

Integration, Provision, and Semantic Annotation of Radiological Imaging Data using Innovative Methods

**MIRACUM DIFUTURE
SYMPOSIUM**
München, 29. Oktober 2024

Björn Schreiweis

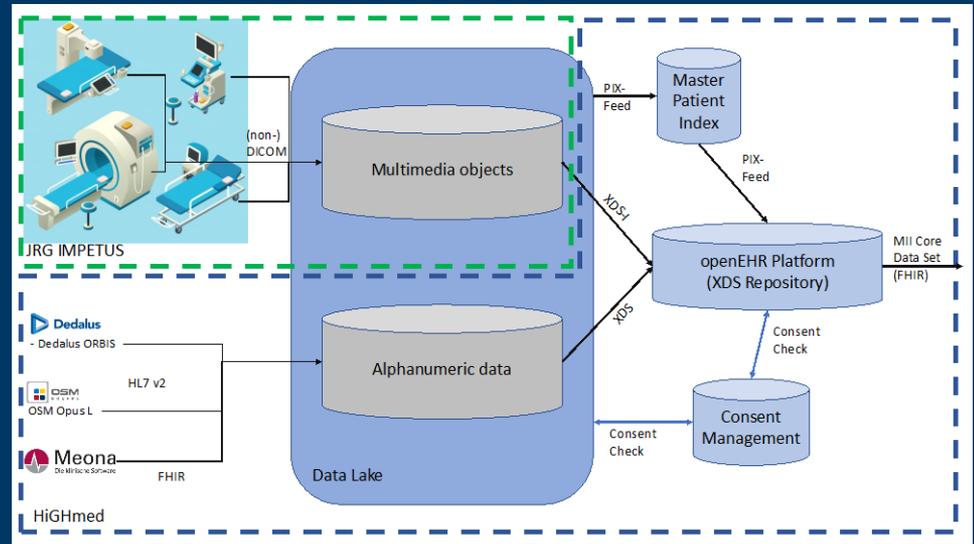
Für die Nachwuchsgruppe
IMPETUS

Nachwuchsgruppe IMPETUS

Integrating **M**ultimedia Objects and **P**ACS Environments in **T**o **U**niversal Knowledge Management **S**ystems



Multimedia goes MeDIC

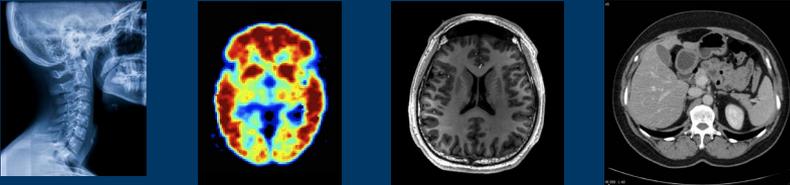


Klinische Bildgebung

Definition:

Klinische Bildgebung beinhaltet die Verwendung verschiedener Technologien um visuelle Repräsentationen des Körperinneren für klinische Analysen und medizinische Interventionen zu liefern.

Beispiele:



Charakteristika:

- Bereitstellung nicht-invasiver, detaillierter Visualisierung interner Strukturen des Körpers ohne exploratorische Operationen

Picture Archiving and Communication System (PACS) und DICOM

Das PACS ist ein System zur sicheren Speicherung und Übermittlung digitaler Bilder sowie klinisch relevanter Befunde.

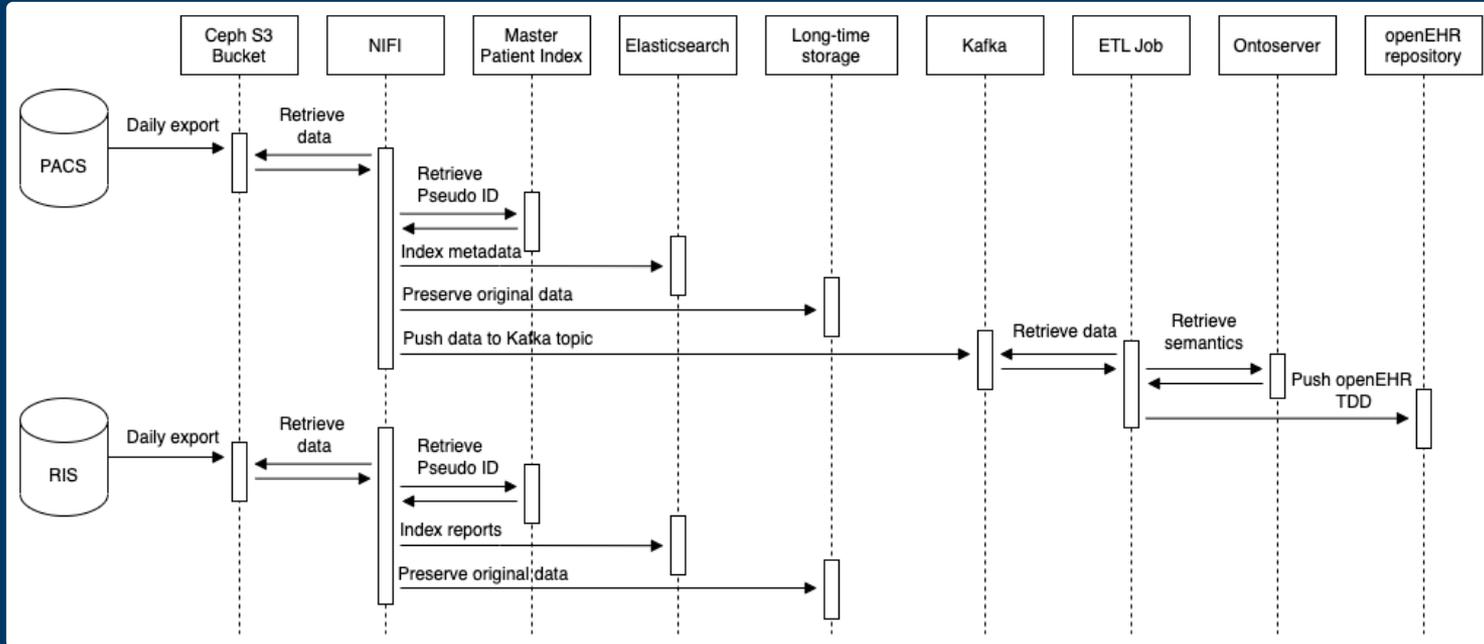
DICOM ist der internationale Standard zu Übermittlung, Speicherung, Abruf, Druck, Verarbeitung und Visualisierung medizinischer Bildinformationen

DICOM File

Metadata - Header

SOPClassUID	1.2.840.10008.5.1.4.1.1.1.7
SOPInstanceUID	2.16.840.1.113662.2.256
StudyDate	19970205
SeriesDate	19970205
AcquisitionDate	19970205
StudyTime	074737.000000
SeriesTime	074737.000000
AcquisitionTime	074737.000000
Modality	CT
ConversionType	WSD
Manufacturer	Picker International, Inc.
InstitutionName	BOSTON MED CENT E.N.C.
ReferringPhysiciansName	OREGAN
StudyDescription	<VYGR> Path: dicom
ManufacturersModelName	POS000
Unknown	70
PatientsName	CT BRONCHOSCOPY
PatientID	000003 M 47
PatientsSex	O
Unknown	122

Integration des PACS des Universitätsklinikum Schleswig-Holstein



Ulrich et al. Large-Scale Standardized Image Integration for Secondary Use Research Projects. *Stud Health Technol Inform.* 2024;310:174-8. doi: 10.3233/SHTI230950.

Problem

- Daten für Forschende bereitstellen
- Sehr detaillierte Anfragen

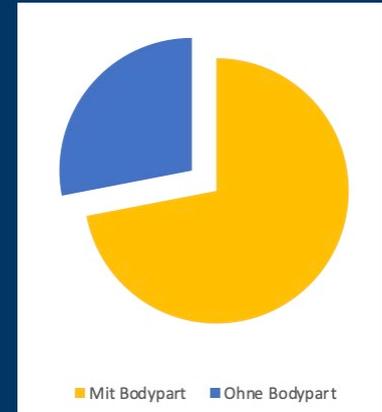
Specific Demands on Data: mindestens 10 Jahre alte CT Aufnahme (Thorax, Abdomen) von 60 jährigen Personen (zum Zeitpunkt der Aufnahme). Wir benötigen mindestens 100 Fälle mit Fraktur pro Frakturart (Wirbelkörper-Fraktur, Hüft-Fraktur).

- Fehlende Informationen zur Identifikation der relevanten Daten
- Körperregion ist relevant
 - DICOM (0018,0015)
 - IMPAX

Modalität Computertomographie (CT)

- Über 7 Millionen CT Serien
 - ~25% fehlende Informationen zur Abfrage

- Studien/Serien Beschreibung ist sehr breit
- Kein verwendbares Muster gefunden



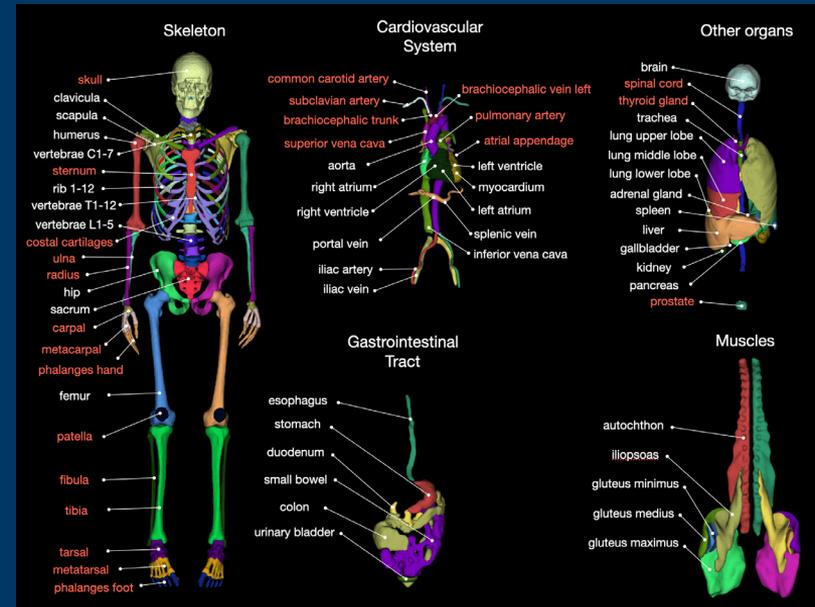
Ulrich et al. Full-Scale Indexing and Semantic Annotation of CT Imaging: Boosting FAIRness. 2024. <https://arxiv.org/abs/2406.15340>

Semantische Indizierung aller(!) CT Studien

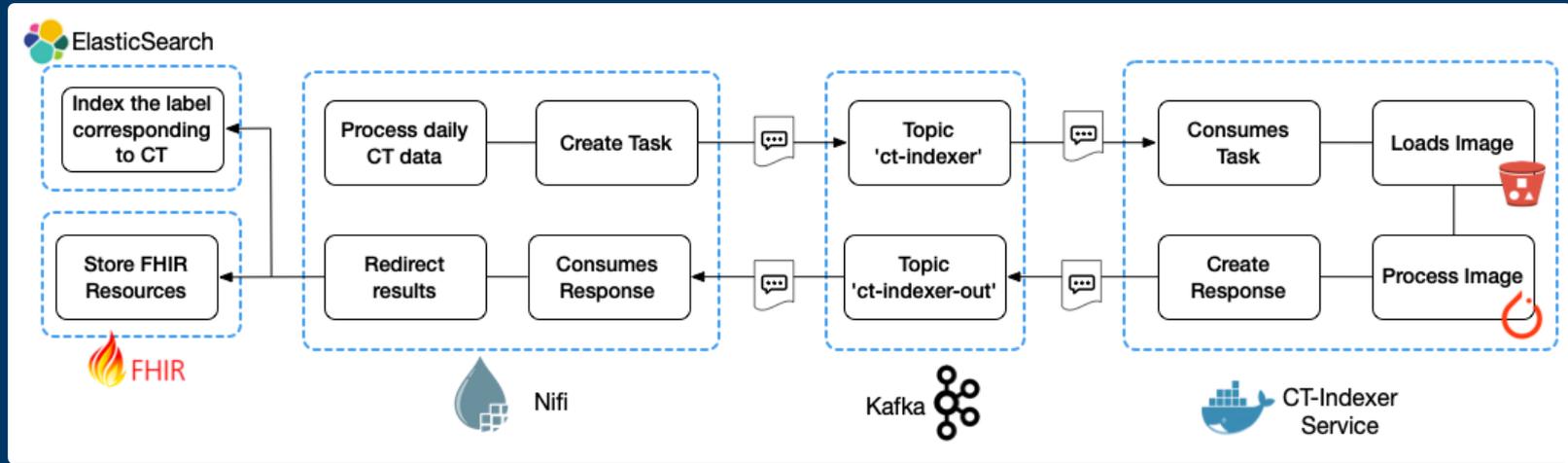
- Nutzung des TotalSegmentator um alle CTs entweder zu indizieren oder zu segmentieren
- Mapping der Label nach SNOMED CT und Radlex
 - Validiert durch einen Radiologen
- 16 Instanzen auf 4 VMs mit je einer GPU
 - 1,3+ Millionen Serien indiziert
 - Absolute GPU Zeit von 1000+ Tagen
 - 250 Bilder / h

Ulrich et al. Full-Scale Indexing and Semantic Annotation of CT Imaging: Boosting FAIRness. 2024.
<https://arxiv.org/abs/2406.15340>

Mögliche Label



CT Indizierungsworkflow



Ulrich et al. Full-Scale Indexing and Semantic Annotation of CT Imaging: Boosting FAIRness. 2024. <https://arxiv.org/abs/2406.15340>

Ergebnisse

Structure	SCT Code	Amount
Heart	80891009	728.000
Left Lung Lobe	44714003	664.000
Liver	10200004	483.000
L1	66794005	657.000
Right Hip	287579007	359.000
Pancreas	15776009	583.000

Nutzung

Specific Demands on Data: mindestens 10 Jahre alte CT Aufnahme (Thorax, Abdomen) von 60 jährigen Personen (zum Zeitpunkt der Aufnahme). Wir benötigen mindestens 100 Fälle mit Fraktur pro Frakturart (Wirbelkörper-Fraktur, Hüft-Fraktur).

1. Finden der korrespondierenden SNOMED CT Codes zu den angefragten anatomischen Strukturen
2. Durchsuchen des Index nach allen Bildern mit den entsprechenden Codes
3. Überprüfung der Einwilligung der Patient:innen zur Datennutzung
4. Abrufen der Bilder aus dem PACS
5. Pseudonymisierung der **DICOM** Header
6. Export der Daten an die Forschenden ☺

Ulrich et al. Full-Scale Indexing and Semantic Annotation of CT Imaging: Boosting FAIRness. 2024. <https://arxiv.org/abs/2406.15340>

Annotation von Röntgenbildern

- KI-Modell trainiert auf öffentlichen Datensätzen
 - IRMA, ROCO und UNIFESP
- Erkennt wichtige Bildinformationen:
 - 45 IRMA-Codes (Röntgenbilder beinhalten Körperrichtungen und Körperteile)
 - 45 Körperteile in SNOMED-CT-Codierung
- Mit einer Genauigkeit von über 90%



Beispiel mit SNOMED CT Vorhersage

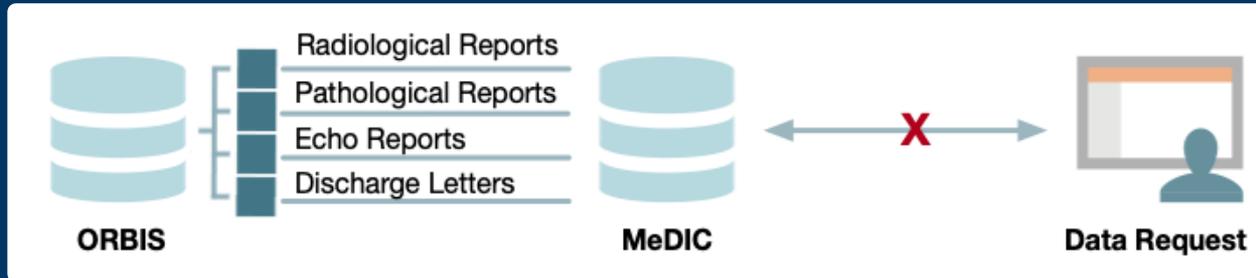
Cheng et al. Instance-level medical image classification for text-based retrieval in a medical data integration center. *Computational and Structural Biotechnology Journal*. 2024;24:434-50. doi:10.1016/j.csbj.2024.06.006

Übernahme von Routinedaten für die klinische Forschung

Wir integrieren täglich Freitexte

- Forschende fordern Informationen an, die im Freitext dokumentiert sind
- Anonymisierung/Pseudonymisieren erforderlich, bevor wir klinische Texte aufgrund der DSGVO bereitstellen können

Requested Data: We need comprehensive insight into the medical documentation, since it is about the determination of a therapy complication that has to be described and is under-recorded - discharge letters, documentation of doctor's visits, radiological, pathological and laboratory data.

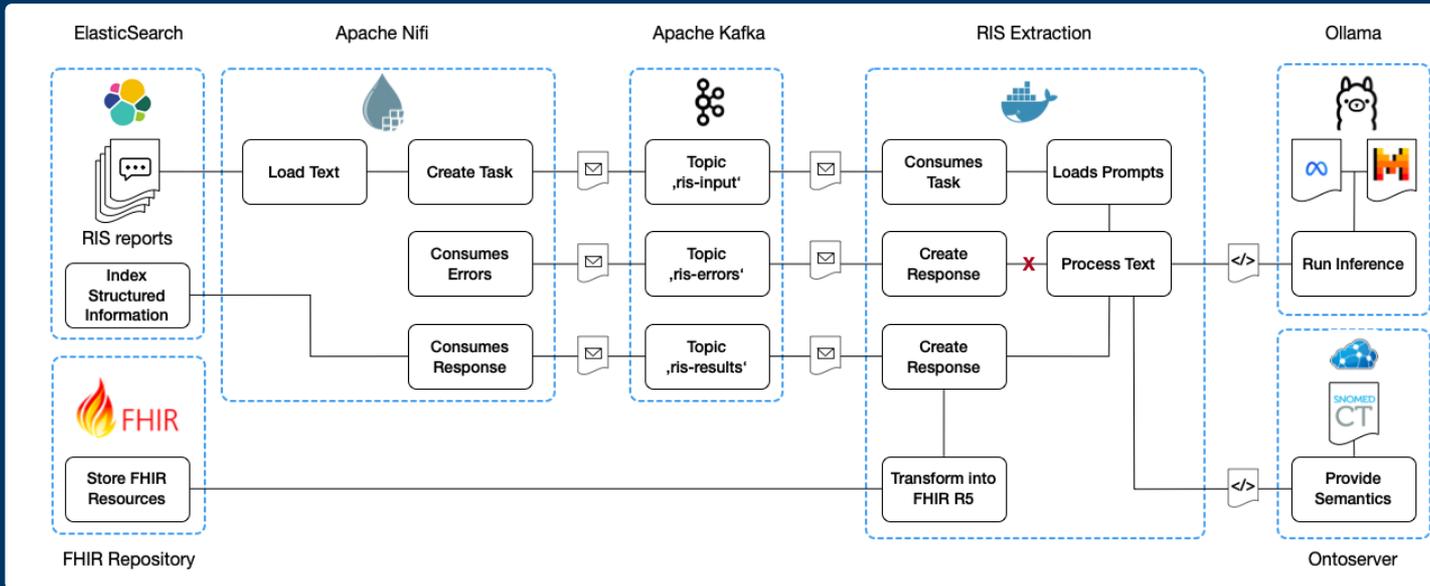


Informationsextraktion

- Suche nach medizinischen Konzepten in den Freitexten
 - Strukturierung von Echokardiographiebefunden
 - Strukturierung von Radiologiebefunden
- Extraktion und Evaluation mit Mediziner:innen
 - Langwierige Bearbeitung aller Entitäten
- Speicherung der benötigten Informationen und der entsprechenden Provenancedaten



Workflow zur Verarbeitung von Freitexten



(Eingereicht als Fullpaper zur Medical Informatics Europe 2025)

Vorbereitung für MII Kerndatensatz Bildgebung

Zwei separate Integrationswege für
die beiden KDS Teile

- Export der Metadaten getrennt von den Bildern
- Auswertung der Texte, da die Befunde ausschließlich im Freitext vorliegen

Alle Tools sind frei verfügbar



CT Indexer



Xray Classifier

Weitere werden folgen...



Prof. Dr. Björn Schreiweis

Christian-Albrechts-Universität zu Kiel
Institut für Medizinische Informatik und Statistik
Universitätsklinikum Schleswig-Holstein
Kaistraße 101 | 24114 Kiel | Deutschland

Tel: +49-431-500 31601

Email: bjoern.schreiweis@uksh.de

Vielen Dank!

