder Datenbedarfsträger), der auf der Suche nach Daten für die Umsetzung eines Anwendungsfalls oder einer Fachaufgabe ist. Die Bedarfsmeldung kann je nach Komplexität eines Anwendungsfalls einen oder mehrere Datensätze (Dateien) betreffen.

3.1.2 Datenquellen identifizieren

Für den Fall, dass ein Dateneigentümer Daten veröffentlichen möchte, bespricht der Datenlotse mit ihm oder dem Datenpfleger die möglichen Datenquellen. Man unterscheidet u.a. zwischen:

- Schnittstellen in Form von Webdiensten oder APIs,

- eine (strukturierte) Sammlung von Datenbankeinträgen, eines Data Warehouses oder im Rahmen von Fachsystemen,
- einzelnen Datensätzen (Dateien) in verschiedenen Formaten (Excel, Shapefiles, Sharepoint-Tabellen etc.)

3.1.3 Anforderungen erfassen

Bei diesem Schritt besprechen der Datenlotse und der Dateneigentümer, was mit den Daten gemacht werden soll (Form der Visualisierung wie Webservice oder 3D, Funktionalitäten wie Filter oder Time Slider etc.), welcher Nutzerkreis erreicht werden soll (Open Data oder eingeschränkter Nutzerkreis) und welche Inhalte benötigt werden (Definition des Datenmodells). Zudem wird die Einhaltung der DSGVO (Datenschutz-Grundverordnung der Europäischen Union) und anderer rechtlicher Vorgaben geprüft.

3.1.4 Visualisierung

Im Rahmen der Visualisierung legen der Datenlotse und der Dateneigentümer zusammen fest, welche Layer und Attribute dargestellt werden sollen und wie die Gestaltung von Punkt-, Linien- und Flächenobjekten erfolgen soll (Größen, Klassifizierung, Farben, Maßstabsbereich etc.). Zudem werden Schreibweisen vereinbart (z.B. strasse oder Straße, Vermeidung von Abkürzungen und Fachbegriffe).

3.1.5 Metadaten erfassen

Die Erfassung der Metadaten für einen Datensatz erfolgt durch den Datenpfleger mittels eines Templates (vgl. Anhang 6.3.1, 6.3.2 und 6.3.3).

Metadaten sind unentbehrlich für die Dokumentation, den Transfer/Datenaustausch sowie die längerfristige Wertsicherung der Daten. Sie beschreiben u.a. Eigenschaften, Definition, Herkunft, Gültigkeit, Genauigkeit, Aktualisierungszyklus sowie Einsatz- und Nutzungsmöglichkeiten von Datensätzen auf unterschiedlichen Aggregationsebenen. Auch die Bündelung von Metadaten über Stadt- oder Ländergrenzen hinaus verbessert den Zugang zu und die gemeinsame Nutzung von Daten (Bsp. MetaVer, verfügbar unter: metaver.de).

Die Metadaten für die erzeugten Dienste (und nicht die gelieferten Datensätze/Dateien/Datenbankeinträgen) werden in Hamburg, Leipzig und München durch die datenhaltende Stelle (GDI/UDP) eingetragen und gepflegt (vgl. 3.1.8 und Anhang 6.3.3). Es ist zu beachten, dass die Metadaten bei Bedarf in den nächsten Schritten aktualisiert werden können und am Ende des Prozesses, wenn die Daten oder Dienste verfügbar und einsatzbereit sind, veröffentlicht werden.

3.1.6 Daten freigeben

Die Dateneigentümer bestätigen je nach Voraussetzung in den Städten (vgl. HmbTG²) die Freigabe der Daten (Internet, Intranet, weitere Beschränkungen/Gruppen) und unterschreiben eine Datenschutzerklärung. Die Datenfreigabe erfolgt in Hamburg und München per Formular (vgl. Anhang 6.2) frühzeitig im Prozess und zusammen mit der Metadatenerfassung. Auch in Leipzig ist ein digitales Formular geplant.

3.1.7 Datenübernahme

Für die Übernahme der Daten muss der Datenpfleger dem Dateningenieur den Datenzugriff ermöglichen. Mittels ETL-Prozessen (z.B. per FME) werden Datenquelle und Datenziel festgelegt und Anpassungen der eingehenden Daten vorgenommen, um die Integrität der zentralen Datenbasis zu erhalten. Es muss zudem die Vollständigkeit und logische Konsistenz (vgl. Handlungsfeld Datenqualität) sichergestellt sein, so dass alle Daten innerhalb der Datenbank aufrufbar und im Sinne des Datenmodells verknüpfbar sind.³ Die Sicherstellung der inhaltlichen Qualität der Daten ist in allen drei CUT-Städten Fachaufgabe bzw. liegt in der Verantwortung des Dateneigentümers.

Werden technische Fehler wie Geometriefehler oder Mappingfehler in den Daten identifiziert, so werden sie dem Datenpfleger übermittelt, der die Fehler in einem Zeitfenster von einigen Tagen im Fachsystem korrigieren kann.

Die ETL-Prozesse für die Datenübernahme werden auf einem Webserver (z.B. FME Server) veröffentlicht und so konfiguriert, dass die Daten bei Bedarf regelmäßig aus der Fachanwendung oder Fachdatenbank ausgelesen und automatisch in der GDI/UDP aktualisiert werden.

3.1.8 Erstellung von Webdiensten

Unter Berücksichtigung der Anforderungen (vgl. 3.1.3) und der Vorgaben zur Visualisierung (vgl. 3.1.4) erstellt der Dateningenieur einen (Geo-)Datendienst (i.d.R. WMS, WFS, OAF, Sensor-Things API) und stellt diesen für den Dateneigentümer zur Abnahme bereit. Darüber hinaus werden die Metadaten veröffentlicht und können im Idealfall durch andere Portale (z.B. GovData, data.europa.eu) gefunden und eingesammelt (geharvestet) werden.

² Hamburgisches Transparenzgesetz

³ Lexikon der Kartographie und Geoinformatik (2001), online verfügbar unter: <https://www.spektrum.de/lexikon/kartographie-geomatik/datenintegration/851>

3.1.9 Test & Veröffentlichung

Vor der Veröffentlichung bzw. Produktivschaltung werden die Ergebnisse wie z.B. Dienste oder Portale noch einmal durch den Kunden und ggf. den Dateneigentümer (oder dem Datenpfleger) getestet. Auch erfolgt die Veröffentlichung nur, wenn die Metadaten vorliegen. Nach der Veröffentlichung können Datennutzer die Dienste nun über verschiedene Portale suchen, via URL aufrufen und je nach Spezifikation (z.B. WFS) und Konfiguration auch auf Dateninhalte zugreifen.

3.2 Weitere Aspekte

Im Folgenden werden die Datenbereitstellung und die Datenaktualisierung, die im Rahmen des Datenintegrationsprozesses bereits kurz thematisiert wurden (vgl. 3.1), näher betrachtet. Zudem wird der für den Datenintegrationsprozess notwendige Zeitrahmen aufgezeigt.

3.2.1 Zielgruppen, Zugangswege und Nutzerberechtigungen

Im Rahmen der Datenbereitstellung innerhalb der jeweiligen UDP/GDI ergeben sich vielfältige Zielgruppen und Zugangswege. Die Schwerpunkte variieren je nach städtischer Strategie und den geltenden Gesetzen. Ein Beispiel hierfür ist Hamburg, wo primär die Daten, nach Transparenzgesetz, vorrangig für das Internet freigegeben werden. Neben der internen Bereitstellung über das Verwaltungsintranet ist auch der Zugang über spezifische Verwaltungsgruppen/Referate vorgesehen. In Leipzig und München liegt die Entscheidung über den Datenzugang bei den Dateneigentümern. Die Daten können für die breite Nutzung außerhalb der Verwaltung oder für die interne Nutzung im Intranet oder nur für bestimmte Gruppen bereitgestellt werden. In beiden Städten existiert ein Open Data Portal, das sämtliche öffentliche städtische Daten durchsuchbar macht. Alle drei Städte bieten Zugang zu geobezogenen Daten sowohl innerhalb der Verwaltung (Intranet oder für spezifische Verwaltungsgruppen/Referate) als auch für die Öffentlichkeit über standardisierte Schnittstellen und vielfältige Formen der Datenvisualisierung an.

In bestimmten Fällen erfolgt die Steuerung der Nutzerberechtigungen über die verschiedenen Aggregationsebenen der Datensätze. Dies impliziert, dass für verschiedene Zielgruppen die Bereitstellung von Datensätzen in unterschiedlichen Detailstufen möglich und ggf. erforderlich ist. Ein Beispiel hierfür ist München, wo Luftbilder in Auflösungen von 20 und 60 cm als Open Data im Internet veröffentlicht werden, während für die interne Nutzung die Auflösungen von bis zu 5 cm zur Verfügung stehen. Auch können einzelne Attribute eines Datensatzes für die Darstellung z.B. im Internet ausgeblendet werden, d.h. auch attributbezogen können Daten in unterschiedlichen Detailstufen bereitgestellt werden.

Jedoch gilt generell in allen drei Städten die Maxime, dass die von der UDP/GDI vorgehaltenen Informationen für alle möglichst gleich sind, da das Ziel das „Aufbrechen der Datensilos“ und nicht die Schaffung neuer exklusiver Datenräume für ausgewählte Nutzer ist.

3.2.2 Daten- Aktualisierungszyklen

Sowohl für die Dateneigentümer/-anbieter als auch für die Datennutzer sind die Aktualisierungszyklen der Daten von Bedeutung. Dabei ist zwischen inhaltlichen Datenaktualisierungen und technischen Datenaktualisierungen im Rahmen des Datenintegrationsprozesses zu unterscheiden. Die Häufigkeit der inhaltlichen Datenaktualisierung variiert je nach Anwendungsfall und Nutzungsbedarf und liegt in der Verantwortung des Dateneigentümers. Die Datenaktualisierung erfolgt dabei entweder individuell in Abstimmung mit dem Dateneigentümer oder gemäß festgelegten Zyklen. Die technischen Datenupdates (die Aktualisierung der Datensätze entsprechend den inhaltlichen Datenupdates), hängen weitgehend von der Effizienz der in dem Kapitel 3.1 genannten Schritte ab.

3.2.3 Turnus Datenintegration

Der Zeitrahmen für den Datenintegrationsprozess variiert je nach Faktoren wie Datensatzkomplexität und Erfahrung des Datenpflegers. Dies kann zwischen wenigen Tagen bis hin zu mehreren Monaten dauern, abhängig von den Anforderungen, Bedingungen und der Zuarbeit aus den Fachabteilungen. Zudem verlangsamt sich der Prozess durch die Auslagerung der Produktivumgebung zum IT-Dienstleister, der jedoch die Ausfallsicherheit besser gewährleisten kann. Die Einbindung automatisierter Schritte in den bestehenden Prozess hat das Potenzial, die Dauer der Datenintegration zu verkürzen und eine schnellere Aktualisierung zu ermöglichen. Dazu gehören z. B. die teilweise automatisierte Metadatenerfassung, die Self-Service-Bereitstellung von Diensten oder eine Kundenportallösung mit assistierten Onboarding- bzw. Aktualisierungsprozess.

4 Dokumentation & Kommunikation des Datenintegrationsprozesses

Die Kommunikation und Dokumentation der Datenintegration stellen eine anspruchsvolle Aufgabe dar, da sie verschiedene Herausforderungen mit sich bringen. Zum einen müssen komplexe technische Aspekte verständlich und präzise an unterschiedliche Zielgruppen vermittelt werden, darunter IT-Experten, Dateneigentümer, -pfleger und -nutzer. Gleichzeitig muss die Dokumentation sowohl technische Details als auch geschäftliche Anforderungen abdecken, was eine Balance zwischen Tiefe und Verständlichkeit erfordert (vgl. Anhang 6.4). Zudem kann

die sich stetig weiterentwickelnde Natur von Technologien und Anforderungen dazu führen, dass Dokumentationen rasch veralten.

Im Rahmen des Datenintegrationsprozesses sind im Zusammenspiel von Kunden, einschließlich Dateneigentümern, Datenlotsen und Datennutzern vor allem die folgenden Herausforderungen zu beachten:

- Eine wichtige Aufgabe besteht darin, den jeweiligen Dateneigentümer zu identifizieren, was oft dem Datenbedarfsträger überlassen wird.
- Die Datenfortführung und -qualität bis hin zur Kontaktaufnahme und Abstimmung mit Herstellern der datenführenden Fachsysteme liegen in der Verantwortung der Dateneigentümer und erfordern oft zeitintensive Aufbereitungen wie Normalisierung seitens der Datenpfleger.
- Der Abstimmungsprozess für Dienstveröffentlichungen kann sich aufgrund hoher Qualitätsansprüche und ausgedehnter Abstimmungs- und Freigaberunden im Fachbereich stark verlängern, was in einigen Fällen bereits 2-3 Jahre in Anspruch genommen hat.
- Die begrenzte Datenkompetenz seitens des Kunden kann zu Schwierigkeiten bei der Kommunikation sowie Abstimmung im Datenintegrationsprozess führen.

Um die Zusammenarbeit mit den Kunden, einschließlich Dateneigentümern, Datenlotsen und Datennutzern, zu optimieren, werden verschiedene Maßnahmen empfohlen:

- Festsetzung von Fristen für die Zulieferung von Daten, Datenkorrekturen, Metadaten und Freigabeerklärungen seitens des Dateneigentümers sowie für notwendige Absprachen und Einrichtung von (automatischen) Erinnerungen.
- Etablierung von Datenlotsen in den jeweiligen Ämtern, die als Multiplikatoren agieren und über umfassende fachliche sowie technische Kompetenz verfügen.
- Durchführung regelmäßiger Treffen mit den Datenlotsen, um den Statusbericht der UDP/GDI zu erörtern und neue Anforderungen zu verstehen und zu sammeln.
- Aktive Identifikation von Anwendungsfällen durch direkte Kundenkontakte in vordefinierten Sitzungen mit städtischen Unternehmen und datenhaltenden Ämtern.
- Definierte Einbindung der GDI/UDP Koordinierungsstelle als Beratung in IT-Projekten mit Geodatenbezug.
- Pflege von Kundenkontakten und Vorstellung erfolgreich umgesetzter Anwendungsfälle auch im Rahmen städtischer Veranstaltungen, um die Zusammenarbeit und das Verständnis zu fördern.
- Koordinierungsstelle UDP/UDZ sollte in der Verantwortung stehen, Daten in den Fachbereichen zu identifizieren und nutzbar zu machen. Dies kann durch persönliche Vernetzung, Beratung in IT-Projekten sowie interne Öffentlichkeitsarbeit realisiert werden.

5 Ausblick

Zukünftige Themen im Zusammenhang mit der Datenintegration bieten interessante Möglichkeiten, und es gibt mehrere Schwerpunkte, die weiter erforscht werden sollten. Zunächst versprechen die Fortschritte bei der Verwaltung von Indikatoren und Zeitreihen (vgl. 3.2) eine Verbesserung der Handhabung und Analyse komplexer zeitlicher Daten. Darüber hinaus wird die Entwicklung von Data Cubes, die für schnellen Zugriff und schnelle Analyse optimiert sind, die datengestützte Entscheidungsfindung revolutionieren. Verbesserte Datenarchivierungs-, Historisierungs- und Versionierungsstrategien werden die Datenintegrität im Laufe der Zeit sicherstellen. Die Automatisierung des gesamten Integrationsprozesses, wie kurz in Abschnitt 3.2.3 diskutiert, ist eine entscheidende Front, die Effizienzgewinne verspricht. Schließlich werden die Integration von Linked Data und die Anwendung der Prinzipien des Data Mesh zu einem stärker dezentralisierten und vernetzten Datenökosystem beitragen und neue Wege der Zusammenarbeit und Erkenntnisse in diesem sich ständig weiterentwickelnden Bereich eröffnen. Auch Hamburg, Leipzig und München arbeiten intensiv daran, ihre Datenintegrationsprozesse mit neuen und innovativen Ideen und Technologien zu erweitern und zu verbessern.

In Hamburg ist gerade ein Kundenportal in Arbeit, das den Datenintegrationsprozess für einfache Datensätze automatisieren soll und den Kunden die Möglichkeit gibt, nach einem angeleiteten Start auch eigenständig Daten zu integrieren oder zu aktualisieren. Darüber hinaus testet Hamburg immer auch neue Schnittstellen (z.B. OAF) und Formate (z.B. Zeitreihen), um diese mit anbieten zu können und damit die GDI/UDP ständig weiter auszubauen und aktuell zu halten.

In Leipzig wird im Oktober 2023 das Open-Data-Portal auf das Metadatenprofil DCAT-AP.de umgestellt. Neben der externen Anbindung an die Open-Data-Portale der übergeordneten Verwaltungsebenen wurde in diesem Zuge auch der Metadatenfluss innerhalb der Stadt Leipzig verbessert: Mit Hilfe einer neu entwickelten Schnittstelle (DCAT-Bridge) können als „Open Data“ klassifizierte Metadateneinträge aus dem internen GeoDatenkatalog der GDI Leipzig nun im Rahmen der Datenintegration automatisiert in den Katalog des Open-Data-Portals übertragen werden.

In München wird derzeit eine PoC Geodaten austauschplattform als Software as a Software (SaaS)-Lösung erprobt, die als zukünftige Arbeitsgrundlage für eine verbesserte Datenintegration im Rahmen des "Digitalen Zwillinges München (DZ-M)" bereitgestellt wird. Die Geodaten austauschplattform ermöglicht es, Daten aus dem DZ-M verschiedenen internen und externen Projektbeteiligten in unterschiedlichen Projekten und Kontexten räumlich und zeitlich begrenzt zur Verfügung zu stellen. Diese Daten sollen direkt in einer Viewing-Komponente verfügbar sein und über entsprechende Dienste bereitgestellt und nahtlos in Drittsysteme integriert wer-

den können, sowohl online als auch offline. Die Versionierung bei der Bereitstellung einer Version erfolgt automatisch. Perspektivisch sollen die Daten und Anwendungen der Geodaten austauschplattform in den Datenkatalog integriert werden.

6 Anhang

6.1 Anhang 1: Data Governance-Rollen im Kontext GDI/UDP-Datenmanagement in drei CUT-Städten

	Hamburg	Leipzig	München
Datenintegrationsrollen	<ul style="list-style-type: none"> - Datenmanager: Koordinieren die Datenbedarfe, kommunizieren mit Datenbedarfsträger und –halten, bereiten die Datenintegration vor (Freigabeerklärung, Metadateneintrag, Vorgaben für die Datenübermittlung, Visualisierung, etc.) - Datenengineers: setzen die Vorgaben um und erzeugen anhand der übermittelten Vorgaben die Dienste (ETL-Prozesse und Diensterzeugung) - IT-Dienstleister: nach Abnahme der Dienste in der Testumgebung durch den Kunden (Datenhalter) erfolgt die Produktivsetzung durch Dataport – Überführung des Dienstes ins Produktivsystem und Veröffentlichung des Datensatzes über die entsprechenden Portale <p>Vergleichbare Personen wie die GDI-L Verantwortlichen in den Behörden und Ämtern der FHH sind bisher nicht etabliert – hier soll zukünftig aber die „Data Governance Rollen“ ansprechbar sein die in allen Organisationseinheiten der FHH etabliert werden sollen (diese haben aber zunächst keinen direkten Bezug zur GDI-HH/UDP)</p>	<ul style="list-style-type: none"> - Die GDI Koordinierungsstelle dient für alle Ämter als erste Anlaufstelle für sämtliche Fragen zum Thema Geodaten und zur Koordinierung des stadtweiten Geodatenmanagements. - Datenmanager:Innen in der GDI-L: integrieren Datensätze, erzeugen Dienste - Administration Fachdatensystem: technische Betreuung einer Fachanwendung, stellt Schnittstellen bzw. Datenbankzugänge bereit = technische Ansprechpartner:Innen in der Datenintegration - GDI-L Verantwortliche: sind Ansprechpartner:Innen und Multiplikatoren der GDI-L in den Ämtern, Pflegen Metadaten, betreiben im Idealfall Datenmanagement in den Ämtern, Kennen die Daten und Ansprechpartner:Innen in den Ämtern 	<ul style="list-style-type: none"> - Das Koordinierungsbüro Geodateninfrastruktur dient für alle Referate als erste Anlaufstelle für sämtliche Fragen zum Thema Geodaten und zur Koordinierung des stadtweiten Geodatenmanagements. - Datenexpert:Innen: integrieren Datensätze, erzeugen Dienste, etc. - IT-Dienstleister: Komponentenverantwortliche – konfigurieren und administrieren die Server nebst Zugriffsrechte gemäß Ticket Zeitbasierter Transportprozess: stellt Daten aus einem Eingangsbereich (kann auch ein Fachverfahren sein) über einen QS-Bereich in den Publikationsbereich ein - Datenredakteur: fachlich Zuständiger für die Daten (Aktualität, Metadaten etc.)

Kunden der GDI/UDP	<ul style="list-style-type: none"> - Datenbedarfsträger: kommuniziert einen Datenbedarf und hat im Idealfall bereits den Datenhalter identifiziert - Datenhalter: verfügt über die Datenhoheit und darf diese veröffentlichen, über den Grad der Veröffentlichung entscheiden und Visualisierungen definieren. Langfristig für die Pflege / Aktualisierung der Daten und Metadaten zuständig. 	<ul style="list-style-type: none"> - Endanwender:Innen (Kunden): nutzen Daten direkt oder über Anwendungen - Datenhalter/-eigentümer: erheben/kaufen/pflegen Daten (Ämter), können den Datensatz fachlich einschätzen = fachliche Ansprechpartner:Innen in der Datenintegration 	<ul style="list-style-type: none"> - Datennutzer:Innen (Kunden): je nach Zugriffsberechtigung können die Daten direkt oder über Anwendungen genutzt werden. - Dateneigentümer:Innen: Verfügt über die Datenkontrolle und ist befugt, diese zu veröffentlichen, den Umfang der Veröffentlichung festzulegen und Visualisierungen zu definieren.
--------------------	---	---	--

6.2 Anhang 2: Freigabeerklärung der FHH

- *Datenfreigabe verwaltungsintern:*
http://geoportal-hamburg.de/urbandataplattform/Freigabeerklaerung_FHH.pdf
- *Datenfreigabe extern:* http://geoportal-hamburg.de/urbandataplattform/Freigabeerklaerung_Extern.pdf

6.3 Anhang 3: Formular Metadatenerfassung

6.3.1 UDP FHH - Handzettel Metadatensatz erfassen

http://geoportal-hamburg.de/urbandataplattform/Handzettel_Metadatensatz_erfassen.pdf

6.3.2 GDI München - Metadatenkatalog Formular

Grundinformationen

SCHEMA Tabelle	<Schema auswählen>	
ÜBERBLICK	Kurzbeschreibung Typ Genauigkeit	<Geometrietyp auswählen> <Lagegenauigkeit auswählen>
Bemerkung:		
VERANTWORTLICHKEIT	Fachl. Zuständigkeit Externe Datenerheber	Wenn ja. Verantwortlicher:
AKTUALITÄT	Update-Zyklus	<Update-Zyklus auswählen>
Stadtarchiv <small>Bewertung durch Stadtarchiv</small>	Archivwürdigkeit	BEGRÜNDUNG

Nutzungshinweise

SCHEMA Tabelle	-<Schema auswählen> 0		
LHM-Intranet	Fachverfahren	<input checked="" type="checkbox"/> alle Zugriffuser (inkl. GeoInfoWeb) [Standard] bestimmte Zugriffuser	
	GeoInfoWeb	<input checked="" type="checkbox"/> alle Nutzer [Standard] bestimmte Organisationseinheiten bestimmte (übergreifende) Gruppe(n)	
	Internet und Datenabgabe	Veröffentlichungen / Präsentationen im Aufgabenbereich der LHM	
		<input type="checkbox"/> als Backend-Daten ohne Präsentation	
		<input checked="" type="checkbox"/> für Präsentationen [Standard]	
Datenabgabe an externe Auftragnehmer		<input checked="" type="checkbox"/> Weitergabe von Daten oder Diensten an Dritte, die im Auftrag der Stadt tätig werden [Standard] Verpflichtungserklärung des Dritten notwendig	
Datenabgabe an Externe, die nicht Auftragnehmer sind		<input type="checkbox"/> im Rahmen der Nutzungsbedingungen Bekanntgabe der Nutzungsbedingungen für den Dritten notwendig	
		<input type="checkbox"/> über den Rahmen der Nutzungsbedingungen hinaus Nutzungsvereinbarung mit Dritten notwendig	
Open Data		<input type="checkbox"/> alle Attribute und Inhalte nutzbar Nutzungsrechte open data	
<input type="checkbox"/> nur bestimmte Attribute und Inhalte nutzbar Nutzungsrechte open data			
weitere Hinweise zur Nutzung	Stichwort	Hinweise	

Attribute

ATTRIBUTE	DATENTYP (Oracle)	BEDINGUNG	BESCHREIBUNG	BEMERKUNG

alter DATENTYP	DATENTYP ArcGIS	DATENTYP ORACLE
Objekt-ID	Objekt-ID	number
Geometrie	Geometry	sdo_geometry
Text (9)	Text	nvarchar2
Text (255)	Text	nvarchar2
Float	Float	float
short	short	number
Datum	DATE	date oder timestamp(6)
BLOB	BLOB	blob

Änderungshistorie

ÄNDERUNG	GEAENDERT_AM	GEAENDERT_VON

6.3.3 Geoportal München – Metadatenkatalog Formular

Berücksichtigung der Profilvergaben der GDI-BY unter dem Aspekt einer möglichst einfachen Erfassung für datenhaltende Stellen.



Formular zur Übergabe von Metadaten

Organisation

Referat

Abteilung / Unterabteilung

Sachbearbeiter

Für interne Bearbeitung

Bitte geben Sie das Kürzel der Abteilung an

Für Rückfragen von GDI

Vertrieb

Titel

(Neben einer zutreffenden fachlichen Benennung der Ressource kann der Titel auch den Geltungsbereich beidlich beinhalten)

Kurzbeschreibung

Bitte beschreiben Sie kurz die Metadaten

Schlüsselwörter

(Schlüsselwörter (oder Schlagwörter) stellen eines der zentralen Metadatenlemente dar. Hierbei können Sie die Auffindbarkeit Ihrer Georesourcen maßgeblich beeinflussen! Keyword „Opendata“ wenn zutreffend. Mindestens ein Schlüsselwort muss „Landeshauptstadt München“ lauten

Thematik

Wichtig! Thematische Klassifizierung von geographischen Daten

6.4 Anhang 4: Dokumentation Datenintegrationsprozess

6.4.1 Beispiel Leipzig: Grafik zum Datenintegrationsprozess im Intranet

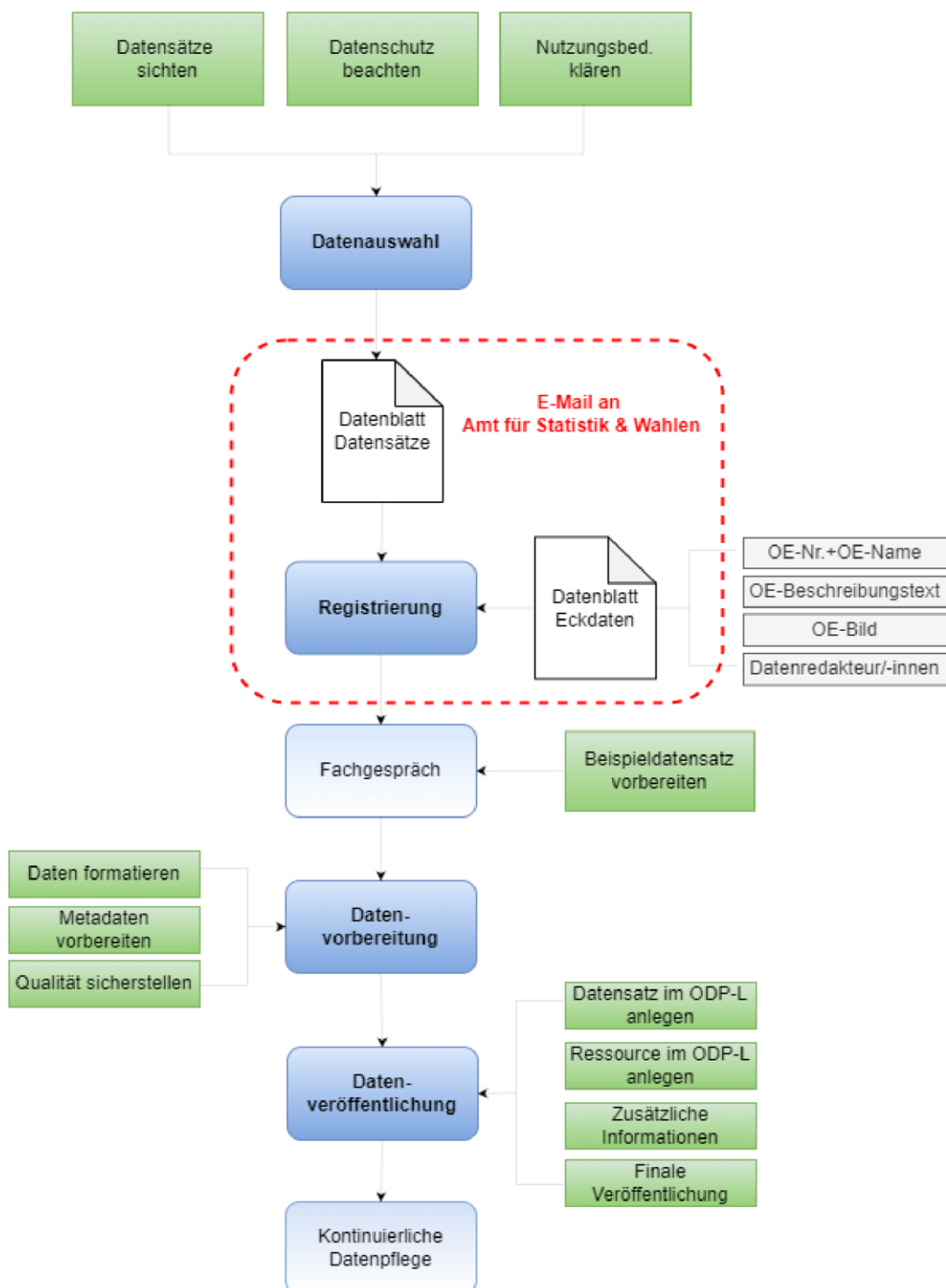


Abbildung 3: Darstellung des Prozesses der Datenveröffentlichung im ODP-L im Intranet (Quelle: Leitfaden Open Data Portal Leipzig, abgerufen am 2.12.2022)

6.4.2 Beispiel Hamburg: Handzettel Datenintegration

http://geoportal-hamburg.de/urbandataplattform/Handzettel_Datenintegration.pdf