

Evolutionenfähiges Post-COVID Datenmodell

BMI, 17. September 2024 in Berlin

AGENDA

1. Wer sind wir?
2. Stakeholder- und Anforderungsanalyse
3. Herausforderungen
4. Das ontologiebasierte Datenmodell
5. Minimal Viable Product
6. Ausblick auf Stufe 2

WER WIR SIND

Edmund Fröhlich,
GF DESAM, Berlin

Deutsche Stiftung für Allgemeinmedizin und Familienmedizin

Gegründet 1973, Forschungs- und Nachwuchsförderung in der Allgemeinmedizin

Vorsitzender Prof. Dr. Ferdinand M. Gerlach

(2012 – 2023 Vorsitzender des Sachverständigenrats zur Begutachtung der Entwicklung im Gesundheitswesen)

als Konsortialführer in Kooperation mit:

- Universitätsklinikum Tübingen, Institut für Allgemeinmedizin, Zentrum für innovative Versorgung, Medic
- Universitätsklinikum Kiel, Institut für Allgemeinmedizin
- Universitätsklinikum Würzburg, Institut für Allgemeinmedizin
- Universität Tübingen, Lehrstuhl Methoden der Medizininformatik
- Hochschule Reutlingen, Reutlingen Research Institute

KOMPETENZ- NETZWERK

Edmund Fröhlich
GF DESAM, Berlin

DESAM-Stiftungsrat, u. a.:

- Thomas Ballast, stellv. Vorsitzender der Techniker Krankenkasse
- Denis Nößler, Chefredakteur der Ärztezeitung
- Dr. med. Klaus Reinhardt, Präsident der Bundesärztekammer
- Dr. med. Annette Rommel, 1. Vorsitzende Kassenärztliche Vereinigung Thüringen
- Prof. Dr. med. Martin Scherer, Direktor des Instituts und der Poliklinik für Allgemeinmedizin und Präsident der DEGAM

Gemeinsame Bürogemeinschaft in Berlin-Mitte:

- Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM)
- Deutsche Gesellschaft für Sozialmedizin und Prävention (DGSMP)
- Netzwerk evidenzbasierte Medizin
- Stiftung Praxissiegel



DESAM ForNet

Edmund Fröhlich
GF DESAM, Berlin



Koordinierung der Initiative Deutscher Forschungspraxennetze

- Gefördert seit 02/2020 durch BMBF
- Ziel ist der Aufbau einer bundesweiten Forschungsinfrastruktur in der hausärztlichen Primärversorgung
- Beteiligt sind aktuell 32 allgemeinmedizinische Universitäts-Institute und ca. 1.400 Praxen
- Leitung Dr. med. Leonor Heinz

Konsortialpartner

Technologie- und Methodenplattform für die vernetzte medizinische Forschung (TMF)

DESAM-Konsortium

Antragstellerin:



**Deutsche Stiftung für
Allgemeinmedizin und
Familienmedizin
(DESAM)**

Edmund Fröhlich (Geschäftsführung)

Beteiligte Institutionen

(DESAM-Unterauftragnehmer):

Universität Tübingen



Institut für
Allgemeinmedizin
und Interprofessionelle
Versorgung
Universitätsklinikum
Tübingen

Prof. Dr. med. Stefanie Joos (Leitung)
Dr. med. Roland Koch (Projektleitung)
Andreas Polanc, M. Sc. PH

Universität Würzburg



**Uniklinikum
Würzburg** UK
Institut für Allgemeinmedizin

Prof. Dr. med. Ildikó Gágyor (Leitung)
Christian Kretschmann (Projektleitung)

Universität Kiel



C | A | U
Christian-Albrechts-Universität zu Kiel
Medizinische Fakultät
Institut für Allgemeinmedizin

Prof. Dr.med. Hanna Kaduszkiewicz (Leitung)
Dr. med. Martin Williamson (Projektleitung)

Hochschule Reutlingen



RRI
Reutlingen Research Institute

Prof. Dr. rer. medic Christian Thies (Leitung)
Arndt Becker
Patrick Schmutz
Philip Stolz

Universitätsklinikum Tübingen



DIZ MEDIZIN
INFORMATIK
INITIATIVE
DATENINTEGRATIONSZENTRUM

Prof. Dr. Carsten Eickhoff (Leitung)
Dr. Michaela Hardt (Projektleitung)
Dr. Dieter Weichart
Benjamin Sailer

Universität Tübingen



ZiV Zentrum
für innovative
Versorgung

Prof. Dr. rer. nat. Nico Pfeifer (Leitung)
Marc Mausch (Projektleitung)

Abschlussbericht

Post Covid Datenmodell
Stufe 1

Stakeholder gruppiert

- Von der Erkrankung betroffene
- Interessierte Öffentlichkeit inkl. Presse / Medien
- Professionalisierte Forschung inkl. Fördergeber
- Nicht-Professionalisierte Forschung / Citizen Science
- Leistungserbringer
- Kostenträger
- u.v.m.
=> liefern/nutzen Daten zu
Therapie, Erkenntnisgewinn, Entscheidungsfindung,...

Insgesamt 71 Stakeholder
In 11 Gruppen identifiziert
Mit fast 200 Fragestellungen

Datenarten (Beispiele)

Kleines Blutbild

Leukozyten	⊖	5	72	Mio/μl
Erythrozyten	⊖	5	5	Mio/μl
MCH	⊖	29	0	pg
MCHC	⊖	33	9	g/dl
MCV	⊖	86	1	fl
Hämoglobin	⊖	16	1	g/dl
Hämatokrit	⊖	47	5	%
Retikulozyten	⊖	10	0	%
Thrombozyten	⊖	209	0	Tsd./μl

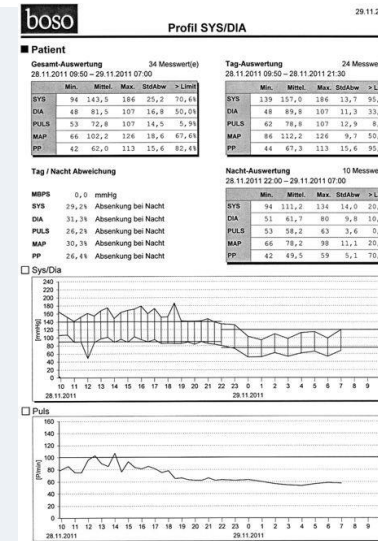
Untersuchungsbefund / Beschwerden: *perf. Aussehen in*
Leber 13,5 g/l in 1,6 d in 1,6 d in 1,6 d
in 1,6 d in 1,6 d in 1,6 d in 1,6 d
in 1,6 d in 1,6 d in 1,6 d in 1,6 d
in 1,6 d in 1,6 d in 1,6 d in 1,6 d

Diagnose: *unbekanntes Abdomen, DD: G-*

Empfehlung / Therapie: *keine weitere Maßnahmen*
keine weitere Maßnahmen
keine weitere Maßnahmen

Mit freundlichen kollegialen Grüßen

[Signature]



da hapert's bei uns. Ich denke
 hapert's genauso dran. Also
 nicht. Nee, ich glaube, es ist
 die Zeit dafür nicht nimmt, (m
 Die Patienten würden's verst
 glaube ich.

125 I: Gibt es denn irgendwas, wa

126 B: Ja, ich weiß schon, diese
 Ich glaube, wenn man einfach
 Kollege gemacht, und der wa
 persönliche Schiene bei Ärz
 hospitiert und das war super,
 wichtig" Ich glaub, sowas mu
 wir machen gerade so ein Ä
 das ist eine ganz wichtige Sa
 Zeit, und ich weiß nicht wofür

	C	D	E	F	G
			Innere Medizin	17.05.-21.05.2021	
			Allgemeinmedizin	12.04.-23.04.2021	
			Chirurgie	26.04.-30.04.2021	
			Gynäkologie	03.05.-07.05.2021	8
			Gynäkologie	07.06.-11.06.2021	4
				12	
Lina	1105160	stu123456@mail.uni-kiel.	x		
Gabriela	1105161	stu123457@mail.uni-kiel.	x	x	07.06.-11.06.2021
Leonie	1105162	stu123458@mail.uni-kiel.	x	x	03.05.-07.05.2021
Gothia	1105163	stu123456@mail.uni-kiel.	x		
Denise	1105164	stu123457@mail.uni-kiel.	x,Ersatz		x,Ersatz
Jann	1105165	stu123458@mail.uni-kiel.de			
Enno	1105166	stu123456@mail.uni-kiel.de			
Charlotte	1105167	stu123457@mail.uni-kiel.	x,Ersatz	x	03.05.-07.05.2021
Kim	1105168	stu123458@mail.uni-kiel.de			x,Ersatz
Maren	1105169	stu123456@mail.uni-kiel.de			
Maximilian	1105170	stu123457@mail.uni-kiel.de			
Benita	1105171	stu123458@mail.uni-kiel.	x	x	03.05.-07.05.2021

Heterogene Speicherung ähnlicher Daten

Praxisverwaltungssystem (PVS) A

PatID	Erste Impfung	Zweite Impfung	COVID Erkrankt	...
1	20210501	20220201	Ja	...
2	20210501	20220201	Nein	...

Studie B

Proband Nr.	Impfdatum	Impf Nr.	...
bgH65Cx	01.05.2021	1	...
bgH65Cx	01.02.2022	2	...

Proband Nr.	PCR Testdatum	PCR Test	...
bgH65Cx	01.05.2021	Pos	...
avF47Kq	01.02.2022	neg	...

Impfregister C

```
<Person ID="1234">  
  <Datum Impfung lfdNr="1">  
    2021-05-01-T-00:00:00.000  
  <\Datum Impfung >  
  <Datum Impfung lfdNr="2">  
    2021-05-01-T-00:00:00.000  
  <\Datum Impfung>  
  <Covid Erkrankung>  
    Ja  
  <\Covid Erkrankung>  
  ...  
<\Person>
```

Unterschiedliche Elementbeziehungen

Heterogene Speicherung ähnlicher Daten

Praxisverwaltungssystem (PVS) A

PatID	Erste Impfung	Zweite Impfung	COVID Erkrankt	...
1	20210501	20220201	Ja	...
2	20210501	20220201	Nein	...

Studie B

Proband Nr.	Impfdatum	Impf Nr.	...
bgH65Cx	01.05.2021	1	...
bgH65Cx	01.02.2022	2	...

Proband Nr.	PCR Testdatum	PCR Test	...
bgH65Cx	01.05.2021	Pos	...
avF47Kq	01.02.2022	neg	...

Impfregister C

```
<Person ID="1234">
  <Datum Impfung lfdNr="1">
    2021-05-01-T-00:00:00.000
  <\Datum Impfung >
  <Datum Impfung lfdNr="2">
    2021-05-01-T-00:00:00.000
  <\Datum Impfung >
  <Covid Erkrankung>
    Ja
  <\Covid Erkrankung>
  ...
<\Person>
```

Unterschiedliche Bezeichner
z.B. auch bei Werten
-1,0,1 ⇔ schlechter, gleich, besser

Heterogene Speicherung ähnlicher Daten

Praxisverwaltungssystem (PVS) A

PatID	Erste Impfung	Zweite Impfung	COVID Erkrankt	...
1	20210501	20220201	Ja	...
2	20210501	20220201	Nein	...

Studie B

Proband Nr.	Impfdatum	Impf Nr.	...
bgH65Cx	01.05.2021	1	...
bgH65Cx	01.02.2022	2	...

Proband Nr.	PCR Testdatum	PCR Test	...
bgH65Cx	01.05.2021	Pos	...
avF47Kq	01.02.2022	neg	...

Impfregister C

```
<Person ID="1234">
  <Datum Impfung lfdNr="1">
    2021-05-01-T-00:00:00.000
  <\Datum Impfung >
  <Datum Impfung lfdNr="2">
    2021-05-01-T-00:00:00.000
  <\Datum Impfung >
  <Covid Erkrankung>
    Ja
  <\Covid Erkrankung>
  ...
<\Person>
```

Unterschiedliche
Datenformate

Heterogene Speicherung ähnlicher Daten

Praxisverwaltungssystem (PVS) A

PatID	Erste Impfung	Zweite Impfung	COVID Erkrankt	...
1	20210501	20220201	Ja	...
2	20210501	20220201	Nein	...

Studie B

Proband Nr.	Impfdatum	Impf Nr.	...
bgH65Cx	01.05.2021	1	...
bgH65Cx	01.02.2022	2	...

Proband Nr.	PCR Testdatum	PCR Test	...
bgH65Cx	01.05.2021	Pos	...
avF47Kq	01.02.2022	neg	...

Impfregister C

```
<Person ID="1234">
  <Datum Impfung lfdNr="1">
    2021-05-01-T-00:00:00.000
  <\Datum Impfung >
  <Datum Impfung lfdNr="2">
    2021-05-01-T-00:00:00.000
  <\Datum Impfung>
  <Covid Erkrankung>
    Ja
  <\Covid Erkrankung>
  ...
<\Person>
```

Bedeutungs-
Unschärfe

Heterogene Speicherung ähnlicher Daten

Praxisverwaltungssystem (PVS) A

PatID	Erste Impfung	Zweite Impfung	COVID Erkrankt	...
1	20210501	20220201	Ja	...
2	20210501	20220201	Nein	...

Impfregister C

```
<Person ID="1234">  
  <Datum Impfung lfdNr="1">  
    2021-05-01-T-00:00:00.000  
  <\Datum Impfung >  
  <Datum Impfung lfdNr="2">  
    2021-05-01-T-00:00:00.000  
  <\Datum Impfung>  
  <Covid Erkrankung>  
    Ja  
  <\Covid Erkrankung>  
  ...  
<\Person>
```

Studie B

Proband Nr.	Impfdatum	Impf Nr.	Depression	...
bgH65Cx	01.05.2021	1	Ja	...
bgH65Cx	01.02.2022	2	Nein	...

Proband Nr.	PCR Testdatum	PCR Test	...
bgH65Cx	01.05.2021	Pos	...
avF47Kq	01.02.2022	neg	...

Erweiterungen im Zuge
neuer Erkenntnisse zu COVID19

Heterogene Speicherung ähnlicher Daten

Praxisverwaltungssystem (PVS) A

PatID	Erste Impfung	Zweite Impfung	COVID Erkrankt	...
1	20210501	20220201	Ja	...
2	20210501	20220201	Nein	...

Jede Quelle hat eigenes
anwendungsspezifisches
Datenmodell (DM) dass sich
ändern kann.

Studie B

Proband Nr.	Impfdatum	Impf Nr.	Depression	...
bgH65Cx	01.05.2021	1	Ja	...
bgH65Cx	01.02.2022	2	Nein	...

Proband Nr.	PCR Testdatum	PCR Test	...
bgH65Cx	01.05.2021	Pos	...
avF47Kq	01.02.2022	neg	...

Impfregister C

```
<Person ID="1234">  
  <Datum Impfung lfdNr="1">  
    2021-05-01-T-00:00:00.000  
  <\Datum Impfung >  
  <Datum Impfung lfdNr="2">  
    2021-05-01-T-00:00:00.000  
  <\Datum Impfung>  
  <Covid Erkrankung>  
    Ja  
  <\Covid Erkrankung>  
  ...  
<\Person>
```

Abschlussbericht

Post Covid Datenmodell
Stufe 1

Anforderungen an Post-COVID DM

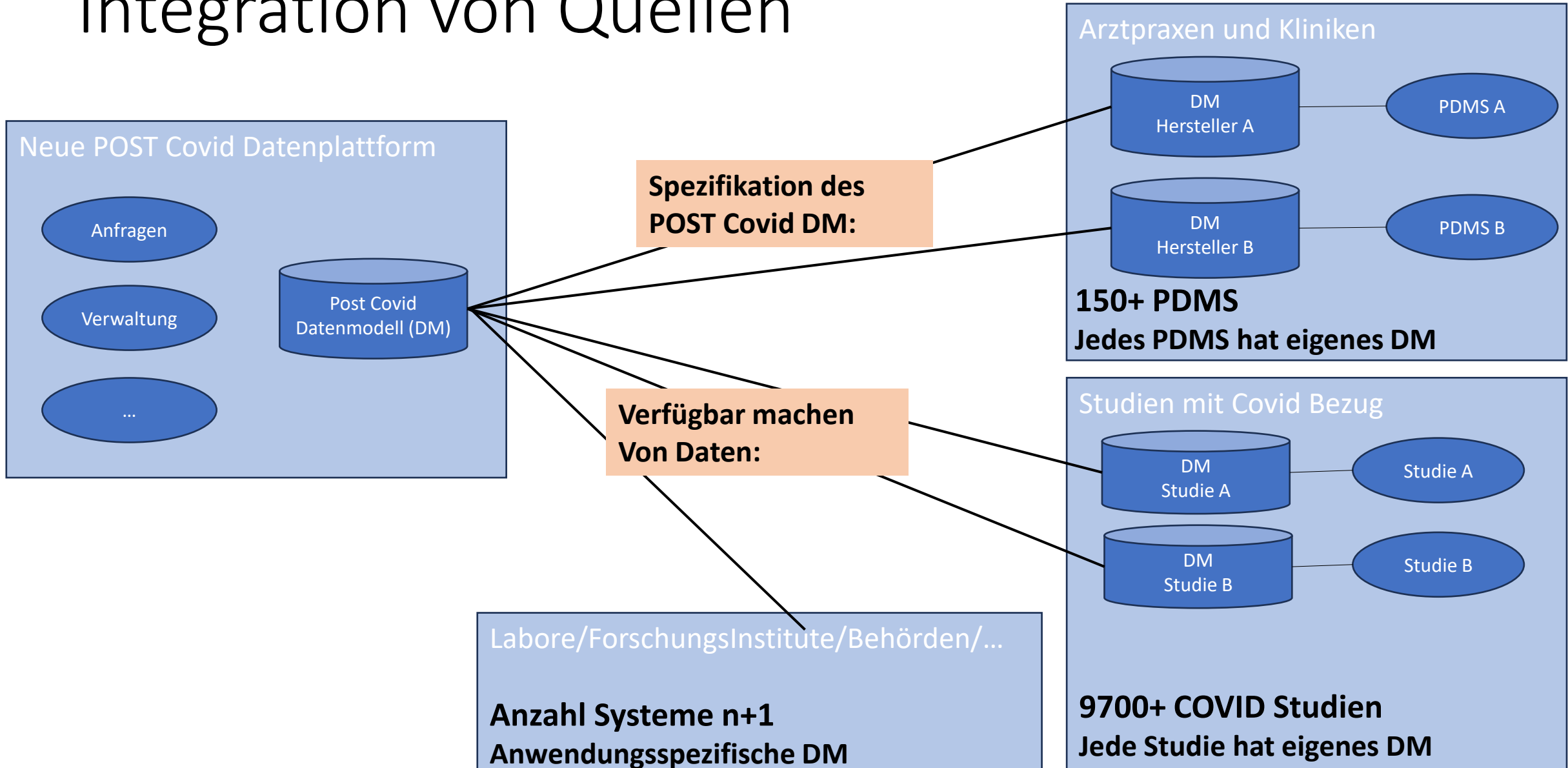
Formale Modelldefinition

- Harmonisierung von Information mit Post-COVID Bezug
 - Einheitliche Bezeichnungen
 - Beziehungen der Datenelemente
- Evolutionsfähigkeit wenn Informationsgrundlage sich ändert

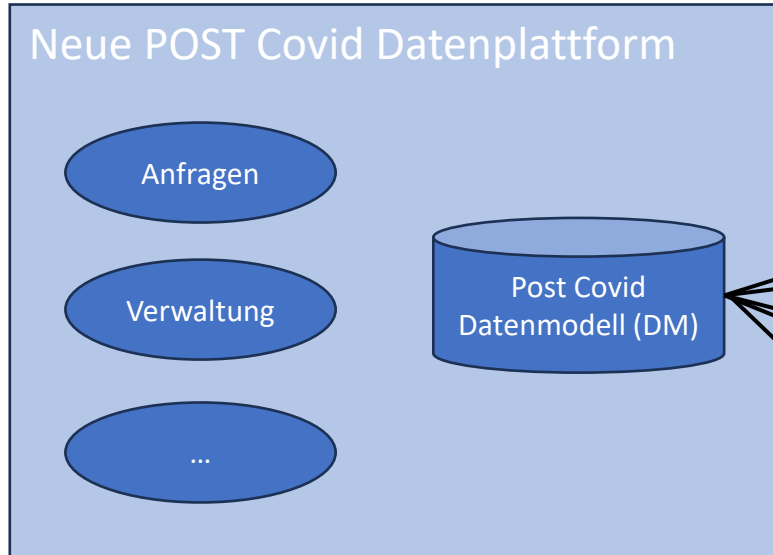
Praktische Umsetzung

- Konkrete Spezifikation mit allen - Post-COVID relevanten - Datenelementen
- Import aus heterogenen Datenquellen ermöglichen
- Anfragen an gespeicherte Daten einheitlich ermöglichen

Integration von Quellen



Integration von Quellen

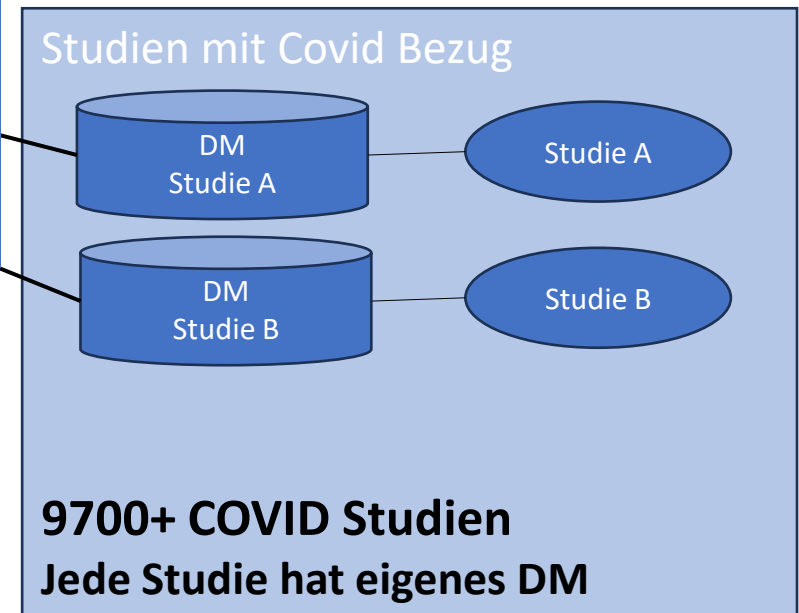
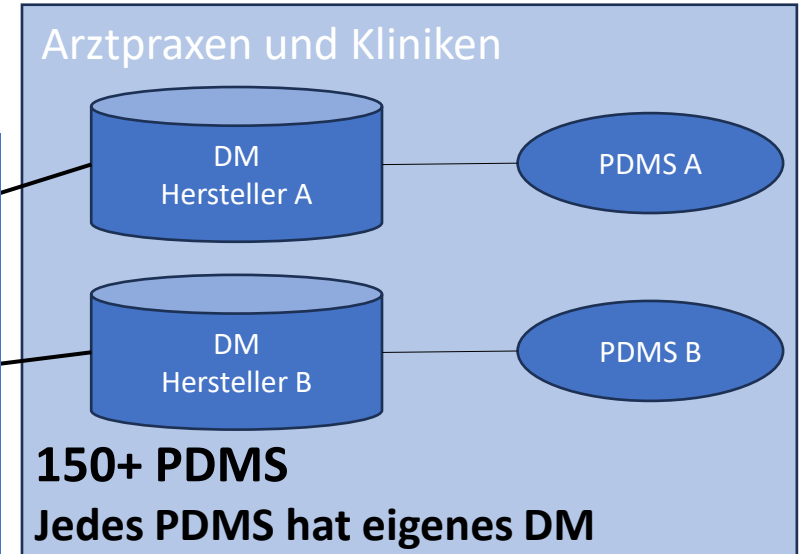


Auf welches PDM-Element muss ein Element aus externem DM abgebildet werden ?

Wie können die Daten aus fremden DM technisch importiert werden ?

Labore/Forschungsinstitute/Behörden/...

Anzahl Systeme n+1
Anwendungsspezifische DM



Abschlussbericht

Post Covid Datenmodell
Stufe 1

Evolutionenfähiges Post-COVID Datenmodell

Im Verlauf notwendige Anpassungen und Erweiterungen des Informationsmodells bedeuten

Änderung:

- Modellspezifikation
- implementierte Werkzeuge

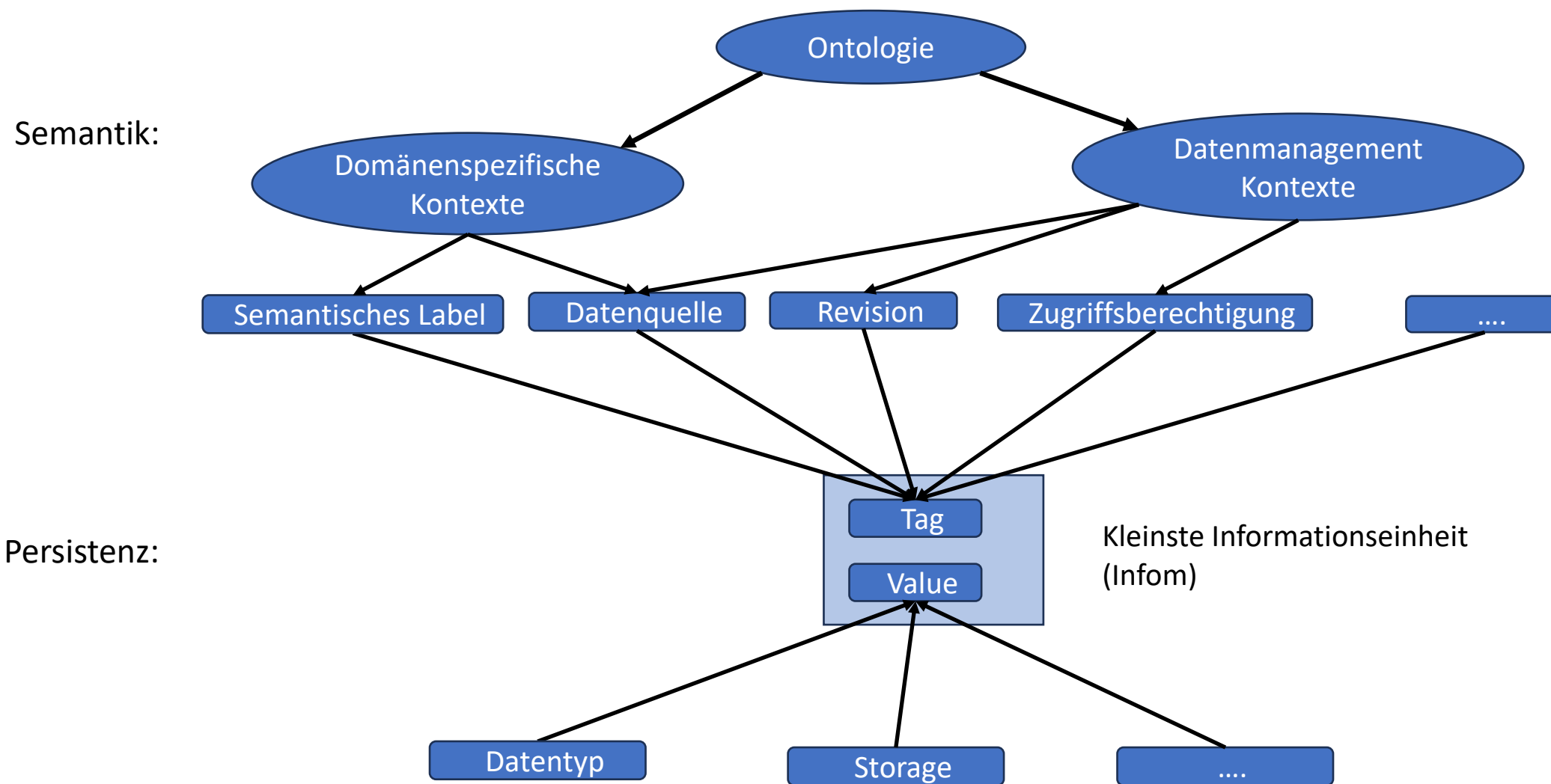
Bedeutet Änderung:

- Spezifikation der Datenelementen **und** deren Beziehung
- Datenspeicherung

=> Anforderung

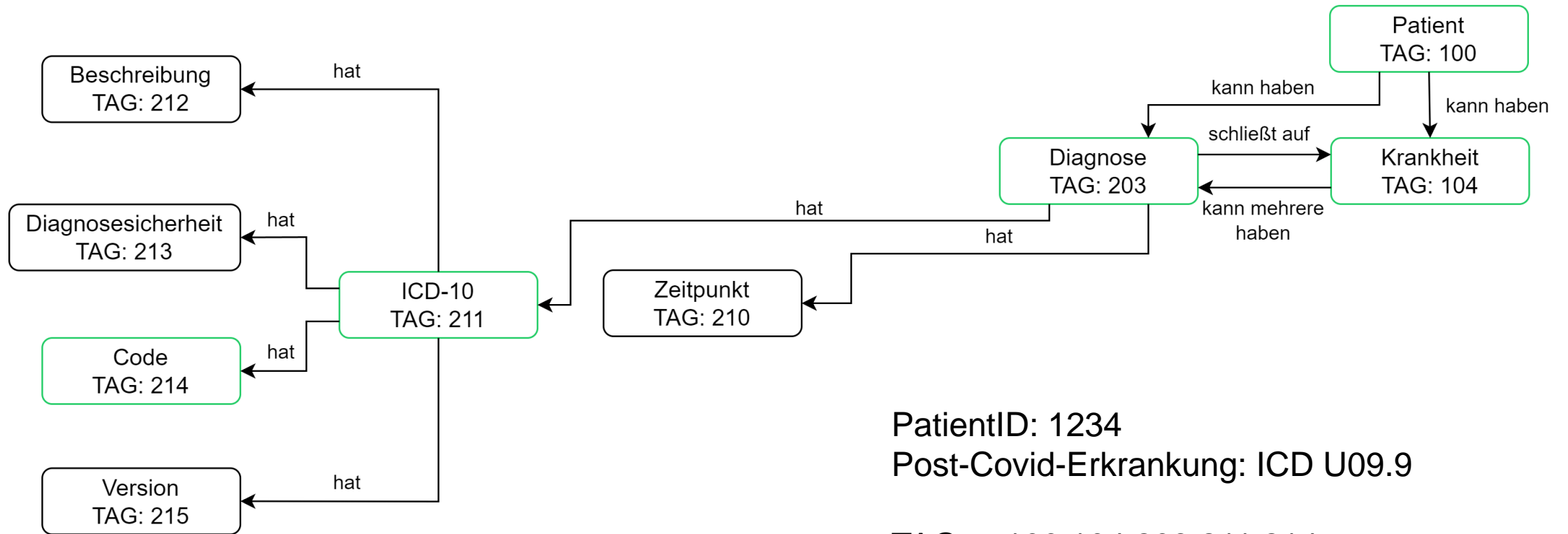
- im laufenden Betrieb möglich **und** schrittweise Modifikation der Spezifikation

Basis für lernfähiges Datenmodell: „Infome“



Basis für lernfähiges Datenmodell: Ontologie

Identifikation von Personen mit Post Covid: Diagnose; Post Covid ja oder nein



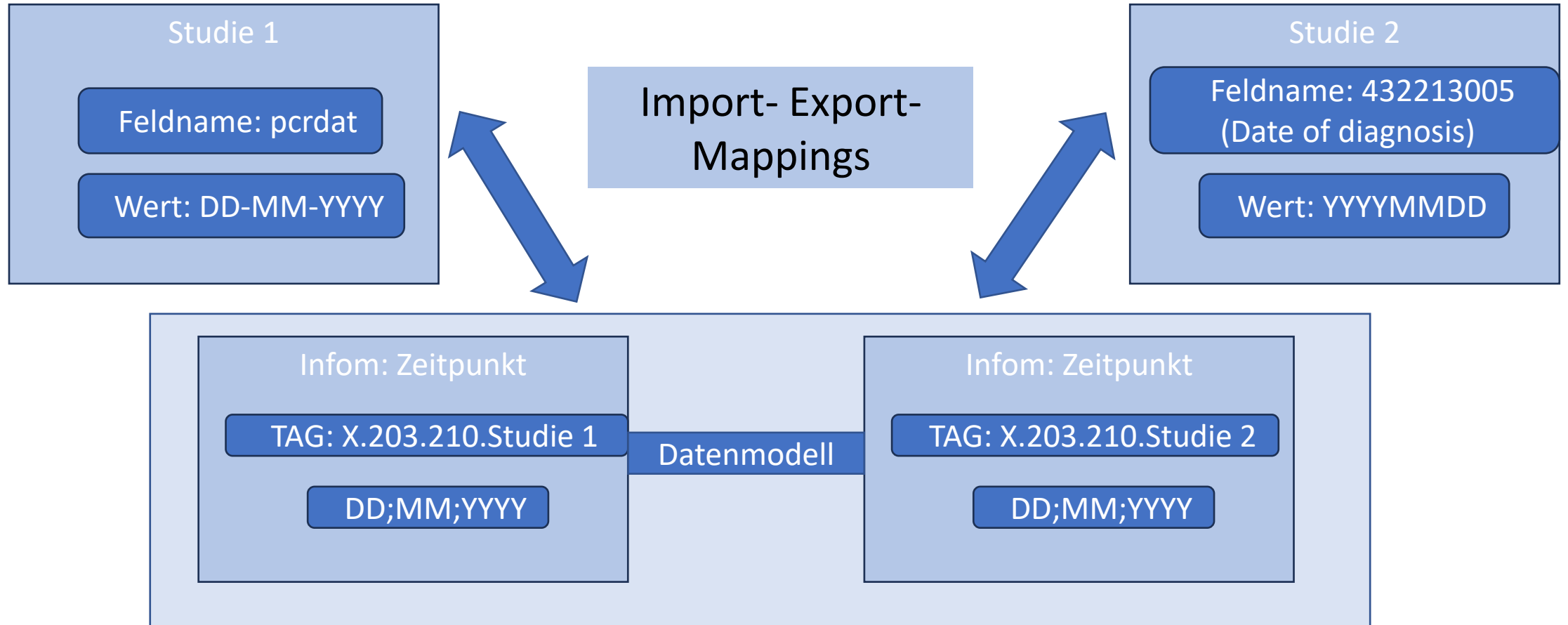
PatientID: 1234

Post-Covid-Erkrankung: ICD U09.9

TAG: 100.104.203.211.214

Value: 1234.X.X.X.U09.9

Mappings „Erstdiagnose“ aus 2 Datenbanken in ein Infom



Abschlussbericht

Post Covid Datenmodell
Stufe 1

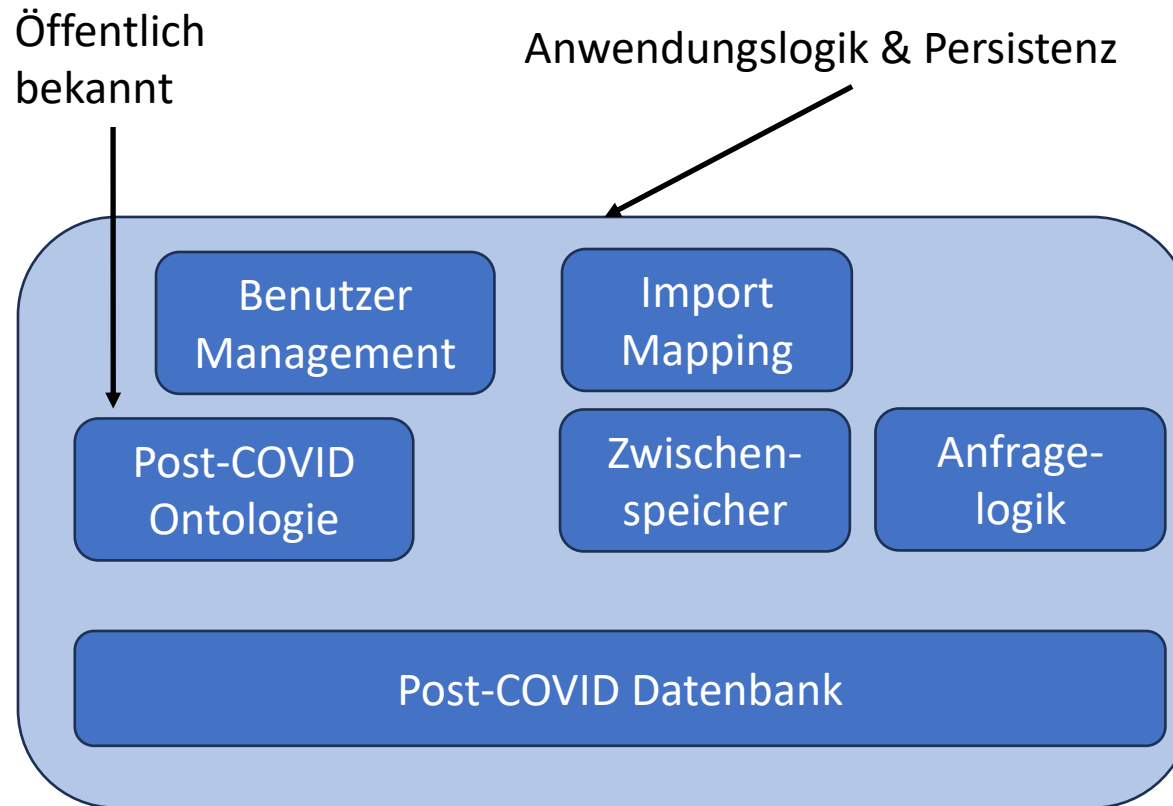
Anwendung des Mappings

- Ontologie vorausgesetzt
- Quellspezifisch definieren
- Integration in externe Prozesse
 - Einwilligungsmanagement
 - Record Linkage
 - Abbilden von Datenschutzoptionen im Datenmodell, etc.
z.B. Anonymisierung durch Auftrennen von Termen
 - usw.

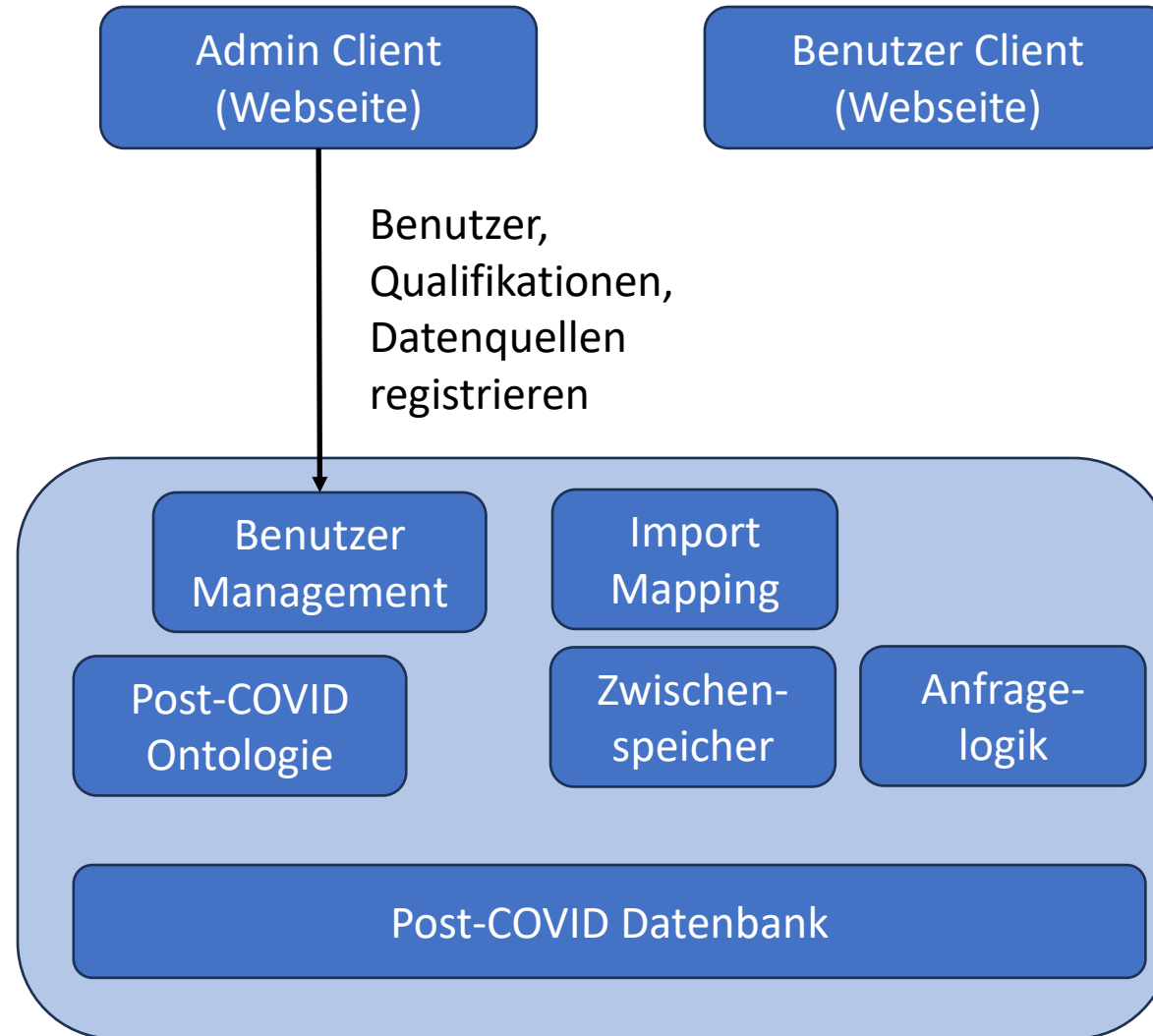
=> Einheitliche Anfragen über Tags möglich

=> Governance Strukturen erforderlich

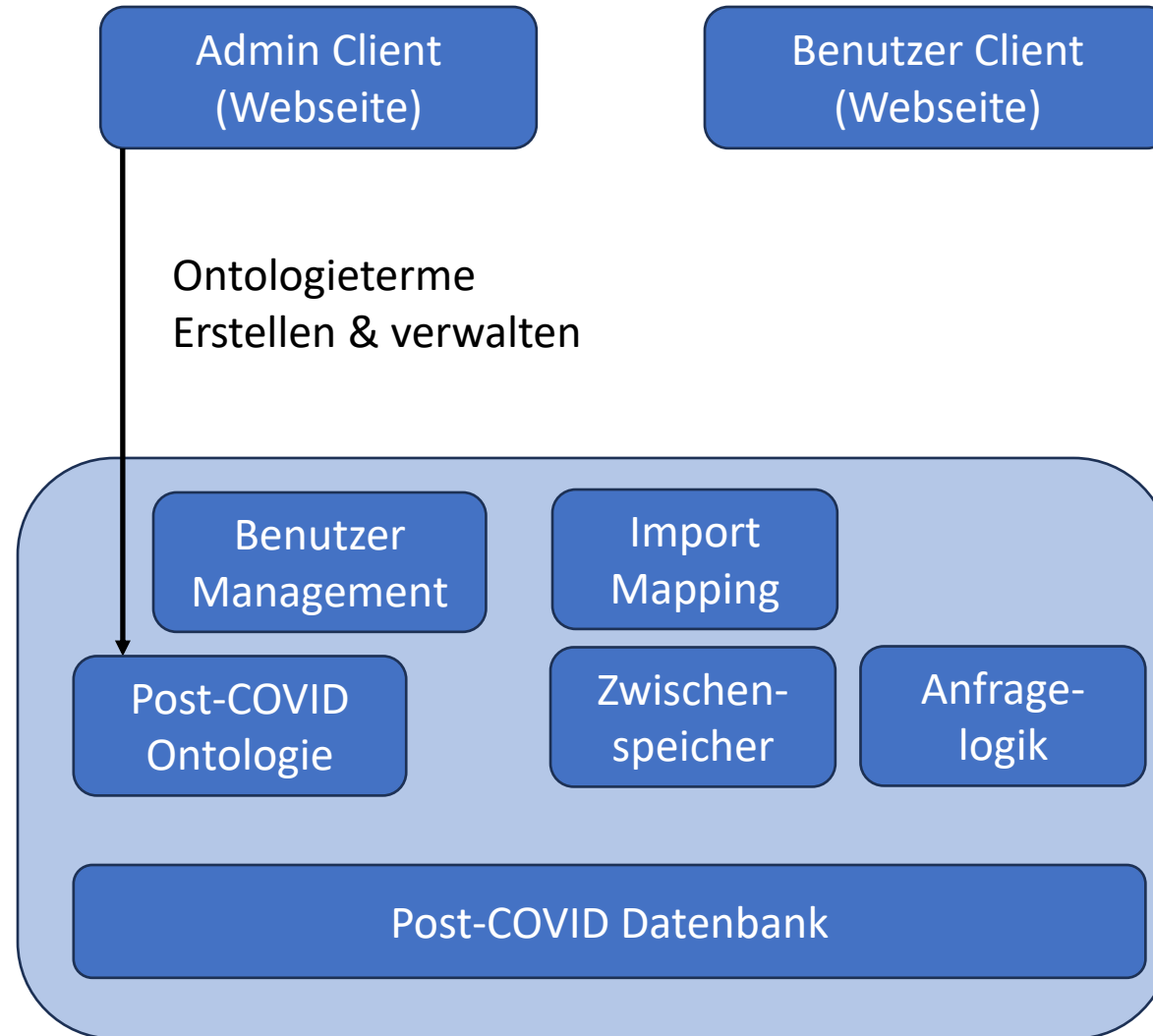
System- komponenten und MVP



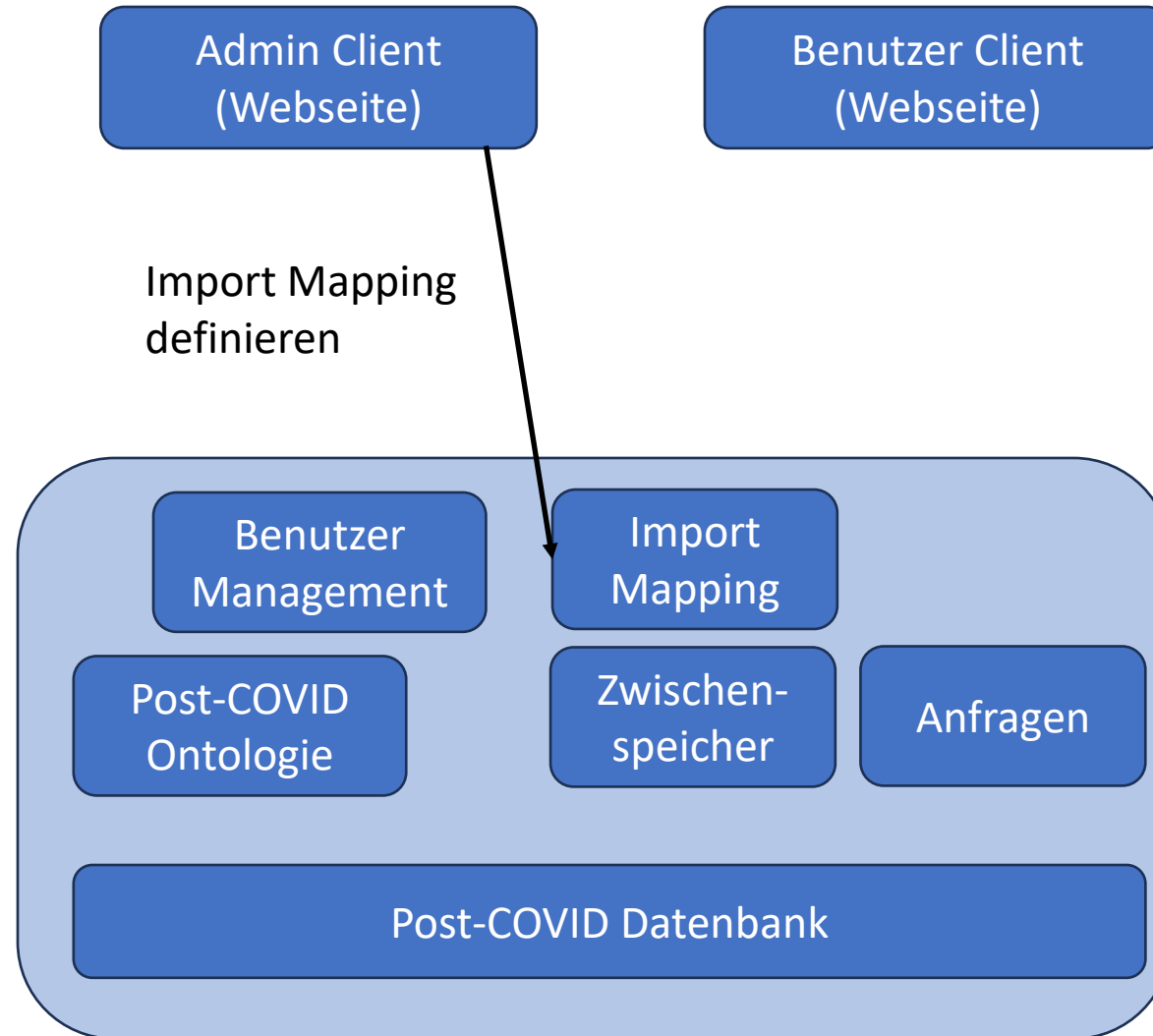
Anwender- Prozesse MVP 1



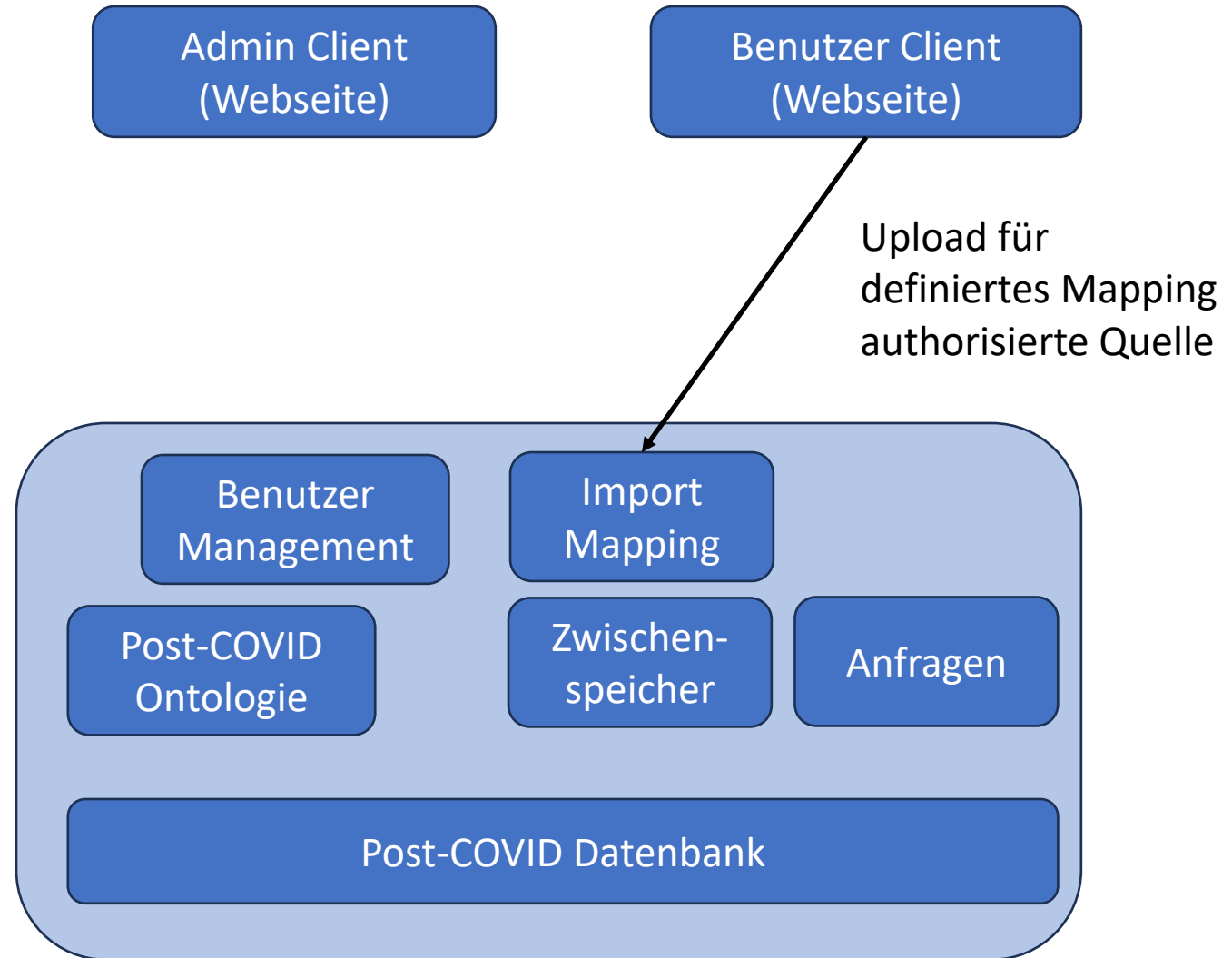
Anwender- Prozesse MVP 2



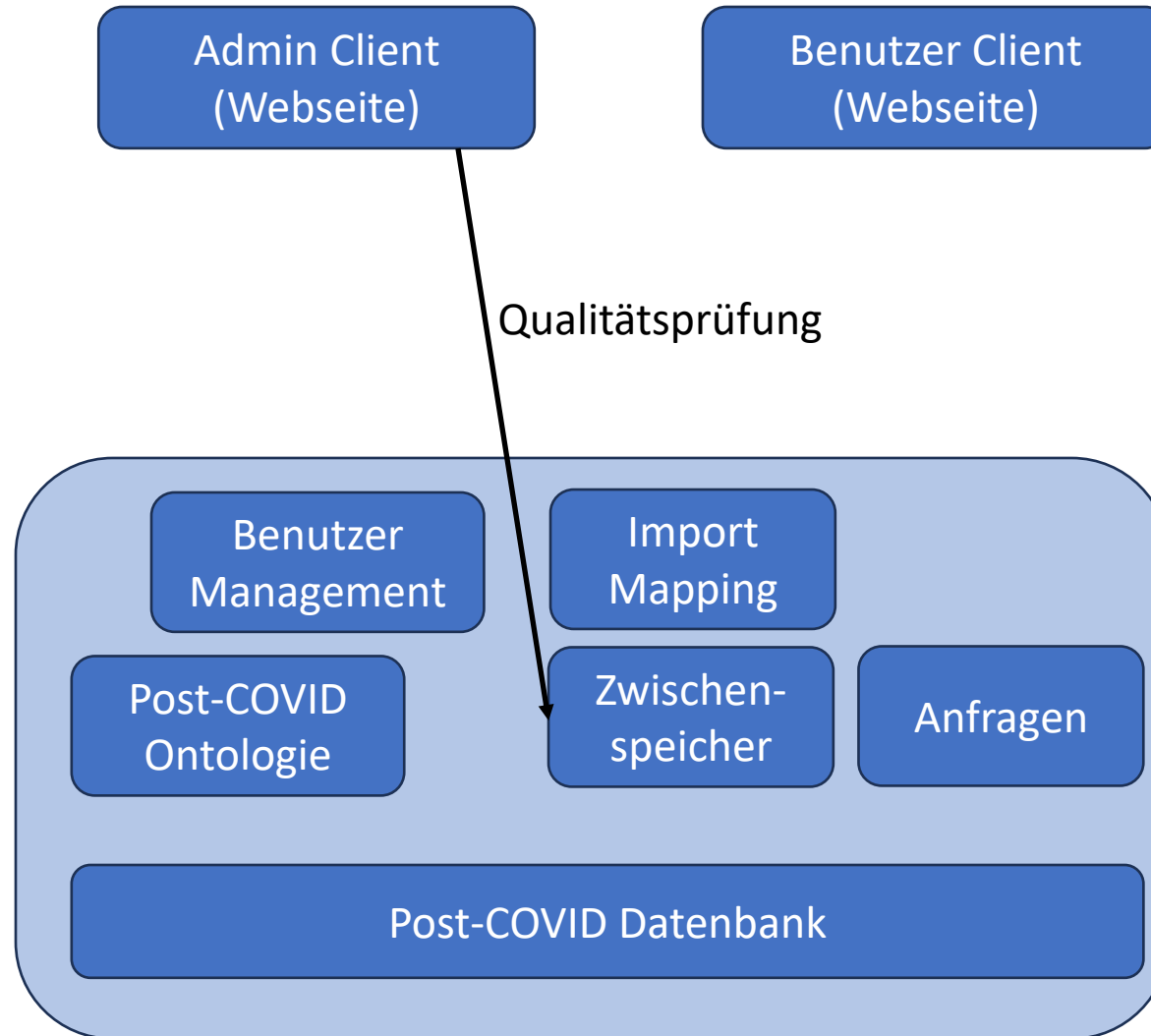
Anwender- Prozesse MVP 3



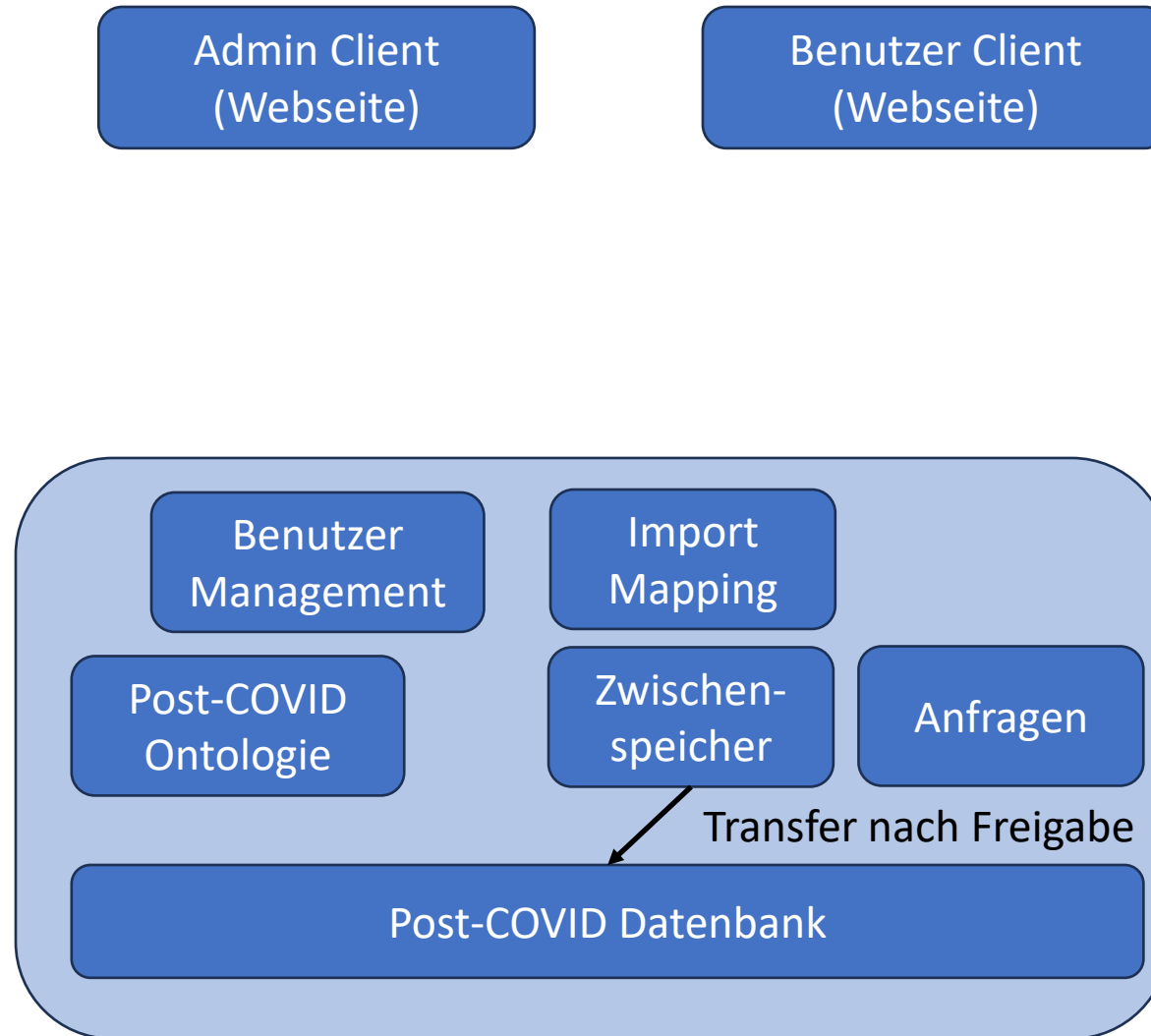
Anwender- Prozesse MVP 4



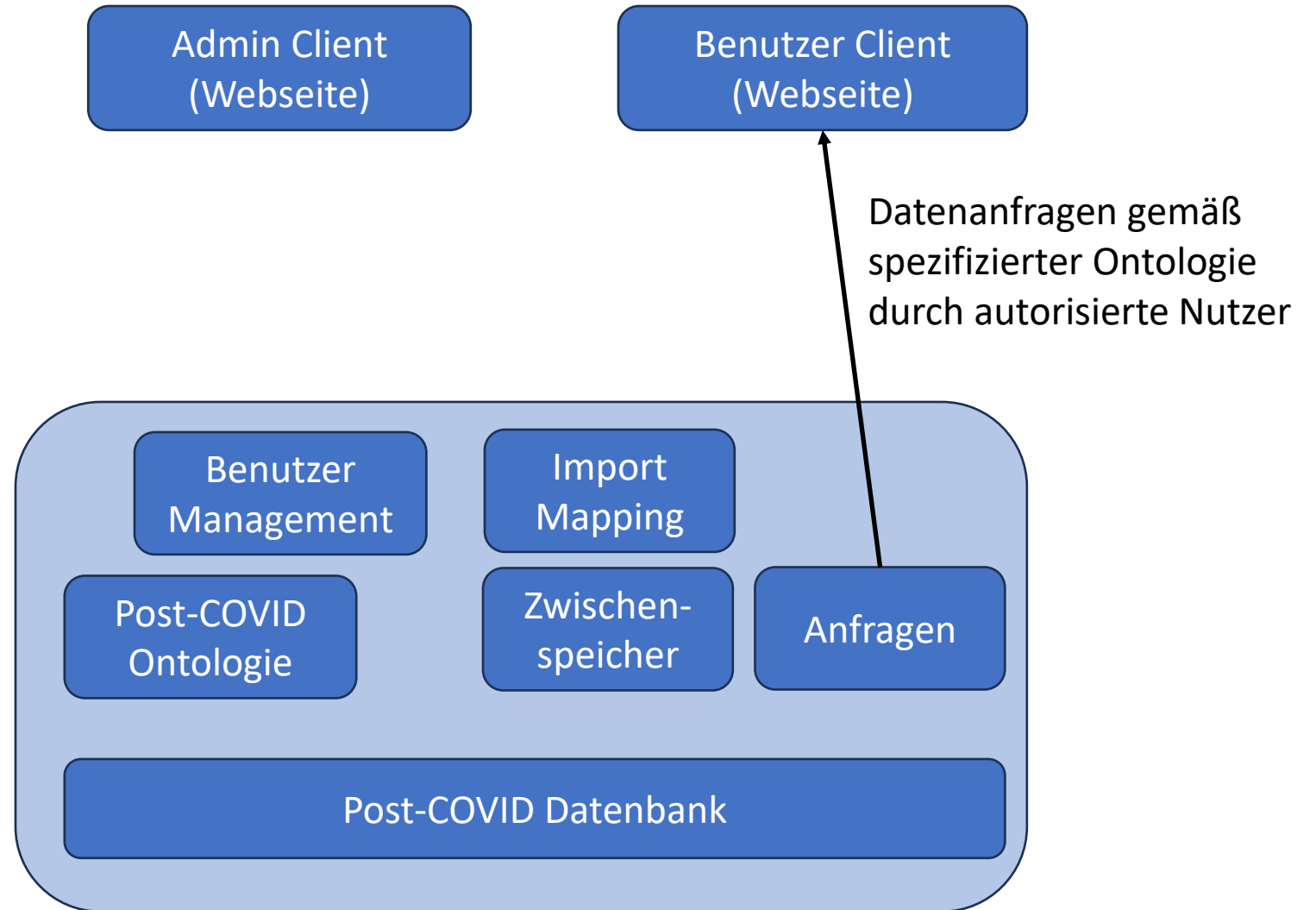
Anwender- Prozesse MVP 5



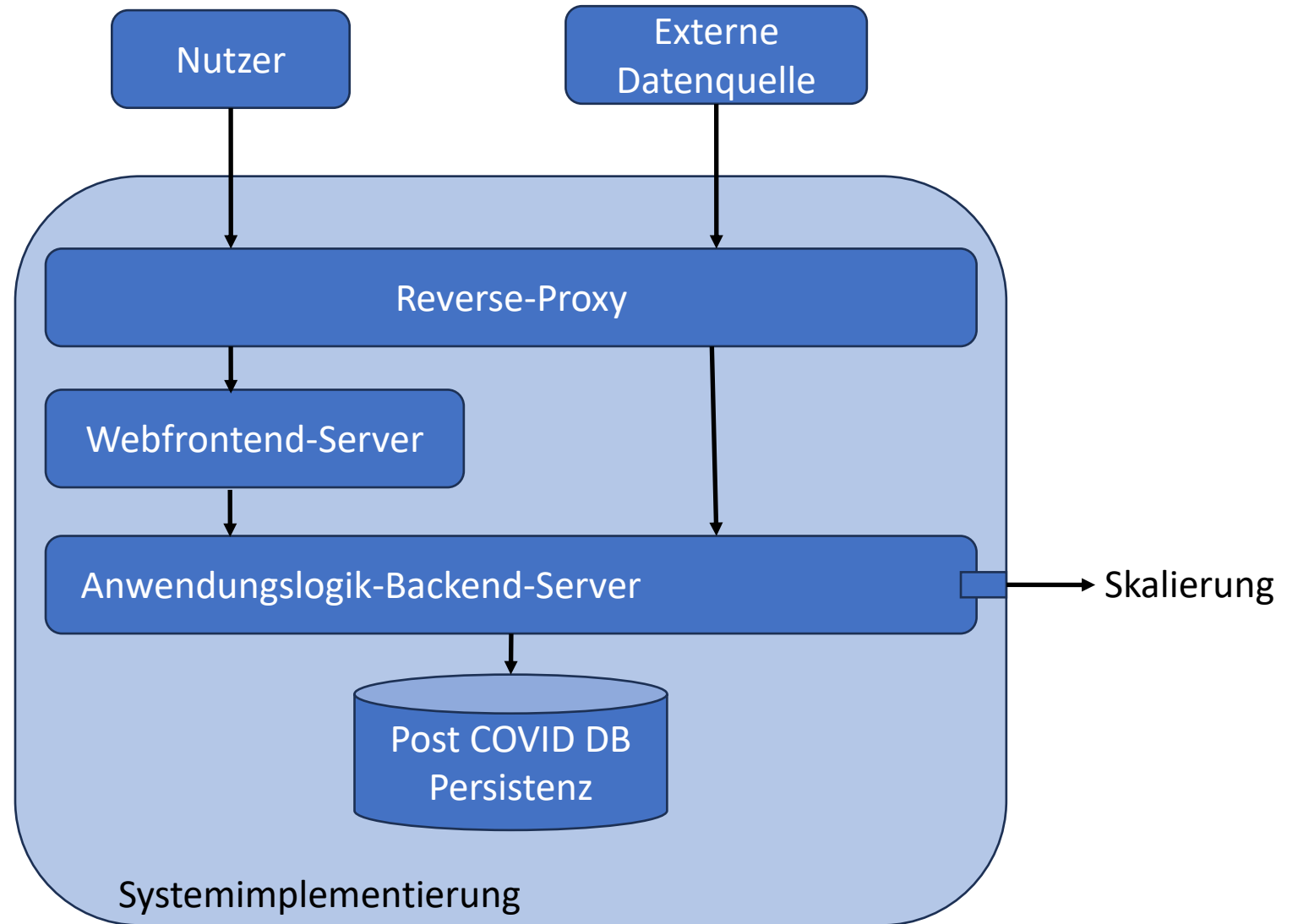
Anwender- Prozesse MVP 6



Anwender- Prozesse MVP 7



System- architektur MVP



Abschlussbericht

Post Covid Datenmodell
Stufe 1

COVID-bezogene Datensätze – Auswahl

- PreVitaCOV (IfA Kiel, Tübingen, Würzburg)
 - NAPKON (teils öffentlich, teils auf Antrag)
 - RKI (teils öffentlich, teils auf Antrag)
 - Charité (auf Antrag)
 - Covid_ShortForm_GP (6 Forschungs-Praxisnetze, IfA Tübingen, Würzburg)
 - Online-Panel zu Post-COVID (Kooperation DESAM-Konsortium und infas360)
- => Grundlage für initiale Ontologie Definition
- => Auswahl wird als Datenmodell für MVP umgesetzt

Abschlussbericht

Post Covid Datenmodell
Stufe 1

Ausblick Stufe 2

- basierend auf ausgewählten COVID-bezogenen Datensätzen
 - Spezifikation einer beispielhaften Ontologie
 - Definition der Mappingmethodik
- Spezifikation der Prozesse 1-7 aus dem MVP
- Spezifikation zugehöriger Governance Strukturen
- Planung der Umsetzung
- Auswahl geeigneter Komponenten
 - Open-Source: Graphdatenbank, WebFramework, etc.
 - Ontologievorlagen

TEAM

Edmund Fröhlich,
GF DESAM, Berlin



Prof. Nico Pfeifer,
Tübingen, Informa-
tiker, COVID Daten-
management und
-analyse



Marc Mausch,
Tübingen (ZIV), Phy-
siker, Datenanalyse u.
Systemarchitektur



Prof. Ildikó Gagyor,
Würzburg, Ärztin, Post-
COVID Expertin



Christian Kretschmann,
Versorgungs-
forscher, Würzburg,
Citizen Science



Prof. Stefanie Joos,
UKT, Ärztin, Klinische
Datenerhebung, Post-
COVID Expertin



Dr. Roland Koch,
UKT, Arzt, Daten-schutz,
Digitali-sierung in der
Medizin



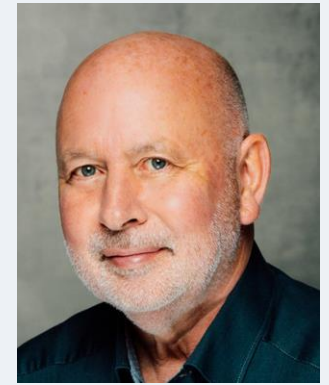
Michaela Hardt,
PhD UKT, Daten-
integrationszentrum
(DIZ)



Prof. Christian Thies,
Reutlingen, Informati-
ker, verteilte medicin-
ische Datensysteme



Dr. Martin Williamson,
Kiel, Hausarzt, Post-
COVID Experte



Edmund Fröhlich, GF
DESAM, wirtschaftl.
u. administrative
Abwicklung

Abschlussbericht

Post Covid Datenmodell
Stufe 1

Vielen Dank