MEDIZIN INFORMATIK INITIATIVE

Federal Ministry of Education and Research

miracum
Medical Informatics in Research and Care in University Medicine

# DIZ & MIRACOLIX

MEDICAL INFORMATICS REUSABLE ECO-SYSTEM OF

OPEN SOURCE LINKABLE AND INTEROPERABLE

SOFTWARE TOOLS – X

Prof. Dr. Martin Sedlmayr

Institut für Medizinische Informatik und Biometrie

Zentrum für Medizinische Informatik

TU Dresden / Universitätsklinikum Dresden

# 1001 Tools

miracum

Reporting Analysen Integration

Connect

Platform

tranSMART FOUNDATION

i2b2

Galaxy PROJECT

OHDSI

eConsent

mirth

talend

MDR

iRODS SMART

Terminologieserver

Audit

elasticsearch logstash kibana

Connect

Infrastructure

openstack

vmware

docker

KVM

hadoop

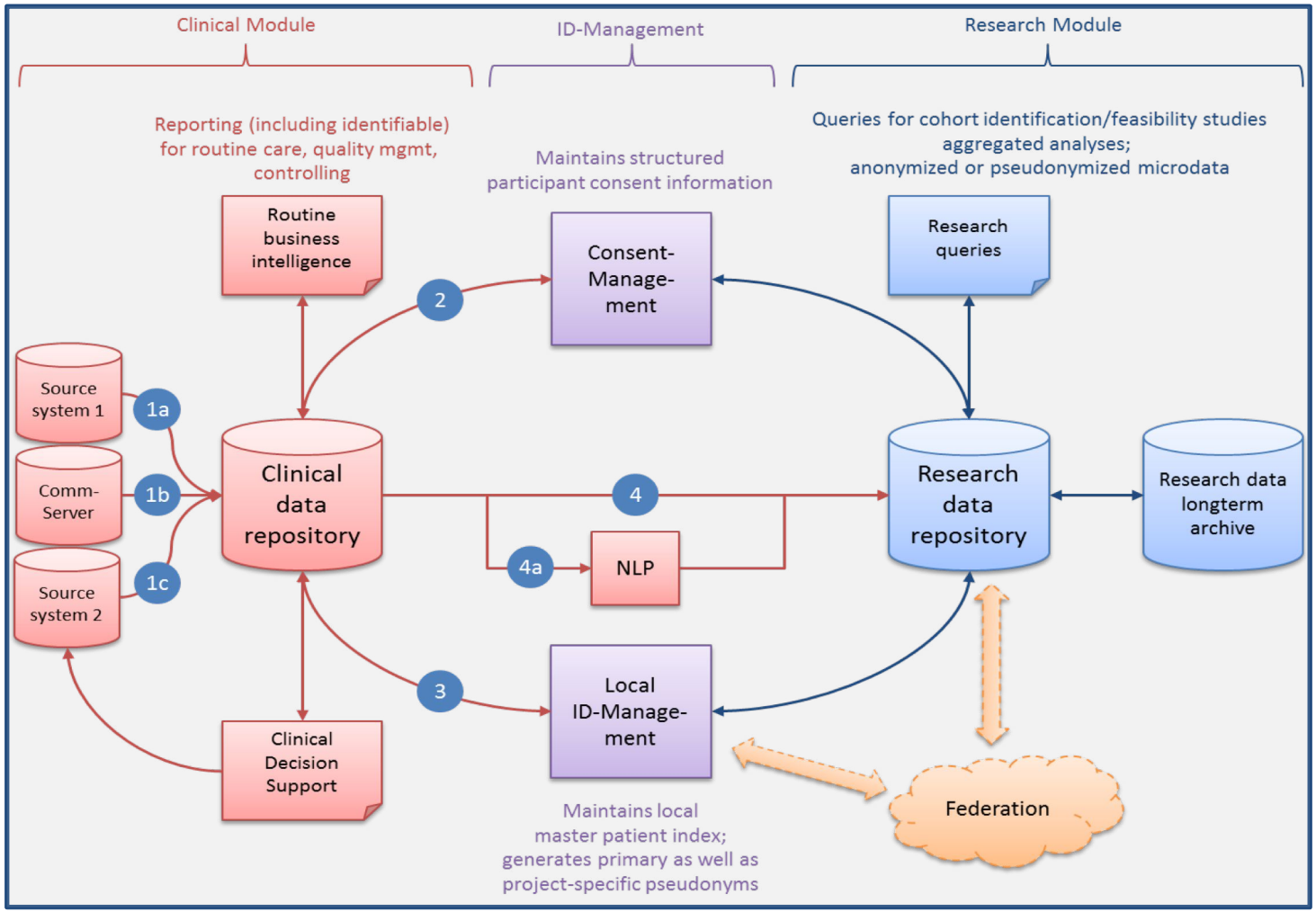
Spark

PostgreSQL

RDF

HBASE

# Die prinzipielle Architektur eines DIZ



# MIRACOLIX
# Das MIRACUM Ökosystem

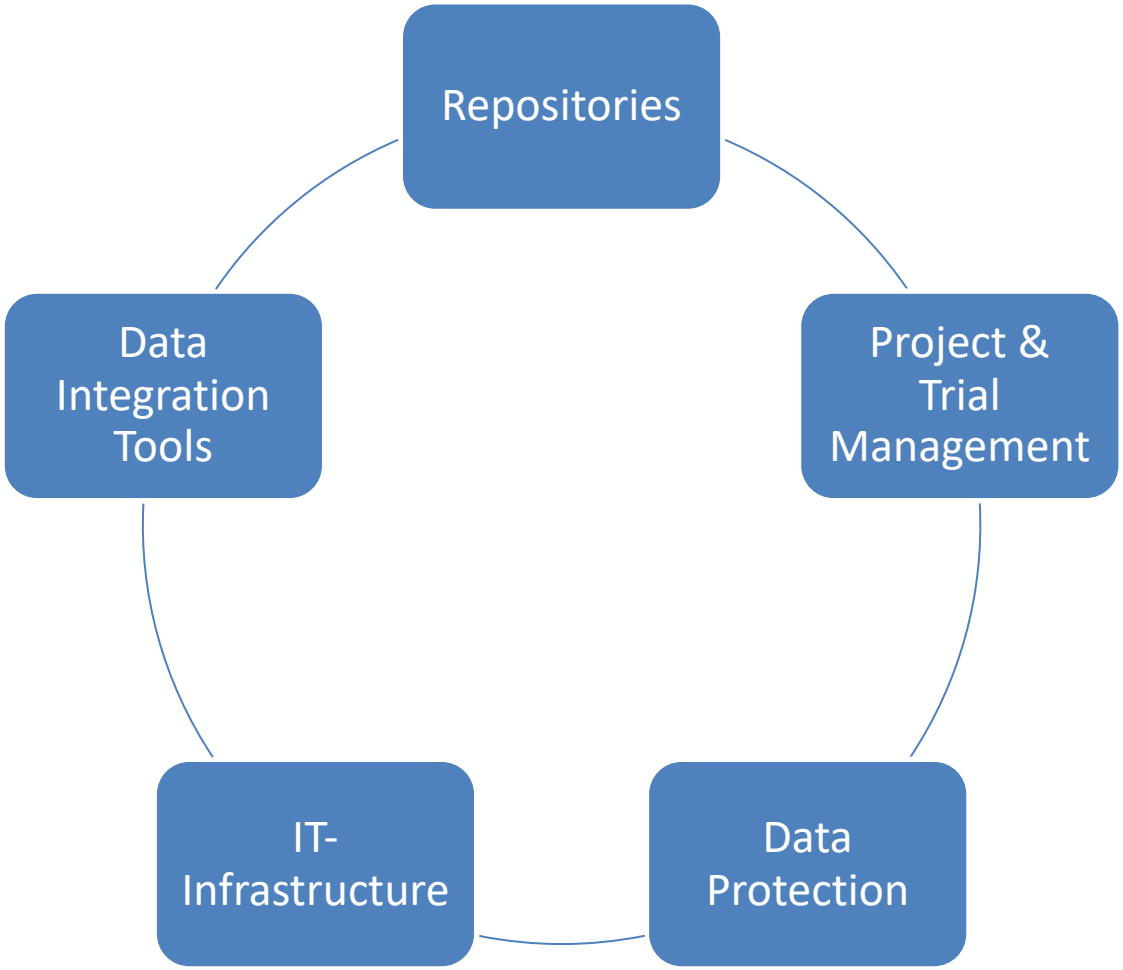
miracum

- Medical Informatics ReusAble eCo-system of Open source Linkable and Interoperable software tools – X

**MIRACOLIX**

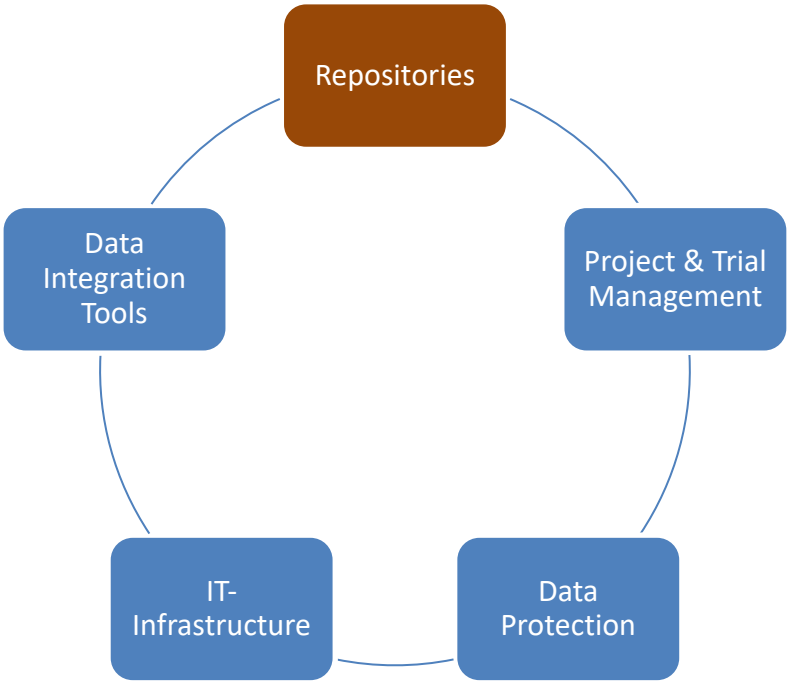
- MIRACOLIX is
  - pragmatic
  - modular
  - reusable
  - open source
  - interoperable
  - federated

# Bereiche


Repositories
Project & Trial Management
Data Protection
IT-Infrastructure
Data Integration Tools

# Bereiche
# Speicherung der Daten

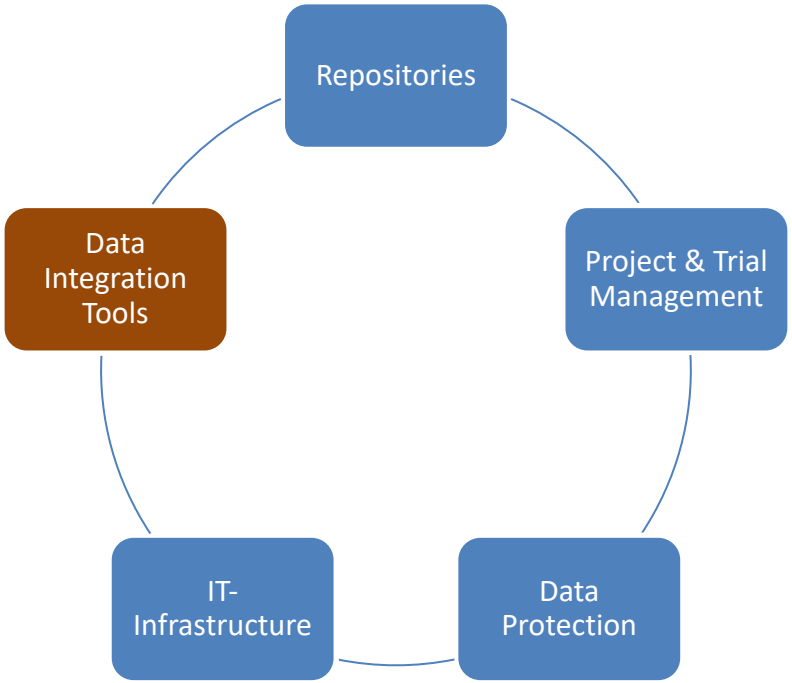
miracum




- Data Repositories, Exploration & Visualisierung
  - i2b2 & IDRT Tools
  - OMOP CDM & OHDSI Tools
  - XNAT
  - tranSMART
  - cBioportal
- Langzeitarchivierung/Sharing
  - iRODS
  - CKAN

# Bereiche
# Erschließung der Daten

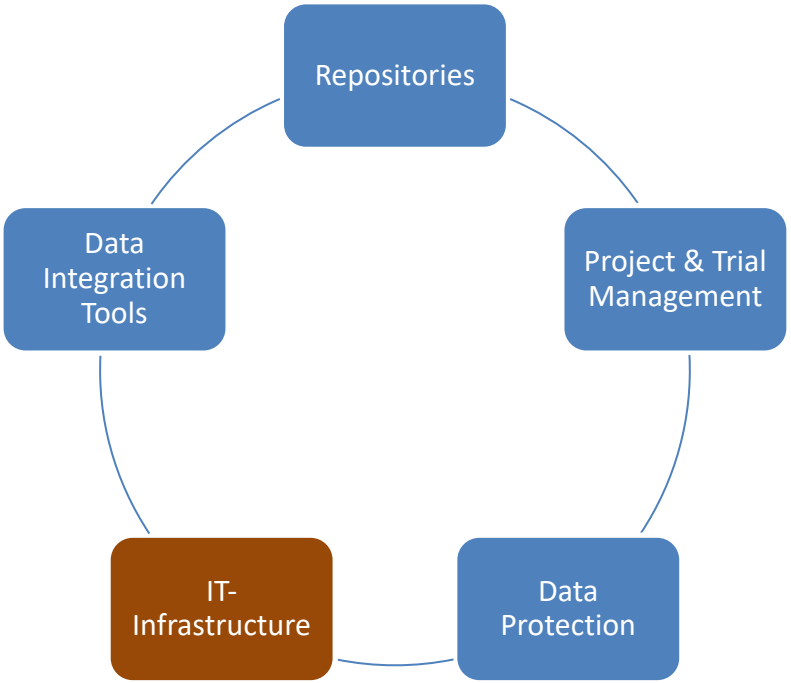
miracum




- MIRACUM Metadata Repository
  - basiert auf Samply.MDR
- ETL Tools
  - Talend Open Studio / pentaho
- Datenharmonisierung/Mapping Tool
  - Erlangen Ontology-based Mapping Tools (GBA / BBMRI ERIC CS IT)
  - MOLGENIS (BiobankConnect)
- Natural Language Processing Tool
  - Averbis Information Discovery Tool
- Connector Komponenten
  - für die Förderation der DIZe
  - basieren auf li2b2, Samply Broker/Connector, OHDSI Arachne

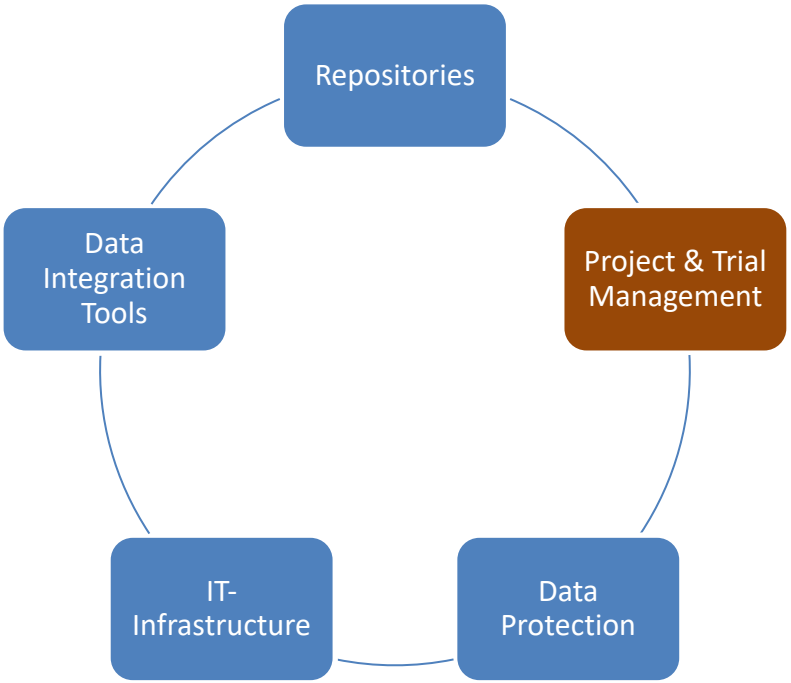
# Bereiche IT-Infrastruktur

miracum



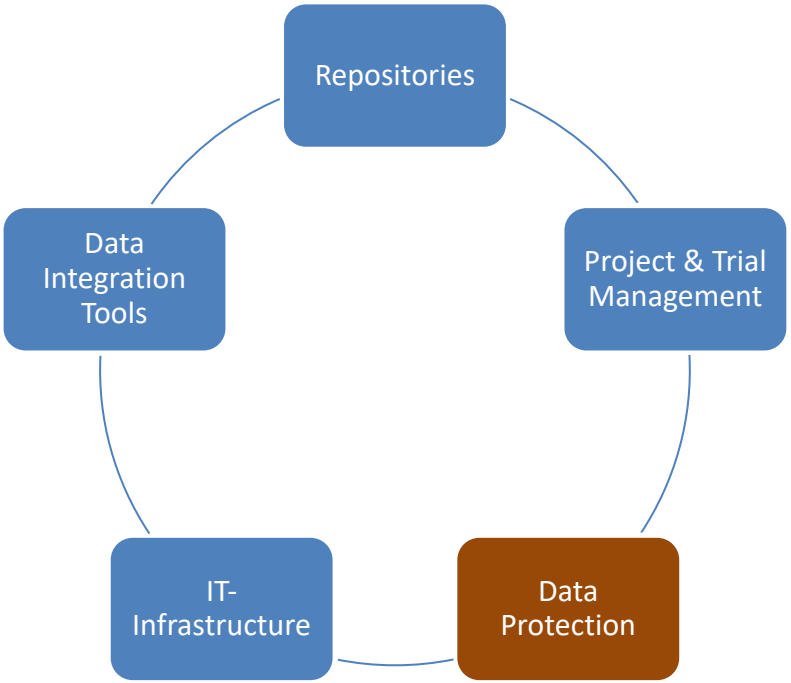

- Förderierte Authentifizierung
  - oAuth
  - existierende Logins via DFN-AAI
- Tools für Entwicklung, Deployment und Monitoring
  - Confluence, JIRA, Bitbucket, ...
  - Docker
- Software Pipelines für die Analyse von omics-Daten
  - teilen, verwenden, lernen
  - Vergleich von Ergebnissen und Qualitätsparametern
  - Galaxy, Galaxy Tool Shed & Docker

# Bereiche Organisation




- Projektverwaltung
  - Atlassian Confluence
  - Arachne
  - ProSKIVE
- Zentrales Projekt-/Studienregister
  - Inkl. aller relevanten Metadaten und Einschlusskriterien
  - SODA (Freiburg)

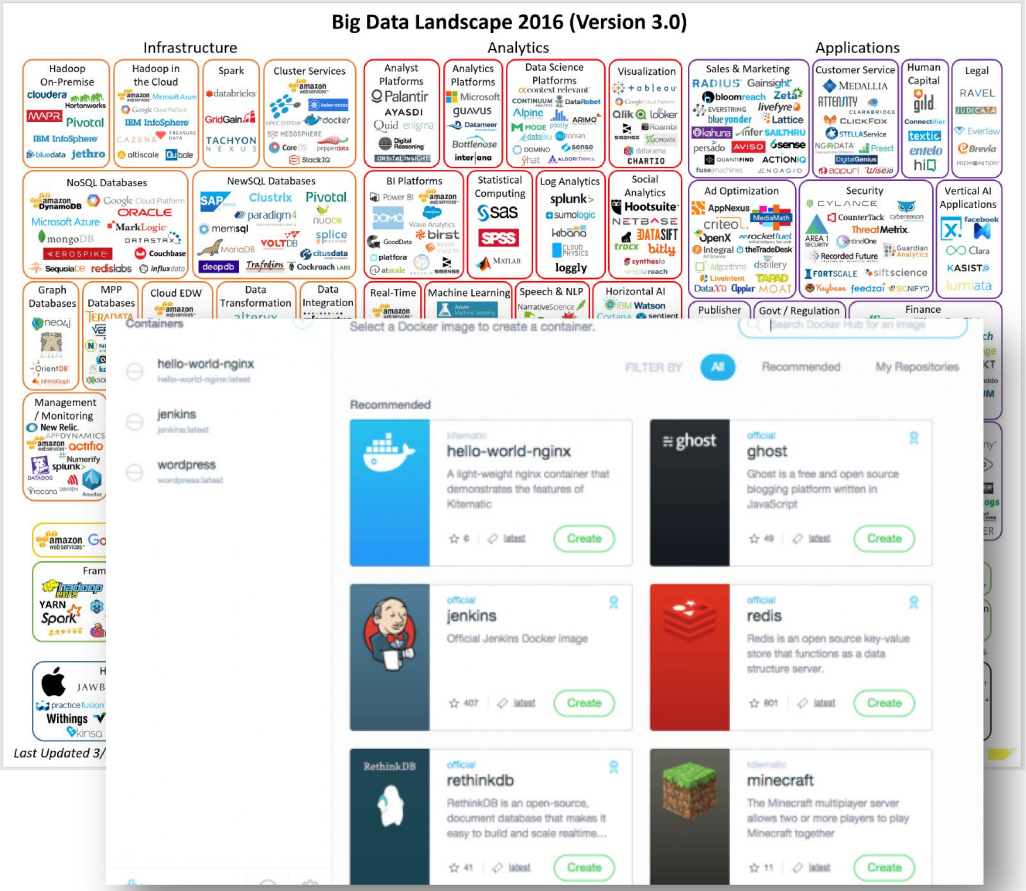
# Bereiche
# Datenschutz und -sicherheit




- ID-Management Tool
  - Mainzelliste, ePIX
- Datenanonymisierung
  - z.B. ARX (http://arx.deidentifier.org/)
- Consent Management Tool
  - gICS (generic informed consent service)

# Deployment

- Kapselung der Services
  - einfache Bereitstellung
  - archivierbar
- Docker
  - Aufbau produktiver Umgebungen
  - Zertifizierung der Images




https://www.datameer.com/blog/big-data-ecosystem/

# Vorgehen
# MIRACUM Competence Center

miracum

- Lead Partner pro Komponente
  - kompetenzgetrieben
  - z.B. Erlangen für i2b2, Dresden für OHDSI, Magdeburg für XNAT, ...

- Verantwortlich für die Entwicklung bzw. Anpassung, stellt Installationskandidaten bereit
- Leistet Support für Partner-DIZe beim Aufbau der Komponenten

- Starke Community

http://clipart-library.com/

# (Einige ☺) Herausforderungen

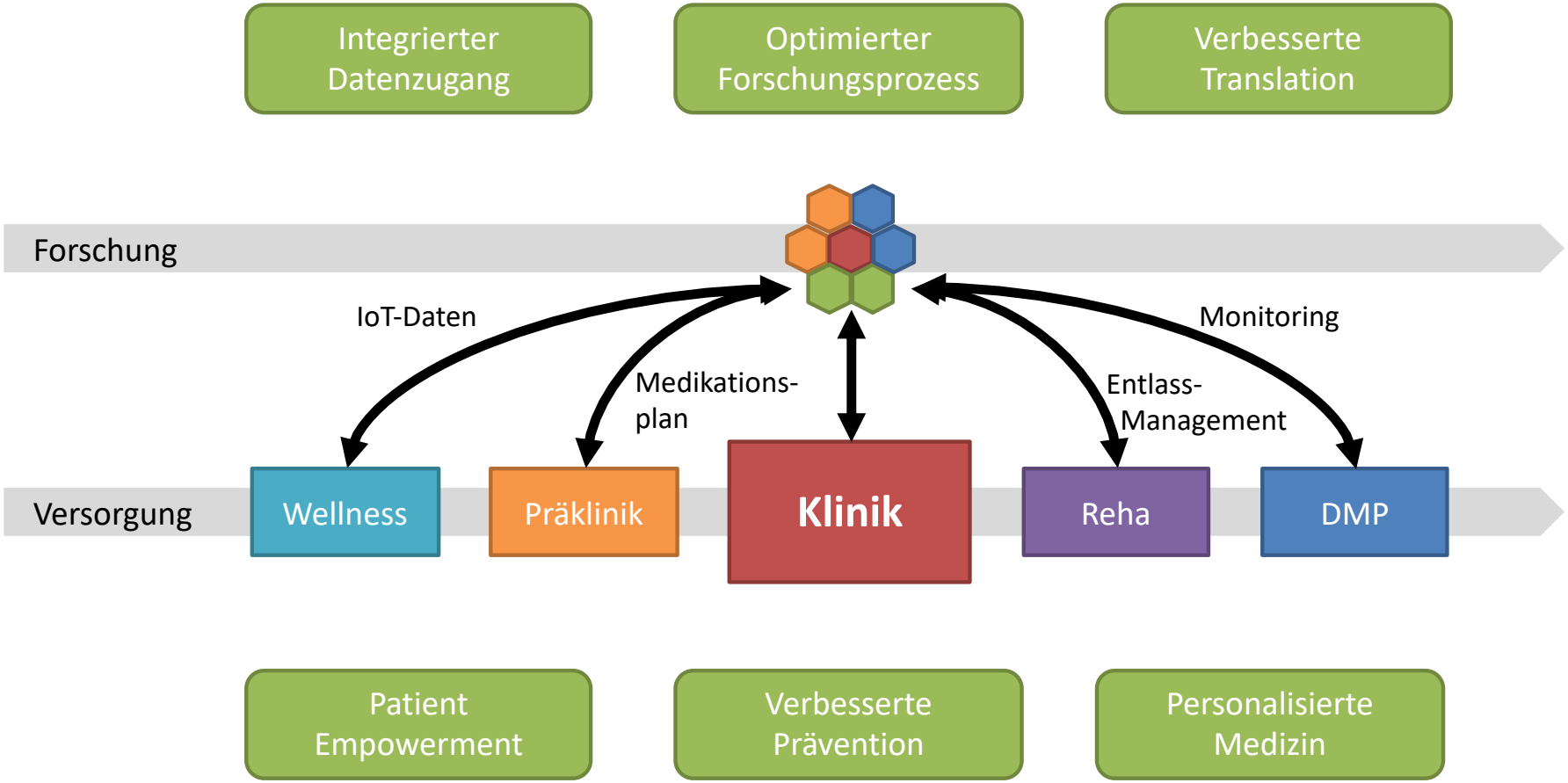
miracum

## Forschung … Versorgung ... Lehre

- Etablierung eines DIZ am Standort
- Heterogenität der Werkzeuge
  - Lizenzen, Releasezyklen, Technologien (docker, FHIR, …)
- Rahmenbedingungen
  - DSGVO, MPG, Zertifizierbarkeit, Auditierbarkeit
- Organisation
  - Zentrales Studienregister
  - Einrichtung UACs
  - Treuhandstellen

# Das DIZ als Kommunikationsdrehscheibe:
*Potenzial und Lust auf mehr*

miracum



# Zusammenfassung

miracum

- MIRACOLIX ist
  - eine dynamische Sammlung von Tools
  - open-source und gut dokumentiert
  - die Referenzimplementierung eines MIRACUM DIZ
- MIRACUM baut auf
  - das Engagement der Competence Center
  - eine aktive Community aus Entwicklern und Anwendern
  - Best Practice Beispiele