

# Das Projekt CODEX des Netzwerks der Universitätsmedizin (NUM)

Miracum Symposium 16.7.2021

Dagmar Krefting,  
Universitätsmedizin Göttingen  
Institut für Medizinische Informatik



**DZHK**  
DEUTSCHES ZENTRUM FÜR  
HERZ-KREISLAUF-FORSCHUNG E.V.

# Zielvorgaben

„COVID-19 Data Exchange Platform“

**Bundesministerin Karliczek:** (26.03.2020)

*An perspektivisch allen Universitätskliniken werden zudem die Daten der behandelten COVID-19-Patienten systematisch erfasst und in einer Datenbank gebündelt. So können die Krankengeschichten oder auch Daten zur Konstitution der Erkrankten zusammengeführt werden.*

**Konzept Nat. Covid-19-Forschungsnetzwerk der UM:** (30.03.2020)

*Die Daten- und Probengewinnung bzw. -sicherung und deren standortübergreifende Zusammenführung [ist] ein entscheidender Faktor für die Forschung.*

**Antwort der Bundesregierung auf parl. Anfrage:** (12.10.2020)

*Außerdem wurde ein Kerndatensatz für Covid-19 in enger Zusammenarbeit mit der Medizininformatik-Initiative und führenden Akteuren des Gesundheitssystems mit international anschlussfähigen Formatvorgaben definiert. Basierend auf dieser Festlegung sollen Covid-19-bezogene Daten über alle Unikliniken hinweg einheitlich gesammelt und für eine spätere Auswertung zusammengeführt werden. [...Es] wurden [...] Vorarbeiten für den Aufbau der Dateninfrastruktur durchgeführt.*



# CODEx

„COVID-19 Data Exchange Platform“

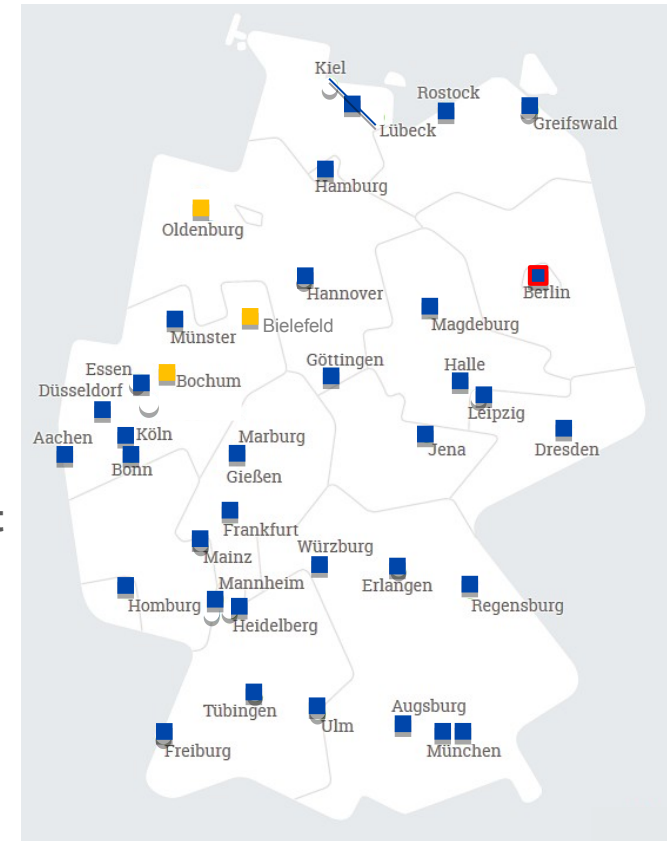
## Ziel

Aufbau einer bundesweit einheitlichen, datenschutzkonformen Infrastruktur zur Speicherung von Covid-19 Forschungsdatensätzen.

- Datenrepositories und Datenerfassungsinstrumente (z.B. eCRF),
- Use & Access-Verfahren und Treuhandstellen

## Mehrwert

- Covid-19-Forschungsdatensätze **verschiedener Komplexität** (klinische Daten und Daten zu Bioproben)
- aus allen deutschen Universitätskliniken, pseudonymisiert der Forschung zur Verfügung stellen.
- **NAPKON: Deep-Phenotyping - Kohortendaten**
- **CODEx-MII: Big Data - Versorgungsdaten**



# Der GECCO Datensatz

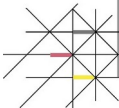
## COVID-19-DATENSATZ

### Der GECCO-Datensatz für die COVID-19-Forschung

Zur Bewältigung der COVID-19-Pandemie und zur besseren Behandlung von COVID-19-Patienten fördert das Bundesministerium für Bildung und Forschung (BMBF) das Netzwerk Universitätsmedizin (NUM) zu COVID-19. Das Netzwerk soll die Daten behandelter COVID-19-Patienten systematisch erfassen und zusammenführen. Der bundesweit einheitliche Datensatz „German Corona Consensus Data Set“ (GECCO) wurde entwickelt, damit Forschende wissenschaftliche Daten zu COVID-19 in Deutschland und die Informationen zur Behandlung von COVID-19-Patienten standardisiert und systematisch erheben und analysieren können. GECCO dient der Wissenschaft damit als gemeinsame Sprache und Arbeitsgrundlage zur Bekämpfung der Pandemie. Der Datensatz wurde in einem transparenten Prozess unter Beteiligung von 50 Expertinnen und Experten aus Unikliniken, Fachgesellschaften und Forschungsinitiativen erarbeitet.

#### Inhaltliche Parameter

1. **Anamnese/Risikofaktoren:** Erfasst wird beispielsweise, ob der Patient/die Patientin Vorerkrankungen wie Herz-Kreislauf- oder Tumorerkrankungen, HIV oder Diabetes hat. Außerdem wird nach Risikofaktoren wie Rauchen gefragt.
2. **Bildgebung:** Erfasst werden Daten bildgebender Verfahren wie z.B. von Röntgenbildern der Lunge.
3. **Demographie:** Es werden z.B. Angaben zum Geburtsdatum, Geschlecht, Gewicht, zur Größe sowie zur ethnischen Zugehörigkeit erhoben.
4. **Epidemiologische Faktoren:** Gefragt wird, ob der Patient/die Patientin in den letzten 14 Tagen vor Beginn seiner/ihrer Beschwerden wissentlich Kontakt mit einer wahrscheinlich oder nachgewiesenermaßen an COVID-19 erkrankten Person hatte.
5. **Komplikationen:** Erhoben wird beispielsweise, ob der Patient/die Patientin eine Lungenembolie, einen Herzinfarkt oder eine Blutstrominfektion hat/hatte.
6. **Krankheitsbeginn/Aufnahme:** Erhoben wird der Zustand bei der Diagnose/Aufnahme des Patienten/der Patientin.
7. **Laborwerte:** Verschiedene Laborwerte werden erfasst wie z.B. Blutwerte.
8. **Medikation:** Gefragt wird nach eingenommenen Medikamenten sowohl zur COVID-Therapie als auch zur Behandlung weiterer Erkrankungen.
9. **Outcome bei Entlassung:** Erfasst wird die Art der Entlassung einer Patientin/eines Patienten, ob sie/er beatmet wurde und ob sie/er mit einem negativen oder positiven Testergebnis entlassen wurde.
10. **Studieneinschluss/Einschlusskriterien:** Als Kriterium für die Teilnahme an der Studie wird erfasst, ob COVID-19 Hauptursache für die Aufnahme





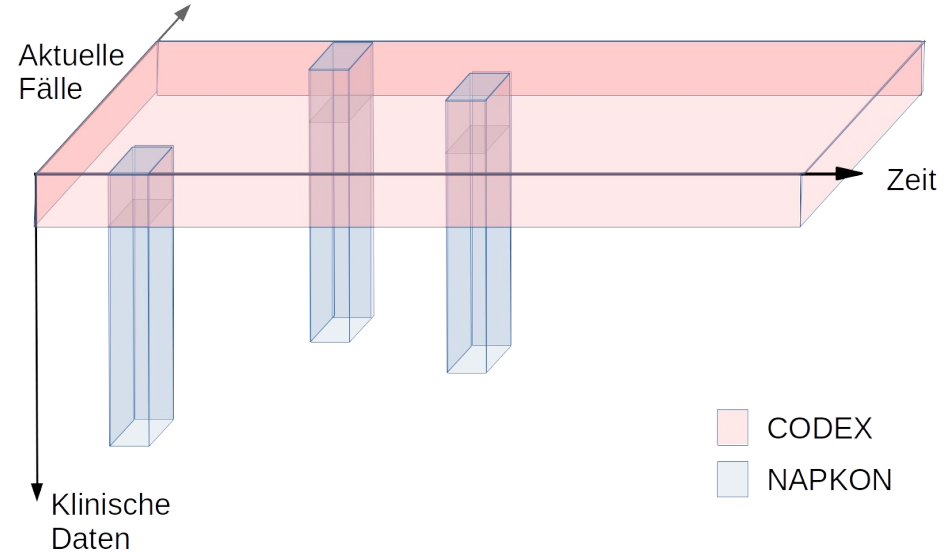
# Deep Phenotyping vs. GECCO

## Deep Phenotyping

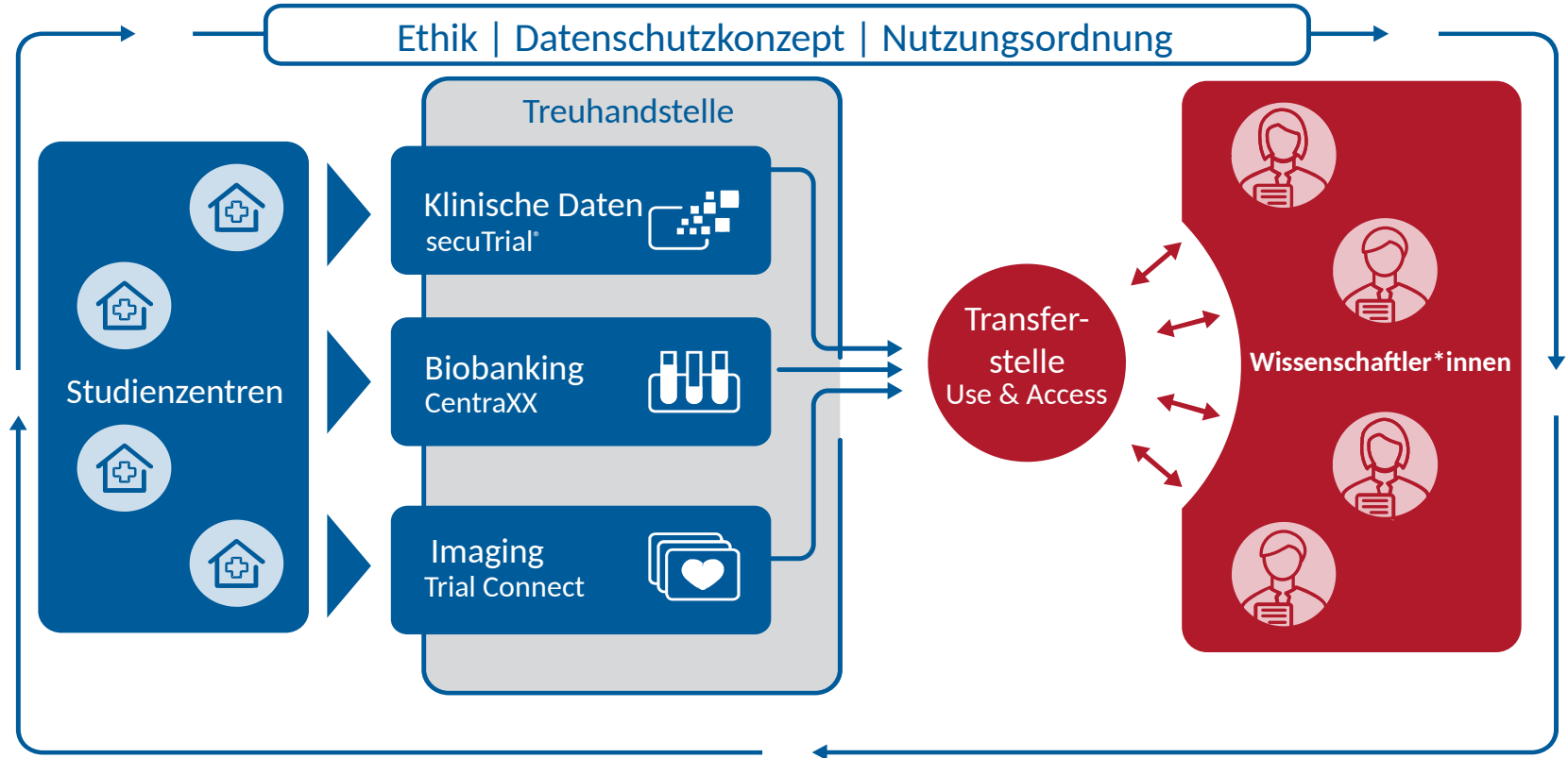
- Kohortenspezifische Datensätze
- Umfangreiche Untersuchungen
- Festes Studienprotokoll mit Follow-ups
- > 5000 Datenitems (HAP)
- Bild- und Biomaterialproben

## GECCO-Datensatz

- Konsensusdatensatz
- Informationen aus der Versorgung
- Erhebung der Informationen bei Bedarf
- 83 Datenitems
- Klinische Daten



# Kohorten - NAPKON powered by DZHK

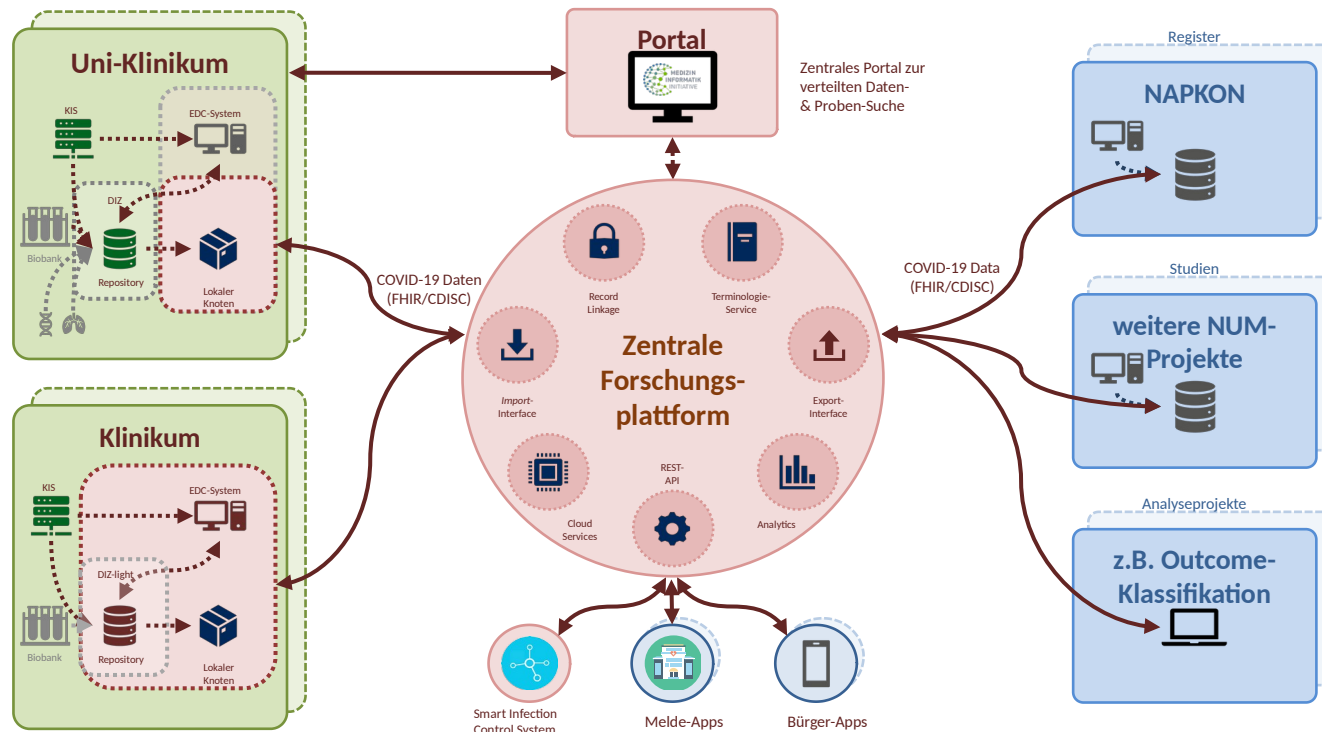


# Schnelle Verfügbarkeit von Routinedaten aus der Krankenversorgung




- **Zeitpunktbetrachtung** und **Verlauf** der Patientenstruktur oder den stationären Behandlungskonzepten
- **Tagesaktuell und kontinuierlich**: Einsatz und Erfolg von **Therapien** und **Medikamenten**
- **Behandlungsrelevante Daten für Ärzt\*innen**
  - Welche Patient\*innen haben die schlechtesten Lungenwerte auf der ICU? Welche Vorerkrankungen haben diese Patient\*innen?
  - Behandlungserfolge bei Patient\*innen mit ähnlichen Vorerkrankungen

# GECCO - CODEX-MII aka CODEX



# Die Medizininformatikinitiative



 MEDIZIN INFORMATIK INITIATIVE

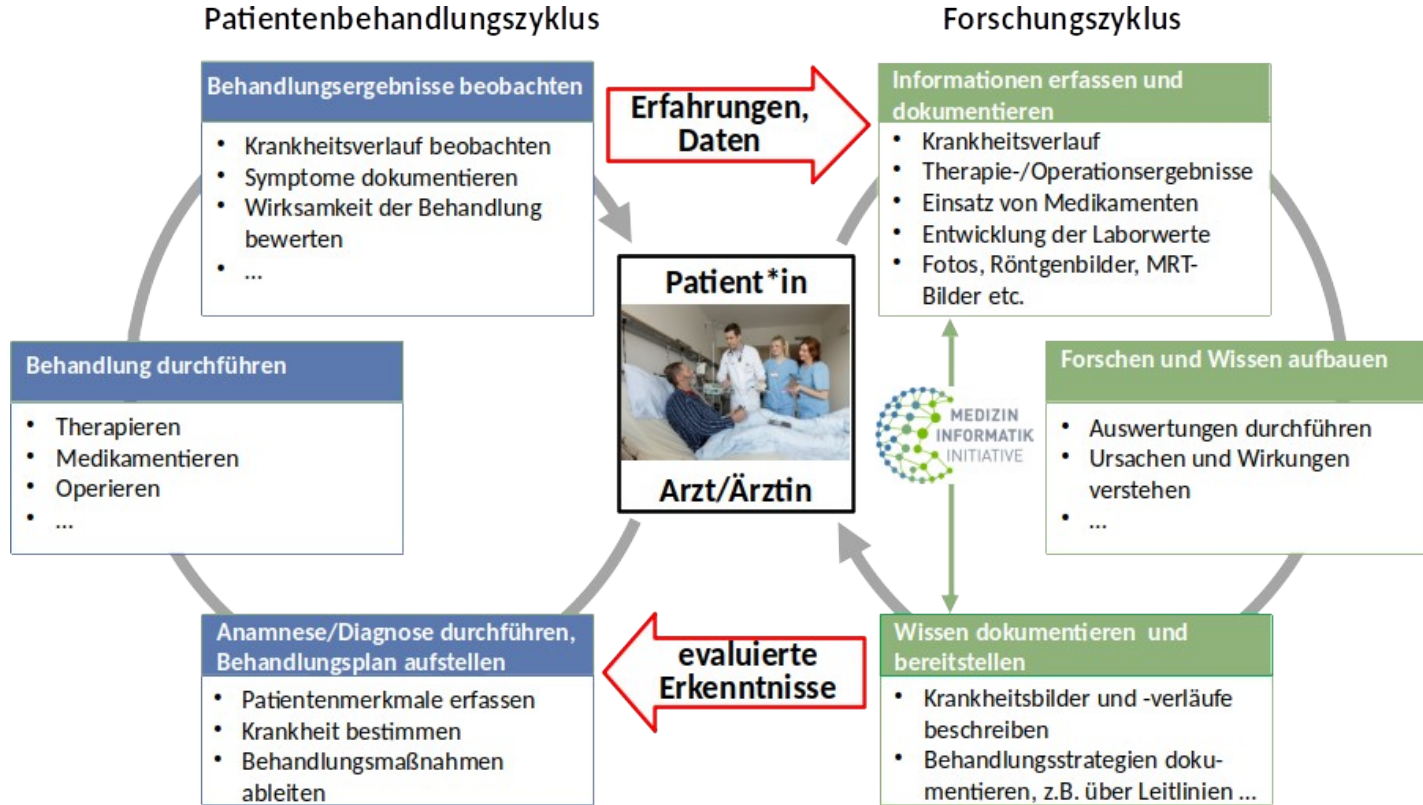
Suchbegriff eingeben  EN

ÜBER DIE INITIATIVE KONSORTIEN ZUSAMMENARBEIT MITMACHEN AKTUELLES

**Daten gemeinsam nutzen.**  
Forschung stärken, Versorgung verbessern.  
Medizininformatik.

Um Daten aus Krankenversorgung und Forschung besser nutzbar zu machen, hat das Bundesministerium für Bildung und Forschung das Förderkonzept Medizininformatik initiiert und plant, hierfür in den kommenden Jahren 150 Millionen Euro bereitzustellen. Die Fördermaßnahme soll die medizinische Forschung stärken und die Patientenversorgung verbessern.

# Die Medizininformatikinitiative - Ziele



# Die Medizininformatikinitiative



## Methode - Datenintegrationszentren

- Anbindung der Versorgungssysteme
- Etablierung von lokalen Prozessen zum Forschungsdatenmanagement

## Mitglieder der Initiative

- 36 universitätsmedizinische Standorte
- 29 Datenintegrationszentren
- Koordinierungsstelle
- mehr als 60

Partner aus Akademie & Industrie

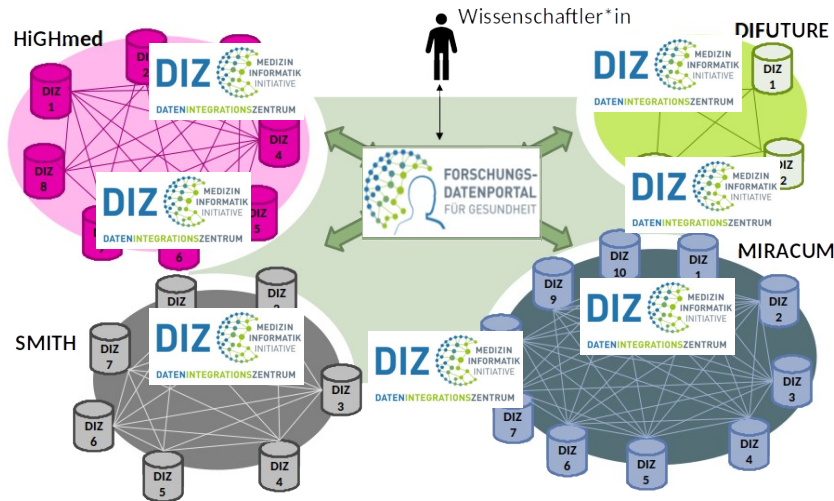


Koordinationsstelle





# Gemeinsame Nutzung von Daten zwischen Krankenversorgung und Forschung



## Regulatorisch-Organisatorisch

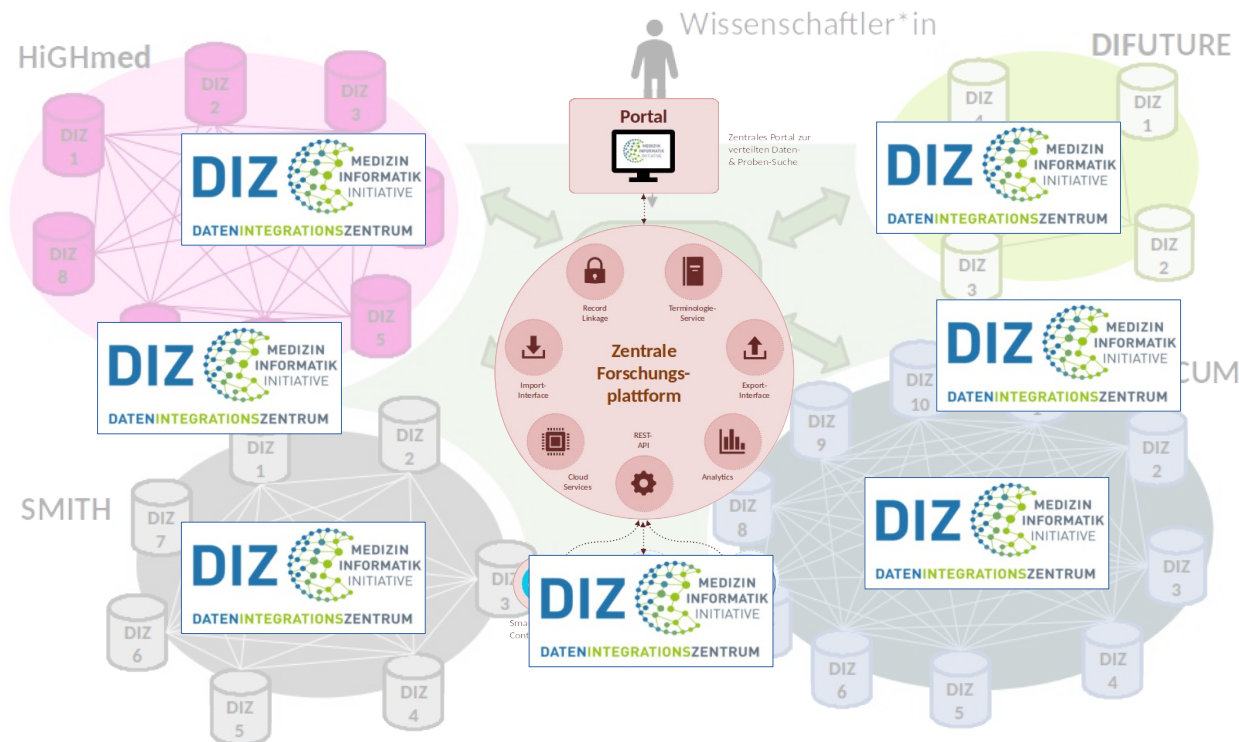
- Broad Consent
- Vertragswerke Datennutzung
- Zentrale Antrags- und Registerstelle (Forschungsdatenportal)
- Konzepte Datenschutz und Datensicherheit

## Technisch

- MII-Kerndatensatz
- HL7-FHIR als Austauschformat

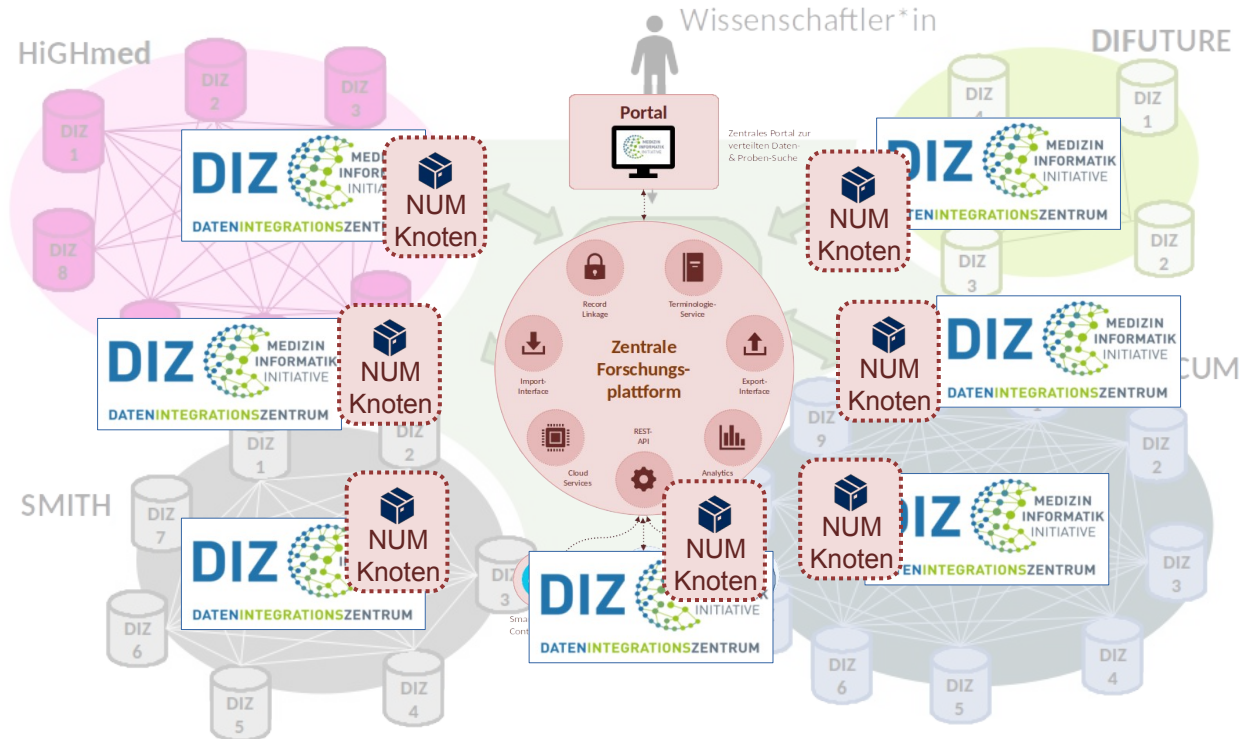


# COVID-19 Data Exchange Platform

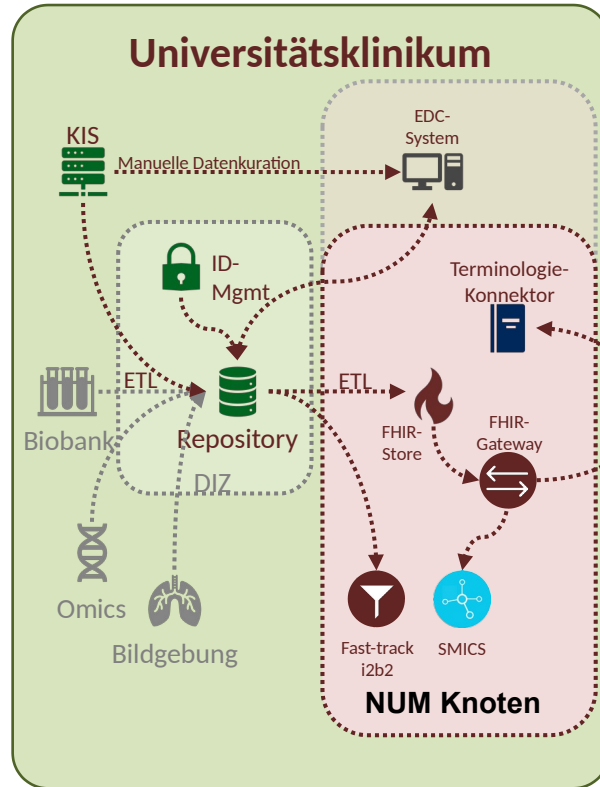




# COVID-19 Data Exchange Platform

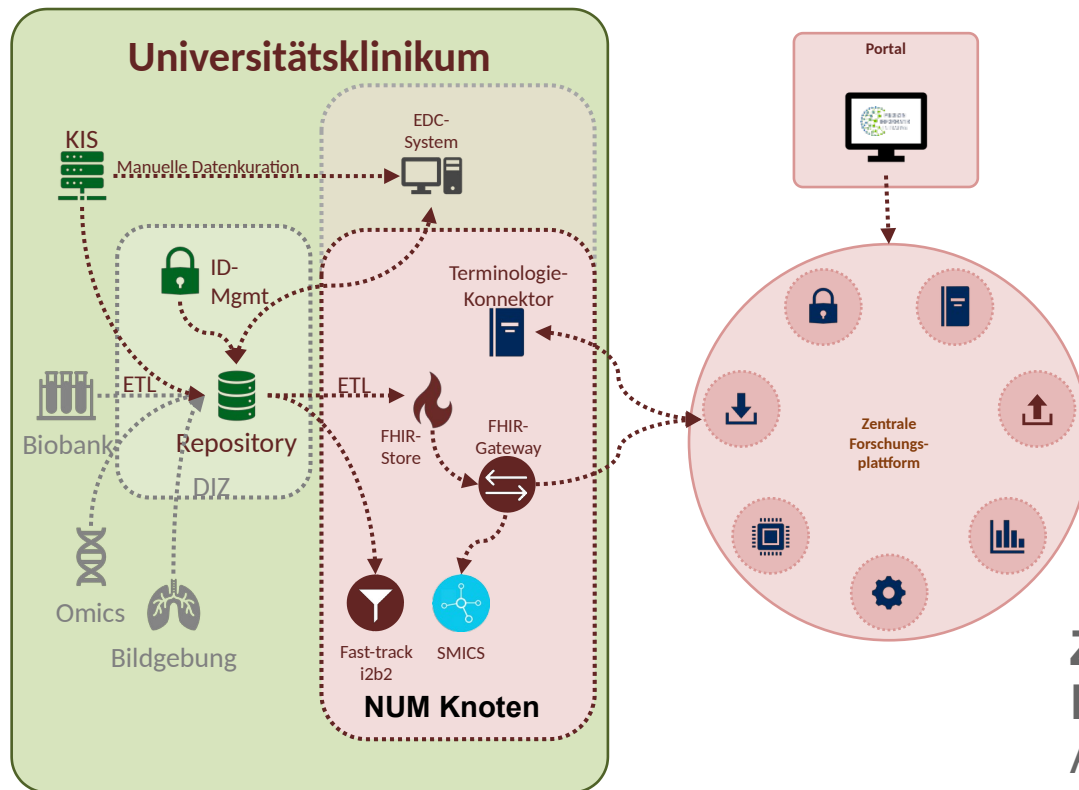


# CODEX Architektur – NUM Knoten



- lokale Datenspeicher an den UK-Standorten für  
→ GECCO-Daten (FHIR-Store)  
→ Fast-Track (i2b2, standard. Datenmodell)
- Wo möglich ETL aus Primärsystemen
- Wo nicht möglich Ergänzung der lokalen Dokumentation (EDC-System)
- tagesgenaue Datenverfügbarkeit
- zeitnahe Abbildung des realen Behandlungsgeschehens
- verteilte Real World Data Analysen
- → Real World Evidenz
- Optional: SmICS, Dashboard

# CODEX Architektur – Zentrale Komponenten



**Zentrale Forschungsdatenrepo**  
Transfer und Bereitstellung  
der GECCO-Daten

**Zentrales Dashboard**  
Transfer und  
Visualisierung  
aggregierter Daten

**Zentrales Feasibilityportal**  
Abfrage lokaler Daten

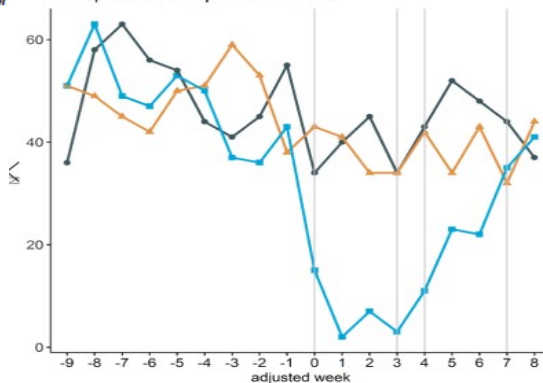
# Ergebnisse auf MII Vorarbeiten (Fast Track)

## Reduced Rate of Inpatient Hospital Admissions in 18 German University Hospitals During the COVID-19 Lockdown

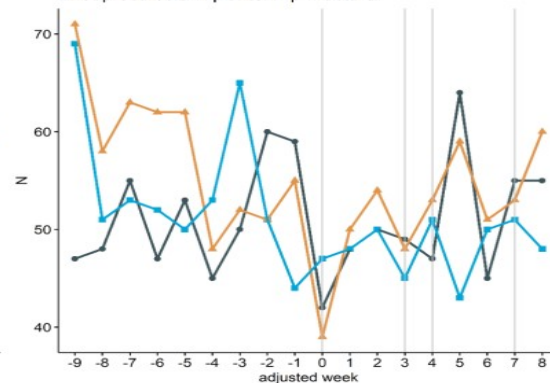
Lorenz A. Kapsner<sup>1,2\*</sup>, Marvin O. Kampf<sup>1</sup>, Susanne A. Seuchter<sup>1</sup>, Julian Gruendner<sup>3</sup>, Christian Gulden<sup>3</sup>, Sebastian Mate<sup>1</sup>, Jonathan M. Mang<sup>1</sup>, Christina Schüttler<sup>3</sup>, Noemi Deppenwiese<sup>3</sup>, Linda Krause<sup>4</sup>, Daniela Zöller<sup>5</sup>, Julien Balig<sup>5</sup>, Timo Fuchs<sup>7</sup>, Patrick Fischer<sup>9</sup>, Christian Haverkamp<sup>9</sup>, Martin Holderried<sup>10</sup>, Gerhard Mayer<sup>6</sup>, Holger Stenzhorn<sup>11,12</sup>, Ana Stolnicu<sup>6</sup>, Michael Storck<sup>13</sup>, Holger Storf<sup>14</sup>, Jochen Zohner<sup>15</sup>, Oliver Kohlbacher<sup>12,15,16,17</sup>, Adam Strzelczyk<sup>18</sup>, Jürgen Schüttler<sup>19</sup>, Till Acker<sup>20</sup>, Martin Boeker<sup>5</sup>, Udo X. Kaisers<sup>21</sup>, Hans A. Kestler<sup>8</sup> and Hans-Ulrich Prokosch<sup>3</sup> on behalf of our MII-Research Group

18 Unikliniken aus allen Konsortien  
~ 1,3 Mio Fälle (Jan – Mai; 2018-2020)

Endoprosthese Implants: Arthrosis



Endoprosthesis Implants: Hip Fracture



Dezentrale  
Komponenten

- Standort-  
übergreifende  
Analysen über  
MII-Kerndatensatz



# Ergebnisse auf MII Vorarbeiten (Fast Track)



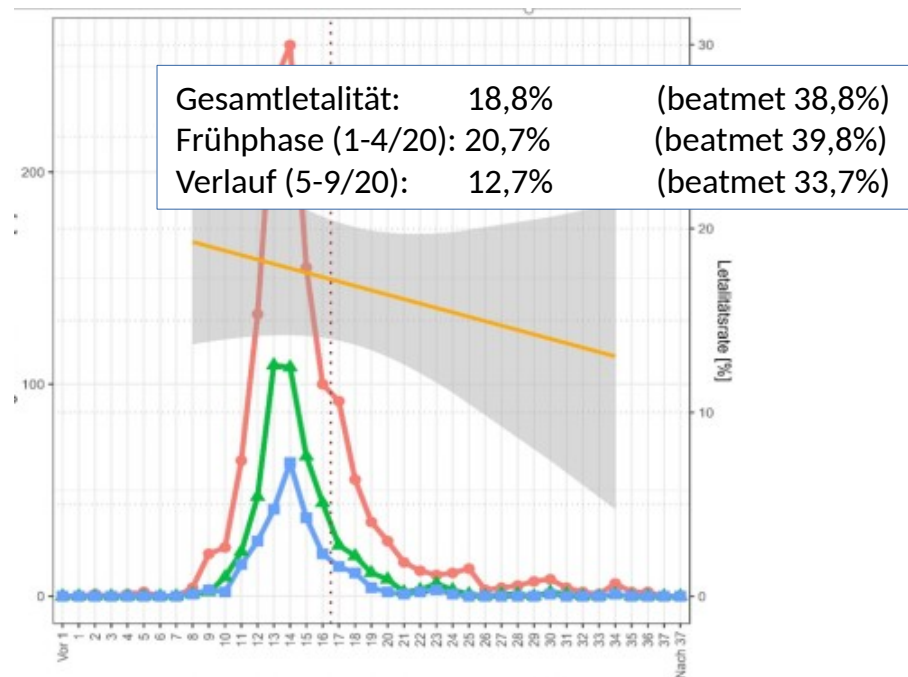
## 14 Unikliniken aus allen Konsortien 1.318 Covid-19 Fälle (Jan – Sept 2020)

### Letalität von Patienten mit COVID-19: Untersuchungen zu Ursachen und Dynamik an Deutschen Universitätsklinika

36	Autoren	43	Prof. Dr.-Ing. Oliver Kohlbacher <sup>5,6</sup>
37	Prof. Dr. med. Dr. h. c. Jürgen Schüttler <sup>*1</sup>	44	Prof. Dr. med. Martin Boeker <sup>4</sup>
38	Jonathan M. Mang, M.Sc. <sup>*2</sup>	45	Prof. Dr. Dr. med. Kai Zacharowski <sup>7</sup>
39	Dr. med. Lorenz A. Kapsner <sup>2,3</sup>	46	Prof. Dr. med. Gernot Rohde <sup>8</sup>
40	Susanne A. Seuchter <sup>2</sup>	47	Julien Balig <sup>9</sup>
41	Prof. Dr. Harald Binder <sup>4</sup>	48	Marvin O. Kampf, M.Sc. <sup>2</sup>
42	Dr. rer. physiol. Daniela Zöller <sup>4</sup>	49	Prof. Dr. med. Rainer Röhrig <sup>10</sup>
		50	Prof. Dr. biol. hom. Hans-Ulrich Prokosch <sup>11</sup>

Dezentrale  
 Komponenten

- Standort-  
 übergreifende  
 Analysen über  
 MII-Kerndatensatz



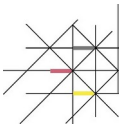


# CODEX - Analysen

Dezentrale Komponenten

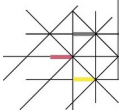
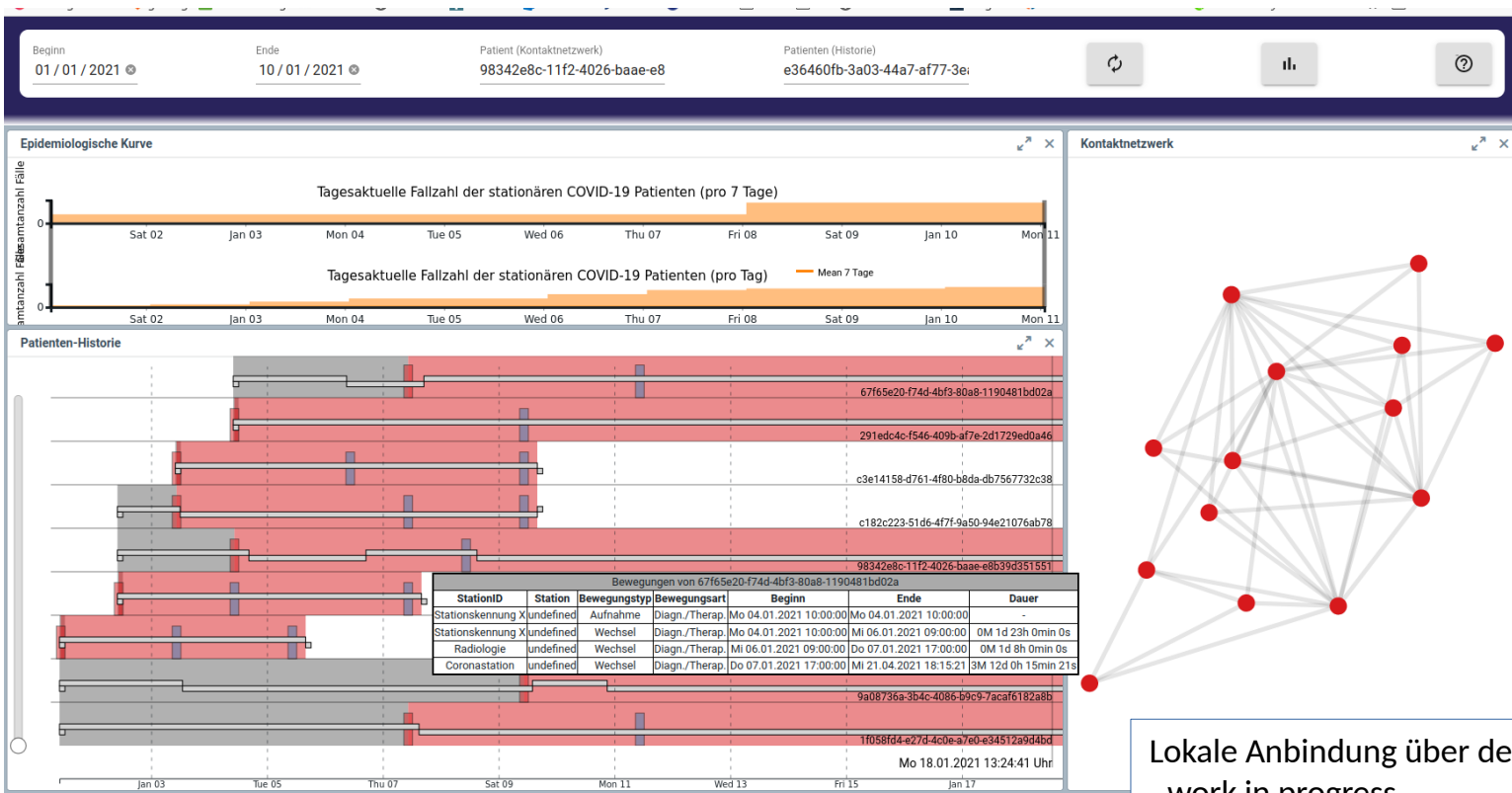
- Standort-übergreifende Analysen über **MII-Kerndatensatz**





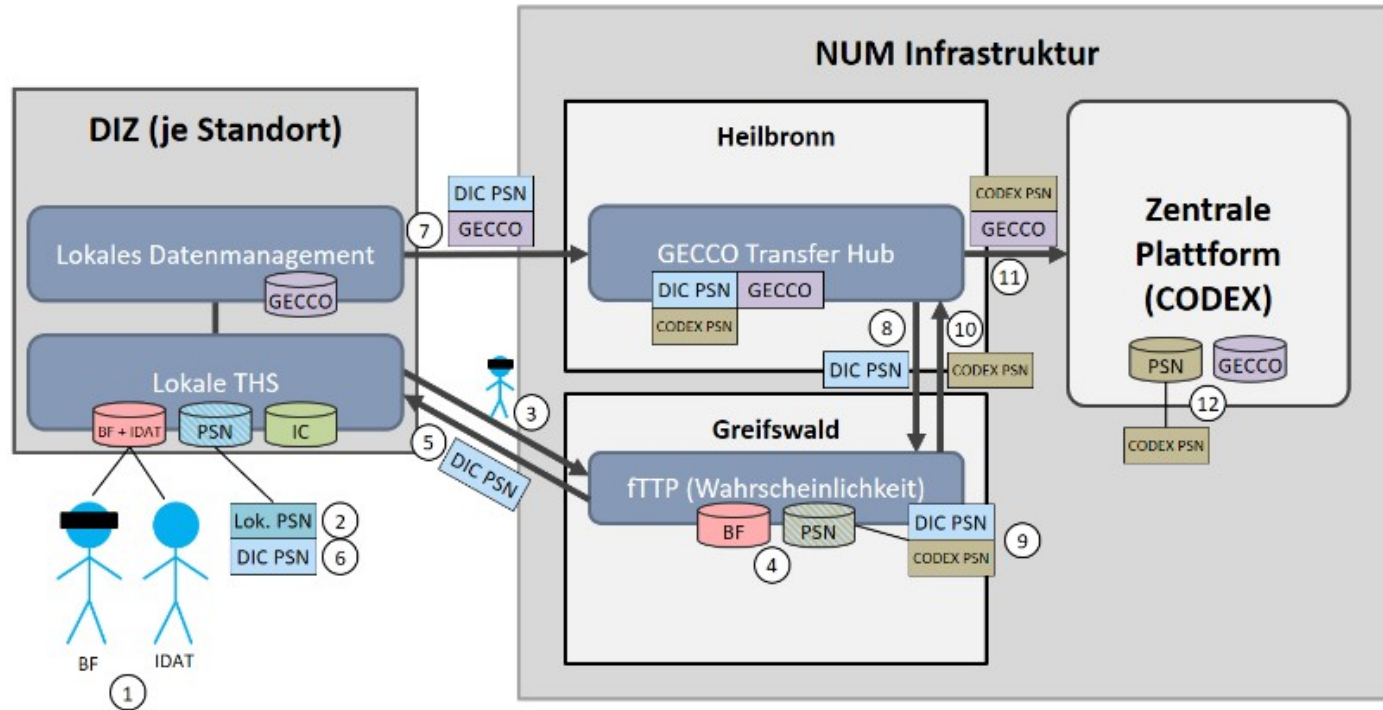


# CODEX – Smart Infection Control System



# CODEX – Datentransfer ins zentrale Repo

⚠ **Achtung:** Stark vereinfachte Darstellung



# CODEX – Hackathon 13.7.2021

Standort	DSF onboarding liegt vor	FHIR-Server/BPE installiert	DSF- Prozesse deployed	Bloomfilter/PSN - Variante 1, 2 oder 3	PING-PONG Test erfolgreich	Kommunikation in Testumgebung	FTTP Zertifikate beantragt	FTTP Anbindung hergestellt	FTTP Anbindung erfolgreich getestet	Datensatz bei zentraler Komponente verarbeitet	Datensatz sinnvoll
Aachen											
Augsburg											
Berlin - Charité	x	x	x	1	x	x	x	x	x	x	
Bonn - UKB		x		2			x				
Dresden	x	x	x		x		x				
Düsseldorf											
Erlangen	x	x	x	2	x	x	x	x	x	x	
Essen	x	x	x	2	x	x	x	x	x	x	
Frankfurt	x	x	x	2	x	x	x	x	x	x	
Freiburg	x	x	x	2	x	x	x				
Gießen - JLU											
Göttingen - UMG	x	x		2	x		x				
Greifswald	x	x		2			x				
Halle - UKH	x	x	x	2	x		x				
Hamburg - UKE	x	x	x	2	x						
Hannover - MBH	x	x	x	2	x	x	x	x	x	x	
Heidelberg	x	x	x	1	x	x	x	x	x	x	
Jena - FSU											
Köln	x	x	x	2	x	x					
Leipzig											
Magdeburg - OVGU											
Mainz - JGU											
Mannheim							x				
Marburg	x	x	x	2	x	x	x	x	x	x	
München - LMU	x	x	x	2	x	x	x	x	x	x	
München - TUM/MRI							x				
Münster - WWU	x	x	x	1	x	x	x	x	x	x	
Regensburg - UKR	x	x	x	2	x	x	x				
Rostock	x	x	x	2	x	x	x	x	x	x	
Saarland (Homburg) - UKS	x	x	x	2	x	x	x				
Schleswig-Holstein - UKSH	x	x		1	x		x				
Tübingen	x	x	x	1			x				
Ulm							x				
Würzburg	x	x			x		x				

Pro Version

# CODEX Herausforderungen

## Regulatorisch-Organisatorisch

- Broad Consent + Z-Modul
- Vertragswerke Datennutzung
- Konzepte Datenschutz und Datensicherheit

## Technisch

- ETL der GECCO-Daten in den FHIR-Store
- Data-Sharing-Framework
- Föderierte Treuhandstelle



CODEX

AP6 REPORTING KW26

a



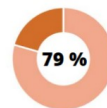
GEFÖRDEBT VOM  
Bundesministerium  
für Bildung  
und Forschung

34 Universitätskliniken nehmen an CODEX teil

Davon liefern bisher

**keine** Standorte Daten an die zentrale Plattform

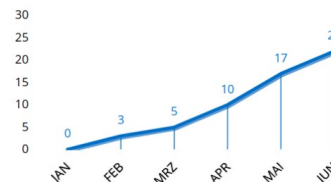
Um dies zu erreichen benötigen die Standorte:



(27/34) haben den NUM Knoten v1 installiert

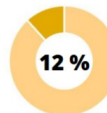
...davon nahmen 70 % an Studien teil

### #1 SOFTWARE



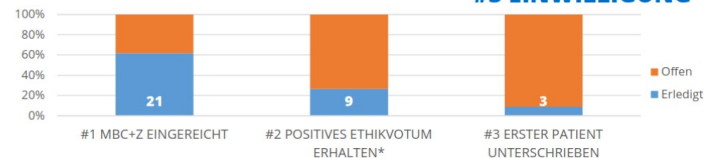
65 %

(22/34) haben bereits den  
NUM Knoten v2 im  
Testbetrieb installiert



haben ihre Betriebs-/Datenschutz- und Sicherheitskonzepte  
geschrieben

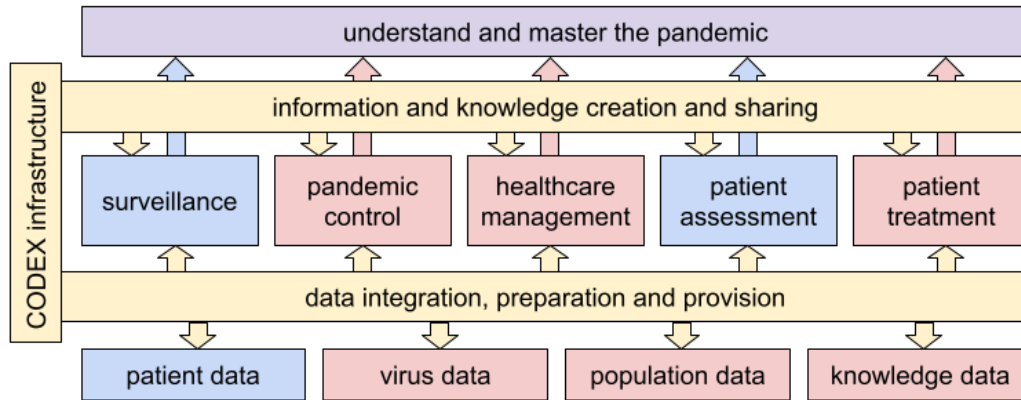
### #2 KONZEPTE



\* oder kann bereits über eigenen Consent Daten nach Zentral liefern

Stand 30.06.2021

# CODEX – Ausblick



## Infrastrukturförderung

- NUM – Routinedatenplattform
- 2022 – 2024
- Betrieb NUM-Knoten, zentrale Plattform, Dashboard, Regulatorik

## Projektförderung CODEX+

- Erweiterung der NUM-RDP
- Generische Komponenten
- Integration andere Infrastrukturen
- Use Cases





Vielen Dank...

...für Ihre Aufmerksamkeit

...und den CODEX-Mitstreiter:innen für unermüdlichen Einsatz unter absurden Rahmenbedingungen



dagmar.krefting@med.uni-goettingen.de - <https://medizininformatik.umg.eu>



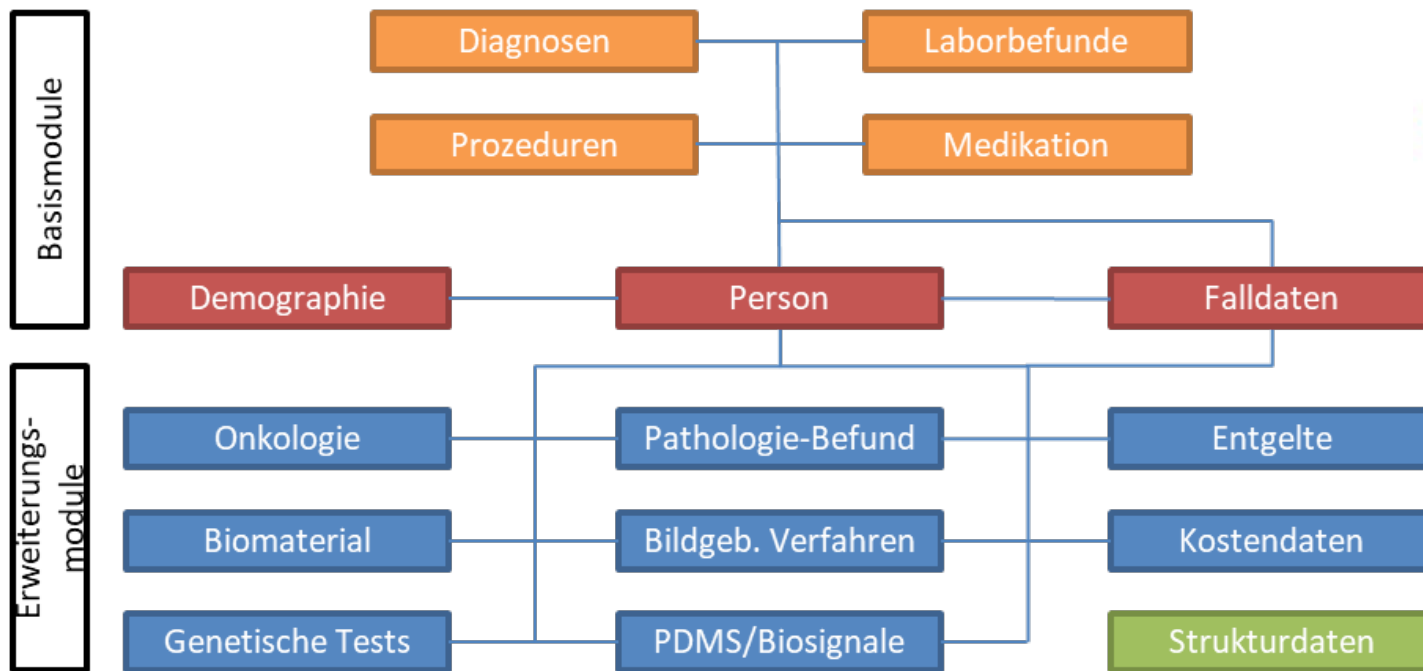


# BACKUP



# Modulare Datenstruktur der MII

## Modularer Kerndatensatz



LOINC®

SNOMED CT



HL7  
FHIR

# Schnittstellen zu den weiteren NUM Projekten

- NAPKON
- B-FAST
- COMPASS
- Racoon
- u.v.a.

CORE Registry Portal

Nicht sicher | 134.76.19.215:8888/patient-explorer

MEDIZIN INFORMATIK INITIATIVE

Übersicht Patienten-Explorer Formulare Clinical Risk Assessment CoV App SMICS ASiC

Patientensuche

Aktuelle Patienten ID: 2b3eff54-833e-448c-9f23-fa4d2fc06cc4 Neues Formular anlegen +

**Verlauf**

- SOFA-Score  
23.05.2020, 21:30 Uhr  
Medizinische Hochschule Hannover
- Intensivmedizinisches Monitoring  
Körpertemperatur  
23.05.2020, 21:29 Uhr
- Intensivmedizinisches Monitoring  
Körpertemperatur  
23.05.2020, 21:29 Uhr
- Intensivmedizinisches Monitoring

**Demographische und Allgemeine Daten**

Geschlecht	weiblich	Größe	177 cm	23.05.2020
Geburtsdatum	06.05.1976	Gewicht	69 kg	23.05.2020

**Covid-19 Anamnese**

Für diesen Patienten wurden noch keine Symptome erfasst.  
Für diesen Patienten wurden noch keine Reisen erfasst.

**Covid-19 relevante Diagnosen, Scores und Laborwerte**

**Diagnose**  
COVID-19 (ICD-10-GM-2020: U07.1): Nein

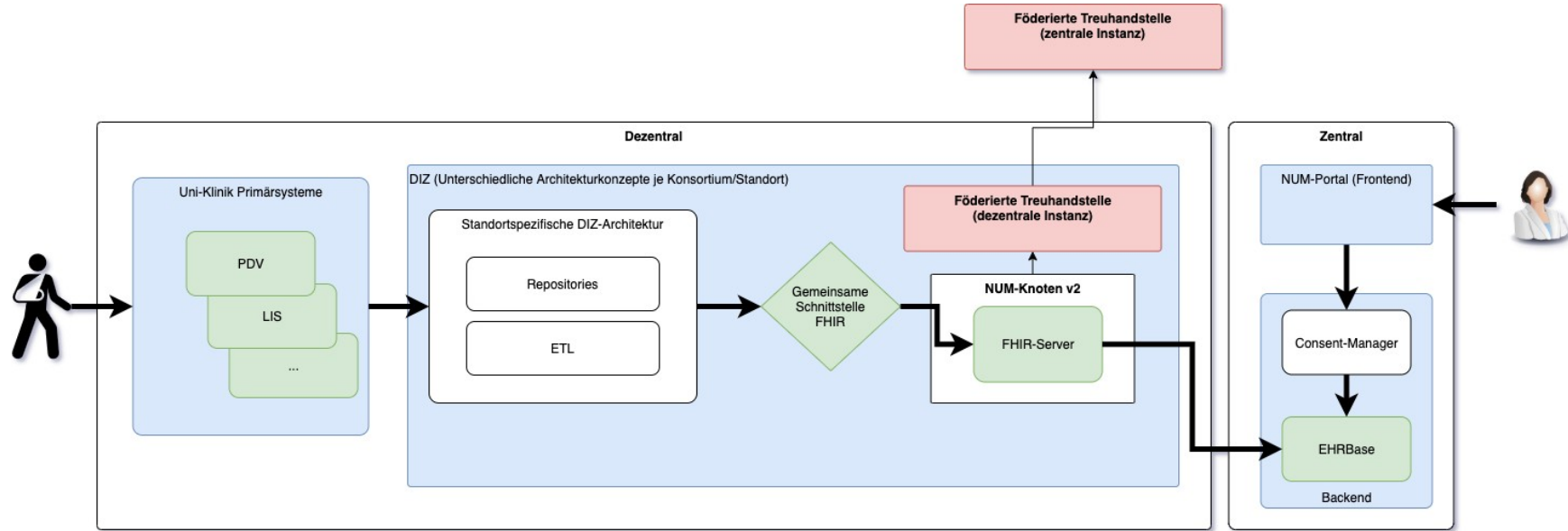
**Laborwerte**

Ferritin	keine Daten vorhanden
CRP	keine Daten vorhanden
LDH	keine Daten vorhanden
Lymphopenie	keine Daten vorhanden

**Scores**

SOFA-Score	14	20.05.2020
qSOFA-Score	keine Daten vorhanden	
Murray-Score	keine Daten vorhanden	
Frailty-Score	keine Daten vorhanden	

# Entwurf – grobes Architekturschaubild



# Entwurf – detailliertes Architekturschaubild

