

Konzeptschreiben: Entwicklung eines Dashboards zur Datenqualitätsüberwachung in der Sarkomversorgung am Kantonsspital Luzern

Sivakumar, Sivanajani

B.Sc. Medizininformatik, Life Science

Betreuung von: Dr. Abdullah Kahraman und Dr. Philip Heesen (LUKS)

Auftraggeber: Dr. Bruno Fuchs und Dr. Philip Heesen

Experte: Dr. Fabian Arnold

Fachhochschule Nordwestschweiz FHNW

Hochschule für Life Sciences HLS

Muttenz, 09.Mai 2024

Zusammenfassung

Im Rahmen dieser Bachelorarbeit wird ein webbasiertes Dashboard zur Überwachung der Datenqualität von Ergebnisdaten bei Sarkom-Patient:innen am Kantonsspital Luzern (LUKS) konzipiert und prototypisch umgesetzt. Dabei werden sowohl patientenberichtete Ergebnisdaten (PROMs) als auch klinisch erhobene Ergebnisdaten (CROMs) berücksichtigt. Beide Datentypen sind essenziell für die Beurteilung von Gesundheitszustand, Therapieerfolg und Versorgungsqualität.

Das geplante Dashboard basiert auf einer mehrdimensionalen Bewertung der Datenqualität, wie sie in der Literatur u. a. von Azeroual (2022) und Pipino et al. (2002) beschrieben wird. Es soll unvollständige, inkonsistente oder unplausible Einträge automatisiert identifizieren und sogenannte Red Flags visuell aufbereiten. Zusätzlich werden Benachrichtigungsmechanismen (z.B. Alerts) integriert.

Die Umsetzung erfolgt in einem modularen Technologiestack (PostgreSQL, FastAPI, React) mit Fokus auf Nutzerfreundlichkeit und klinische Relevanz. Der Zeitplan der Arbeit folgt einem agilen Vorgehen mit sieben Sprints. Vergleichbare Projekte, wie das „RA PRO Dashboard“ von Schmajuk et al. (2024), dienen als methodische Orientierung.

Die zentrale Fragestellung lautet:

Wie kann ein webbasiertes Dashboard die Vollständigkeit und Qualität von PROM- und CROM-Daten bei Sarkom-Patient:innen am Kantonsspital Luzern in Echtzeit überwachen und nutzerfreundlich visualisieren?

Inhaltsverzeichnis

Zusammenfassung

Inhaltsverzeichnis

1	Einleitung	1
2	Ziele und Zielgruppen	2
3	Beschreibung des Konzepts	3
3.1	Vergleichbare Ansätze	3
3.2	Nutzen des Dashboards	4
3.3	Risiken und Herausforderungen	4
3.4	Zeitplan und Ressourcen	4
3.4.1	Visueller Sprint-Zeitplan	5
3.4.2	Tabellarischer Zeitplan	6
3.5	Erfolgskriterien	7
4	Schlussfolgerung und Ausblick	8
5	Literaturverzeichnis	9

1 Einleitung

Am Kantonsspital Luzern (LUKS) werden im Rahmen der Sarkomversorgung regelmässig Ergebnisdaten erhoben. Diese umfassen sowohl patientenberichtete Outcome-Daten (PROMs), die direkt von den Patient:innen über strukturierte Fragebögen erfasst werden, als auch klinisch erhobene Outcome-Daten (CROMs), die von Ärzt:innen und medizinischem Personal dokumentiert werden.

Beide Datentypen spielen eine zentrale Rolle bei der Beurteilung von Gesundheitszustand, Therapieerfolg und Versorgungsqualität. Damit daraus verlässliche Erkenntnisse abgeleitet werden können, ist eine kontinuierliche Überwachung der Datenqualität erforderlich. Fehlende oder unplausible Angaben können wissenschaftliche Auswertungen verzerren, wenn sie nicht erkannt und berücksichtigt werden.

Für die Bewertung der Datenqualität hat sich in der Forschung ein mehrdimensionaler Ansatz etabliert. Dieser wird unter anderem von Azeroual (2022) beschrieben der Dimensionen wie Vollständigkeit, Korrektheit, Konsistenz und Plausibilität als zentrale Indikatoren zur strukturierten Bewertung klinischer Daten hervorhebt. Solche Dimensionen ermöglichen eine systematische und vergleichbare Beurteilung der Datenqualität – sowohl aus technischer als auch aus inhaltlicher Perspektive.

Im Rahmen dieser Bachelorarbeit soll ein webbasiertes Dashboard zur Überwachung und Visualisierung der Datenqualität entwickelt werden. Es fokussiert sich auf die Erkennung unvollständiger oder auffälliger Einträge, die Identifikation sogenannter „Red Flags“ sowie die Bereitstellung automatisierter Benachrichtigungen.

Die zentrale Fragestellung lautet: Wie kann ein webbasiertes Dashboard die Vollständigkeit und Qualität von PROM- und CROM-Daten bei Sarkom-Patient:innen am Kantonsspital Luzern in Echtzeit überwachen und nutzerfreundlich visualisieren?

2 Ziele und Zielgruppen

Ziel dieser Arbeit ist die Konzeption und prototypische Umsetzung eines Dashboards zur Echtzeitüberwachung der Datenqualität von PROMs und CROMs. Dabei werden folgende Teilziele verfolgt:

- Literaturrecherche zu Best Practices im Bereich der Datenqualitätssicherung
- Entwicklung eines modularen Datenqualitätskonzepts für klinische Daten
- Auswahl geeigneter Technologien und Visualisierungstools
- Aufbau eines Dashboards (Prototyp) zur Darstellung relevanter Metriken
- Umsetzung einfacher Prüfregele (z.B. Vollständigkeit, Ausreissererkennung)
- Definition und Simulation von Benachrichtigungsmechanismen

Die Zielgruppe umfasst klinisch tätige Ärzt:innen, Datenmanager:innen und Forschende am LUKS. Das System ist für den internen Gebrauch bestimmt und wird nicht von Patient:innen verwendet.

3 Beschreibung des Konzepts

Das Dashboard basiert auf der internen PostgreSQL-Datenbank des LUKS, in der PROMs und CROMs pseudonymisiert gespeichert und regelmässig aktualisiert werden. Die Datenbank befindet sich derzeit im Aufbau durch die Firma Renuo und wird in der Kalenderwoche 19 fertig sein (Stand 16.05.2025).

Folgende konzeptionelle Komponenten sind vorgesehen:

- Ein Kriterienkatalog zur Beurteilung der Datenqualität (z.B. Vollständigkeit, Konsistenz, Eindeutigkeit, Plausibilität)
- Definition typischer Red Flags (z.B. fehlende Einträge, unrealistische Werte)
- Entwicklung eines Visualisierungskonzepts für eine nutzerfreundliche Darstellung

Die Auswahl und Strukturierung dieser Dimensionen orientiert sich an Konzepten der Datenqualitätsforschung, wie an den Arbeiten von Pipino et al. (2002), die Vollständigkeit, Konsistenz, Genauigkeit und Plausibilität als zentrale Dimensionen operationalisiert haben (Pipino et al., 2002).

Die technische Umsetzung erfolgt in einem flexiblen Webstack, basierend auf React (Frontend) und FastAPI (Backend). Ein Hosting über Google Cloud ist möglich. Beispielhafte Funktionen des Dashboards:

- Übersicht über fehlende Fragebogeneinträge (PROMs) und fehlende klinische Befunde (CROMs)
- Aggregierte Auswertungen zur Vollständigkeit je Datentyp
- Identifikation von Fällen mit Red Flags, z. B. fehlender Bildgebung (CROM) oder unplausible Selbsteinschätzungen (PROM)

Die Auswahl geeigneter Technologien sowie ein Überblick zu deren Funktionen und Einsatz im geplanten Dashboard sind im separaten Dokument "Technologien und Visualisierungstools" beschrieben.

3.1 Vergleichbare Ansätze

Ein vergleichbares Vorhaben wurde von Schmajuk et al. (2024) realisiert. In ihrer Arbeit wurde ein sogenanntes EHR-Sidecar-Dashboard („RA PRO Dashboard“) entwickelt, das patientenberichtete Outcome-Daten bei rheumatoider Arthritis systematisch erfasst, überwacht und in die klinische Entscheidungsunterstützung integriert. Der Entwicklungsprozess erfolgte schrittweise, von der Definition relevanter Metriken über die Backend-Integration bis zur nutzerfreundlichen Visualisierung im Klinikalltag.

Besonders relevant für die vorliegende Arbeit sind die folgenden Aspekte:

- Trennung zwischen PROM-Logik und klinischem System („Sidecar“-Architektur)
- Kontinuierliche Datenqualitätssicherung (z.B. Erkennung fehlender Werte)
- Visualisierung individueller und aggregierter Trends zur Entscheidungsunterstützung

Obwohl sich das RA PRO Dashboard auf eine andere Patient:innengruppe bezieht, liefert es wertvolle Impulse für den strukturierten Aufbau eines datenqualitätsorientierten Dashboards in der Sarkomversorgung.

3.2 Nutzen des Dashboards

Ein Dashboard zur Datenqualitätsüberwachung ermöglicht:

- Frühzeitige Erkennung von Lücken und Auffälligkeiten in PROM- und CROM-Daten
- Direkte Rückmeldung an die verantwortliche Datenmanager:innen (z.B. durch Alerts)

3.3 Risiken und Herausforderungen

- Verzögerungen im Datenbankaufbau: Die produktive Datenbank ist aktuell noch im Aufbau (Stand 16.05.2025).
- Die Definition sinnvoller Prüfregeln (z.B. für Ausreisser) erfordert medizinische Fachexpertise.
- Visualisierungen müssen sowohl für medizinisches als auch wissenschaftliches Personal verständlich sein.
- Abhängig von der Toolauswahl kann es eine technische Einarbeitungsphase geben, insbesondere bei der Anbindung an bestehende Systeme oder der Konfiguration des Datenzugriffs.

3.4 Zeitplan und Ressourcen

Start der Bachelorarbeit: 05.05.2025

Betreuer: Dr. Abdullah Kahraman und Dr. Philip Heesen (LUKS)

Auftraggeber: Dr. Bruno Fuchs und Dr. Philip Heesen

Experte: Dr. Fabian Arnold

Die Arbeit ist auf 15 Wochen angelegt und orientiert sich an einem iterativen, agilen Vorgehen in Sprints. Der Zeitplan kann bei Bedarf agil angepasst werden, z.B. bei Verzögerungen im Datenbankaufbau oder Technologieentscheidungen.

Die Abbildung 1 zeigt den geplanten zeitlichen Ablauf der Bachelorarbeit in Form eines Gantt-Diagramms. Die Arbeit ist in sieben Sprints gegliedert, die sich an einem agilen Vorgehensmodell orientieren. Sprint 0 umfasst vorbereitende Aktivitäten wie das Kick-off-Meeting und die Einrichtung der Projektinfrastruktur. In den folgenden Sprints werden zentrale Meilensteine wie die Literaturrecherche, die Entwicklung eines modularen Datenqualitätskonzepts, die technische Umsetzung (Backend, Frontend) sowie die Validierung und Dokumentation abgebildet. Der Zeitplan berücksichtigt Pufferzeiten sowie iterative Abstimmungen mit Betreuenden. Die Darstellung dient der strukturierten Projektsteuerung und ermöglicht eine transparente Übersicht über Aufgaben, Zuständigkeiten und Abhängigkeiten.

3.4.1 Visueller Sprint-Zeitplan

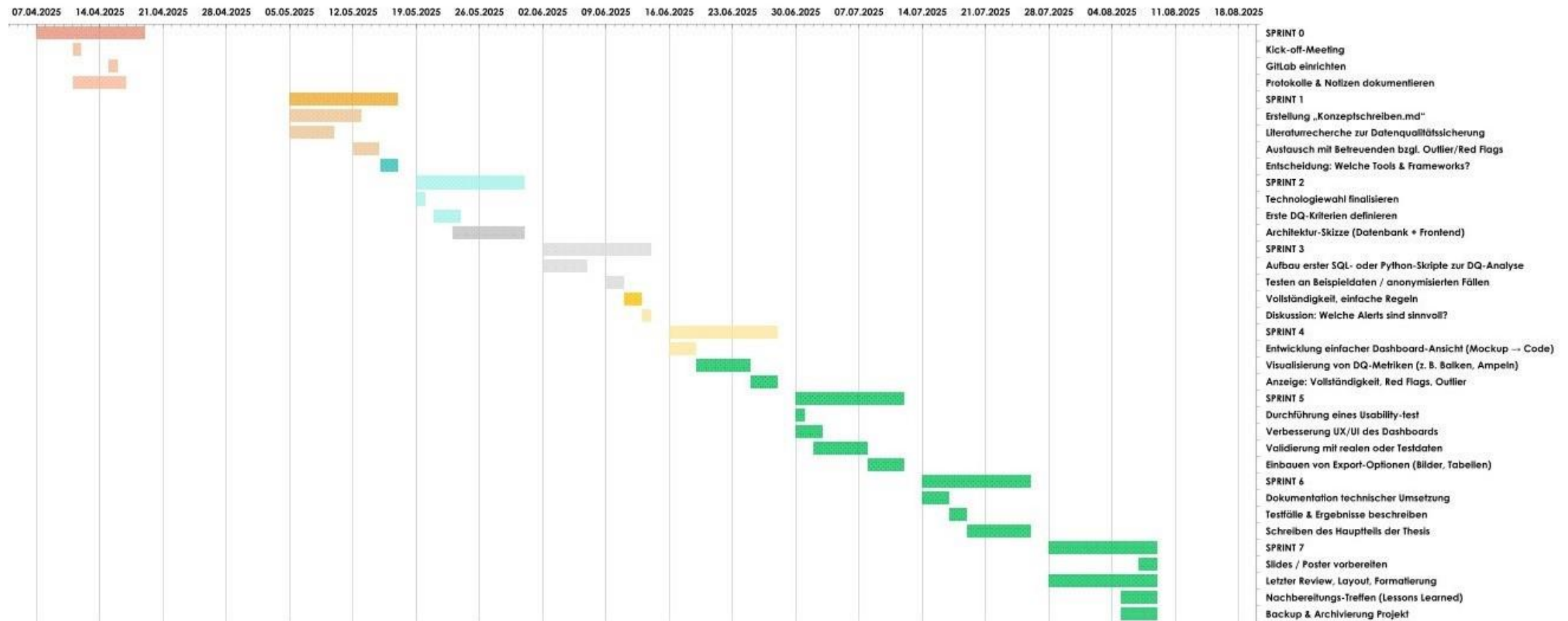


Abbildung 1 Visueller Sprint-Zeitplan der Bachelorarbeit – Entwicklung eines Datenqualitäts-Dashboards.

3.4.2 Tabellarischer Zeitplan

Die Tabelle 1 und Tabelle 2 zeigen die zeitliche Planung der einzelnen Arbeitspakete innerhalb der sieben Sprints der Bachelorarbeit. Für jede Phase werden die zugehörigen Aufgaben, deren Start- und Enddaten sowie die geschätzte Dauer in Tagen angegeben. Die Einteilung folgt einem iterativen Vorgehensmodell mit klar definierten Funktionstypen.

Tabelle 1 Strukturierter Zeitplan der Bachelorarbeit zur Entwicklung eines Datenqualitäts-Dashboard.

Funktionstyp	Aufgabenname	Startdatum	Enddatum	Dauer [Tage]
Setup & Orientierung	SPRINT 0	07.04.2025	18.04.2025	12
	Kick-off-Meeting	11.04.2025	11.04.2025	1
	GitLab einrichten	15.04.2025	15.04.2025	1
	Protokolle & Notizen dokumentieren	11.04.2025	16.04.2025	6
Literatur & Konzept	SPRINT 1	05.05.2025	16.05.2025	12
	Erstellung „Konzeptschreiben.md“	05.05.2025	12.05.2025	8
	Literaturrecherche zur Datenqualitätssicherung	05.05.2025	09.05.2025	5
	Austausch bzgl. Outlier/Red Flags	12.05.2025	14.05.2025	3
	Entscheidung: Welche Tools & Frameworks	15.05.2025	16.05.2025	2
Tool-Setup & Architektur	SPRINT 2	19.05.2025	30.05.2025	12
	Technologiewahl finalisieren	19.05.2025	19.05.2025	1
	Erste DQ-Kriterien definieren	21.05.2025	23.05.2025	3
	Architektur-Skizze (Datenbank + Frontend)	23.05.2025	30.05.2025	9
Prototyp Backend / DQ-Checks	SPRINT 3	02.06.2025	13.06.2025	12
	SQL- oder Python-Skripte zur DQ-Analyse	02.06.2025	06.06.2025	5
	Testen an Beispieldaten	09.06.2025	10.06.2025	2
	Vollständigkeit, einfache Regeln	11.06.2025	12.06.2025	2
	Alerts diskutieren und festlegen	13.06.2025	13.06.2025	1

Tabelle 2 Fortsetzung des Zeitplans – Entwicklung, Evaluation und Abschlussphase der Bachelorarbeit.

Funktionstyp	Aufgabenname	Startdatum	Enddatum	Dauer [Tage]
Frontend / Dashboard-Prototyp	SPRINT 4	16.06.2025	27.06.2025	12
	Dashboard-Ansicht (Mockup → Code)	16.06.2025	18.06.2025	3
	Visualisierung von DQ- Metriken	19.06.2025	24.06.2025	6
	Anzeige: Vollständigkeit, Red Flags, Outlier	25.06.2025	27.06.2025	3
Evaluation & Optimierung	SPRINT 5	30.06.2025	11.07.2025	12
	Usability-Test	30.06.2025	30.06.2025	1
	Verbesserung UX/UI	30.06.2025	02.07.2025	3
	Validierung mit Testda- ten	02.07.2025	07.07.2025	6
Dokumentation & Vorbereitung	SPRINT 6	14.07.2025	25.07.2025	12
	Dokumentation techni- scher Umsetzung	14.07.2025	16.07.2025	3
	Testfälle & Ergebnisse beschreiben	17.07.2025	18.07.2025	2
	Schreiben Hauptteil Thesis	19.07.2025	25.07.2025	7
Abgabe & Abschluss	SPRINT 7	28.07.2025	08.08.2025	12
	Slides / Poster vorberei- ten	07.08.2025	08.08.2025	2
	Letzter Review, Layout, Formatierung	28.07.2025	08.08.2025	12
	Lessons Learned / Nachbereitungs-Treffen	05.08.2025	08.08.2025	4
	Backup & Archivierung	05.08.2025	08.08.2025	4

3.5 Erfolgskriterien

- Vollständige, nachvollziehbare Dokumentation der Anforderungen und Konzepte
- Entwicklung eines funktionierenden Dashboard-Prototyps mit echten oder simulierten Daten
- Validierung durch Feedback von Betreuenden und potenzielle Nutzer:innen

4 Schlussfolgerung und Ausblick

Durch die Konzeption und prototypische Umsetzung eines Dashboards zur Echtzeitüberwachung wird ein praxisnahes Werkzeug geschaffen, das nicht nur technische Prüfungen ermöglicht, sondern auch die Sichtbarkeit und Reaktionsfähigkeit im Umgang mit unvollständigen oder fehlerhaften Daten verbessert.

Das Dashboard vereint Datenqualitätsdimensionen aus der Literatur (u. a. nach Azeroual (2022); Pipino et al. (2002); Schmajuk et al. (2024)) mit konkreten klinischen Anforderungen aus der Sarkomversorgung am Kantonsspital Luzern. Der modulare Aufbau sowie die geplante visuelle Darstellung von Metriken ermöglichen eine zielgerichtete Nutzung durch Ärzt:innen, Datenmanager:innen und Forschende.

In Zukunft kann das System durch zusätzliche Funktionalitäten erweitert werden, etwa:

- Maschinelles Lernen zur Anomalieerkennung, um auch komplexe Datenmuster oder versteckte Fehlerquellen automatisiert zu identifizieren.

Langfristig bietet das Dashboard das Potenzial, als generisches Instrument für die Datenqualitätsüberwachung in weiteren Fachbereichen oder Institutionen eingesetzt zu werden. Damit leistet diese Arbeit nicht nur einen Beitrag zur lokalen Datenqualitätssicherung, sondern liefert eine konzeptuelle Vorlage für skalierbare Lösungen in der klinischen Forschung.

5 Literaturverzeichnis

- Azeroual, O., 2022. Untersuchungen zur Datenqualität und Nutzerakzeptanz von Forschungsinformationssystemen: Framework zur Überwachung und Verbesserung der Qualität von Forschungsinformationen, Research. Springer Vieweg, Wiesbaden.
- Pipino, L.L., Lee, Y.W., Wang, R.Y., 2002. Data quality assessment. Commun. ACM 45, 211–218. <https://doi.org/10.1145/505248.506010>
- Schmajuk, G., Nasrallah, C., Berrean, B., Prugh, J., Wilson, C., Hamblin, A., Young, C., Jacobsohn, L., Kay, J., Li, J., Kersey, E., Subash, M., Murray, S., Yazdany, J., 2024. A step-by-step roadmap for the development and deployment of an electronic health record sidecar application that tracks patient outcomes: The RA PRO dashboard. DIGITAL HEALTH 10, 20552076241288739. <https://doi.org/10.1177/20552076241288739>