

DS Mission Cancer

Data Steward für interoperables Management klinischer Daten innerhalb der Mission Cancer an der JKU

Programm / Ausschreibung	Implementing EU Missions, IMPA 2024, Data Stewards 2025	Status	laufend
Projektstart	01.11.2025	Projektende	30.04.2028
Zeitraum	2025 - 2028	Projektlaufzeit	30 Monate
Keywords	Mission Cancer, Datenmanagement, FAIR Data Principles, vernetzte Forschungsinfrastruktur, Datenintegration, ACCN		

Projektbeschreibung

Die personalisierte Krebsmedizin bietet großes Potenzial zur Verbesserung von Prävention, Diagnose und Therapie. Ihre Umsetzung ist jedoch mit erheblichen Herausforderungen, insbesondere im Bereich des Datenmanagements, verbunden. Die dafür notwendigen Daten – etwa molekulare Diagnostikdaten, Bildgebungsdaten, klinische Routinedaten, patientenberichtete Ergebnisse (PROMs, PREMs) – werden in unterschiedlichen Systemen dokumentiert, in verschiedenen Formaten und Qualitäten. Sie sind oft nur unzureichend dokumentiert, fragmentiert oder institutionell nicht zugänglich. Der steigende Umfang und die zunehmende Komplexität der Daten erschweren die systematische Nutzung zusätzlich. Hinzu kommen mangelnde Interoperabilität zwischen Datenquellen innerhalb und auch zwischen Institutionen. Der Bedarf an hochwertigen, standardisierten Datensätzen für moderne Analyseverfahren, insbesondere für KI-basierte Methoden, sowie für nationale und europäische Forschungsinitiativen (z. B. ACCN, AHDRH, UNCAN.eu) wächst.

Das vorliegende Projekt adressiert diese Herausforderungen durch die Etablierung und institutionelle Verankerung der Rolle eines Data Stewards für die medizinische Fakultät an der Johannes Kepler Universität Linz, wo sich das koordinierte Forschungsdatenmanagement erst im Aufbau befindet. Neben den bereits bestehenden CCCs in Graz, Wien und Innsbruck ist in Linz ein weiteres CCC in Planung. Dabei werden rasch wachsende Molekular- („Multi-Omics“) der PatientInnen in Einheit mit individuellen Therapie- und Outcome-Informationen in einem immer komplexeren, ML-/AI-basierten Experten-System abgebildet. Dadurch können zukünftige Patient*innen mit ihrem persönlichen Molekularprofil einen immer besser passenden „Digitalen Zwilling“ finden, zu dem es zunehmend belastbarere Therapieempfehlungen gibt. Der Data Steward soll den Aufbau des CCCL und des Experten-Systems begleiten und fungiert dabei als zentrale Schnittstelle zwischen Forschung, klinischer Versorgung und IT und unterstützt die strukturierte, qualitätsgesicherte und langfristige Nutzbarmachung der genannten Daten.

Ziel ist es, isolierte, heterogene Datensätze aus unterschiedlichen Quellen mit unterschiedlicher Qualität anhand der FAIR-Prinzipien sowie anerkannter internationaler Standards in eine integrierte, standardisierte Datenstruktur zu überführen, um sie für Forschung, klinische Anwendungen und innovative Technologien optimal nutzbar zu machen. Dabei sind ebenso ethische und regulatorische Anforderungen zu berücksichtigen. Die Einführung eines Data Stewards für die medizinische Fakultät der JKU im Bereich der Krebsforschung ist daher nicht nur eine Reaktion auf den wachsenden Bedarf an Datenmanagement, sondern ein wesentlicher Bestandteil der digitalen und wissenschaftlichen Strategie zur Verbesserung

der standortübergreifenden Datenintegration und -nutzung. Dieser Prozess soll mithilfe dieses Projektes gestartet werden, wobei Ausbildung, Weiterbildung und die Initiierung eines österreichweiten Netzwerks von Data Stewards, um eine institutionsübergreifende Zusammenarbeit in der datengetriebenen onkologischen Forschung zu fördern, ein integraler Bestandteil ist. Nach Abschluss des Projektes ist geplant den Data Steward entfristet in den Personalstand der Bibliothek zu überführen, um die Kontinuität und Nachhaltigkeit zu gewährleisten.

Abstract

Personalized cancer medicine offers great potential for improving prevention, diagnosis and therapy. However, its implementation is associated with considerable challenges, particularly in the area of data management. The data required for this - such as molecular data, imaging data, routine clinical data, patient-reported outcomes (PROMs, PREMs) - are documented in different systems, in different formats and qualities. They are often inadequately documented, fragmented or inaccessible. The increasing volume and complexity of the data makes systematic use even more difficult. In addition, there is a lack of interoperability between data sources within and between institutions. There is a growing need for high-quality, standardized data sets for modern analysis procedures, especially for AI-based methods, as well as for national and European research initiatives (e.g. ACCN, AHDRH, UNCAN.eu).

This project addresses these challenges by establishing and institutionally embedding the role of a data steward for the Faculty of Medicine at the Johannes Kepler University Linz, where coordinated research data management is currently being established. In addition to the existing CCCs in Graz, Vienna and Innsbruck, a further CCC is planned in Linz. Here, rapidly growing molecular data ("multi-omics") from patients will be mapped together with individual therapy and outcome information in an increasingly complex ML/AI-based expert system. This will enable future patients to find a more and more suitable "digital twin" with their personal molecular profile, for which increasingly reliable therapy recommendations can be made. The data steward will support the development of the CCCL and the expert system, acting as a central interface between research, clinical care and IT and supporting the structured, quality-assured and long-term utilization of the aforementioned data.

The aim is to transfer isolated, heterogeneous data sets from different sources of varying quality into an integrated, standardized data structure based on the FAIR principles and international standards in order to make them optimally usable for research, clinical applications and innovative technologies. Ethical and regulatory requirements must also be taken into account. The introduction of a data steward for the JKU Faculty of Medicine in the field of cancer research is therefore not only a response to the growing need for data management, but also an essential component of the digital and scientific strategy to improve cross-site data integration and use. This process is to be started with the help of this project, whereby training, further education and the establishment of an Austria-wide network of data stewards to promote cross-institutional cooperation in data-driven oncological research is an integral part. At the end of the project, it is planned to take on the data steward as full-time library staff in order to ensure continuity and sustainability.

Projektpartner

- Universität Linz