

STCS L3 Layer von Rohdaten zu finalen L3 Layern

Studiengang: MAS Data Science

Länder haben sich organisiert, um grenzüberschreitend Daten zum Transplantationsverlauf zu sammeln und eine optimale Verfügbarkeit von Spenderorganen zu erreichen. Täglich arbeiten Mediziner und Forscher daran, die Transplantationsmedizin zu verbessern. Durch den länderübergreifenden Zusammenschluss wird der Datenpool grösser; die Rahmenbedingungen für die Forschung werden verbessert.

Ausgangslage und Zielsetzung

Die Schweizer Transplantationskohorten Studie (STCS) ist ein vergleichbarer Zusammenschluss aller Schweizer Transplantationsspitäler, mit dem Auftrag, spitalübergreifend Daten für die Forschung zu sammeln.

Eine Hauptaufgabe des Datenzentrums (DC) ist es, Daten für Analysezwecke bereitzustellen. Die Kohorte hat zum Ziel, durch die Nutzung der Studienergebnisse die Transplantationsmedizin zu verbessern. Durch die Bereitstellung der Daten liefert das DC die Basis einer jeden Studie und ist somit an der Zielerreichung mitverantwortlich. Die Aufgabe des DC besteht darin, qualitativ hochwertige, gut strukturierte Daten zu generieren. Sauber generierte Tabellen für die Forschung aufzubereiten ist Thema dieser Masterarbeit.

Umsetzung

Die Umsetzung der Arbeit lässt sich grob in zwei Teile gliedern. Einem technischen Teil und einem Teil, der die Validierung beinhaltet.

Im technischen Teil wird ein Entity-Relationship (ER)-Modell in einem SAS Programm umgesetzt. Die Tabellen werden automatisch aus dem elektronischen Codebook (eCb) heraus generieren. Das bedeutet, dass Anpassungen im eCb – Erweiterungen und Ände-

rungen gleichermaßen – durch das Programm ausgeführt werden, ohne dass der SAS Code verändert werden muss. Bestehende Einzelprogramme werden im Zuge der Masterarbeit zu einem einheitlichen Gesamtprogramm zusammengefügt. Neu erarbeitet wird das ER – Processing, zur Erstellung der ER Tabellen. Es ist als Baustein entwickelt und in das Gesamtkonzept integriert. Resultat dieses Programms sind Draft L3 – Tabellen.

Im zweiten Teil, dem Validierungsteil, sind die Draft L3 – Tabellen mittels des agilen Projektmanagement SCRUM im Team diskutiert, Vorschläge eingebracht und Veränderungen erörtert worden. Die Lösungsvorschläge sowie deren Umsetzung münden in den finalen Tabellen.

Ergebnis

Die Daten der STCS durchlaufen nun einen harmonisierten, kaskadenartigen Prozess. Basierend auf den unstrukturierten Rohdaten durchlaufen die Daten verschiedene Qualitätschecks und erfahren strukturelle Veränderungen. In jedem Verarbeitungsschritt werden die Daten qualitativ hochwertiger und benutzerfreundlicher und werden final in den strukturierten ER Tabellen abgelegt



Juliane Rick