

Helmholtz Metadata Collaboration Hub Health @ DKFZ

Lucas Kulla^{1,2,3,4}, Hamideh Haghiri^{1,2}, Anand Deshpande^{1,2}, Klaus Maier-Hein^{1,2}, Marco Nolden^{1,2}

¹ Division of Medical Computing, Deutsches Krebsforschungszentrum (DKFZ) Heidelberg

² Helmholtz Metadata Collaboration (HMC) - Hub Health

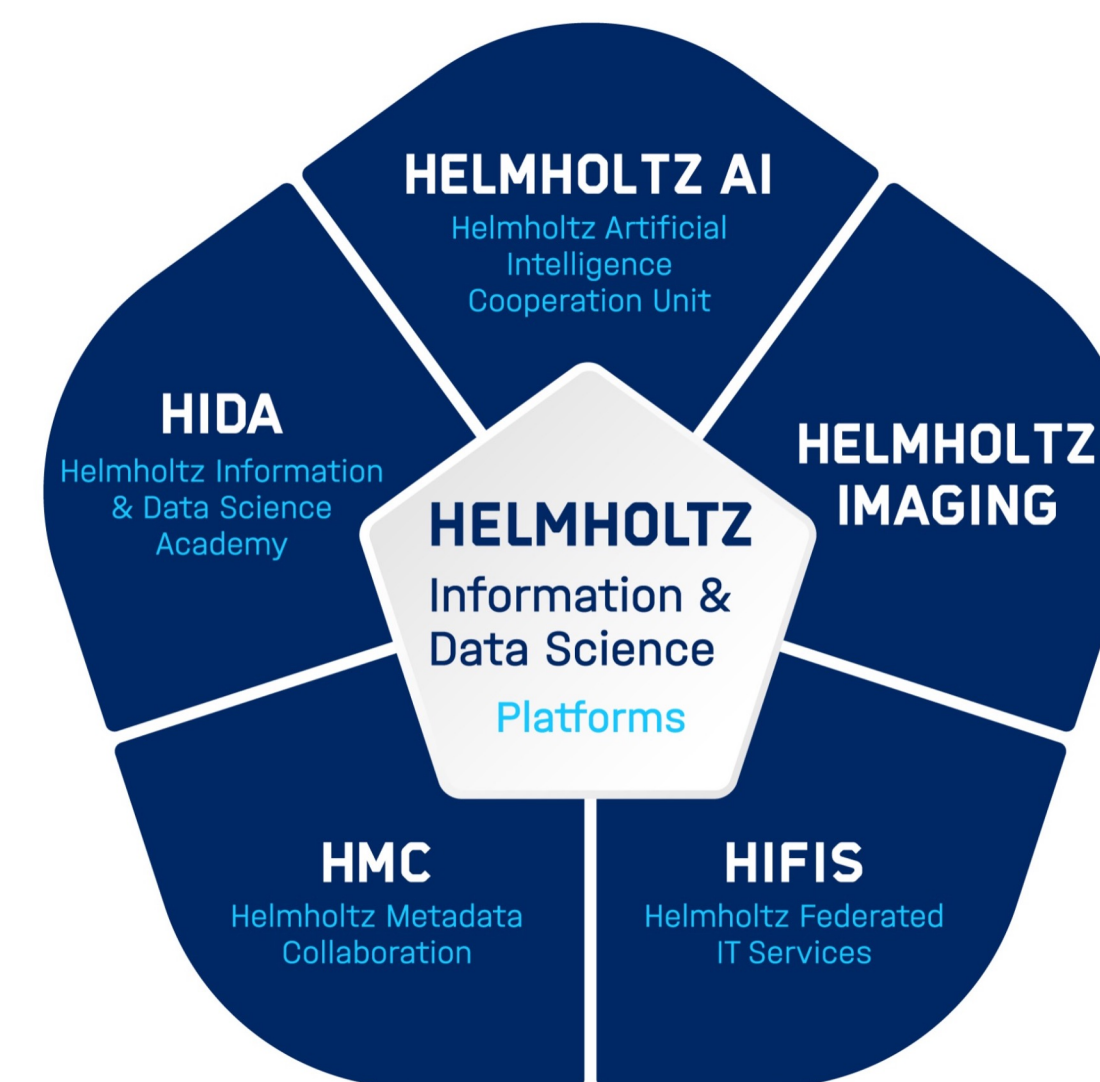
³ National Center for Tumor Diseases (NCT), Heidelberg, a partnership between DKFZ and University Hospital Heidelberg, Germany

⁴ NCT Digital Oncology Program

HMC Hub Health

A unit of the Helmholtz Metadata Collaboration, located completely at DKFZ, working on the FAIR Data Space for Helmholtz

- Contact point for Health-related metadata questions from the community
- Bridge to the Helmholtz Information and Data Science Framework (HIDS)
- Collaborations with life science NFDIs, Medical Informatics Initiative (MI-I), Network University Medicine (NUM), National Centre for Tumor Diseases (NCT)
- Consulting for HMC projects. Next call: €400k (50%), Opening ~March 2026, Closing ~May 2026
- Contact hmc-health@dkfz.de



nfdi Nationale Forschungsdaten Infrastruktur

MEDIZIN INFORMATIK INITIATIVE

NUM Network of University Medicine

NCT NATIONAL CENTER FOR TUMOR DISEASES HEIDELBERG

Community Dialogue & Consulting with Infrastructures

Priority Fields for Harmonization

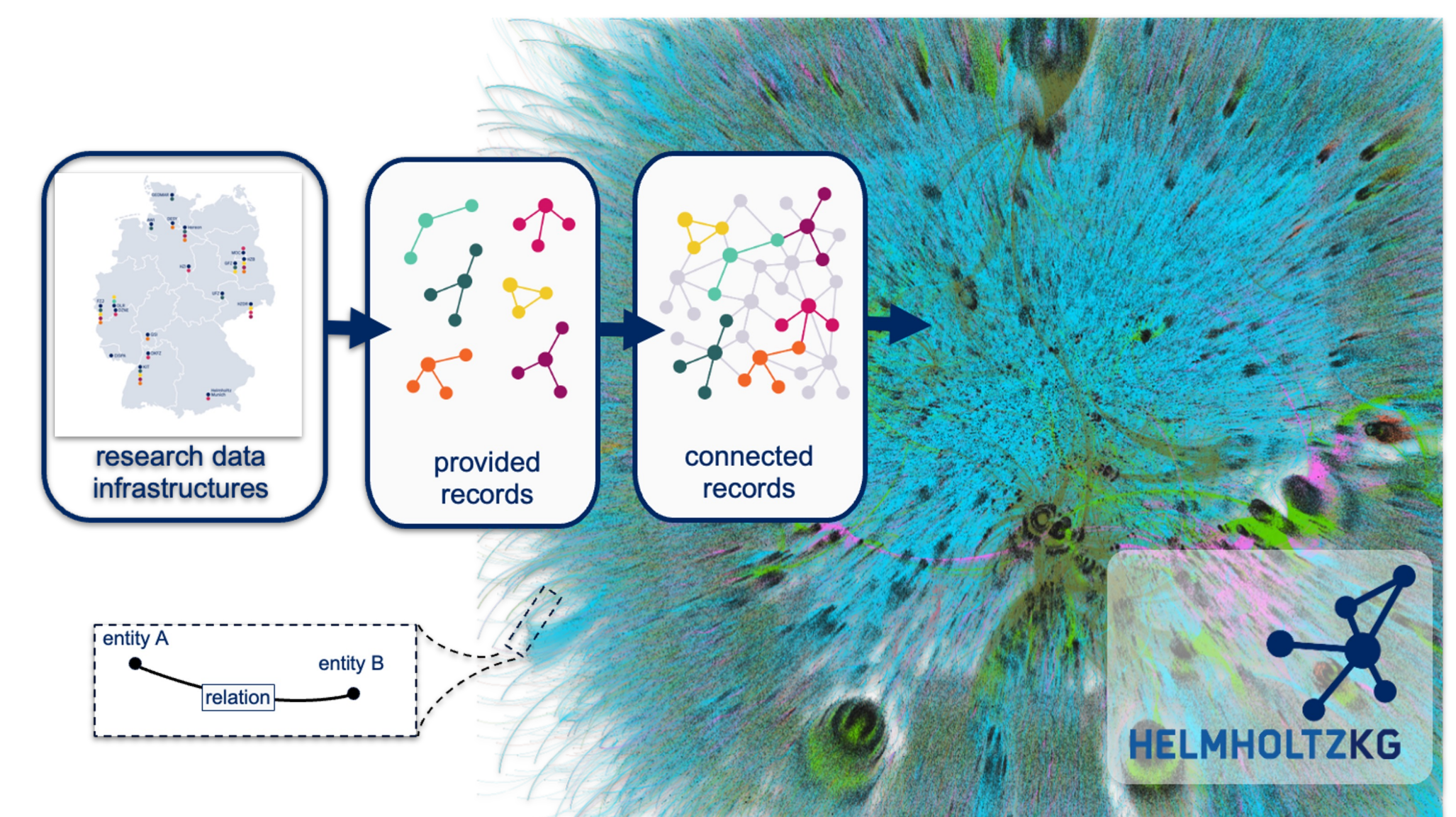
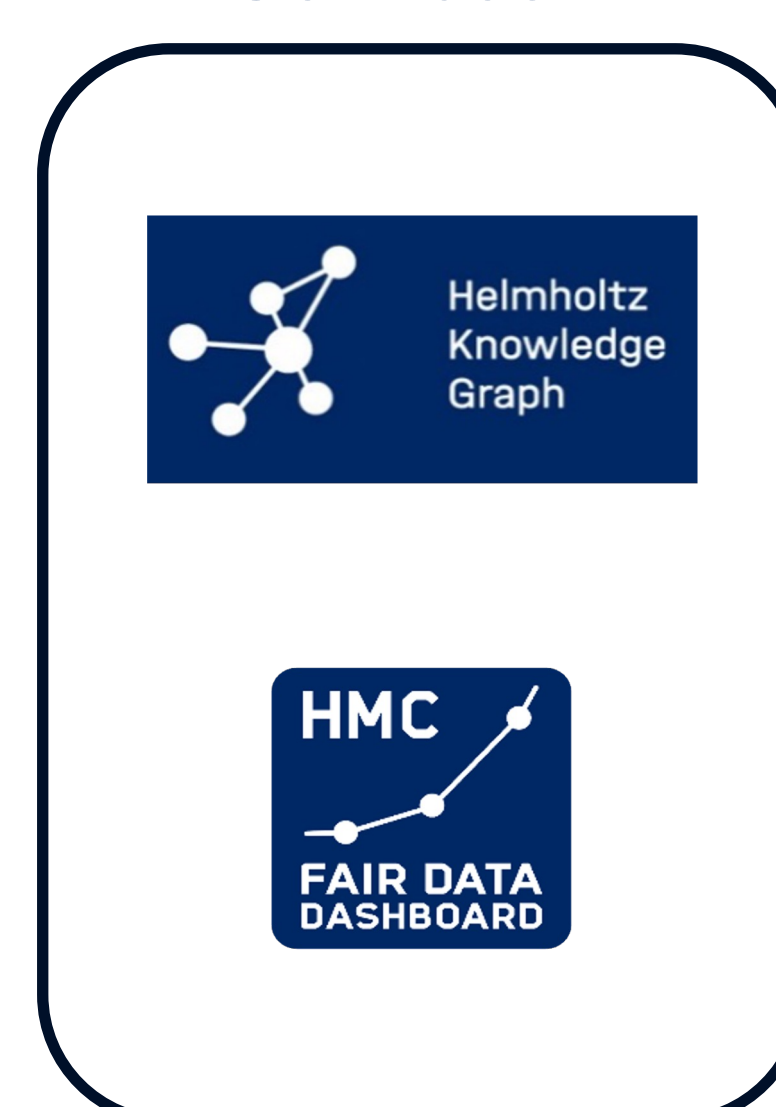
Expected Impact for HMC Services



Gap Analysis
Workshop
One-on-One Consulting



Valid Identifiers
ISO Publication Year
Consistent
Resource Type

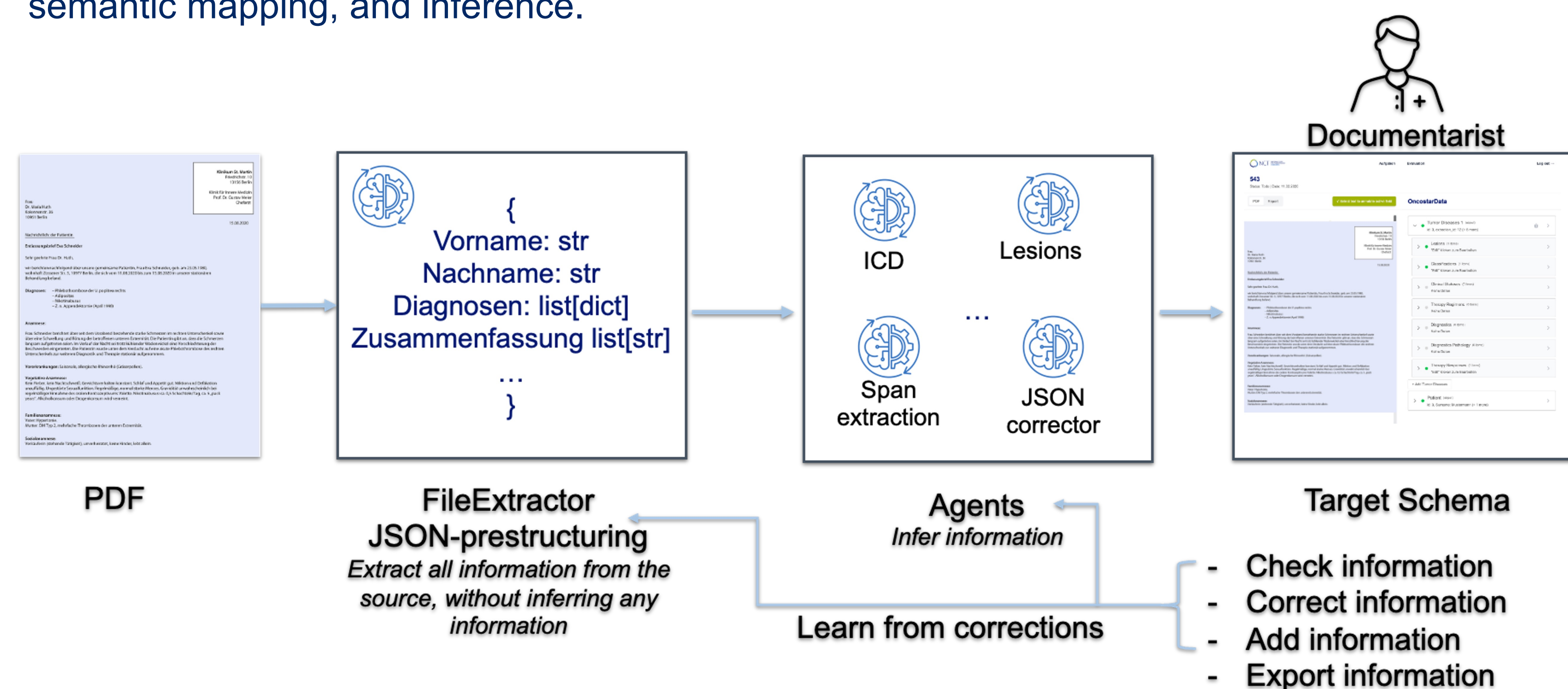


Approach: cross-unit working group (TF Harmony) that engages with repository maintainers directly

Data Extraction & Curation Framework

Approach: Manual curation of patient data from paper-based medical documents into structured clinical formats.

Task: Automated transformation into a predefined target schema via information extraction, semantic mapping, and inference.



Framework is developed and implemented within the Digital Oncology Group at NCT Heidelberg, in close collaboration with clinical partners from the department of Translational Medical Oncology.

Community Activities

HIDS Course Portfolio

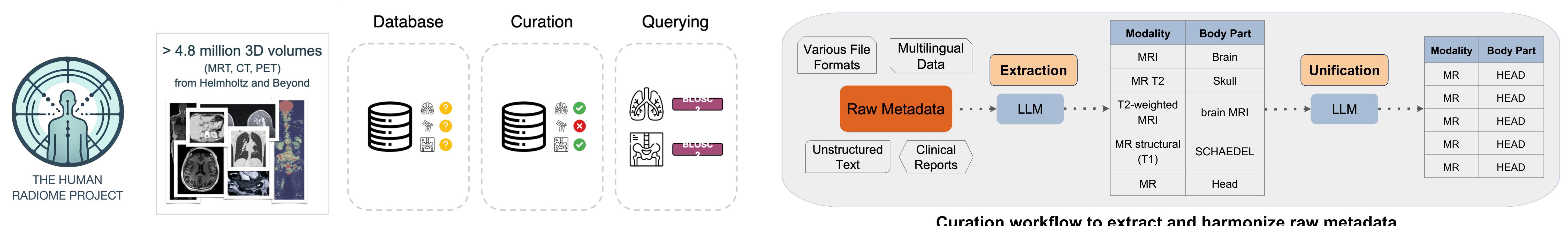
Fundamentals of Scientific Metadata - for Health
15.-16.01.2026 - ON-SITE

<HMC> HELMHOLTZ Metadata Collaboration

Register here!

HMC Conference
28.-30.04.2026, Heidelberg

Support for Public Data Acquisition and Metadata Curation Pipeline for Foundation Model Development



Curation workflow to extract and harmonize raw metadata.