

Leuchtturm KI & Bioinformatik

Das Leuchtturmprojekt baut eine Competence Unit für föderierte KI & bioinformatische Analysen (KIBCU) auf. Wir etablieren IT-Infrastrukturen und methodische Kompetenzen, um das Potenzial verteilter onkologischer Datensätze zu nutzen.

Speaker: Prof. Dr. Hans-Ulrich Prokosch, Erlangen

Konzept

BZKF IT Strategie

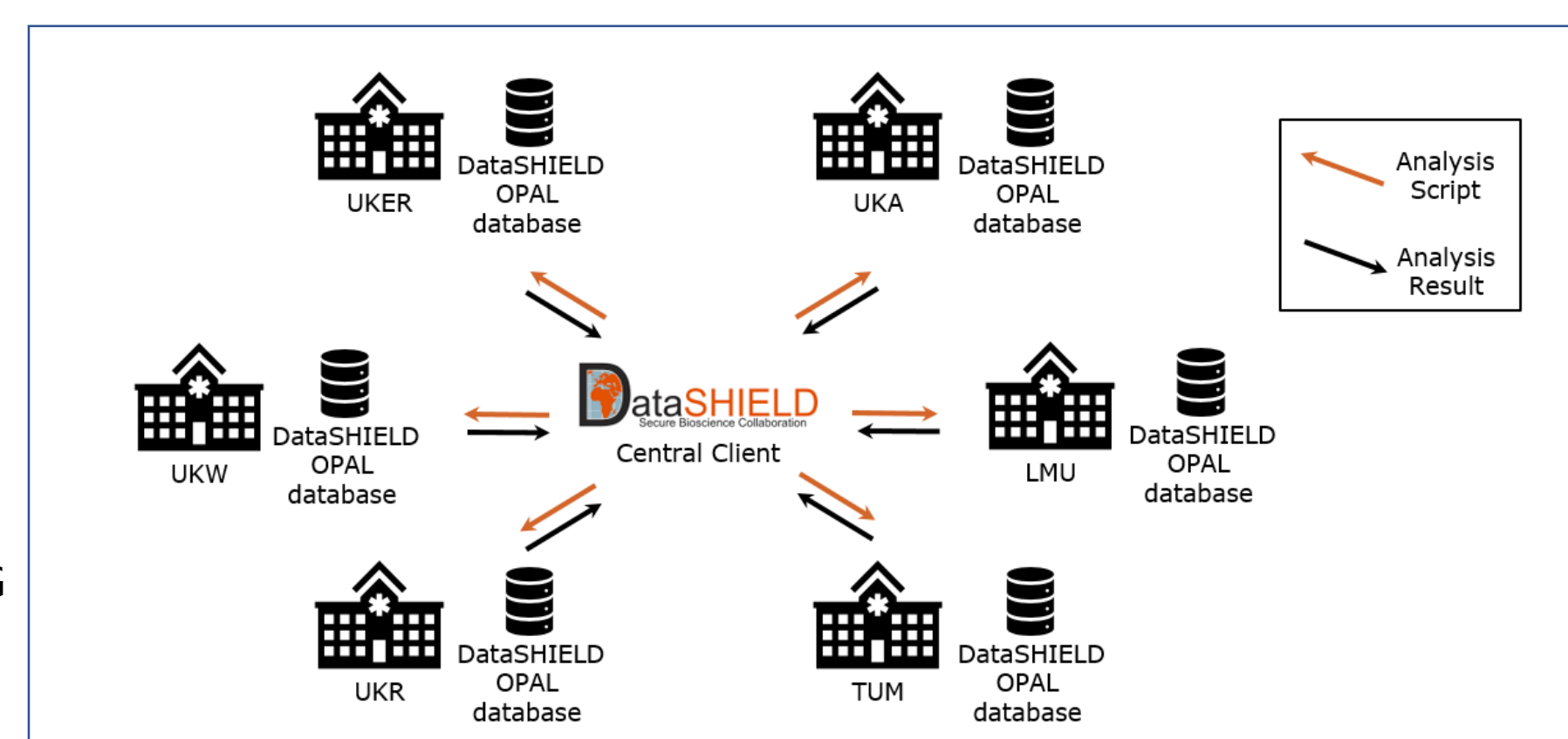
Die IT-Strategie des BZKF basiert auf Konzepten der deutschen Medizininformatik-Initiative mit dem Ziel „Real World Daten“ aus der Patientenversorgung lokal an den Universitätskliniken zu halten und die Analyse zu den Daten zu bringen (föderierte Analyse und föderiertes maschinelles Lernen). Daher kooperieren wir eng mit den bereits an den sechs BZKF-Standorten etablierten Datenintegrationszentren (DIZ) und den standardisierten FHIR-Daten zu Patientendiagnosen, Verfahren, Laborergebnissen und Medikamenten. Um diese Datenspeicher zu erweitern, wurden verschiedene IT-Projekte im BZKF initiiert:

- » **Onkologische Real World Data Platform**
- » **IT support für Molekulare Tumor Boards**
- » **Strukturierte Radiologische Befunde**

KIBCU Workpackages 2024/2025

- **DA1:** Aufbau von Modulen zur Analyse komplexer OMICS Daten
- **DA2:** Integration der histologischen Bilddaten in die Befunde und gemeinsame Nutzbarmachung
- **DA3:** Use Case mit BZKF SG AYA
- **UC1:** Zusammenarbeit mit der BZKG SG Prostata
- **UC2:** Zusammenarbeit mit der BZKF SG Lebertumore
- **WP1:** Regelungen und Prozesse zur Governance der Datennutzung (1/24 – 9/24)
- **WP2:** Datenqualitätsanalysen (1/24-12/25)
- **WP3:** Unterstützung der BZKF-Forscher bei der Entwicklung neuer Modelle des maschinellen Lernens (1/24 – 12/25)
- **WP4:** Entwicklung neuer KI Modelle
- **WP5:** Toolbox für verteiltes maschinelles Lernen
- **WP6:** Kommunikation und fortlaufende Weiterbildung (1/24 – 12/25)

Meilensteine



1. Lokale IT-Infrastrukturen (DataSHIELD) für föderierte/s Analysen/maschinelles Lernen wurden erfolgreich etabliert
2. Entwicklung einer Pipeline für Datenübertragung, Pseudonymisierung und Mapping, um grundlegende onkologische Daten im FHIR Format in die DataSHIELD-Datenbank zu bringen; diese Routine wurde allen BZKF-Standorten zur Verfügung gestellt.
3. Gründung einer BZKF ETL Taskforce zur Aufgaben- und Wissensteilung im Weiterentwicklungsprozess der Pipeline
4. Integration von Datenqualitäts- und Plausibilitätschecks in die Pipeline
5. Gemeinsame Veröffentlichung zuverteilten Analysen mittels eigener Pipeline „onco-analytics-on-fhir“
6. 2. BZKF-Summer School



Langfristige Ziele

- » **Sicherstellen einer Governance-Struktur für Datennutzung und Datenaustausch im Einklang mit dem BZKF-Direktorium**
- » **Datenqualitätsanalyse und kontinuierliches Feedback zur Datenverbesserung**
- » **Unterstützung von Studiengruppen und medizinischen Forschern bei der Bioinformatikanalyse und AI-Entwicklungen**
- » **Verfolgung der Entwicklung und Analyse föderierter KI-Modelle basierend auf der etablierten Infrastruktur**
- » **Entwicklung neuer Algorithmen für föderiertes maschinelles Lernen auf verschiedenen Datentypen**
- » **Kontinuierliche BZKF-weite Bildung und Schulung in föderierten maschinellen Lernmethodologien**

Autoren: Prokosch HU, Braren R, Klauschen F, Raffler J, Rückert D, Spang R