

Mehr Forschung wagen

Warum Daten das neue Öl sind und was das mit Forschung zu tun hat – von Ralph Szymanowsky



Spätestens mit der aktuellen Pandemie ist auch dem Letzten klar geworden, dass es überaus wichtig ist, in Europa unabhängige Forschung zu betreiben. Das ist jedoch sehr viel leichter gesagt als getan – wie vieles, an das man sich heranwagen müsste.

Mit der Aufforderung zum Handeln/Forschen sind gleich mehrere Fragen verbunden:

Welche Plattform und welche Daten benötigt man dafür? Geht das konform mit dem Datenschutz oder begibt man sich auf dünnes Eis? Wer kann das – nur die Großen, oder ist es auch für Häuser der Regelversorgung und Fachkliniken sinnvoll? Wie kann der Patient in die Verwendung seiner Daten einwilligen, z. B. in Form einer Datenspende?

Der nachfolgende Artikel wird versuchen, eine Antwort auf einige dieser Fragen zu geben und grundsätzlich in das Thema einzuführen.

Seit geraumer Zeit wird der Spruch „Daten sind das neue Öl“ in der HealthCare-IT-Branche nicht mehr nur geraunt, sondern vielmehr von den Dächern gepfiffen. Der Ursprung des Zitats ist ein Artikel im Eco-

nomist von 2017. Die wirtschaftliche Bedeutung von Informationen wurde in diesem Artikel mit der von Öl für die sogenannte Old Economy gleichgesetzt. Dabei sind es nicht nur die Bedeutung und der Preis, sondern auch die Gewinnung und Aufbereitung, die durchaus interessante Parallelen aufzuweisen haben.

Daten sind das neue Öl

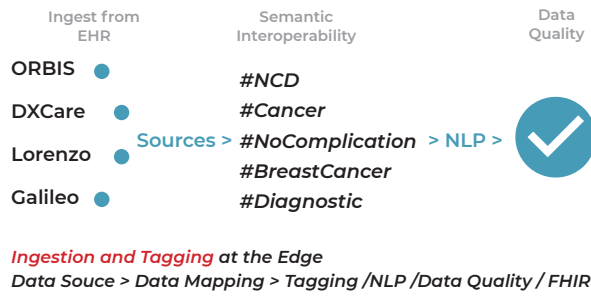
Öllagerstätten können schwer zugänglich und der Bodenschatz somit kaum verwertbar sein. Die Förderung des Rohöls bedeutet nicht automatisch ein hochwertiges Produkt, sondern bis zu diesem ist eine aufwendige Verarbeitung erforderlich – ganz abgesehen von unterschiedlichen Qualitäten des Rohöls und letztlich auch der Daten. Hand aufs Herz: Wie schätzen Sie die Qualität Ihres „Datenöls“ ein? Eine schwierige Frage, vor allem weil man nur punktuell mit der Qualität der Daten in Berührung kommt. Zumindest solange diese nicht gefördert und aufbereitet sind, ist eine realistische Einschätzung schwierig. Die Nutzer von BI-Systemen werden sich möglicherweise zurücklehnen und meinen, sie hätten das Thema im Griff und den Datenschatz quasi in der Hand – dem muss man leider widersprechen. BI-Systeme sind angetreten, um administrative Daten insbesondere für das Controlling zugänglich zu machen und bieten darüber hinaus gute Möglichkeiten, Prozesse und betriebswirtschaftliche Kennzahlen wie Kosten und Ergebnis transparent zu machen. Die Analyse insbesondere medizinischer Zusammenhänge erfordert Informationen zum Labor, der Medi-

kation usw. – und zwar nicht zu deren Kosten, sondern zum Behandlungskontext, in dem diese zum Patienten dokumentiert wurden.

Über Jahre hinweg wurde umfassendes Wissen über verschiedenste Krankheitsbilder gesammelt. Die Dokumentation wurde akribisch angelegt und enorme Summen wurden für die Diagnostik ausgegeben. Nichtsdestotrotz sind die daraus resultierenden Erkenntnisse meist auf den Erfahrungsgewinn einzelner Behandler und die Genesung einzelner Patienten begrenzt. Der Aufbau des ORBIS-Datenmodells ähnelt dem einer Bibliothek: Man findet die Werke einzelner Autoren und Genres schnell, allerdings wird es kompliziert, wenn sich die Frage auf Inhalte von Werken bezieht, wie zum Beispiel: „alle Romane, in denen von Epidemien die Rede ist“. Dann ist man weitgehend auf die Erfahrung und das Erinnerungsvermögen des Bibliothekars angewiesen. Entsprechende Suchalgorithmen? Fehlanzeige – was mit Blick auf die Aufgabe von Bibliotheken auch nachvollziehbar ist. Das Tagging fügt entsprechende Vermerke an, so dass eine schnelle Suche nach den Inhalten möglich ist, ohne im Suchprozess die gesamten Texte durchsuchen zu müssen.

Die erforderlichen Informationen sind in ORBIS bei einer Vielzahl von Kunden umfassend dokumentiert und der Umfang wächst ständig weiter – die entscheidende Frage ist, wie diese nutzbar gemacht werden können.

Edge Ecosystem



Analytics Workbench

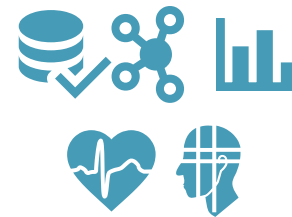
Edge Data Store

T4C Edge Explore for In-Hospital Research
Secure Research at the Edge (without Cloud)

Secure Research Hub



Feasibility Studies for the Researcher
A Portal to export Data and request the patient record



T4C Portal

Metadata Observatory

T4C Gravity is a secure Research Hub
A multi-source hub for researchers to find candidates

Ein leistungsfähiges Wissenspotenzial

Die Dedalus-Vision ist eine ganzheitliche Betrachtung von Daten in einem Medical Data Lake – inklusive Labor, Diagnostik, Befunddokumentation, Medikamenten, Monitoring, EKG –, um nicht nur Kosten, Prozesse und Ressourceneinsatz zu optimieren, sondern auch Erkenntnisse für die klinische Qualität generieren zu können. Die Auswertung aller verfügbaren Daten sollte Erkenntnisse zur Verbesserung der Behandlung, Evaluierung von Behandlungsrichtlinien und Aussagen zur Behandlungsqualität möglich machen.

Als technologische Grundlage wurde dabei bewusst auf FHIR gesetzt. Diese Technologie unterstützt den Datenaustausch zwischen Softwaresystemen im Gesundheitswesen über sogenannte „Ressourcen“.

Ziel dabei ist die Schaffung einer medizinischen Datenbasis zur Charakterisierung von Patientenpopulationen anhand verschiedener Parameter wie

- demografische Informationen,
- Krankheitsverläufe,
- Labor,
- Medikation,
- Morbidität,
- Behandlungen und Behandlungsabfolgen.

Die an OMOP angelehnte herstellerunabhängige Organisation der Daten richtet sich dabei an der medi-

zinen Analyse aus und ist personenzentriert aufgebaut; alle klinischen Ereignistabellen sind mit der Tabelle/Ressource Person verknüpft.

1. Die Daten-Generierung

Den aktuell abgebildeten Ressourcen entsprechend werden die zugehörigen Rohdaten per SQL aus den KIS-Systemen extrahiert. Die Extraktion der Daten aus ORBIS ist als Standard bereits integriert. Die umfangreiche Erfahrung von TIP HCe in der Anbindung entsprechender Systeme ermöglicht außerdem die Integration weiterer KIS/KAS/EMR.

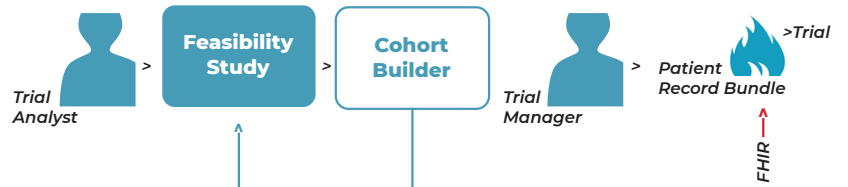
2. Die Daten-Transformation

Die Überleitung der Daten in FHIR-Formate ist zum einen Voraussetzung für die flexible Analyse der Daten und zum anderen eine Möglichkeit, zukünftig direkt Daten aus

Der Business Flow

CRO

T4C Gravity



T4FC Edge



Hospital



den verschiedenen Systemen via FHIR übernehmen zu können.

Das Tagging und Staging der Daten ist Teil der Datenanreicherung für analyserelevante Kommentierungen der übernommenen Rohdaten.

Es reicht nicht aus, in Befunden nach Schlagworten oder prägnanten Formulierungen zu suchen, da deren Kontext sich aus dem Gesamttext ergibt und für die korrekte Suche nicht ausreichend ist. Erst das Tagging reichert die Daten zum Patienten automatisch um hierarchisch strukturierte Tags an und bietet über die semantische Klassifizierung systemübergreifende Auswertungsmöglichkeiten.

3. Die Daten-Analyse

Die Analyse der Daten wird dem Anwender über Dashboards ermöglicht, dabei besteht an verschiedenen Stellen die Möglichkeit des Exports von Detaildaten. Die einfache Definition und Kombination von Filtern wird über die GUI unterstützt. Die Dashboards können intern bereitgestellt und mit wenigen Schritten auf spezielle Krankheitsbilder adaptiert werden. Je nach Berechtigung der Hospital Researcher können die Informationen anonymisiert oder dem Patienten zuordenbar zur Verfügung gestellt werden.

4. Die Datenanonymisierung

Der Patientenbezug der FHIR-Objekte wird in diesem Schritt aufgelöst und durch eine Collection-ID ersetzt. Von den Freitexten aus Befunden und Briefen werden nur die Tags zur Recherche angeboten; damit wird eine Identifikation des Patienten ausgeschlossen.

5. Die Datenbereitstellung

Im Rahmen der Forschungsaktivi-

täten werden ausschließlich anonymisierte Meta-Daten in einem T4C-Portal bereitgestellt. Die potenzielle Patientenkohorte für Studien kann über Filter auf den aggregierten Daten erstellt und dem Hospital Trial-Koordinator übermittelt werden. Die Anwendung der Ein- und Ausschlusskriterien gibt dem Hospital Trial-Koordinator die Möglichkeit, die betreffenden Patienten zu identifizieren.

Diese Rolle entscheidet, ob man mit dem forschenden Unternehmen zusammenarbeiten will und je nach Anforderung und Einverständnis des Patienten weitere Daten zur Verfügung stellt oder die Anfrage ablehnt.

Der Ausblick: Woran arbeiten wir?

Die Nutzung gescannter Dokumente durch das Tagging der enthaltenen Texte als eine weitere mögliche Informationsquelle ist in der Erprobung und würde zur Vervollständigung der Akte beitragen. Das Verfahren des Taggings und Stagings muss auf einzelne Krankheitsbilder optimiert werden. Insbesondere dabei steht die Zusammenarbeit mit entsprechenden Fachleuten im Fokus. Diese Zusammenarbeit führt zu fachlich fundierten Verfahren, die später allen Kunden zur Verfügung stehen werden. Daten sind essenziell – ihre Auswertung kann Leben retten.

Ein häufiger Wunsch von Patienten ist es, über neue Therapiemöglichkeiten informiert zu werden. Universitäre Einrichtungen können mit der Lösung klinische Forschung besser vorantreiben. Nicht-universitären Häusern eröffnet sich die Möglichkeit, Patienten durch die Bereitstellung anonymisierter Daten die Teilnahme an Studien zu ermöglichen.

Der Schutz der persönlichen Daten

und die Entscheidungshoheit des Patienten haben höchste Priorität.



Medizin als empirische Wissenschaft lebt von den Patientendaten. Im digitalen Zeitalter unternehmen wir große Anstrengungen, um diese Daten so einfach wie früher die Papierakte, nur ungleich schneller für die Forschung verfügbar zu machen. So können wir rasch den richtigen Probanden die Teilnahme an klinischen Studien anbieten, Bioproben im Labor untersuchen und auch seltene Erkrankungen klinisch erforschen."

Dr. med. Thorsten Hornung
Universitätsklinikum Bonn