



# Data Sharing Framework Community Projekt

## Einführung in das DSF-Community Verbundprojekt

MIRACUM-DIFUTURE Kolloquium – 20.02.2024

Maximilian Kurscheidt für die DSF-Community

SPONSORED BY THE



Federal Ministry  
of Education  
and Research

# Inhalt und Lernziele

---

## Inhalt:

- Rolle der Data Sharing Framework (DSF) Community in der Medizin Informatik Initiative (MII)
- Einführung in das DSF als Infrastrukturkomponente der MII
- Verteilte Data Sharing Prozesse für verschiedene Anwendungsfälle
- DSF Community Ansatz

## Lernziele:

- Grundverständnis über die Funktionsweise des DSF und der verteilten Prozesse
- DSF als Infrastrukturkomponente für eigene Data Sharing Vorhaben nutzen
- Teilnahmemöglichkeiten und Aufbau der DSF Community

# DSF Community Verbundprojekt

- Modul 2B Verbundprojekt der Medizininformatik Initiative gefördert durch das Bundesministerium für Bildung und Forschung

- **Verbundpartner:**

- GECKO Institut für Medizin, Ökonomie und Informatik, Hochschule Heilbronn (Projekt-Lead)
- Institut für Medizinische Informatik, Universitätsklinikum Heidelberg
- LIFE – Leipziger Forschungszentrum für Zivilisationserkrankungen, Universitätsklinikum Leipzig

- **Team:**

- Prof. Dr. Christian Fegeler (GECKO)
- Hauke Hund (GECKO)
- Maximilian Kurscheidt (GECKO)
- Simon Schweizer (GECKO)
- Christoph Zilske (GECKO)
- Reto Wettstein (IMI)
- Alexander Kiel (LIFE)
- Mathias Rühle (LIFE)
- Jan Böhringer (GECKO)
- Amir Kannout (GECKO)
- Mohamad Khalil (GECKO)

- **Community Entwickler:**

- u.a. Florian Seidel, Julian Gründner, Simon Mödinger, Alexander Twirdik, Patrick Werner



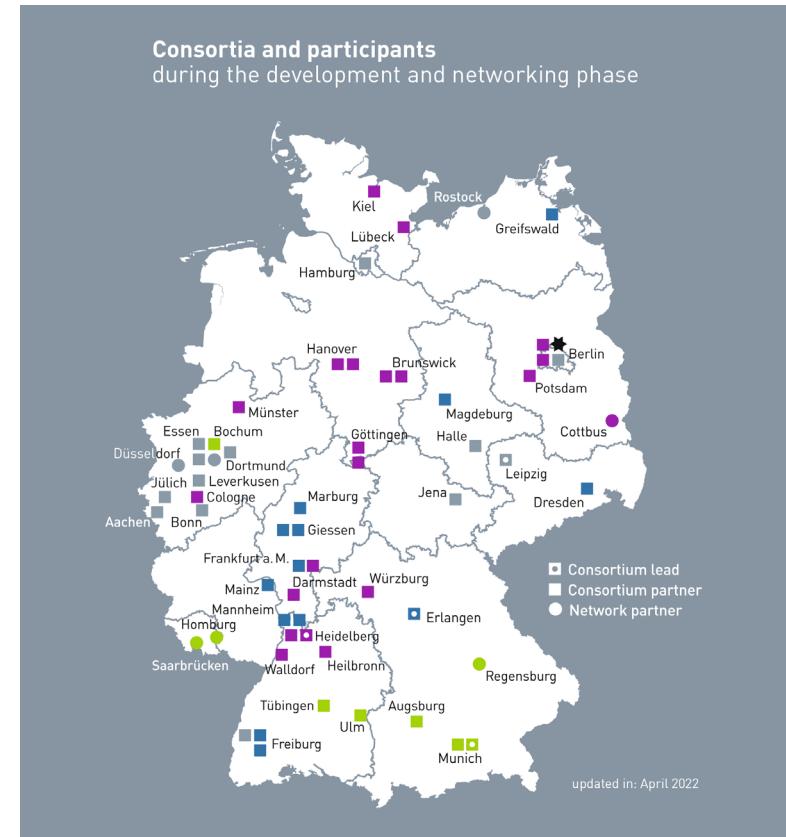
SPONSORED BY THE



Federal Ministry  
of Education  
and Research

# Überblick Medizin Informatik Initiative

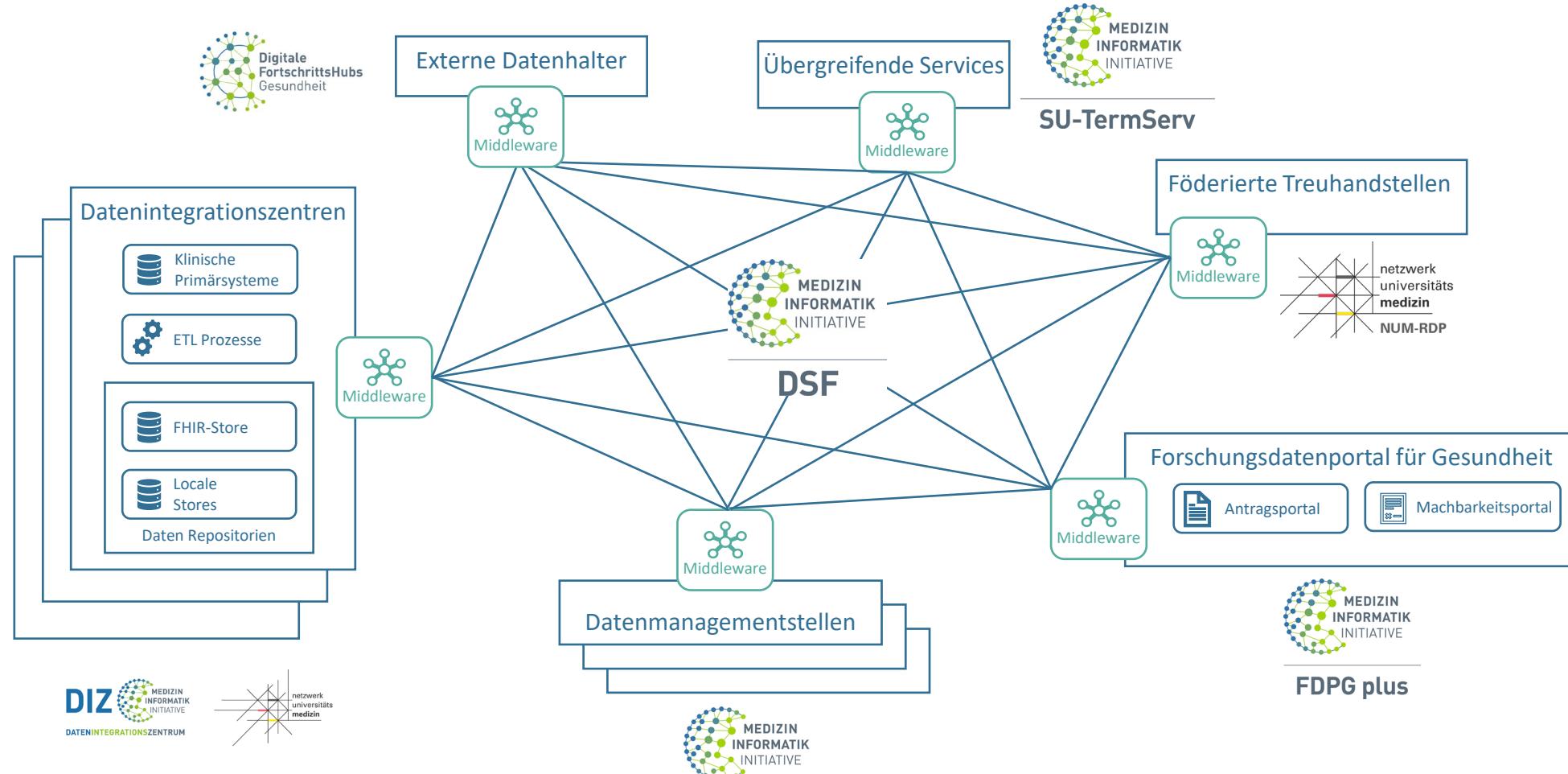
- Ziel:
  - Aufbau einer föderierten Forschungsdateninfrastruktur für sekundäre Nutzung von Routine Daten
  - Automatisierte Prozesse für die Kohortenselektion, die Datenselektion, -transformation und -Ausleitung
  - Netzwerk von medizinischen Datenintegrationszentren (DIZ) wurden an deutschen Unikliniken aufgebaut
- Herausforderungen:
  - Daten einzelner DIZ reichen für viele Forschungsfragen nicht aus  
→ Forschende müssen Daten dezentral anfragen können
  - DIC sind heterogen in ihren Datenmodellen, Primärsystemen, IT-Architekturen und lokalen Gesetzgebungen
  - Datenhaltung verbleibt grundsätzlich Dezentral bei den DIZ  
→ Projektspezifische Datenzusammenführung über Datenmanagementstellen (DMSt)
  - Hoher Grad an automatisierten Prozessen ohne zentrale Speicherung von Daten



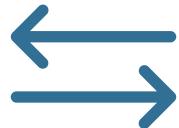
DIFUTURE  
HiGHmed  
MIRACUM  
SMITH  
Coordination office



# MII Projekt Übersicht



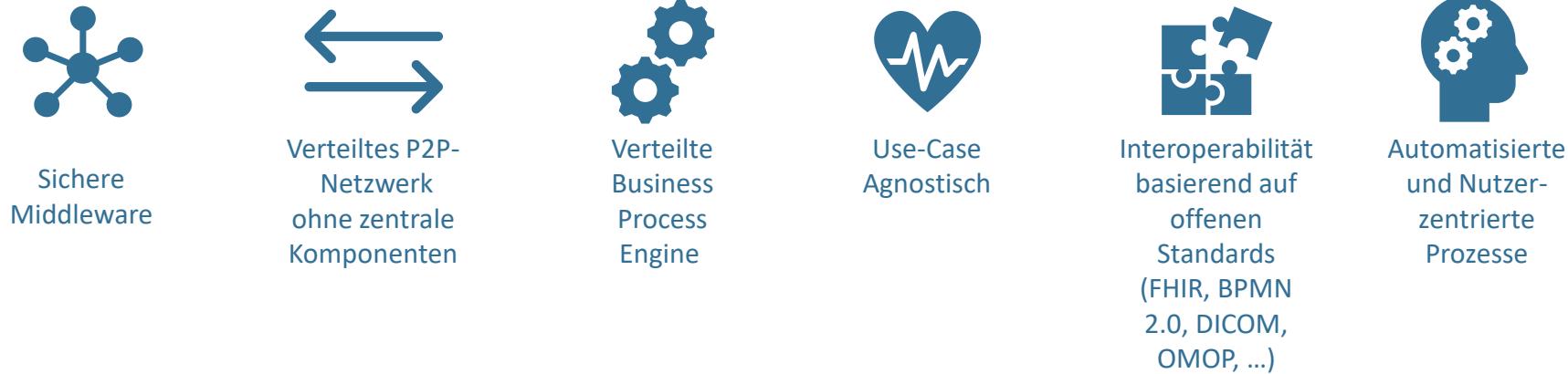
**Organizations-  
übergreifendes  
Data Sharing**



**Verteilte Prozesse**

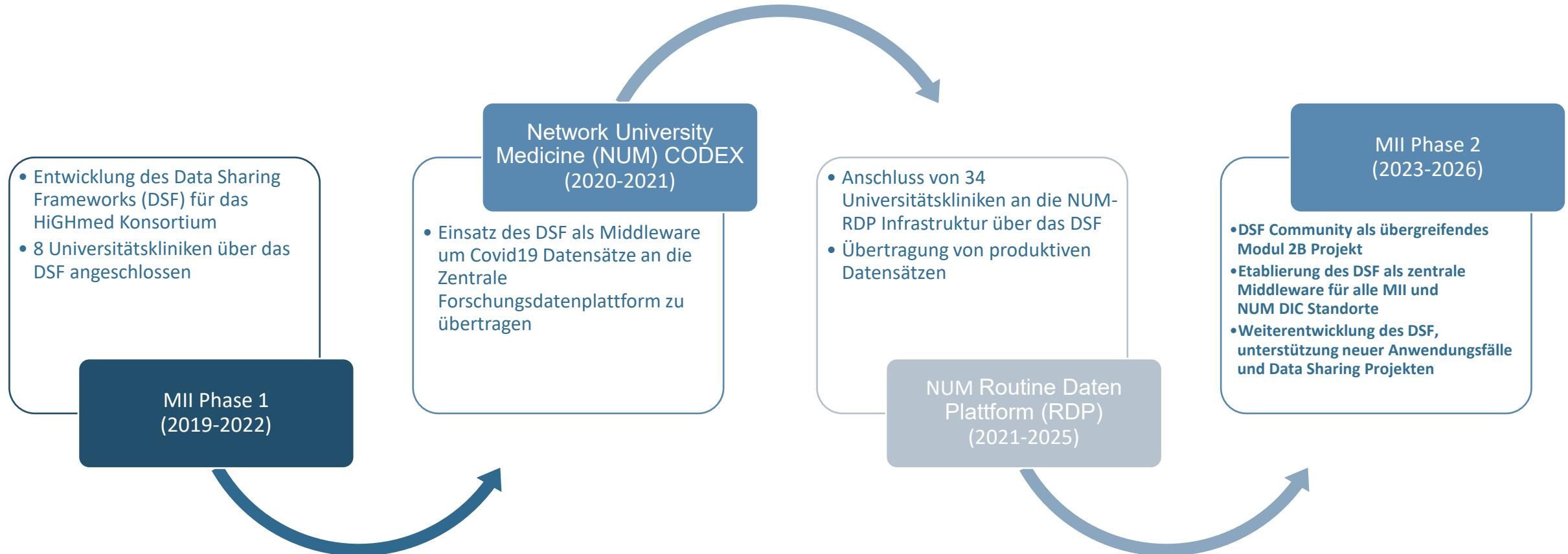
**Ziel:** Syntaktische-, Semantische und  
Prozessinteroperabilität herzustellen

# Data Sharing Framework – Lösung für organizationsübergreifendes Data Sharing



**Das DSF ist ein Konzept für eine sichere Middleware zur Verteilung von Data-Sharing Prozessen, die darauf abzielt, in mehreren verteilten DIZ gespeicherte Daten zu extrahieren, zusammenzuführen, zu pseudonymisieren und bereitzustellen.**

# Meilensteine des DSF in mehreren Projekten



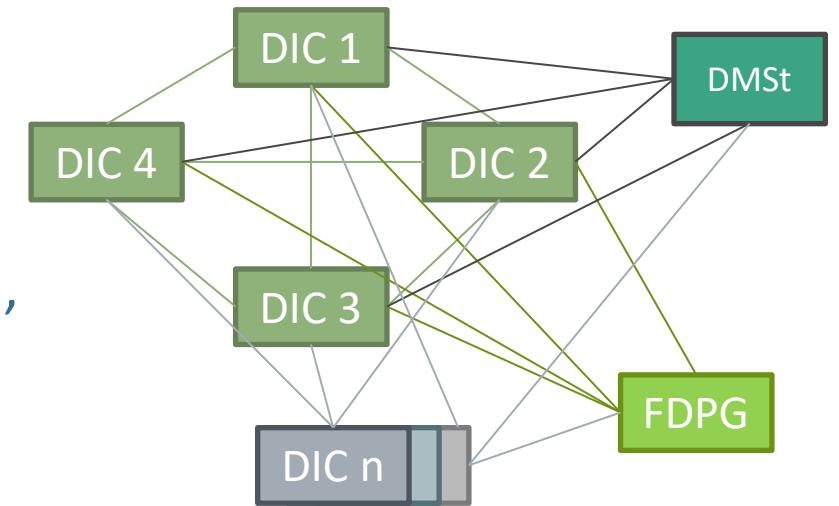
## Die übergreifenden Ziele der DSF-Community sind:

---

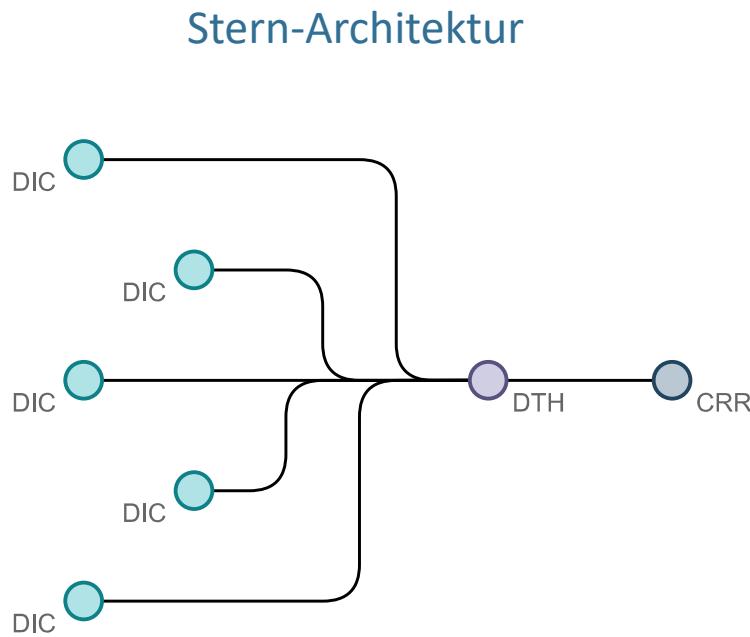
1. Weiterentwicklung der DSF-Kernkomponenten in einem getakteten qualitätsgesicherten Release-Schema
2. Grundlegenden Anwendungsfall der MII für die gemeinsame Datennutzung zusammen mit den anderen Modul 2b-Projekten wie FDPG+ und TRANSIT weiterzuentwickeln
3. Ein Qualitätssicherungsverfahren für DSF Prozess Plugins zu etablieren
4. Aufbau der DSF-Community als standort- und use-case übergreifendes Netzwerk von Entwicklern und Anwendern in der MII

# Anforderungen an das DSF

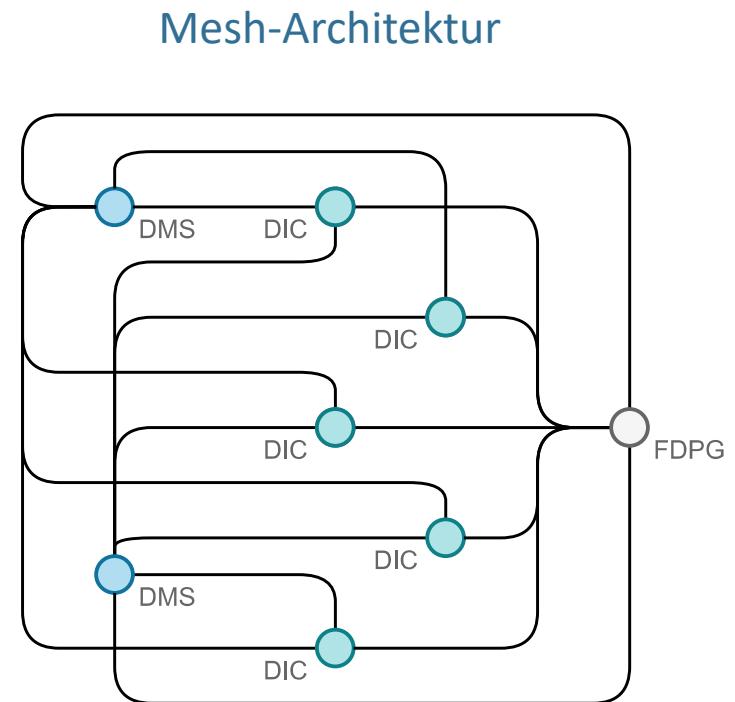
- Verteilte Architektur mit unterschiedlichen Akteuren
- Sichere dezentrale Kommunikationsinfrastruktur, die:
  - Skaliert mit steigenden Kommunikationsbeziehungen
  - Prozesse und Instanzen orchestriert
  - Ausführungslogik von Programmlogik trennt
  - Automatisierte und Nutzer-zentrierte Prozessschritte
  - An heterogene Strukturen angebunden werden können
- Nutzung von standardisierten Kommunikationsprotokollen, Informationsmodellen und offenen Standards
  - HL7 FHIR
  - BPMN 2.0
  - X509 Client- und Server Zertifikate
- Datenschutzkonform zu EU GDPR, nationalen und regionalen Gesetzgebungen



# Flexible Peer-2-Peer Architektur

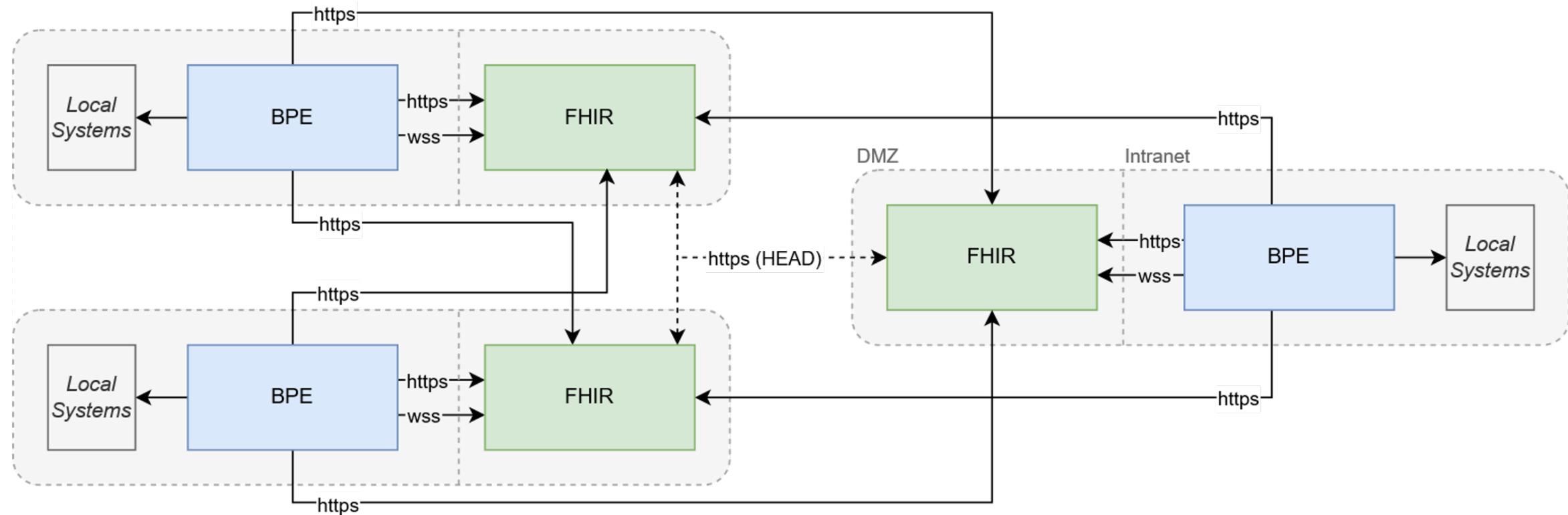


Bildet einen Use-Case ab



Ergibt sich aus der Nutzung  
durch mehrere Use-Cases

# Architektur und Funktionsweise des DSF

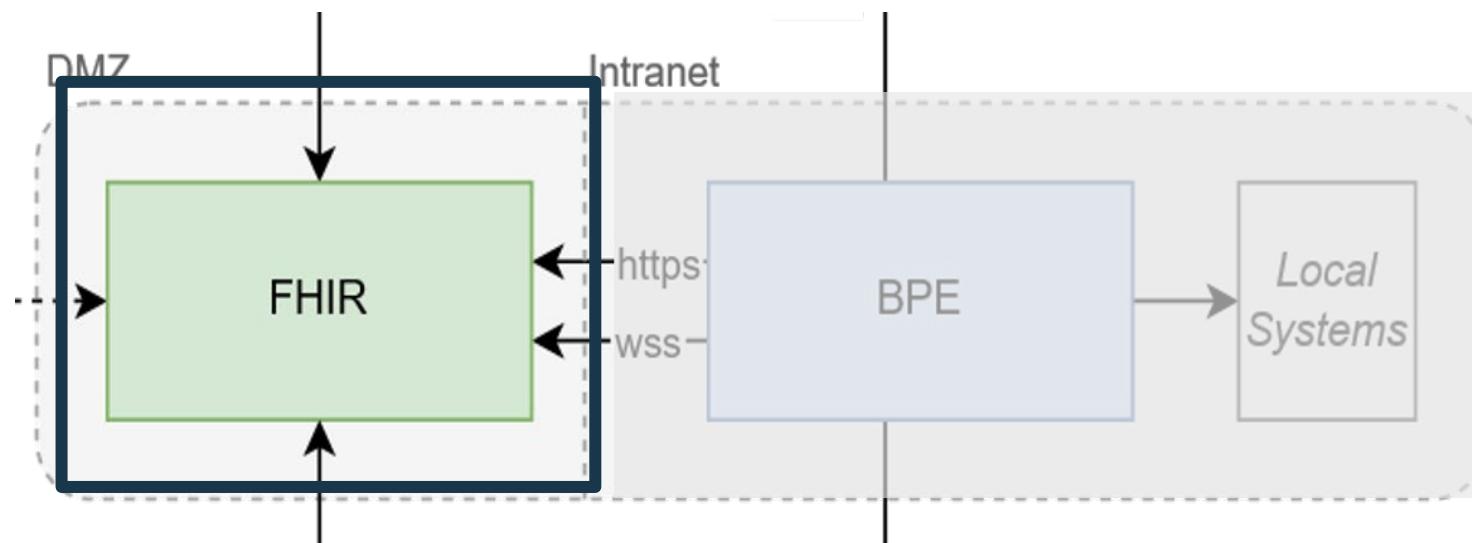


# DSF FHIR Mailbox

## “Mailbox”

- Kommunikation / Auditing
- Deployed in demilitarized Zone (DMZ)
- Keine Persistierung von medizinischen Daten

## *FHIR Task Ressourcen für Prozess Management*



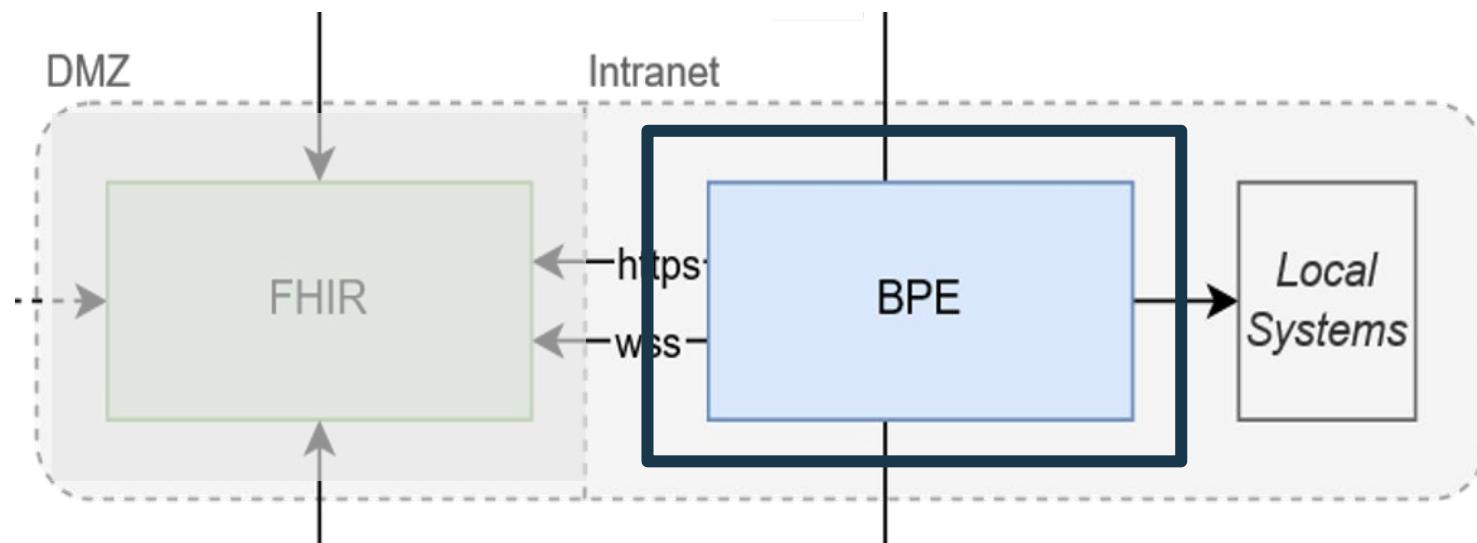
# DSF Business Process Engine (BPE)

## *Führt Prozesse aus*

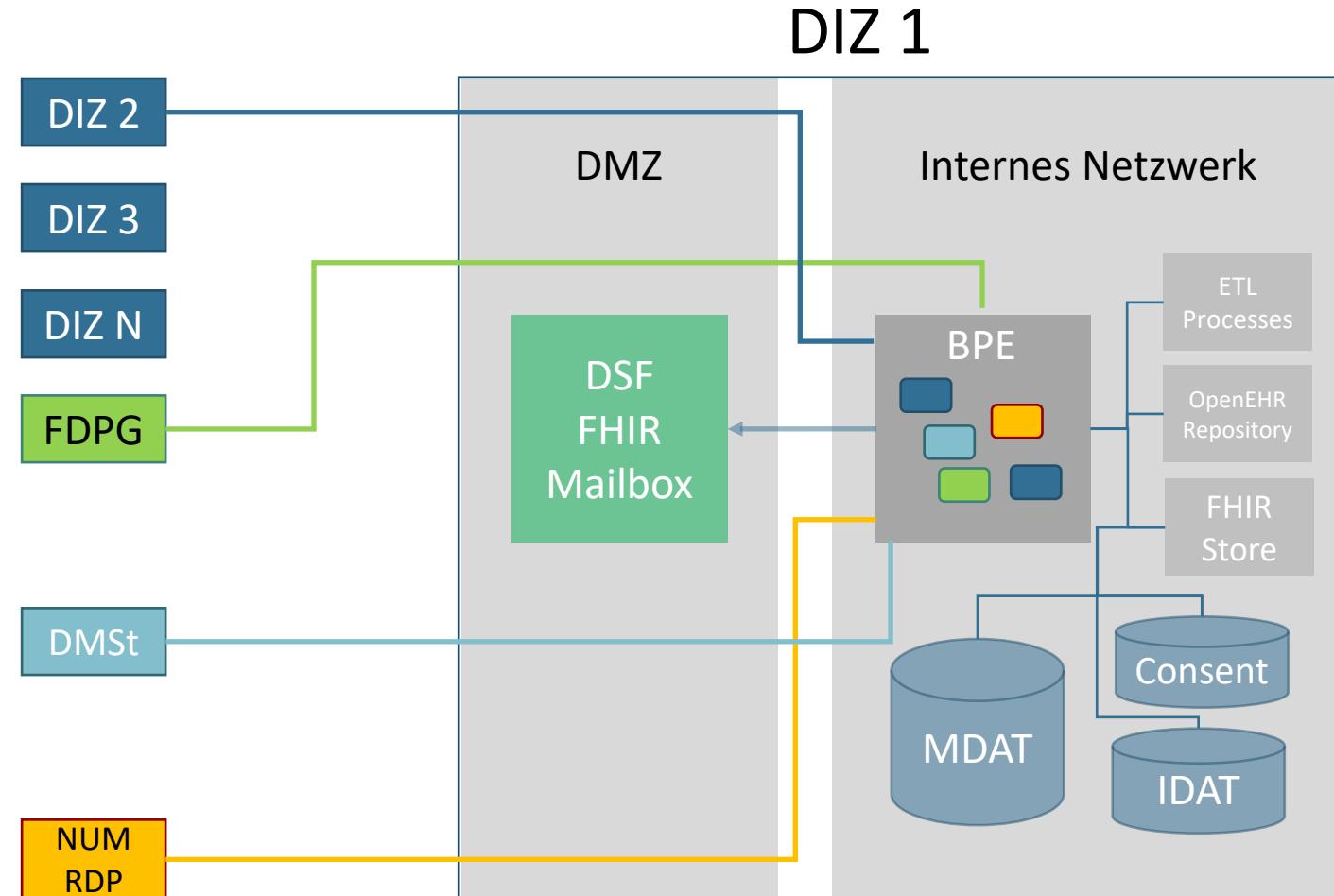
- Deployed im internen Netz einer Organisation
- Parallel Ausführung von Prozessen
- “Plugin” Konzept

## *Ausführung getriggert durch*

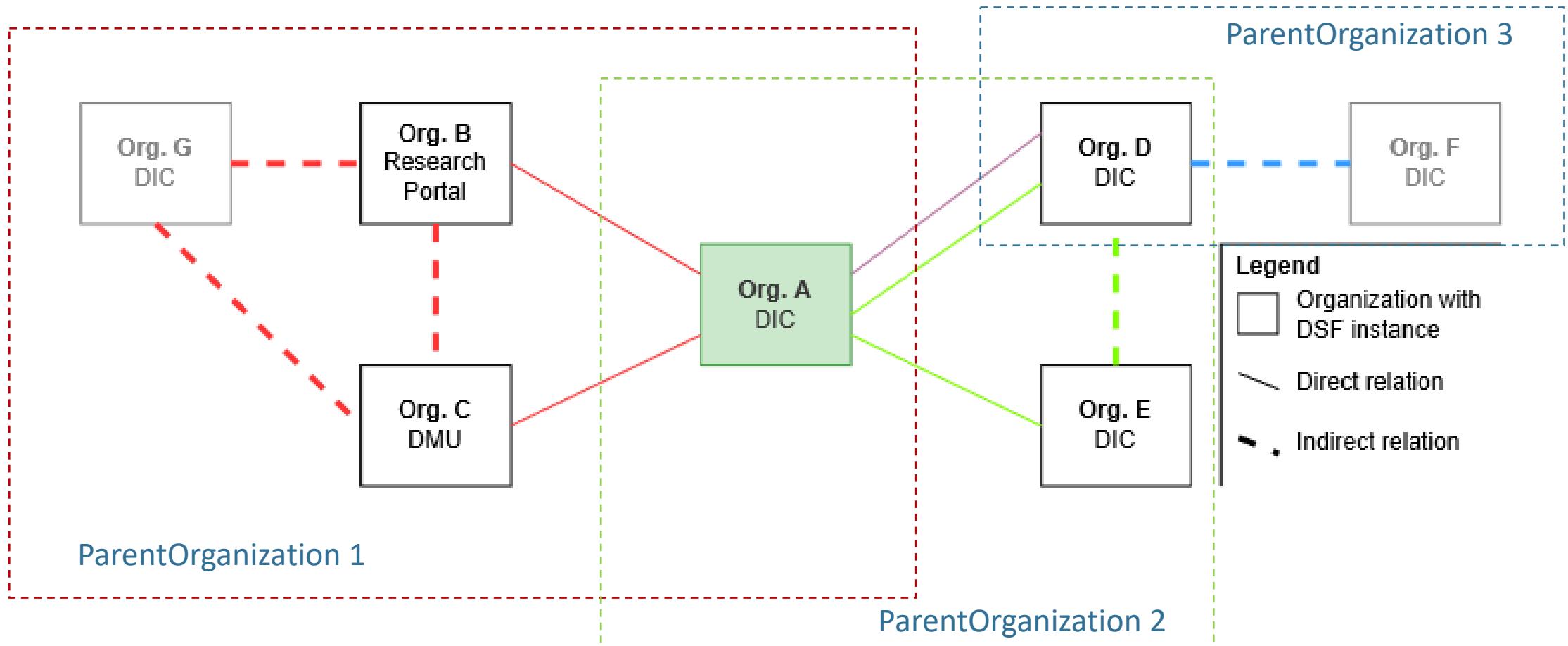
- Task Ressourcen | FHIR Websocket Subscription
- Timer, System Starts



# DSF Architektur: Parallele Plugins & Anbindung an interne Systeme



# DSF Allow-List Konzept: Authentifizierung und Autorisierung



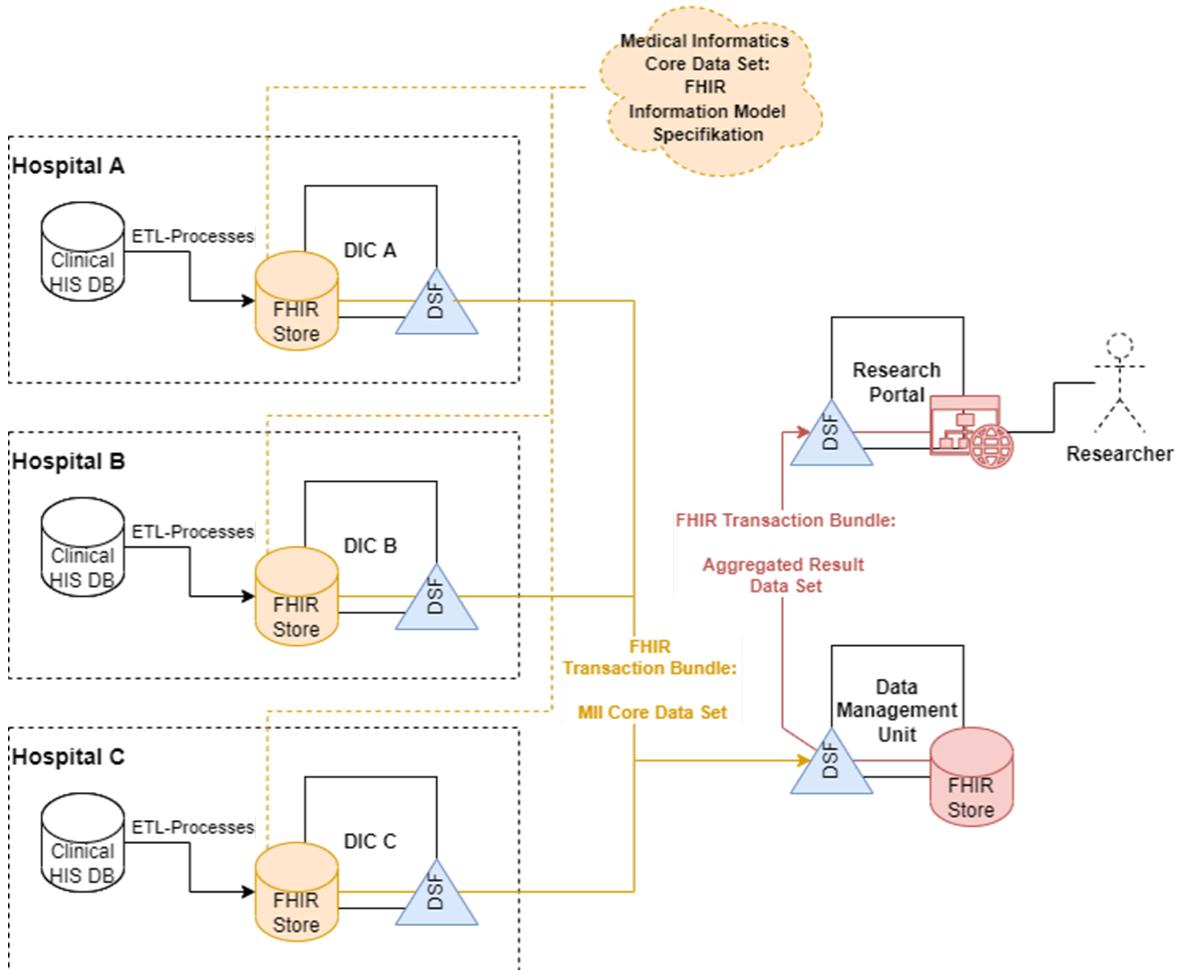
# Verteilte Prozesse und Prozess Plugins

---

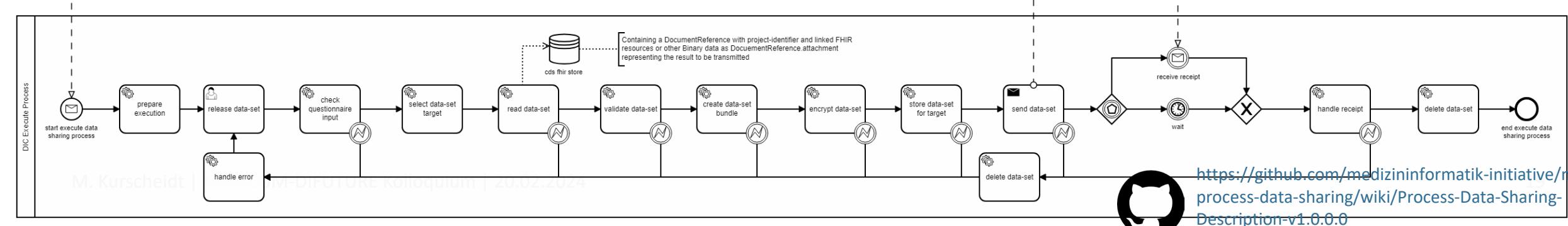
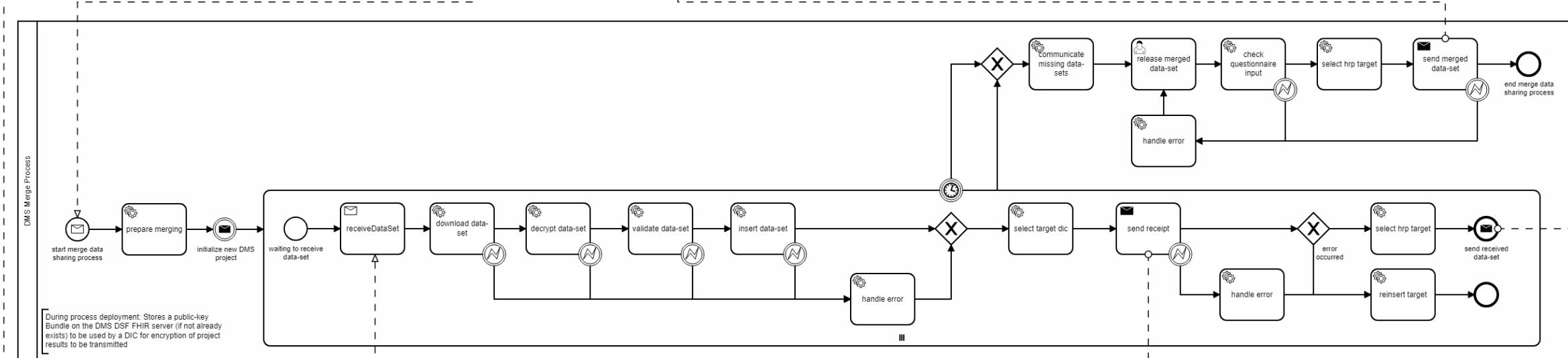
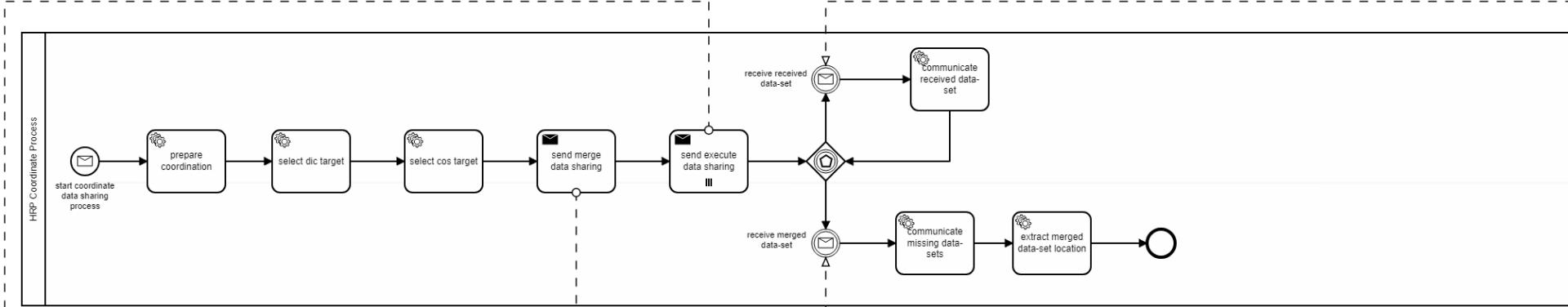
- Verteilte Prozesse werden als Prozess Plugins für das DSF entwickelt
  - Die verteilten Prozesse werden in BPMN modelliert
  - Die Prozess Plugins werden über die BPE des DSF ausgeführt
  - Über FHIR Task Ressourcen werden die Prozesse Organisationsübergreifend gestartet
- MII Prozesse die parallel über ein DSF ausgeführt werden können
  - FDPG Machbarkeitsanfragen
  - Datenausleitungsprozess der MII
  - Datenausleitungsprozess des NUM-RDP
  - FDPG Schaufenster
  - ...
- Weitere verteilte Prozesse und Use-Cases können für das DSF implementiert werden
  - DSF Infrastruktur an allen DIZ implementiert in TEST- und PROD-Systemen

## MII Data Sharing Prozess

- Das Forschungsdatenportal für Gesundheit koordiniert die Datenausleitung von DIZ zu DMSt
- Der MII Data Sharing Process besteht aus folgenden Teilprozessen
  - Coordinate Process – FDPG
  - Execute Process – DIZ
  - Merge Process – DMSt



# DATA SHARING FRAMEWORK

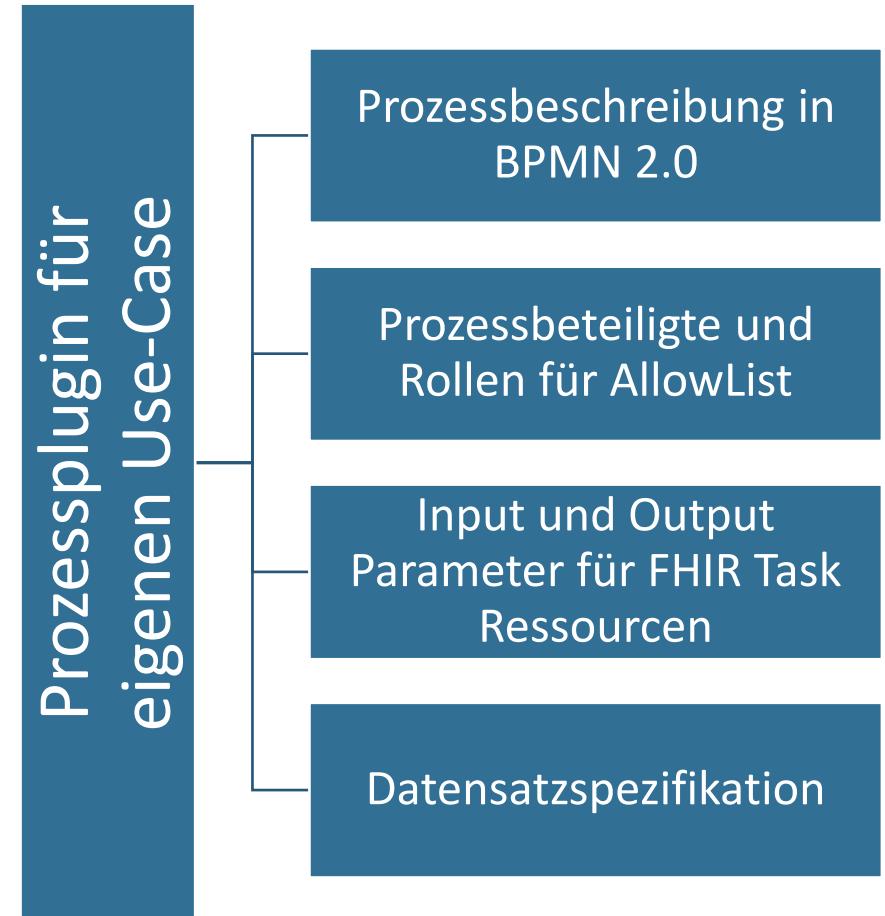



# DSF als Infrastrukturkomponente für eigene Data Sharing Vorhaben nutzen



- DSF ist in allen DIZ als Infrastrukturkomponente als Test- und Produktiv-Umgebung ausgerollt
- Bestehende Prozessplugins können in unterschiedlichen Kontexten und mit wechselnden Beteiligten wiederverwendet werden
- Flexible „Payloads“ können über das DSF übertragen werden
- Gründe für eigenes Prozessplugin
  - Prozesslogik weicht von den bestehenden Prozessen ab
  - Modul 3 Projekte weichen von den Standard FDPG Prozessen ab
  - Neue Prozessrollen müssen definiert werden
  - Bestehende Prozesse erlauben bestimmte Payloads nicht

→ Sprecht die DSF Community gerne für Beratung und Fragen an!



## DSF Community: Teilnahmemöglichkeiten



Aufbau einer Community aus DSF-Nutzern und Entwickler



Austausch mit der Community über Zulip, GitHub oder „in-persona“



Fragen von anderen Nutzern beantworten



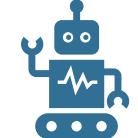
Feature Requests stellen



Testen neuer DSF Releases und Prozess-plugins



Code und Dokumentation beisteuern



Bugs melden und Fixen



Neue Prozessplugins und Tools Entwickeln

- Das DSF und die meisten Prozessplugins sind unter Apache 2.0 Lizenz als Open-Source Software veröffentlicht
- Eigene verteilte Prozesse können Entwickelt werden und auf der verteilten DSF Infrastruktur ausgeführt werden
  - Flexible Anbindung an weitere Systeme, Oberflächen etc.
- Jede\*r ist herzlich eingeladen teil der DSF-Community zu werden!

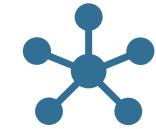
# DSF Spring School für Prozess Plugin Entwickler und Hackathon



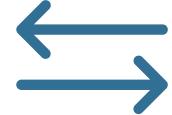
- Findet am **12-14.03.2024** in Heilbronn statt
- Kostenlose Teilnahme (ohne Unterkunft und Anreise)
- Anmeldung bis 29.02.2024 über <https://eveeno.com/dsf-spring-school-2024>
- Programm
  - **Dienstag, 12.03. - Mittwoch, 13.03. Vormittag:**
    - **Track 1: Fachliche Diskussionen** – Dieser Track richtet sich an Domänenexpert\*innen und fokussiert sich auf Konzepte und Diskussionen zu unseren Themenschwerpunkten
    - **Track 2: Praktische Einführung in die DSF-Prozesspluginentwicklung** – Ideal für Entwickler, die lernen möchten, wie DSF-Prozessplugins für neue Use-Cases oder lokale Anforderungen entwickelt werden.
  - **Mittwoch, 13.03. Nachmittag und Donnerstag, 14.03.:**
    - **Praktische Sessions in Kleingruppen**, z.B. Entwicklung von Prozessplugins, Erarbeitung von Prozessmodellen und Spezifikationen, Entwicklung von Implementation Guides, Tools, Pipelines, ... und Ihre weiteren Ideen.



# Data Sharing Framework Community - „Take-Home Messages“



Sichere  
Middleware



Verteiltes P2P-  
Netzwerk  
ohne zentrale  
Komponenten



Verteilte  
Business  
Process  
Engine



Use-Case  
Agnostisch



Interoperabilität  
basierend auf  
offenen  
Standards (FHIR  
& BPMN 2.0)



Automatisierte  
und  
Nutzerzentriert  
e Prozesse

1. Organisationsübergreifendes Data-Sharing = Verteilte Prozesse
2. Prozess Plugins implementieren Features = Flexibilität der Infrastruktur
3. AllowList Konzept = Authorisierung und Authentifizierung
4. 1 DSF Instanz kann N Use-Cases parallel unterstützen

## Community Ansatz – Hilfe zur Selbsthilfe!

- Aufbau und Teilnahme an der Open-Source Community
- Breite Verbreitung des DSF als Infrastrukturkomponente in den DIZ und darüber hinaus
- Neue Anwendungsfälle können durch eigene Prozess-Plugins implementiert werden

# Publikationen



## Publikationen

- H. Hund, R. Wettstein, C.M. Heidt, C. Fegeler,  
**Executing Distributed Healthcare and Research Processes – The HiGHmed Data Sharing Framework,**  
Stud Health Technol Inform, 278 (2021) 126-133, doi:10.3233/SHTI210060
- R. Wettstein, H. Hund, I. Kobylnski, C. Fegeler, O. Heinze,  
**Feasibility Queries in Distributed Architectures – Concept and Implementation in HiGHmed,**  
Stud Health Technol Inform, 278 (2021) 134-141, doi:10.3233/SHTI210061
- C.M. Heidt, H. Hund, C. Fegeler,  
**A Federated Record Linkage Algorithm for Secure Medical Data Sharing,**  
Stud Health Technol Inform, 278 (2021) 142-149, doi:10.3233/SHTI210062
- R. Wettstein, H. Hund, C. Fegeler, O. Heinze,  
**Data Sharing in Distributed Architectures – Concept and Implementation in HiGHmed,**  
Stud Health Technol Inform, 283 (2021) 111-118, doi:10.3233/SHTI210548
- H.-U. Prokosch, T. Bahls, M. Bialke, J. Eils, C. Fegeler, J. Gruendner, B. Haarbrandt, C. Hampf, W. Hoffmann, H. Hund, et al.  
**The COVID-19 Data Exchange Platform of the German University Medicine,**  
Stud Health Technol Inform, 294 (2022) 674-678, doi:10.3233/SHTI220554
- R. Wettstein, T. Kussel, H. Hund, C. Fegeler, M. Dugas, K. Hamacher,  
**Secure Multi-Party Computation Based Distributed Feasibility Queries – A HiGHmed Use Case,**  
Stud Health Technol Inform, 296 (2022) 41-49, doi:10.3233/SHTI220802
- H. Hund, R. Wettstein, C. Hampf, M. Bialke, M. Kurscheidt, S.T. Schweizer, C. Zilske, S. Mödinger, C. Fegeler,  
**No Transfer Without Validation: A Data Sharing Framework Use Case,**  
Stud Health Technol Inform, 302 (2023) 68-72, doi:10.3233/SHTI230066
- H. Hund, R. Wettstein, M. Kurscheidt, S.T. Schweizer, C. Zilske, C. Fegeler,  
**Interoperability Is a Process- The Data Sharing Framework,**  
Stud Health Technol Inform, 310 (2024) 28-32, doi:10.3233/SHTI230921



# Nehmt Kontakt zur DSF-Community auf!



<https://dsf.dev>



<https://github.com/datasharingframework/dsf>



[DSF-Gecko@hs-heilbronn.de](mailto:DSF-Gecko@hs-heilbronn.de)



<https://mii.zulipchat.com>

Maximilian Kurscheidt

[Maximilian.Kurscheidt@hs-heilbronn.de](mailto:Maximilian.Kurscheidt@hs-heilbronn.de)

GECKO Institut für Medizin, Ökonomie und Informatik,  
Hochschule Heilbronn