Clinical Communication Platform (CCP-IT)

Datenschutzkonzept

Projekt	Clinical Communication Platform (CCP-IT) im Deutschen Konsortium für Translationale Krebsforschung (DKTK)
Autoren	Andreas Borg und Martin Lablans Universitätsmedizin der Johannes Gutenberg-Universität Mainz Institut für Medizinische Biometrie, Epidemiologie und Informatik 55101 Mainz
Träger	Deutsches Krebsforschungszentrum Im Neuenheimer Feld 280 69120 Heidelberg
Version	~~8. Januar~~ 10. Oktober 2014 // VORSCHLAG zur Konsentierung am 4.11.2014

JG|U UNIVERSITÄTSmedizin. MAINZ

Inhalt

1. Einleitung

1.1 Zielsetzung

Vor dem Hintergrund des Wissens um das komplexe Zusammenwirken von individueller genetischer Disposition, Lebensstil und Umweltfaktoren für die Entstehung und den Verlauf von Krebserkrankungen verlangt die heutige Krebsforschung die Beobachtung von Krankheitsverläufen, individuellen Lebensgewohnheiten, und Umweltbedingungen über lange Zeiträume. Nur durch die Forschung in standortübergreifenden Verbünden mit ausreichenden Fallzahlen kann auch die Ursächlichkeit individueller genetischer Dispositionen erforscht werden.

Das DKTK hat sich zur Aufgabe gemacht, durch den Aufbau von effizienten translationalen Forschungseinheiten für anwendungsnahe Krebsforschung an bundesweit vernetzten Partnerstandorten die Erforschung der Entstehungsmechanismen von Krebserkrankungen sowie die Entwicklung optimierter Instrumente für eine gezielte Behandlung und für eine frühzeitige Erkennung von Krebserkrankungen voranzutreiben.

Die CCP-IT stellt dafür eine IT-Infrastruktur zur Verfügung, die es ermöglicht, vor allem Fallzahlen zu definierten Krankheitsbildern standortübergreifend zu ermitteln, aber auch Probanden für klinische Studien zu rekrutieren oder bereits vorliegende Daten für Forschungsfragestellungen zur Auswertung anzufordern. Dazu werden an den beteiligten Standorten Daten aus der klinischen Krebsdokumentation sowie zu dort vorhandenen Biomaterialproben erhoben und je nach Möglichkeit in einer zentralen Datensammlung oder lokal an den Standorten gesammelt. Um diese Daten für die Forschung nutzbar zu machen, stellt die CCP-IT eine zentrale Suchschnittstelle bereit, die es Forschern ermöglicht

- abzuschätzen, ob für ein Forschungsvorhaben im DKTK genügend Probanden mit den betrachteten Eigenschaften auffindbar sind,
- zu ermitteln, welche Institutionen passende Patienten behandeln oder behandelt haben und
- Anfragen zur Nutzung von medizinischen Daten und Biomaterialproben dieser Patienten für Forschungsvorhaben zu stellen.

Eine Besonderheit gegenüber anderen Forschungsverbünden besteht darin, dass auch Bestandsdaten, die vor der Einrichtung des DKTK im Behandlungskontext erhoben wurden, für die Forschung nutzbar gemacht werden sollen. Dieser Nutzung sind enge datenschutzrechtliche Grenzen gesetzt, weil sie zunächst nicht durch eine Patienteneinwilligung gedeckt ist. Andererseits kann dadurch die sonst nötige Wartezeit zwischen Beginn der Datenerhebung und dem Erreichen eines für Forschungszwecke hinreichend großen Datenbestandes erheblich reduziert werden. Darüber hinaus ermöglicht der Einbezug von Altdaten die Evaluation von Therapiefortschritten. Das Ziel dieses Datenschutzkonzepts ist, in dieser Situation, in der Anforderungen bezüglich des Datenschutzes besonders stark mit Wünschen hinsichtlich der Nutzbarmachung für die medizinische Forschung konkurrieren, einen wirksamen und den rechtlichen Anforderungen genügenden Datenschutz sicherzustellen.

1.2 Überblick über die Datenverarbeitung

In der CCP-IT werden Daten von Tumorpatienten, die an den teilnehmenden Kliniken (auch als „Standorte“ bezeichnet) behandelt werden, erhoben und verarbeitet. Sie werden zum größten Teil aus vorhandenen Datenverarbeitungssystemen (z.B. Krankenhausinformationssysteme und Software zur Tumordokumentation), in die zentralen Komponenten der CCP-IT eingebracht. Darüber hinaus erlaubt die Komponente „Lokales Datenmanagement“ (siehe Abschnitt 2.1) die manuelle Eingabe von Daten zu Biomaterialproben.

Grundsätzlich teilen sich die erhobenen Daten in medizinische und identifizierende Daten auf, die im Folgenden in Anlehnung an die TMF[1]-Datenschutzkonzepte als MDAT und IDAT bezeichnet werden. Die erhobenen MDAT umfassen die sogenannten Meldedatensätze *MDS-K* (klinische Daten, wie zum Beispiel codierte Diagnosen und Tumor-Klassifikationen) und *MDS-B* (Daten zu Biomaterialproben). Weiterhin ist vermerkt, an welchen Studien des DKTK (in der Regel eine) der Patient[2] teilnimmt, sowie welche Experimente mit seinen Biomaterialproben bereits durchgeführt werden.[3] Die IDAT enthalten demografische Daten (~~z.B. Name, Geburtsdatum,~~ genaue Definition siehe Abschnitt „Kontrollnummern-Erzeuger“, S. 8), ~~Geschlecht~~), die eine eindeutige Identifikation des Patienten erlauben. Genauere Informationen zum Umfang dieser Daten finden sich in den Darstellungen der folgenden Abschnitte sowie im Anhang.

1.3 Rechtsgrundlage

In Hinblick auf die Rechtsgrundlage sowie die Prozesse der Datenverarbeitung ist zwischen folgenden Patientengruppen zu unterscheiden:

1. Patienten, die der Verwendung ihrer Daten im DKTK aktiv zugestimmt haben („DKTK-Patienten“). Hier ist die informierte Einwilligung des Patienten (siehe Abschnitt 6.1) Rechtsgrundlage der Datenverarbeitung. Auf der Grundlage dieser Einwilligung können MDAT in eine zentrale Komponente (die sogenannte *MDS-Datenbank*) in pseudonymisierter Form exportiert werden und können dort über die *zentrale Suche* von DKTK-Forschern durchsucht werden.
2. Patienten, die der Verwendung ihrer Daten nicht explizit zugestimmt haben („nicht-DKTK-Patienten“). Generell betrifft dieser Fall insbesondere Bestandsdaten, die aus der Zeit vor Errichtung des DKTK stammen und damit wesentlich zahlreicher als DKTK-Daten sind. MDAT dieser Patienten sollen in eine lokale Komponente, den *Brückenkopf*, in pseudonymisierter Form importiert und dort gespeichert werden. Der Brückenkopf steht unter lokaler Kontrolle der behandelnden Einheit des jeweiligen Standorts, und auch nur dort kann (mit erheblichem technischen Aufwand) mithilfe des Pseudonyms auf die Identität des Patienten geschlossen werden. Auf diesen lokal verbleibenden MDAT können zwar im Rahmen einer sogenannten *dezentralen Suche* Suchanfragen von außerhalb des Standorts durchgeführt werden, die Ergebnisse dieser Suchanfragen sind aber im Rahmen der CCP-IT nur nach manueller Freigabe durch den Standort, der die Daten besitzt, für den Anfragenden sichtbar. Rechtsgrundlage sind hier die am Standort anwendbaren landes- und bundesrechtlichen Datenschutzbestimmungen. (vgl. Abschnitt 6.2).

In beiden Fällen werden IDAT durch ein geeignetes Verfahren so verschlüsselt, dass eine exakte Rückführung auf die Eingabedaten durch einen Dritten praktisch unmöglich ist, und nur in dieser faktisch anonymen Form in einer zentralen Datenbank gespeichert. In Fall (2) ist darüber hinaus der verwendete Schlüssel nur dem jeweiligen Standort bekannt, um bestimmte Angriffsszenarien zusätzlich zu erschweren. Eine umfassende Beschreibung der Prozesse zur Verarbeitung der IDAT findet sich im Abschnitt 11.

Nutzer bzw. Empfänger von Daten sind Forscher des DKTK[4]. Aus deren Sicht gibt es, analog zu den zwei betroffenen Patientenkollektiven, zwei verschiedene Zugriffswege (eine detaillierte Beschreibung der Komponenten und Prozesse wird in den folgenden Abschnitten gegeben):

[1] Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.

[2] „Patient“, „Arzt“ und ähnliche Begriffe bezeichnen in diesem Dokument Funktionsrollen und meinen Personen jeglichen Geschlechts. Auf eine gendergerechte Formulierung wurde aus sprachlichen Gründen verzichtet.

[3] Letzteres dient dazu, die mehrfache Erhebung von Daten zu vermeiden.

[4] Siehe Abschnitt 4.2, „Teilnehmende Forscher“, für die genaue Bestimmung dieses Personenkreises.

- Pseudonymisierte Datensätze von DKTK-Patienten, die in der *Zentralen MDS-Datenbank* gespeichert sind, können von Forschern nach vorgegebenen Kriterien direkt durchsucht werden. Hier wird direkt ein Ergebnis in Form eines zusammenfassenden Überblicks über die gefundenen Datensätze zurückgeliefert.
- Anfragen zu Datensätzen von nicht-DKTK-Patienten können über die *Dezentrale Suche* formuliert werden. Diese Suchanfragen durchsuchen keine zentrale Datenbank, sondern werden an die Standorte übermittelt. Der Anfragende sieht zunächst kein Suchergebnis, sondern die die auf die Suchkriterien passenden Datensätze werden nur dem Dateninhaber am Standort angezeigt. Dieser hat dann die Möglichkeit, die Anfrage auf inhaltliche Kriterien und rechtliche Zulässigkeit zu prüfen und, falls beides positiv ausfällt, manuell auf diese Anfrage zu antworten. Diese Art der Suche kommt am ehesten einer klassischen schriftlichen Anfrage gleich, nur dass ihre Begutachtung und Beantwortung durch technische Hilfsmittel unterstützt werden.

1.4 Träger

Träger des Vorhabens ist das Deutsche Krebsforschungzentrum (DKFZ), Heidelberg.

2. Datenverarbeitende Komponenten

2.1 Brückenkopf

Der Brückenkopf dient dazu, die Daten eines Standorts in ein DKTK-kompatibles Format zu überführen und für die anderen Komponenten nutzbar zu machen. Seine Aufgaben sind:

- *Datenharmonisierung*: Daten werden im Brückenkopf so harmonisiert und dupliziert gespeichert, dass sie von den übrigen Komponenten der CCP-IT verstanden werden.
- *Sichtbarmachung für das DKTK*: z.B. für zentrale und dezentrale Suche.
- *Einhaltung von Datenhoheit*: Der Brückenkopf erlaubt dem Standort eine Teilnahme an der CCP-IT auch ohne „Upload auf Verdacht“ seiner patientenbezogenen Daten an eine externe Stelle, was Datenschutz und Datenhoheit fördert.

Der Brückenkopf besteht aus den folgenden, lokal in der Klinik installierten, Softwarekomponenten:

- *Lokales Datenmanagement*: Bereitet die in den lokalen Primärsystemen prinzipiell vorliegenden, aber unterschiedlich strukturierten Datenbestände für eine Nutzung im DKTK auf. Bietet für Biomaterialbanken an den Standorten eine rudimentäre Proben- und Materialverwaltung samt Nutzer- und Gruppenverwaltung und Formularhandling. Diese Komponente entspricht funktional und technisch weitgehend einem Clinical Data Warehouse.
- *Teiler*: Leistet eine kontrollierte Freigabe der Datenbestände des lokalen Datenmanagements zur Nutzung durch Projekte des DKTK. Dabei kommen zwei sich ergänzende „Teilmethoden“ zum Einsatz: Eine zentrale Suche gibt sofort erste Ergebnisse aus, dann folgt eine langsamere, aber dafür umfassendere dezentrale Suche.
- *Lokales Identitätsmanagement*: Stellt für die Pseudonymisierung sowohl von DKTK- als auch von nicht-DKTK- Patienten eine einheitliche Schnittstelle bereit.

Die im Brückenkopf gespeicherten MDAT können prinzipiell alle Elemente des einheitlichen onkologischen Basisdatensatzes der Arbeitsgemeinschaft Deutscher Tumorzentren (http://www.tumorzentren.de/onkol-basisdatensatz.html) sowie Daten zu Biomaterialproben umfassen.

Diese Komponenten stehen unter Kontrolle des jeweiligen Zentrums, das heißt, die in diesen Komponenten gespeicherten Daten stehen weiter unter der Hoheit der Institution, in der sie erhoben wurden. Gegebenenfalls ist der

Zugriff auf die behandelnde Einheit, z.B. die Fachabteilung, einzuschränken (vgl. auch Abschnitt 7, „Lokale Umgebung“).

2.2 Zentrales Identitätsmanagement

Pseudonymisierung ist ein zur Aufrechterhaltung eines hohen Datenschutzniveaus notwendiger Schritt, um den Patienten vor Rückidentifizierung zu schützen. Anstelle seiner identifizierenden Daten (IDAT) treten Pseudonyme. In der CCP-IT kommen folgende Arten von Pseudonymen zum Einsatz:

- *S#ID*: Ein lokaler Identifikator, der nur für den Standort # eindeutig ist und keine standortübergreifende Verknüpfung von Daten eines Patienten erlaubt.
- *DKTK#ID*: Ein lokaler Identifikator, der nur für den Standort # eindeutig ist. Im Gegensatz zur S#ID können DKTK#IDs, die an verschiedenen Standorten zu einem Patienten existieren, in der zentralen Patientenliste einander zugeordnet werden.
- *MDS-ID*: Ein Identifikator, der nur innerhalb der zentralen MDS-Datenbank eindeutig ist. In der zentralen Patientenliste kann die MDS-ID den DKTK#IDs eines Patienten zugeordnet werden.

Im Falle einer geeigneten Einwilligung wird die jeweils standortspezifische DKTK#ID, erzeugt, welche auch (mithilfe der zentralen Patientenliste) eine Verfolgung des Patienten über Institutionsgrenzen hinweg erlaubt. Im Fall einer fehlenden DKTK-Einwilligung dürfen Klarnamen den Behandlungskontext des Standorts nicht verlassen; es wird dann ein lediglich die nur lokal vergleichbare *S#ID*, generiert.

Das zentrale Identitätsmanagement ermöglicht eine datenschutzgerechte Zusammenführung („Record Linkage“) der von mehreren Standorten gesendeten patientenbeziehbaren Daten. Dazu werden drei Methoden/Werkzeuge kombiniert, vgl. folgende Unterabschnitte und Abbildung 1.

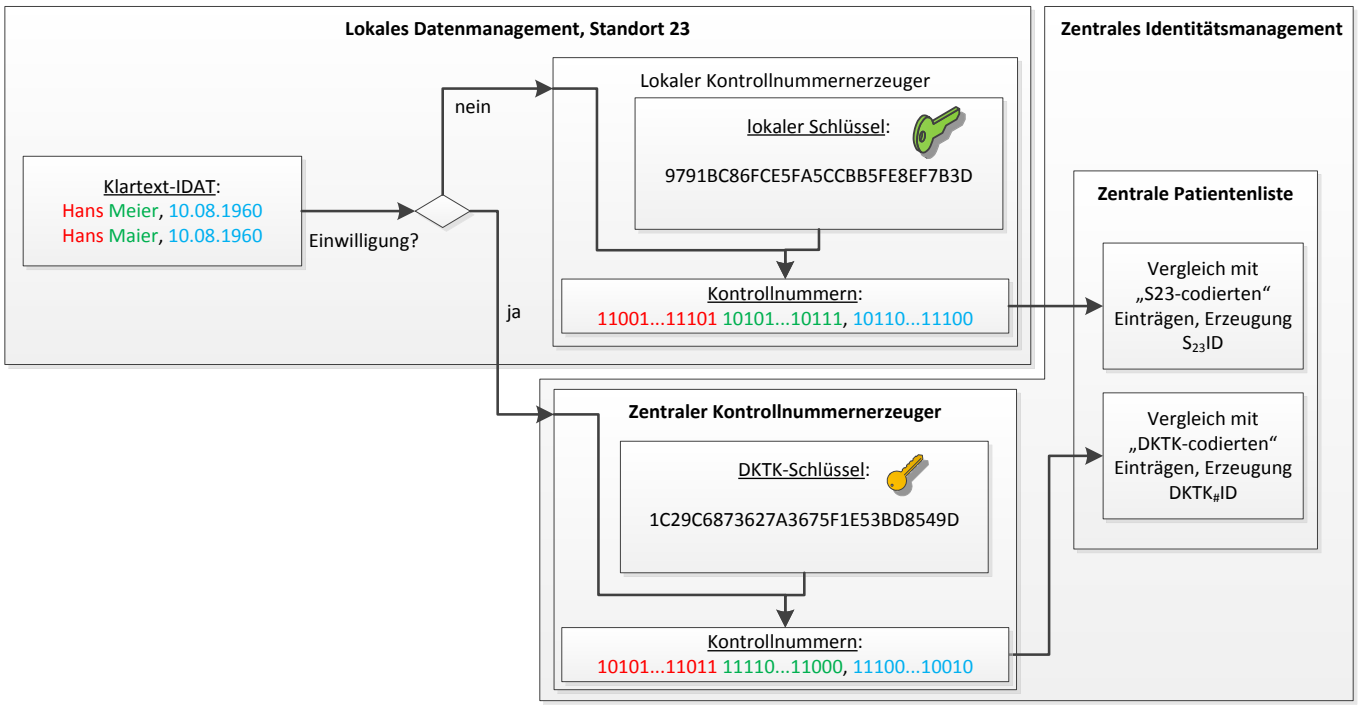

Abbildung 1 - Beispielhaftes Record Linkage-Verfahren.[5]

[5] „Einwilligung? → nein“ bedeutet hier, dass noch keine DKTK-Einwilligung eingeholt wurde. Bei explizitem Widerruf gehen von vornherein keine Daten in die CCP-IT ein.

Kontrollnummern-Erzeuger

Wandelt Identifikationsdaten (Klartext-IDAT) durch eine spezielle Einwegverschlüsselung in unlesbare und nicht rückführbare, aber immer noch gewichtet vergleichbare Zeichenketten („Kontrollnummern“[6]) um. Um Wörterbuchattacken zu vermeiden, wird dabei ein Keyed-Hash Message Authentication Code verwendet, wobei der verwendete Schlüssel (im folgenden „Geheimnis“) für jede Instanz des Kontrollnummern-Erzeugers eindeutig und nur dort bekannt ist.[7] Daraus folgt die Voraussetzung eines organisatorisch unabhängigen Betriebs dieses Moduls.

Die an den Kontrollnummern-Erzeuger gesendeten IDAT bestehen aus folgenden Attributen:

- Vorname
- Nachname
- Frühere Namen (z.B. Geburtsname bei Namensänderung durch Heirat)
- Geburtsdatum, aufgetrennt in die Komponenten Tag, Monat und Jahr
- Staatsangehörigkeit
- Geschlecht

Für die Pseudonymisierung von DKTK-Patienten kommt ein zentraler Kontrollnummern-Erzeuger zum Einsatz. Als weitere Sicherheitsmaßnahme (z.B. gegen Angriffe mit Probeverschlüsselungen), werden die dort erzeugten Kontrollnummern nicht an den Brückenkopf zurückgegeben, sondern direkt an die zentrale Patientenliste übermittelt. Die DKTK#ID wiederum wird nie dem zentralen Kontrollnummern-Erzeuger bekannt gemacht (Details siehe Abschnitt 3.2). Bei nicht-DKTK-Patienten erfolgt die Kontrollnummernerzeugung im lokalen Identitätsmanagement.

Patientenliste

Vergibt für eine Gruppe von Kontrollnummern eines Patienten einen Personenidentifikator (DKTK#ID, S#ID, MDS-ID). Ihr Kontrollnummern-Matcher kann dabei Patienten selbst bei abweichender Schreibweise und im Fall der DKTK#ID auch aus unterschiedlichen Standorten wiedererkennen.

Manuelles Linken

Eine Schnittstelle erlaubt einem Menschen, Ergebnisse des automatischen Matching zu überprüfen und ggfls. zu korrigieren, d.h. Duplikate zusammenzuführen oder fälschlicherweise zusammengeführte Datensätze zu trennen. Hierfür werden die Matchgewichte (Vergleichswerte zwischen den einzelnen Attributen von zu prüfenden Patienten) angezeigt und es besteht die Möglichkeit, zur Entscheidungsfindung auf medizinische Daten zuzugreifen. Diese Komponente ist deshalb beim Betreiber der zentralen MDS-Datenbank, wo ohnehin MDAT von DKTK-Patienten gespeichert sind, angesiedelt.

2.3 Zentrale MDS-Datenbank

Die zentrale MDS-Datenbank nimmt mithilfe einer Webschnittstelle die von den Exportern der Teiler verschickten MDAT von DKTK-Patienten entgegen und verwahrt sie in einer Datenbank. Die Speicherung erfolgt zusammen mit der MDS-ID, um die Zuordnung verschiedener Datensätze eines Patienten zueinander möglich zu machen. Die MDAT umfassen ~~folgende definierte~~ Meldedatensätze ~~(vgl. genaue Definition im Anhang):~~zweier Klassen:

[6] Streng genommen sind das also keine Nummern. Idee und Name wurden (in angepasster Form) aus dem Bereich epidemiologischer Krebsregister übernommen.

[7] Für technische Details siehe: Schnell R, Bachteler T, Reiher J: Privacy-preserving record linkage using Bloom filters. BMC Medical Informatics and Decision Making 2009, 9:41. http://www.biomedcentral.com/1472-6947/9/41

- *MDS-K*: Klinische Daten aus der Tumordokumentation. Diese beinhalten Daten zum Patienten, zum Primärtumor, zur Primärtherapie, zum Ansprechen und zum Vitalstatus. Der MDS-K besteht aus einer Teilmenge des einheitlichen onkologischen Basisdatensatzes der Arbeitsgemeinschaft deutscher Tumorzentren (ADT-Basisdatensatz)[8].
- ~~MDS-B: Metadaten zu Biomaterialproben.~~
- *MDS-B*: Metadaten zu Biomaterialproben. Der MDS-B umfasst zum einen Attribute, die lokal das Auffinden und Zuordnen einer Probe zu einem Patienten ermöglichen (z.B. durch Proben- und Patienten-ID), zum andern allgemeine Informationen zur Beschreibung des Biomaterials (z.B. Gewebetyp, Probenart).

Autorisierte Nutzer können in einer Suchmaske mittels Suchkriterien aus MDS-K und MDS-B Anfragen auf diesem Datenbestand ausführen, diese wird im Abschnitt 3.4 („Zentrale Suche“) beschrieben. Im Rahmen der Protokollierung von Suchvorgängen können identifizierende Daten des zugreifenden Forschers gespeichert werden.

2.4 Suchbroker für die dezentrale Suche

Der Suchbroker für die dezentrale Suche stellt eine Schnittstelle zur Formulierung von Suchanfragen zur Verfügung und verwaltet diese Anfragen. Er verarbeitet keine personenbezogenen Daten von Patienten. Personenbezogene Daten von zugreifenden Benutzern können im Rahmen der Protokollierung gespeichert werden.

2.5 Metadata Repository

Das Metadata Repository (MDR) speichert die Bedeutung (Semantik) sämtlicher im DKTK verwendeten (Nutz-) Datenelemente. Es bietet ein kontrolliertes Vokabular (Syntax) und kann maschinenlesbare, strukturierte Aussagen über Datenelemente machen, bspw. konzeptuelle Domänen oder Wertebereiche. Hier sind auch die ~~in diesem Konzept genannten und im Anhang definierten~~ Meldedatensätze definiert. Da das MDR keine personenbezogenen Daten verarbeitet, wird innerhalb dieses Datenschutzkonzepts nicht weiter darauf eingegangen.

[8] www.tumorzentren.de/onkol-basisdatensatz.html

3. Datenverarbeitende Prozesse

3.1 Import in Brückenkopf

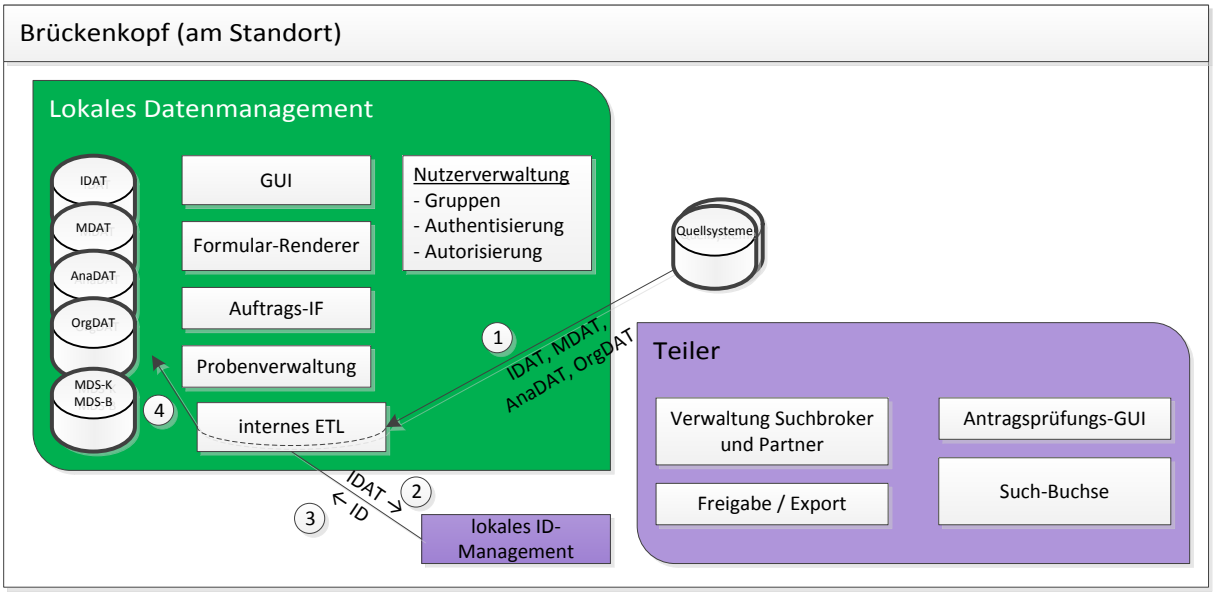

Abbildung 2 - Import von Daten in den Brückenkopf.[9]

Abbildung 2 zeigt, wie Daten aus Quellsystemen eines Standorts in den Brückenkopf importiert werden:

1. Identifizierende, medizinische und Probendaten werden aus mehreren Quellsystemen durch einen ETL-Prozess extrahiert.
2. Identifizierende Daten werden mithilfe des lokalen Identitätsmanagements mit einem primären Patientenidentifikator erster Stufe versehen (Details siehe Abschnitt 3.2, „Pseudonymisierung“).
3. Der ETL-Prozess ordnet den Identifikator dem Datensatz zu.
4. Daten werden im lokalen Datenmanagement abgelegt.

Die Speicherung von identifizierenden Daten zusammen mit klinischen oder Probendaten kann unkritisch sein, solange diese (wie im Brückenkopf) den Standort nicht verlassen. Sollte an einem Standort die lokale Zusammenführung von klinischen und Biobankdaten ausgeschlossen sein, können mehrere Brückenkopf-Instanzen zum Einsatz kommen, die dann unter getrennter lokaler Hoheit stehen. An einem Standort stehen dann mehrere Brückenköpfe, z.B. einer für die Daten der Tumordokumentation und einer für die der Biobank, vgl. Abbildung 3.

[9] Beispiele: IDAT = Vorname, Nachname, Geburtsdatum, DKTK-Einwilligung; MDAT = klinische Daten (Obermenge des MDS-K); AnaDAT = Daten zur Bioprobe (Obermenge des MDS-B); OrgDAT = Verwaltungsdaten zur Probe

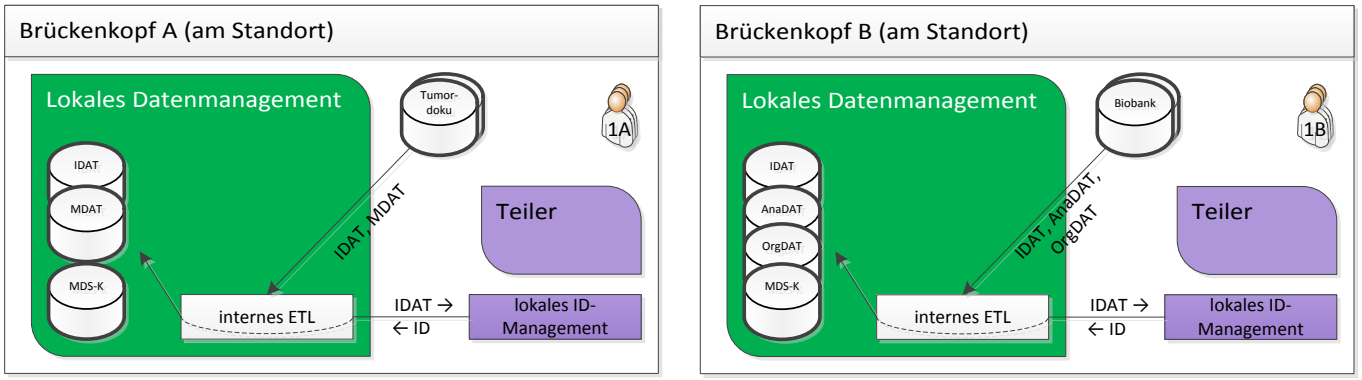

Abbildung 3 - Beispiel für Verwendung mehrerer Brückenköpfe an einem Standort. Brückenkopf A erhält nur die Daten der Tumordokumentation, Brückenkopf B nur die der Biobank. Beide werden von denselben Systembetreuern gewartet wie ihre Quellsysteme.

3.2 Pseudonymisierung

Aus Datenschutzgründen entstehen zwei verschiedene Datenflüsse, je nachdem ob für den zu pseudonymisierenden Patienten eine DKTK-spezifische Patienteneinwilligung vorliegt („DKTK-Patient“) oder nicht. Die Fälle unterscheiden sich wie folgt:

- Fall *mit Einwilligung*: Es wird ein lokales, aber mithilfe der zentralen Patientenliste verbundweit verknüpfbares Pseudonym ($DKTK_{\#}ID$) erzeugt, das eine Zuordnung des Patienten auch über Standortgrenzen hinweg erlaubt.
- Fall *ohne Einwilligung*: Es dürfen keine IDAT den Standort verlassen. Es wird ein lokales Pseudonym ($S_{\#}ID$) erzeugt, d.h. eines, das nur innerhalb des Standorts zugeordnet werden kann. Man kann es etwa dazu verwenden um passendes Biomaterial zu einem klinischen Datensatz zu finden. Es ist aber nicht möglich, einen Patienten über Standortgrenzen hinweg zuzuordnen.

Abbildung 4 zeigt, wie die Pseudonymisierung im Rahmen des ETL-Prozesses genutzt wird um ein Pseudonym zu erhalten.

1. Das lokale Identitätsmanagement erhält die Klartext-IDAT eines Patienten.
2. Klartext-IDAT werden übertragen an den Kontrollnummern-Erzeuger…
   a. im Fall *mit Einwilligung*: …des zentralen Identitätsmanagements.
   b. im *Fall ohne Einwilligung*: …des lokalen Identitätsmanagements.
3. Der jeweilige empfangende Kontrollnummern-Erzeuger errechnet aus den Klartext-IDAT und seinem Geheimnis Kontrollnummern (KN).
4. nur im Fall *mit Einwilligung*: Der zentrale KN-Erzeuger übermittelt die (DKTK-weit vergleichbaren) Kontrollnummern an die Patientenliste. Diese erstellt ein temporäres Kontrollnummern-Ticket (KNTKT) und gibt dies an das lokale Identitätsmanagement zurück.
5. Das lokale Identitätsmanagement übermittelt die erzeugten Kontrollnummern, oder, im Fall *mit Einwilligung* das eben ~~erhaltenen~~erhaltene Ticket KNTKT, an die Patientenliste des zentralen Identitätsmanagements. Durch die Verwendung des Tickets bleiben die DKTK-weit vergleichbaren KN dem lokalen Identitätsmanagement verborgen.
6. Das zentrale Identitätsmanagement gleicht die erhaltenen Kontrollnummern mit den bestehenden ab (KN-Matcher) und gibt im Falle eines Treffers eine bestehende ID ($DKTK_{\#}ID$ bzw. $S_{\#}ID$) zurück oder erzeugt eine neue mithilfe des Pseudonym-Generators.

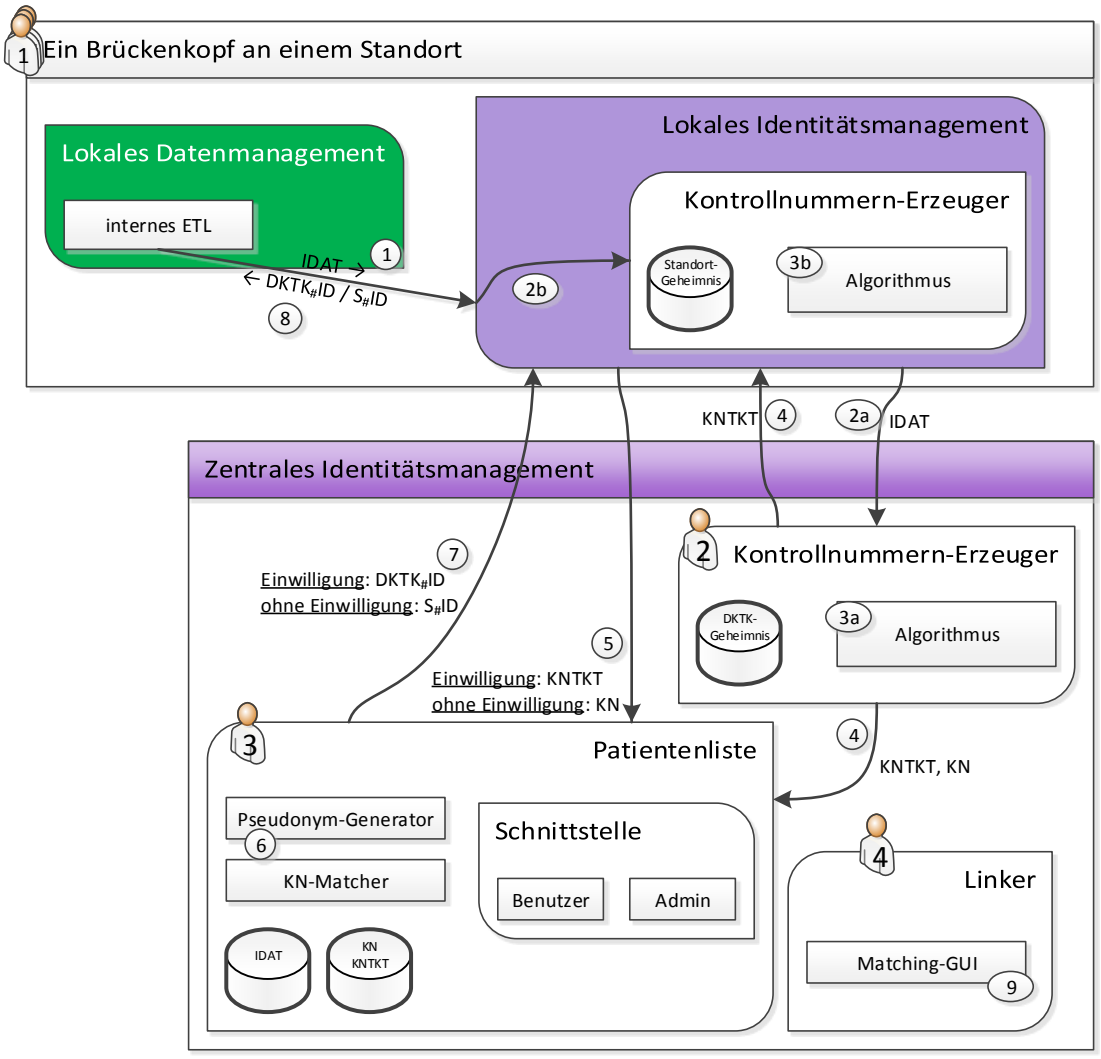

Abbildung 4 - Erzeugung lokaler und verbundweiter Pseudonyme.

7. Das lokale Identitätsmanagement erhält als Antwort auf seine Anfrage die ID. Es erfährt nicht, ob der Patient bereits bekannt war.
8. Die ID wird als Antwort auf die ursprüngliche Anfrage zurückgegeben. Es ist nicht erkennbar, ob der Patient bereits bekannt war. Es ist aber erkennbar, ob es sich um eine lokale S#ID oder um eine verbundweit verknüpfbare DKTK#ID handelt.
9. (optional) Sollte die Patientenliste sich nicht sicher sein, ob ein Patient schon in der Liste steht oder nicht, erfährt das der für das Record Linkage verantwortliche Administrator (④). Er entscheidet manuell auf Basis von Matchgewichten (aber keinen IDAT!) und – im Fall *mit Einwilligung* – Daten in der CCP-IT (v.a. MDAT), ob es sich um denselben Patienten handelt.

Erläuterung: Geheimnisse und Gültigkeit der Pseudonyme

Die Fälle *mit Einwilligung* und *ohne Einwilligung* unterscheiden sich genau in diesem Geheimnis: Das DKTK-weite Geheimnis, das genau dem Kontrollnummern-Erzeuger des zentralen Identitätsmanagements bekannt ist (Fall *mit Einwilligung*), erlaubt beim Matchen durch die zentrale Patientenliste eine Zuordnung des Patienten über Standortsgrenzen hinweg. Die lokalen Geheimnisse hingegen, die genau den Kontrollnummern-Erzeugern der lokalen Identitätsmanagement-Instanzen in allen Brückenköpfen desselben Standorts bekannt sind (Fall *ohne Einwilligung*), erlauben nur einen Abgleich der Patienten desselben Standorts. Das ist nötig um z.B. zu einem gegebenen klinischen

Datensatz das passende Biomaterial zu finden. Lokale Kontrollnummern werden, obgleich sie nur lokal Sinn ergeben, durch die zentrale Patientenliste abgeglichen. Das erspart die lokale Installation einer Patientenliste, wahrt aber weiterhin den Datenschutz, da die Kontrollnummern per Konstruktion eine Reidentifikation außerhalb des Standorts nahezu unmöglich machen.

3.3 Upload in zentrale MDS-Datenbank

Der Upload in die zentrale MDS-Datenbank erfolgt automatisch nach einem festen Zeitintervall (in der Regel täglich). Dabei werden aktuelle (d.h. seit dem letzten Upload hinzugekommene) medizinische Daten von DKTK-Patienten gemäß der Meldedatensätze MDS-K und MDS-B vom Teiler aus dem lokalen Datenmanagement ausgelesen und über eine sichere HTTP-Verbindung in die zentrale MDS-Datenbank exportiert.

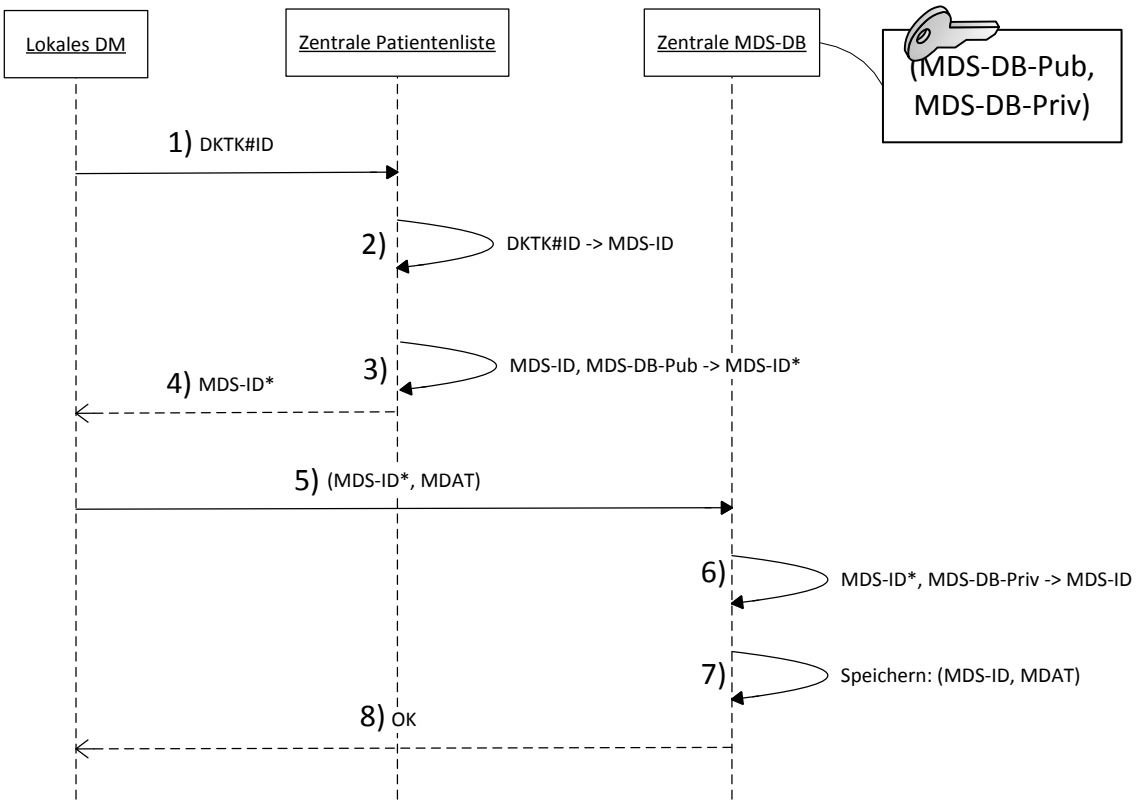

Abbildung 5 - Umpseudonymisierung beim Upload in die zentrale MDS-Datenbank

Um die richtige Zuordnung der Daten in der MDS-Datenbank zu gewährleisten, werden diese beim Export mit Hilfe der zentralen Patientenliste umpseudonymisiert, d.h. die DKTK#ID wird durch die zugehörige MDS-ID ersetzt. Dazu kommt ein asymmetrisches Verschlüsselungsverfahren zum Einsatz, wobei der private Schlüssel nur der zentralen MDS-Datenbank bekannt ist. Der Ablauf ist wie folgt (vgl. das Sequenzdiagramm in Abbildung 5):

1. Der Teiler übermittelt die DKTK#ID des betreffenden Patienten an die zentrale Patientenliste.
2. Die zentrale Patientenliste bestimmt die zugehörige MDS-ID.
3. Die zentrale Patientenliste verschlüsselt die MDS-ID asymmetrisch mit dem öffentlichen Schlüssel (*MDS-DB-Pub*) der zentralen MDS-Datenbank.
4. Die verschlüsselte MDS-ID (*MDS-ID**) wird an das lokale Datenmanagement zurückgegeben.
5. Verschlüsselte MDS-ID und MDAT werden vom lokalen Datenmanagement an die zentrale MDS-Datenbank übermittelt.

6. Die zentrale MDS-Datenbank entschlüsselt die MDS-ID mit ihrem privaten Schlüssel (*MDS-DB-Priv)*.
7. Die zentrale MDS-Datenbank speichert MDS-ID und MDAT.
8. Die erfolgreiche Bearbeitung wird dem lokalen Datenmanagement zurückgemeldet.

Um Wörterbuchattacken zu vermeiden, wird durch Randomisierung im Verschlüsselungsalgorithmus sichergestellt, dass die wiederholte (d.h. zeitlich hintereinander erfolgende) Verschlüsselung der gleichen MDS-ID verschiedene Chiffrate erzeugt.

Die übermittelten Daten können von den Administratoren von Brückenkopf und MDS-Datenbank im Rahmen von Wartungsarbeiten (vgl. Abschnitt „Zugriff durch Systemadministratoren“) eingesehen werden. Ansonsten erfolgt durch den Upload selbst keine Sichtbarmachung, die Daten stehen mit dem Upload aber für die Zentrale Suche (siehe unten) zur Verfügung.

3.4 Zentrale Suche

Mit der zentralen Suche können DKTK-Forscher den Datenbestand der zentralen MDS-Datenbank durchsuchen, um Standorte zu ermitteln, an denen Daten und Proben von DKTK-Patienten vorhanden sind, die für ein Forschungsvorhaben relevant sein könnten. Die zentrale MDS-Datenbank stellt dafür ein Webformular bereit, in dem Attribute der Meldedatensätze nach vorgegebenen Werten oder per Freitextsuche durchsucht werden können. Mehrere Suchattribute können durch logische Operatoren frei kombiniert werden.

Nach Ausführen der Suchanfrage ~~enthält~~erhält der anfragende Forscher maximal folgende Informationen[10]:

- Die Anzahl der Patienten bzw. Proben, die die Suchkriterien erfüllen.
- Die Altersverteilung der gefundenen Patienten in 10-Jahres-Intervallen.
- Die Geschlechtsverteilung (Anzahl Patienten männlichen / weiblichen / unbekannten Geschlechts).
- Die Verteilung auf die Standorte.
- Kontaktdaten der Standorte, an denen die Daten erhoben wurden.

Die Datensätze selbst werden nicht übermittelt. Ein eventueller Zugriff auf Daten durch den anfragenden Forscher sowie die datenschutzrechtliche Grundlage dafür wird zwischen diesem und dem jeweiligen Dateneigentümer (=Standort) verhandelt und findet außerhalb der zentralen MDS-Datenbank statt.

3.5 Dezentrale Suche

Wie die zentrale Suche dient die dezentrale Suche dem Auffinden von „passenden“ Patienten und Proben für ein Forschungsvorhaben, berücksichtigt im Gegensatz dazu aber auch Patienten ohne DKTK-Einwilligung. Da bei diesen angenommen werden muss, dass keine Rechtsgrundlage für die Übermittlung von Daten aus dem zuständigen Standort heraus gegeben ist, erfolgt hier keine automatische Übermittlung von Suchergebnissen.

Das Suchformular der dezentralen Suche wird vom Suchbroker für die dezentrale Suche bereitgestellt. Sie erlaubt nicht nur die Suche nach den Meldedatensätzen, sondern allen im Metadata Repository (vgl. 2.5) abgelegten Begriffen; zusätzlich sind Ergänzungen im Freitext (Prosa) möglich. Der anzufragende Datensatz ist hier also prinzipiell unbeschränkt. Die Suchanfrage wird zunächst gespeichert und dem anfragenden Forscher lediglich die Speicherung der Suchanfrage mitgeteilt. Die Teiler der Standorte rufen in regelmäßigen Abständen neu hinzugekommene Suchanfragen vom Suchbroker ab und ermitteln, welche Datensätze im lokalen Datenmanagement den Suchkriterien entsprechen. Der Inhalt der Anfrage sowie die gefundenen Datensätze können an jedem Standort von einer dazu

[10] Derzeit ist nur der erste Punkt (Anzahl Patienten bzw. Proben) umgesetzt, die weiteren Punkte geben den maximalen Umfang eventueller Erweiterungen wieder.

berechtigten Person eingesehen werden. Diese kann nun den anfragenden Forscher kontaktieren, um eine mögliche Weitergabe von Daten oder Proben zu vereinbaren. Dieser Vorgang erfolgt wiederum außerhalb der CCP-IT und die damit verbundenen datenschutzrechtlichen Fragen müssen im Einzelfall von den beteiligten Personen geklärt werden.

4. Organisatorische Rahmenbedingungen

Die datenverarbeitenden Personen und Institutionen sowie Datenempfänger in der CCP-IT verteilen sich auf die Betreiber der zentralen Komponenten sowie die am DKTK teilnehmenden Forscher.

4.1 Betrieb der Komponenten

Der Betrieb der Brückenköpfe erfolgt durch die im DKTK vertretenen Partnerstandorte:

- Berlin: Charité Comprehensive Cancer Center
- Dresden: Universitätsklinikum Dresden, Technische Universität Dresden, Helmholtz-Zentrum Dresden-Rossendorf, Max-Planck-Institut für Molekulare Zellbiologie und Genetik
- Essen / Düsseldorf: Westdeutsches Tumorzentrum am Universitätsklinikum Essen, Heinrich-Heine-Universität Düsseldorf
- Frankfurt / Mainz: Universitäres Centrum für Tumorerkrankungen Frankfurt, Johann Wolfgang Goethe-Universität Frankfurt, Georg-Speyer-Haus - Chemotherapeutisches Forschungsinstitut in Frankfurt am Main, Krankenhaus Nordwest Frankfurt, Universitätsmedizin Mainz
- Freiburg: Tumorzentrum Ludwig Heilmeyer - Comprehensive Cancer Center Freiburg, Albert-Ludwigs-Universität Freiburg, Universitätsklinikum Freiburg, Max-Planck-Institut Freiburg
- Heidelberg: Deutsches Krebsforschungszentrum (Kernzentrum), Universitätsklinikum Heidelberg, Ruprecht-Karls-Universität Heidelberg, Nationales Centrum für Tumorerkrankungen, Paul-Ehrlich-Institut (assoziierter Partner)
- München: Klinikum der Universität München, Klinikum rechts der Isar der TU München
- Tübingen: Universitätsklinikum Tübingen, Eberhard-Karls-Universität

~~Den~~Der Betrieb der zentralen Komponenten ~~der CCP-IT übernehmen ausgewählte Standorte, die in Absprache mit dem Lenkungsgremium des DKTK bestimmt werden. Dabei wird zum Zweck der informationellen Gewaltenteilung der organisatorisch unabhängige Betrieb von zentraler MDS-Datenbank,~~erfolgt an folgenden Stellen:

- Zentrale Patientenliste: Abteilung Medizininformatik, Institut für medizinische Biometrie, Epidemiologie und ~~zentralem~~Informatik, Universitätsmedizin der Johannes Gutenberg-Universität, Mainz.
- Zentraler Kontrollnummernerzeuger ~~gewährleistet (vgl. Abschnitt 4.1, „Informationelle Gewaltenteilung“).~~: Dezernat 7 – Informations- und Kommunikationstechnologie (DICT), Universitätsklinikum Frankfurt.
- Zentrale MDS-Datenbank und zentrale Suche: Abteilung Theoretische Bioinformatik, Deutsches Krebsforschungszentrum, Heidelberg.

4.2 Teilnehmende Forscher

Teilnehmende Forscher sind die Personen, die über die zentrale und die dezentrale Suche Anfragen an das System stellen können. Generell können alle Mitglieder der DKTK-Standorte als teilnehmende Forscher die CCP-IT nutzen, wobei jeder Standort selbst entscheidet, welche seiner Mitglieder eine Zugangsberechtigung erhalten (siehe auch Abschnitt 5.2, „Authentifizierung“).

Wissenschaftler, die nicht Mitglieder eines DKTK-Standorts sind, können auf Antrag vom Ausschuss für Datenschutz eine Zugangsberechtigung erhalten. Diese ist angemessen zu befristen.

4.3 Zugriff durch Systemadministratoren

Die in der CCP-IT gespeicherten Daten können prinzipiell von den Administratoren der verwendeten IT-Infrastruktur eingesehen werden. Zugriffe auf die Daten durch Administratoren dürfen nur erfolgen, wenn dies zur Erfüllung ihrer Aufgaben zwingend erforderlich ist. Alle Administratoren sind auf diesen Grundsatz und auf ihre Pflicht zur Verschwiegenheit hinzuweisen[11].

4.4 Ausschuss für Datenschutz

Vom Lenkungsausschuss des DKTK wird ein Ausschuss für Datenschutz eingesetzt. Dieser erfüllt insbesondere folgende Aufgaben:

- Prüfung und Bewilligung von Anträgen externer Forscher[12] für die Nutzung der CCP-IT (zentrale und dezentrale Suche).
- Prüfung und Bewilligung von Anträgen auf Export medizinischer Daten für externe Forschungsprojekte.
- Prüfung und Bewilligung von Anträgen auf die Benachrichtigung betroffener Patienten über Forschungsergebnisse.

Darüber hinaus ist der Ausschuss für Datenschutz erster Ansprechpartner für datenschutzrechtliche Angelegenheiten.

Der Ausschuss für Datenschutz wird so besetzt, dass jeder der DKTK-Standorte darin vertreten ist. Zu den Mitgliedern zählen mindestens:

- Ein Arzt~~[13]~~, der vorwiegend in der Behandlung ~~von~~ betroffener Patienten tätig ist.
- Ein Wissenschaftler, der mit den in der CCP-IT verwalteten Daten (oder Daten ähnlichen Typs) forscht.
- Ein Datenschutzbeauftragter oder ein mit dem Thema Datenschutz vertrauter Jurist.

Zusätzlich kann ein Vertreter der Entwickler der CCP-IT in beratender Funktion hinzugezogen werden.

5. Maßnahmen zum Datenschutz

5.1 Informationelle Gewaltenteilung

Konsequent durchgeführt wird eine informationelle Gewaltenteilung, angedeutet durch die nummerierten Personenbilder (④ in Abbildung 4). Das bedeutet, dass die Komponenten mit unterschiedlichen Nummern logisch, physikalisch und organisatorisch getrennt voneinander laufen, was die Gefahr eines Datenlecks verringert:

- Das zentrale Identitätsmanagement (②, ③) wird getrennt betrieben von den übrigen zentralen Komponenten der CCP-IT (④). So kann jemand, der in der CCP-IT auf klinische oder ~~BMB-Daten~~Biomaterialdaten Zugriff hat, diese keinen realen Patienten zuordnen.
- Innerhalb des zentralen Identitätsmanagements wird der Kontrollnummernerzeuger (②; zu schützen ist hier das verbundweite Geheimnis) getrennt betrieben von der Patientenliste (③). So ist sichergestellt, dass die in der Patientenliste gespeicherten Kontrollnummern keinen direkten Rückschluss auf die Identität des Patienten zulassen.
- Selbiges gilt natürlich für lokale Geheimnisse.

[11] Dies sollte in der Regel im Rahmen des Arbeitsverhältnisses an der zuständigen Institution ohnehin geschehen sein.
[12] D.h. Personen, die nicht Mitglied eines DKTK-Standorts sind.
~~[13] Gemeint sind entsprechende Personen beiden Geschlechts; auf eine neutrale Formulierung wurde aus sprachlichen Gründen verzichtet.~~

Für die konkreten Betreiber der zentralen Komponenten vgl. 4.1.

5.2 Authentifizierung

Authentifizierung von Benutzern

Die Authentifizierung von teilnehmenden Forschern (im Folgenden auch „Benutzer“) gegenüber der CCP-IT erfolgt über Benutzername und Passwort gegenüber einem zentralen Authentifizierungsdienst, der vom DKFZ betrieben wird. Die Prüfung von Identität und Berechtigung von Benutzern~~, die Mitglied eines DKTK-Standorts sind, obliegt dem~~ erfolgt dabei auf Standort ~~selbst.~~ oder Projektebene. Dazu ~~ernennt jeder~~ wird pro DKTK-Standort bzw. -Projekt eine zuständige Person ernannt, welche die Anträge zum Zugriff auf die CCP-IT entgegennimmt, Identität und Berechtigung prüft und dann dem DKFZ die freizuschaltenden Personen mitteilt.

Im Fall von externen Forschern prüft der Ausschuss für Datenschutz Identität und Berechtigung und veranlasst die Freischaltung durch das DKFZ.

Die Freischaltung von Benutzern des Brückenkopfs erfolgt direkt durch den jeweiligen Standort. Dabei sind lokale Regelungen des Datenschutzes (zum Beispiel Sichtbarkeit bestimmter Patienten in bestimmten Abteilungen) zu berücksichtigen.

Authentifizierung von Komponenten

Zugriffe einer CCP-IT-Komponente auf eine andere über das Internet erfolgen nur nach erfolgreicher Authentifizierung, d.h. nicht nur die Berechtigung (Autorisierung), sondern auch die Identität der zugreifenden Komponente wird geprüft.

5.3 Maßnahmen in der IT-Infrastruktur

Sicherheit der gespeicherten Daten

Alle in den zentralen Komponenten der CCP-IT erhobenen Daten werden auf verschlüsselten Festplattenpartitionen gespeichert. Der zugehörige Schlüssel befindet sich jeweils auf einem getrennten Medium pro Server (z.B. Papier, USB-Stick). Dieses Medium wird nur während des Mount- bzw. Bootvorgangs benötigt und wird ansonsten sicher verwahrt. Nur der Administrator des jeweiligen Servers hat Zugriff auf ‚sein‘ Schlüsselmedium. Der Schlüssel kann nicht errechnet werden. Alle Server befinden sich in Rechenzentren, die über eine personengebundene Zugangskontrolle (zum Beispiel per Chipkarte ~~für jeweils berechtigte Personen~~) verfügen.

Sicherheit der Kommunikation

Die Vertraulichkeit der Kommunikation zwischen den Komponenten wird durch folgende Maßnahmen sichergestellt:

- Die Kommunikation zwischen den Komponenten erfolgt grundsätzlich über verschlüsselte Verbindungen (HTTPS). Die dafür eingesetzten Schlüssel und Zertifikate sind so zu erstellen, dass sie den aktuell anerkannten Anforderungen entsprechen (z.B. Schlüssellänge).
- Durch Firewalls ist sichergestellt, dass die Server, auf denen die zentralen Komponenten laufen, nur über diejenigen Protokolle und Ports erreichbar sind, die für die Kommunikation mit Benutzern oder anderen Komponenten erforderlich sind (in der Regel HTTPS-Verbindungen). Der administrative Zugang ist auf das Intranet des Betreibers beschränkt.
- Alle Kommunikationsvorgänge zwischen dem Brückenkopf und zentralen Komponenten werden vom Brückenkopf initiiert. Der Brückenkopf kann dadurch hinter einer Firewall oder einem Proxyserver betrieben werden, ohne über eine öffentliche Adresse aus dem Internet erreichbar zu sein.

Protokollierung

Es erfolgt eine Protokollierung der Zugriffe von Forschern auf die Komponenten sowie zwischen den Komponenten untereinander. Das Protokoll enthält mindestens:

- Die Identität der zugreifenden Person oder Komponente.
- Datum und Uhrzeit des Zugriffs.
- Den Inhalt des Zugriffs (die übermittelten Daten, ggfls. aggregiert) oder Informationen, aus denen dieser rekonstruiert werden kann (z.B. Verweis auf einen Datenbankeintrag o.ä.). Davon ausgenommen ist die Übertragung von IDAT an den Kontrollnummerngenerator.

Das Protokoll wird zusammen mit den Nutzdaten des entsprechenden Servers gespeichert und zwischen einem und sechs Monaten aufbewahrt. Die aufgezeichneten Daten ~~dürfen~~werden nur ~~im Rahmen der technischen Administration (insbesondere zur Fehlersuche)~~für folgende Zwecke verarbeitet und eingesehen ~~werden.~~:

- Im Rahmen der technischen Administration (insbesondere zur Fehlersuche).
- Zur Aufdeckung möglicher Missbrauchsfälle.
- Zur Erstellung anonymisierter Nutzungsstatistiken.

6. Wahrung von Betroffenenrechten

6.1 Aufklärung und Einwilligung

Im Falle von DKTK-Patienten ist die informierte Einwilligung (Volltext siehe Anhang) Rechtsgrundlage der Datenverarbeitung. Mit der Einwilligung erklärt sich der Patient insbesondere dazu bereit, dass

- seine ~~identifizierende~~identifizierenden Daten an das zentrale Identitätsmanagement übermittelt werden,
- medizinische Daten gemäß MDS-K und MDS-B an die zentrale MDS-Datenbank übermittelt werden und
- diese Daten von Forschern des DKTK gemäß der Funktionsweise der Zentralen Suche durchsucht werden können.

Mit Einholen der Einwilligung wird der Patient über sein Recht auf Auskunft und Widerruf informiert.

6.2 Rechtsgrundlage bei nicht-DKTK-Patienten

Die Erhebung und Speicherung der Daten von Patienten, die nicht in die Nutzung ihrer Daten durch das DKTK zugestimmt haben, erfolgt nur in der behandelnden Institution. Dennoch reicht der Behandlungsvertrag nicht als Rechtsgrundlage für diese Datenverarbeitung aus. Ähnlich wie zum Beispiel bei einem Clinical Data Warehouse verfolgt die Speicherung in diesem Fall nämlich nicht mehr den ursprünglichen Zweck der medizinischen Versorgung, sondern dient der medizinischen Forschung. Hier sind die jeweiligen landesrechtlichen Regelungen mit den entsprechenden Ausnahmetatbeständen zu prüfen. Sollten diese im Einzelfall eine Verwendung von Bestandsdaten aus dem Behandlungskontext für die medizinische Forschung zulassen, können diese auch ohne Einwilligung verwendet werden. Ebenso ist die Speicherung von Daten von nicht-DKTK-Patienten im Brückenkopf unbedenklich, sofern sichergestellt ist, dass diese Daten auch nach der Speicherung im Brückenkopf nur von der behandelnden Einheit eingesehen werden können (vgl. Abschnitt 7, „Lokale Umgebung“).

Falls es für einen der beteiligten Standorte keine spezialgesetzlichen Ermächtigungsregelungen für die Verwendbarkeit der Daten aus dem Behandlungskontext zu Forschungszwecken (z.B. Landeskrankenhausregelungen) gibt, so kann eine Erhebung der Daten von nicht-DKTK-Patienten auf Basis der Forschungsklauseln des jeweiligen Landesdatenschutzrechts (für öffentliche Stellen der Länder) oder des Bundesdatenschutzgesetzes (für Stellen in privater Trägerschaft) möglich sein. Die Regelungen sehen eine Verarbeitung personenbezogener Daten auch ohne

Einwilligung vor, wenn „dies zur Durchführung wissenschaftlicher Forschung erforderlich ist, das wissenschaftliche Interesse an der Durchführung des Forschungsvorhabens das Interesse des Betroffenen an dem Ausschluss der Erhebung, Verarbeitung und Nutzung erheblich überwiegt und der Zweck der Forschung auf andere Weise nicht oder nur mit unverhältnismäßigem Aufwand erreicht werden kann~~“.~~“[14]. Im Fall der nicht-DKTK-Patienten ist aus folgenden Gründen davon auszugehen, dass diese Anforderungen erfüllt werden:

- Die Wichtigkeit der translationalen Krebsforschung ist in der Fachwelt anerkannt. Ihre Förderung durch das BMBF und andere öffentliche Stellen belegt das große öffentliche Interesse an den Forschungszielen des DKTK. Im Sinne des Gesetzes ist das „wissenschaftliche Interesse an der Durchführung des Forschungsvorhabens“ hoch anzusetzen.
- Die Daten dieser Patienten verlassen nicht die Institution, die sie ohnehin erhoben und gespeichert hat. Der Personenkreis, der diese Daten einsehen kann, ändert sich durch die Speicherung im Brückenkopf also nicht, lediglich die Zweckbindung an die Behandlung entfällt. Die Pseudonymisierung der Daten erschwert zudem die Reidentifizierung des Patienten aus den Daten im Brückenkopf. Das Interesse des Patienten daran, diese Datenverarbeitung zu verhindern, kann deshalb im Vergleich zum wissenschaftlichen Interesse als gering angesehen werden.
- ~~Ohne den Einbezug~~Wegen der immer stärkeren Zergliederung der Forschung im Bereich der Onkologie, mit zahlreichen molekular definierten Subgruppen, wird in Zukunft weder ein Standort alleine über die für Forschungsprojekte ausreichende Menge an Daten ~~der nicht-DKTK-~~verfügen, noch wird die Anzahl der Patienten ~~würde es unverhältnismäßig lang dauern, bis das durch die CCP-IT erfasste Patientenkollektiv~~mit DKTK-Einwilligung in absehbarer Zeit groß genug für die wissenschaftlichen Ziele des DKTK ~~ist~~sein. Die Bedingung, dass „der Zweck der Forschung auf andere Weise nicht oder nur mit unverhältnismäßigem Aufwand erreicht werden kann“ ist damit ebenso als erfüllt anzusehen.

Sofern möglich, sollte von nicht-DKTK-Patienten, deren Daten im Brückenkopf gespeichert werden, in der Folge eine DKTK-Einwilligung eingeholt werden (in der Regel beim nächsten Behandlungstermin). Falls der Patient diese ablehnt, sind die bereits gespeicherten Daten zu löschen. Daten von Patienten, die der Verwendung ihrer Daten zu Forschungszwecken generell widersprochen haben, dürfen nicht in den Brückenkopf importiert werden.

6.3 Auskunft über gespeicherte Daten

DKTK-Patienten haben das Recht, Auskunft über die in der CCP-IT über sie gespeicherten Daten zu erhalten. Der Antrag auf Auskunft ist schriftlich an die behandelnde Klinik zu stellen. Diese wendet sich zunächst unter Nennung der DKTK#ID an den Betreiber der zentralen Patientenliste. Dieser vergibt für den Auskunftsvorgang eine eindeutige, nicht-sprechende[15] Fallnummer und meldet diese an die anfragende Klinik zurück. Die Anfrage wird nun, unter Nennung der Fallnummer und der MDS-ID, aber ohne Nennung der DKTK#ID, an den Betreiber der zentralen MDS-Datenbank weitergeleitet. Dieser erstellt einen menschenlesbaren Ausdruck der Daten (mit Ausnahme der MDS-ID) und stellt diese unter Nennung der Fallnummer in einem versiegelten Umschlag der zuständigen Klinik zu. Diese kann nun anhand der Fallnummer den Patienten identifizieren und ihm den Ausdruck aushändigen.

Die in der zentralen Patientenliste gespeicherten Kontrollnummern liegen nicht in menschenlesbarer Form vor und können prinzipiell nicht in die ursprünglich eingegebenen Klartextdaten zurücktransformiert werden (vgl. Abschnitt Pseudonymisierung, S. 11). Eine Auskunft über diese Daten findet deshalb nur auf ausdrücklichen Wunsch des Patienten statt. In Antwort auf den allgemeinen Antrag auf Auskunft wird der Patient darüber informiert, dass eine

[14] §13, Abs. 2, Bundesdatenschutzgesetz. Diese Stelle ist beispielhaft erwähnt, im konkreten Fall gelten ggfls. entsprechende Regelungen der Landesdatenschutzgesetze. Da an dieser Stelle nicht alle möglichen lokal gültigen Ausnahmetatbestände erfasst werden können, ist eine Prüfung vor Ort erforderlich.

[15] D.h. für Dritte ist kein Rückschluss auf Daten oder Pseudonyme des Patienten möglich.

Mitteilung der IDAT aufgrund der Chiffrierung impraktikabel ist und auf sein Recht, dennoch einen Ausdruck dieser Daten anzufordern, hingewiesen.

6.4 Widerruf, Löschung, Anonymisierung

Patienten haben das Recht, die Einwilligung in die Verarbeitung ihrer Daten in der CCP-IT zu widerrufen. Der Widerruf ist schriftlich an einen der Standorte (in der Regel die behandelnde Klinik) zu richten. Der betroffene Patient kann mit dem Widerruf zusätzlich die vollständige Löschung seiner Daten beantragen. Fehlt dieser Antrag, so erfolgt eine Anonymisierung.

Nach Prüfung des Widerrufs wird der Antrag auf Löschung oder Anonymisierung zunächst an den Betreiber der Patientenliste weitergeleitet. Der Patient wird dabei über die DKTK#ID des betreffenden Standorts identifiziert. Zwecks Löschung oder Anonymisierung der Daten in der zentralen MDS-Datenbank wird ~~dessen~~deren Betreiber vom Betreiber der zentralen Patientenliste die MDS-ID des Patienten mitgeteilt. Im Falle einer Löschung werden alle dem Patienten zugeordneten Datensätze in Patientenliste und zentraler MDS-Datenbank gelöscht. Im Falle der Anonymisierung werden die Datensätze in der zentralen Patientenliste gelöscht und in den Datensätzen in der MDS-Datenbank die MDS-ID des Patienten durch ein zufälliges Pseudonym ersetzt. Durch den Algorithmus zur Pseudonymerzeugung ist sichergestellt, dass die Pseudonyme eines gelöschten oder anonymisierten Patienten am jeweiligen Standort nicht mehr für neue Patienten verwendet werden.

Die Löschung bzw. Anonymisierung ist von den zuständigen Betreibern zeitnah, maximal innerhalb von 14 Werktagen, vorzunehmen[16]. Der Betreiber der Patientenliste informiert wiederum alle Standorte über der Löschung oder Anonymisierung. In den Standorten werden eventuell lokal gespeicherte Daten des Patienten gelöscht, oder (bei Anonymisierung) die DKTK#IDs durch Standortpseudonyme ersetzt[17]. Dieser Vorgang wird protokolliert und dem Betreiber der Patientenliste bestätigt. Der Abschluss der Löschung oder Anonymisierung wird von den Betreibern von MDS-Datenbank und Patientenliste dem Standort, an dem der Widerruf eingegangen ist, mitgeteilt und von diesem dem Patienten schriftlich bestätigt.

6.5 Regelungen für nicht-DKTK-Patienten

Daten von Patienten ohne Einwilligung für das DKTK liegen unter der Hoheit des behandelnden Standorts. In Bezug auf die Patientenrechte sind also die lokalen Regelungen des Standorts zu berücksichtigen. Auf Anweisung eines Standorts können aber die in der Patientenliste gespeicherten Kontrollnummern zu einem solchen Patienten gelöscht werden. Für den Ablauf der Löschung gelten die Regelungen für DKTK-Patienten entsprechend.

6.6 Dauer der Speicherung

Die erhobenen Daten bleiben in den zentralen Komponenten der CCP-IT gespeichert, so lange es für sie eine sinnvolle wissenschaftliche Verwendung im Rahmen der Patienteneinwilligung gibt. Falls die Daten nicht mehr in der vorgesehenen Form genutzt werden können (z.B. falls die CCP-IT außer Betrieb genommen wird), prüft der Ausschuss für Datenschutz, ob eine Rechtsgrundlage für eine anderweitige Verwendung der Daten, gegebenenfalls in anonymisierter Form, besteht. Falls diese Prüfung negativ ausfällt, sind die zentralen Daten zu löschen. Hinsichtlich der nur lokal in den Brückenköpfen gespeicherten Daten ist die Entscheidung über den weiteren Umgang mit den Daten von jedem Standort individuell zu treffen, da diese Daten möglicherweise weiter für Behandlungszwecke oder eigene Forschungsvorhaben genutzt werden dürfen.

[16] Die meist impraktikable Löschung oder Anonymisierung in Datensicherungen ist verzichtbar, sofern die Sicherungen nur durch den zuständigen Systemadministrator eingesehen werden können und alte Sicherungen regelmäßig gelöscht werden.
[17] Der Patient wird also im Sinne der CCP-IT von einem DKTK-Patienten zu einem nicht-DKTK-Patienten.

7. Lokale Umgebung

Bei Installation und Betrieb der Brückenköpfe sind die lokalen Bestimmungen des Standorts hinsichtlich Datenschutz und -sicherheit zu beachten. Generell werden die für den Betrieb der zentralen Komponenten genannten Maßnahmen (siehe Abschnitt 5.3, „Maßnahmen in der IT-Infrastruktur“) empfohlen. Da die Komponenten des Brückenkopfes über das Intranet der jeweiligen Einrichtungen kommunizieren, kann hier auf die Verwendung verschlüsselter Verbindungen verzichtet werden.

Sofern der Brückenkopf nicht von der behandelnden Einheit selbst betrieben wird, ist durch Zugriffsrechte sicherzustellen, dass deren Datenhoheit gewahrt bleibt. Die Regelungen hinsichtlich administrativer Zugriffe (Abschnitt 4.3) gelten entsprechend.

7.1 Lokale Bestimmungen für den Standort [Standort]

[Sollte es abweichende Bestimmungen für die lokal an Ihrem Standort laufenden Komponenten oder Prozesse geben, können Sie diese hier einsetzen. Andernfalls löschen Sie den Abschnitt.]

Anhang

1. Patienteneinwilligung DKTK

~~[Einwilligung wird hier eingefügt, sobald bereitgestellt durch Trial Board]~~

Die angehängte Mustererklärung wurde vom CCP-Büro erstellt und dient als Vorlage, die von jedem Standort an lokale Erfordernisse angepasst wird. Diese tatsächlich eingesetzte Einwilligungserklärung erhalten Sie beim Standortvertreter in der AG CCP-IT.

Autor der hier eingebundenen Mustererklärung ist das CCP-Büro; zuständiger Ansprechpartner ist:

Dr. Jennifer Braun
Wissenschaftliche Projektkoordination

Deutsches Konsortium für Translationale Krebsforschung (DKTK)
CCP-Büro

Universitätsklinikum Frankfurt
Med. Klinik II
Haus 33, Zi. 208
Theodor-Stern-Kai 7
60590 Frankfurt am Main

Tel.: +49 (0)69 / 6301 84237
Fax: +49 (0)69 / 6301 7463
E-Mail: j.braun@dkfz.de

2. Votum der AG Datenschutz der TMF

Die Arbeitsgruppe Datenschutz der Technologie- und Methodenplattform für die vernetzte medizinische Forschung (TMF e.V.) hat zu diesem Datenschutzkonzept in der Fassung vom 8. Januar 2014 das angehängte Votum ausgesprochen.