# Schlußbericht zum Projekt „Datentreuhänderschaft bei Biobanken“

hier: Teilprojekt 3, „Auditierungsverfahren und -kriterien”

Schlußbericht 05.fm/02. März 2009

*Prof. Dr.-Ing. Norbert Luttenberger*
*AG Kommunikationssysteme*
*Institut für Informatik*
*Christian-Albrechts-Universität zu Kiel*
*24098 Kiel*

| **Zuwendungsempfänger** | **Förderkennzeichen** |
|---|---|
| AG Kommunikationssysteme<br>Institut für Informatik<br>Christian-Albrechts-Universität zu Kiel<br>24098 Kiel | 01 GP 0613 |
| **Vorhabenbezeichnung**<br>Forschungskooperation ELSA – Datentreuhänderschaft Biobanken<br>Teilprojekt 3: Auditierungsverfahren und -kriterien | |
| **Laufzeit des Vorhabens**<br>15. Dez. 2006 – 31. Aug. 2008 | |
| **Berichtszeitraum**<br>15. Dez. 2006 – 31. Aug. 2008 | |

# 1 Kurzdarstellung

## 1.1 Aufgabenstellung

Das Projekt „Datentreuhänderschaft bei Biobanken – Auditierungsmethoden und -kriterien“ (*Biobank Data Custodianship\Audit Methodology and Criteria*, abgekürzt bdc\Audit) ist ein Forschungsprojekt im Rahmen der vom BMBF seit 1997 geförderten Forschung zu ethischen, rechtlichen und sozialen Aspekten (ELSA) der molekularen Medizin. In diesem Kontext hat sich das Verbundprojekt bdc\Audit mit der datenschutzrechtlichen Auditierung von Biobanken auseinandergesetzt. Im speziellen hat das Teilprojekt 3 (TP-3) des bdc\Audit-Projekts, über das hier berichtet wird, die Aufgabe übernommen, die für die datenschutzrechtliche Auditierung erforderlichen „Auditierungsverfahren und -kriterien” in praxistauglicher Form zu bestimmen und dem Unabhängigen Landeszentrum für Datenschutz Schleswig-Holstein (ULD) zur Verfügung zu stellen. Das ULD ist Projektpartner im bdc\Audit-Verbundprojekt

und hat dort das Teilprojekt 2 „Rechtliche Ausgestaltung der Datentreuhänderschaft für Biobanken" bearbeitet. Im folgenden klären wir kurz die Begriffe *Biobank*, *Datentreuhänder* und *Auditierung*, um die Aufgabenstellung klarer herausarbeiten zu können.

Biobanken sind Sammlungen von Proben menschlicher Körpersubstanzen (Zellen, Gewebe, Blut und DNA [letztere als materieller Träger genetischer Information]), die mit personenbezogenen klinischen, medizinisch-biologischen, genetischen, genealogischen, soziodemographischen, Umwelt-bezogenen und/oder Lebensstil-bezogenen Daten über ihre Spender verknüpft sind bzw. verknüpft werden können. Sie haben also einen Doppelcharakter als Proben- und Datensammlungen. Unter Anwendung von statistischen Verfahren der medizinischen Epidemiologie werden Biobanken heute in einer Vielzahl von Forschungsverbünden genutzt, um Ursachen von Krankheiten aufklären und bessere Therapien entwickeln zu können. Insbesondere die Pharmakogenetik und die Pharmakogenomik sind in hohem Maße auf solche mit Gesundheitsinformationen verknüpfte Sammlungen biologischer Materialien angewiesen. Diese können nicht nur in der Forschung, sondern auch im Zusammenhang mit Arzneimittelzulassungsverfahren zum Einsatz kommen.

Patienten und Probanden („Spender"), die Körpermaterialien und Daten zur Verfügung stellen, haben ein unstrittiges Recht auf den Schutz ihrer personenbezogenen Daten vor Mißbrauch. In dem Moment, in dem Geber Proben und Daten an einen Forschungsverbund abgeben, übernimmt der Forschungsverbund (bzw. ein dazu verpflichteter Akteur in diesem Verbund) de facto die Rolle eines Datentreuhänders, der durch bestimmte Maßnahmen sicherstellt, daß ein solcher Mißbrauch nach bestem Wissen und Gewissen ausgeschlossen ist. Von vielen Biobanken und Forschungsverbünden würden deshalb klare und allgemeinverbindliche Regeln für die Etablierung von Datenschutzmaßnahmen als hilfreich empfunden werden. Sie könnten zum einen den Erklärungsbedarf bei der Rekrutierung von Gebern verringern. Zum anderen würden sie den Forschungsverbund von der Aufgabe einer eigenen Gestaltung der Datentreuhänderschaft entlasten.

Ein wichtiges Problem besteht darin, daß Biobanken und ihre Nutzung außerordentlich heterogen sind. Biobanken unterscheiden sich hinsichtlich ihrer Zielsetzung, der Art des gespeicherten Materials, der Trägerschaft, der Größe, der Art und des Umfangs der gespeicherten Daten sowie der Lagerungsdauer der Proben. Darüber hinaus ist es schon heute Praxis und zeichnet sich für die Zukunft zunehmend ab, daß Biobanken nicht nur von denjenigen Wissenschaftlern bzw. Unternehmen genutzt werden, die die Materialien ursprünglich gewonnen und eingelagert haben, sondern daß die Proben auch von verschiedenen Nutzern zu unterschiedlichen Zwecken untersucht und ggf. auch verkauft werden. Auszugehen ist dabei auch davon, daß Forschungsverbünde, die an der Nutzung solcher Sammlungen interessiert sind, sich nicht nur innerhalb des vergleichsweise einheitlich geregelten Rechtsraumes Europa konstituieren, sondern auch international und europäische Grenzen überschreitend kooperieren.

Vor diesem Hintergrund erscheint es wenig sinnvoll, ein einheitliches Datentreuhänderverfahren zu definieren und sogar zu implementieren, das für alle existierenden und geplanten Verbünde anwendbar ist. Deshalb wurde im Rahmen des hier vorgeschlagenen bdc\Audit-Verbundprojekts ein anderer Weg verfolgt: Aufbauend auf einer systematischen Analyse derzeit existierender und sich im Aufbau befindlicher Biobanken sollten durchgängige, nachvollziehbare und praxisorientierte Verfahren und Kriterien für die **datenschutzrechtliche Auditierung** von Datentreuhänderschaftsverfahren entwickelt werden. Damit sind die Betreiber von Biobanken frei in der Gestaltung der jeweils zweckdienlichen Maßnahmen zur Schutz der persönlichen Daten ihrer Geber, sie erhalten jedoch eine klare Anleitung, wie sie ihre jeweiligen Maßnahmen datenschutzrechtlich auditieren lassen können.

Unter **Auditierung** wird im allgemeinen die Überprüfung eines Konzepts oder eines Systems durch eine unabhängige, einschlägige Stelle verstanden. Die datenschutzrechtliche Auditierung im speziellen wird durch den § 43 Abs. 2 des Datenschutzgesetzes des Landes Schleswig-Holstein ermöglicht: „Öffentliche Stellen können ihr Datenschutzkonzept durch das Unabhängige Landeszentrum für Datenschutz prüfen und beurteilen lassen." Eine solche, dem Gesetz entsprechende Auditierung wird vom ULD angeboten und ist auch schon erfolgreich für das GENOMatch-System der Bayer Schering Pharma AG unter Federführung durch den Leiter des TP-3 erfolgreich in drei Schritten durchgeführt worden.

Im Rahmen des bdc\Audit-Verbundprojekts sollten nun die bereits vorhandenen Auditierungsverfahren und -kriterien mit Hilfe wissenschaftlicher Methoden speziell für die Auditierung von Datentreuhänderschaftsverfahren in Biobank-Forschungsverbünden weiterentwickelt und – wo möglich – formalisiert werden. Zusätzlich sollen rechtliche Empfehlungen zur Arbeitsweise von Auditierungsstellen im Umfeld der Biobank-Forschung erarbeitet werden (TP-2). Diese Verfahren und Kriterien können – quasi als Nebeneffekt – aber auch als *Best Practice*-Regeln für die Entwicklung von angepaßten *Standard Operating Procedures* für ein Datentreuhänderschaftsverfahren benutzt werden und kommen damit auch dem Biobank-Betreiber entgegen. Bei der Inanspruchnahme eines entsprechenden Auditierungsverfahrens können die Biobank-Betreiber damit der Öffentlichkeit demonstrieren, daß der Schutz personenbezogener Geberdaten unter Einhaltung sachgerechter Kriterien nach methodischer Überprüfung durch eine unabhängige Stelle vollzogen wird.

## 1.2 Projektvoraussetzungen

Im bdc\Audit-Verbundprojekt haben drei Partner kooperiert:

1. Forschungsschwerpunkt „Biotechnik, Gesellschaft und Umwelt" (FSP BIOGUM),
   Forschungsgruppe Medizin/Neurobiologie, Universität Hamburg
   Prof. Dr. rer. nat. Regine Kollek
   TP-1: Analyse und Klassifikation von Biobanken: Strukturen – Elemente – Prozesse
2. Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein (ULD), Kiel
   Dr. jur. Thilo Weichert
   TP-2: Rechtliche Ausgestaltung der Datentreuhänderschaft für Biobanken
3. Arbeitsgruppe (AG) Kommunikationssysteme des Instituts für Informatik,
   Christian-Albrechts-Universität zu Kiel (CAU Kiel)
   Prof. Dr.-Ing. Norbert Luttenberger
   TP-3: Biobank-Datentreuhänderschaft – Auditierungsverfahren und -kriterien

Jedem der drei Projektpartner stand für die Projektlaufzeit von zwei Jahren ein wissenschaftlicher Mitarbeiter zur Verfügung. Im TP-3 konnte die volle Laufzeit von zwei Jahren aus technischen Gründen nicht voll ausgenutzt werden: Der wissenschaftliche Mitarbeiter, der für TP-3 eingestellt wurde, war vom 15.12.2006 bis zum 31.08.2008 – also nur 20,5 Monate – für TP-3 tätig. Bei TP-3 kamen zu den Personalmitteln noch Sachmittel für Reisen und für eine studentische Hilfskraft hinzu. Die studentische Hilfskraft hatte die Aufgabe, für den Aufbau einer Web-Präsenz zu sorgen.

Mit Projektbeginn wurde ein wissenschaftlicher Beirat berufen, der die Arbeit des gesamten bdc\Audit-Verbundprojekts begleiten sollte. Diesem Beirat gehörten u.a. die folgenden Personen an:

- Herr Graf (Universitätsklinikum des Saarlands, Homburg/Saar),
- Herr Pickardt (TÜV Nord, Berlin),
- Herr Reischl (Bayer Schering Pharma AG, Berlin),
- Herr Thorun (Verbraucherzentrale Bundesverband, Berlin),
- Frau Wellbrock (HessDSB, Wiesbaden),
- Herr Simitis (Frankfurt/Main)

Dem Beirat und dem Förderer wurden Zwischenergebnisse des Projekts auf einem Workshop am 12.11.2007 vorgestellt. Beiratsmitglieder nahmen auch am Abschlußworkshop des Projekts am 04.07.2008 teil.

Bereits vor dem offiziellen Start des bdc\Audit-Projekts wurde im Rahmen der Jahrestagung der Gesellschaft für Informatik 2006 in Dresden ein Workshop zum Thema „Elektronische Datentreuhänderschaft – Anwendungen, Methoden, Grundlagen" durchgeführt. Auf diesem Workshop wurde in insgesamt sieben Beiträgen über unterschiedliche Datentreuhänderschaftsverfahren bei Biobanken berichtet.

## 1.3 Planung und Ablauf des Vorhabens

Plangemäß wurden von TP-3 die folgenden Arbeitsschritte durchgeführt:

1. Mitwirkung bei der Durchführung von Interviews mit Biobank-Betreibern:
   Empirische Analyse der Vorgehensweisen einer repräsentativen Menge von Biobanken
   (Die Verantwortung für die Interviewdurchführung lag bei TP-1.)
2. Methodik für die Modellierung der Arbeitsweise von Biobanken:
   Auswahl des Modellierungsansatzes, Bestimmung geeigneter Modellkategorien, Auswahl von Modellierungswerkzeugen
3. Biobank-Modellierung:
   Entwicklung eines prozeßorientierten Modells für eine idealtypische Biobank, das unterschiedliche Datentreuhänderschaftsverfahren erkennbar macht
4. Auseinandersetzung mit den „Gemeinsamen Kriterien für die Prüfung und Bewertung der Sicherheit von Informationstechnik" (*Common Criteria*, CC, vgl. http://www.bsi.bund.de/cc/):
   Überprüfung der Kombinierbarkeit der datenschutzrechtliche Auditierung mit der Vorgehensweise, die in den CC für die Prüfung und Bewertung der Sicherheit von Informationssystemen vorgesehen ist
5. Anonymisierungs- und Pseudonymisierungsverfahren:
   Entwicklung von Kriterien zur Bewertung unterschiedlicher Anonymisierungs- und Pseudonymisierungsverfahren

Zusätzlich wurde von TP-3 eine Webpräsenz für das bdc\Audit-Projekt eingerichtet.

Da die Punkte 1 – 3 mehr Zeit benötigten als im Antrag vorgesehen war, und da nicht die volle Laufzeit von zwei Jahren zur Verfügung stand, konnte der Pkt. 5 der o.a. Liste nicht abschließend behandelt werden. Außerdem konnten die im Antrag zusätzlich aufgeführten Punkte:

- Analyse von Zugriffskontrollmechanismen für Datentreuhänderschaftsverfahren und
- Entwicklung von IT-Werkzeugen zur Unterstützung der datenschutzrechtlichen Auditierung

nicht bearbeitet werden.

## 1.4 Wissenschaftlicher und technischer Stand

Der wissenschaftliche und technische Stand, an den angeknüpft werden konnte, läßt sich wie folgt charakterisieren:

- Der Nationale Ethikrat hat sich in zwei Publikationen mit dem Thema Biobanken auseinandergesetzt.[1] In seiner *Stellungnahme* legt der Nationale Ethikrat neben einigen wichtigen ethischen Prinzipien auch einige allgemeine Grundsätze für den Schutz der genetischen Daten von Spendern dar. Es wird u.a. ausgeführt, daß Proben und Daten, die „rechtmäßigerweise gewonnen" und sachgerecht anonymisiert bzw. pseudonymisiert worden sind, von der Forschung uneingeschränkt benutzt werden können, und zum anderen wird darauf hingewiesen, daß die Einhaltung der Datenschutzbestimmungen dem Datenschutzbeauftragten obliegt. Allerdings diskutiert der Nationale Ethikrat das Thema „datenschutzrechtliche Auditierung" nicht. Eine solche Diskussion wird auch in der *Tagungsdokumentation* nicht geführt, in der u.a. eine Beitrag von Rita Wellbrock vom Büro des hessischen Datenschutzbeauftragten enthalten ist, in dem sie einige datenschutzrechtliche Prinzipien, die beim Betrieb von Biobanken zu beachten sind, darstellt.
- Das US-amerikanische *National Cancer Institute* hat 2007 ein Papier *Best Practices for Biospecimen Resources* vorgelegt, in dem Richtlinien für gutes Arbeiten mit Bioproben festgelegt werden. Die Schweizerische Akademie der Medizinischen Wissenschaften (SAMW) legte bereits 2006 ein Papier „Biobanken: Gewinnung, Aufbewahrung und Nutzung von menschlichem biologischem Material" vor und beschreibt dort gute Praxis für die Arbeitsweise von Biobanken in der Schweiz. Keines der Papiere befaßt sich jedoch mit der datenschutzrechtlichen Auditierung von Biobanken oder gar der prozeßorientierten Modellierung von Biobanken oder den anderen von TP-3 behandelten Themen.
- In dem von der Europäischen Kommission geförderten Projekt *European Privacy Seal* (EuroPriSe), an dem sich 9 Stellen (u.a. das ULD) aus 8 EU-Mitgliedsstaaten beteiligen, werden Fragen der Vorgehensweise bei der datenschutzrechtlichen Auditierung intensiv diskutiert. Der von diesem Projekt vorgelegte *EuroPriSe Criteria Catalogue* (abrufbar unter https://www.european-privacy-seal.eu/criteria) stellt eine „Checkliste" dar, die dem „Auditierungswilligen" wichtige Hinweise gibt, mit welchen Fragen er sich auseinandersetzen muß, wenn er in den Prozeß der Auditierung eintritt. Naturgemäß geht dieser Katalog allerdings nicht speziell auf Biobanken ein. Außerdem kann es durchaus als problematisch empfunden werden, daß diesem Katalog für den Auditierungswilligen nicht ein Handbuch zugeordnet ist, aus dem der Auditierer seine Handlungsanleitungen bezieht.

1. Nationaler Ethikrat: *Biobanken für die Forschung – Stellungnahme*. Nationaler Ethikrat, 2004. Und: Nationaler Ethikrat: *Tagungsdokumentation Biobanken – Jahrestagung des Nationalen Ethikrates 2002*. Nationaler Ethikrat, 2003.

- Der *Telematikplattform für Medizinische Forschungsnetze (TMF) e.V.* verfolgt als „Hauptziel … die Verbesserung der Organisation und Infrastruktur für die vernetzte medizinische – klinische, epidemiologische und translationale – Forschung." Dabei steht nach Bekunden des TMF e.V. auch der „Aufbau von IT-Infrastruktur" im Fokus der Arbeit (http://www.tmf-ev.de). Man hat den Ansatz gewählt, den Mitgliedern eine mehr oder weniger einheitliche Struktur für den Aufbau ihrer IT und damit de facto auch für die Gestaltung technischer Datenschutzmaßnahmen vorzugeben. Diese wird einerseits modellhaft beschrieben[2], andererseits werden bestimmte technische Komponenten angeboten. Wie persönliche Gespräche ergaben, wird dieser Ansatz von einigen Mitgliedern als wenig hilfreich empfunden, da er der zu beobachtenden Heterogenität der existierenden Biobanken nicht gerecht wird.
- In drei Schritten wurde in den Jahren von 2003 – 2008 das GENOMatch-System der Bayer Schering Pharma AG vom ULD datenschutzrechtlich auditiert. Die drei Auditierungs(teil)anträge wurden vom Projektleiter des TP-3 entwickelt, beim ULD eingebracht und gegenüber dem ULD vertreten. Das GENOMatch-System bildet die technisch-logistische Infrastruktur für die pharmakogenetische Forschung der Bayer Schering Pharma AG und umfaßt eine Menge von datenschutzrelevanten Komponenten. Neben einer Vielzahl von technischen Komponenten müssen auch eine Menge von *Standard Operating Procedures* zum GENOMatch-System hinzugezählt werden, da Datenschutz erst im Zusammenwirken von technischen und organisatorischen Maßnahmen wirkungsvoll entfaltet werden kann. Hinweise zum GENOMatch-System finden sich in diversen Veröffentlichungen[3] und in den sog. Kurzgutachten des ULD[4]. Das GENOMatch-System wird in der Fachwelt nach wie vor als der „Rolls Royce" unter den Biobank-Managementsystemen angesehen – sicherlich eine Folge der Tatsache, daß dieses System entwickelt wurde, als bereits diverse Erfahrungen mit solchen Systemen vorlagen.

Vor diesem Hintergrund wird im bdc\Audit-Antrag der Sachstand deshalb wie folgt zusammengefaßt: „Bislang allerdings konzentrierte sich die fachliche und gesellschaftliche Diskussion um Biobanken und deren Nutzung vorrangig auf die mit der Probensammlung verbundenen ethischen und teilweise auch rechtlichen Aspekte. Technische, organisatorische und strukturelle Fragen in Zusammenhang mit Etablierung, Betrieb und Nutzung von Biobanken werden in der öffentlichen und biomedizinisch- bzw. bioethisch-fachlichen Diskussion bislang nur am Rande thematisiert, obwohl hier ein deutlicher Klärungs- und Handlungsbedarf besteht." Die mit der Auditierung verbundene Offenlegung der internen Strukturen und Abläufe des GENOMatch der Bayer Schering Pharma AG darf nach wie vor als „Sonder-

2. Reng, M., Debold, P., Specker, C., Pommerening, K.: *Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin*. Berlin 2006 (TMF-Schriftenreihe Bd. 1)
3. Luttenberger, N., Reischl, J., Schröder, M., Stürzebecher, C.S.: *Datenschutz in der pharmakogenetischen Forschung – eine Fallstudie*. DuD Datenschutz und Datensicherheit 28 (2004) 6, 356–363.

   Luttenberger, N., Stürzebecher, C.S., Reischl, J., Schröder, M.: *Der elektronische Datentreuhänder.* DIGMA Zeitschrift für Datenrecht und Informationssicherheit 5, 1 (3/2005), 24–29.

   Reischl, J., Schröder, M., Luttenberger, N., Petrov, D., Schümann, B., Ternes, R., Stürzebecher, C-S.: *Pharmacogenetic Research and Data Protection—Challenges and Solutions*. The Pharmacogenomics Journal (2006), 1–9.

   Luttenberger, N., Kollek, R., Reischl, J., Stürzebecher, C.-S.: *Design of Individual Donor Feedback Processes in Biobank Research*. In: Hochberger, Chr., Liskowsky, R. (Ed.): Informatik 2006, Informatik für den Menschen, Band 1, LNI P-93 (Springer), 722–728.
4. Auffindbar unter https://www.datenschutzzentrum.de/audit/register.htm

fall" gelten. Allerdings kann der bereits oben erwähnte, durch die Partner des bdc\Audit-Verbundprojekts initiierte Workshop „Elektronische Datentreuhänderschaft – Anwendungen, Methoden, Grundlagen" auf der Jahrestagung der Gesellschaft für Informatik 2006 in Dresden getrost als Meilenstein bezeichnet werden, weil hier erstmals im Detail „technische, organisatorische und strukturelle Fragen in Zusammenhang mit Etablierung, Betrieb und Nutzung von Biobanken" in der wissenschaftlichen Öffentlichkeit diskutiert wurden.

Alle diese Aktivitäten haben aber noch nicht die *Vorgehensweise* bei der datenschutzrechtlichen Auditierung von Biobanken angesprochen. Hier sollte das bdc\Audit-Verbundprojekt ansetzen: Durch das Herausarbeiten ihrer methodischen Grundlagen sollte der Auditierung zu einer breiteren Akzeptanz verholfen und ihr ein größeres Gewicht verliehen werden. Die Tatsache, daß vom BMI am 4.9.2008 der „Entwurf eines Gesetzes zur Regelung des Datenschutzaudits …"[5] vorgelegt und von der Bundesregierung am 10.12.2008 verabschiedet wurde, gibt dem Anliegen des Projekts eine zusätzliche Rechtfertigung.

### 1.5 Zusammenarbeit mit anderen Stellen

Während der Projektlaufzeit erfolgte eine Zusammenarbeit mit den folgenden Stellen:

- Durchführung von Interviews mit verschiedenen Biobanken, s.u.
- Informationsaustausch mit dem TMF e.V.
- Kooperation mit der Fa. Bayer Schering Pharma AG
- Informationsaustausch mit Prof. Dr. Ulrich Sax, Georg-August-Universität Göttingen, Universitätsmedizin Göttingen, Abteilung Medizinische Informatik

## 2 Eingehende Darstellung

### 2.1 Erzielte Ergebnisse

Die im TP-3 erzielten Ergebnisse werden im folgenden gegliedert nach den in Pkt. 1.3 genannten Arbeitsschritten dargestellt.

#### 2.1.1 Interviews mit Biobank-Betreibern

Um zu einer möglichst einheitlichen Sicht auf den tatsächlichen Betrieb von existierenden Biobanken zu kommen und damit eine praktikable Grundlage für ihre datenschutzrechtliche Auditierung zu gewinnen, nahm ein Vertreter des TP-3 des bdc\Audit-Verbundprojekts an den Interviews u.a. mit den folgenden Biobanken teil (Die Gesamtverantwortung für die Auswahl der Interviewpartner und für die Durchführung der Interviews lag bei TP-1):

- popgen: Populationsgenetisches Forschungsprojekt des Nationalen Genomforschungsnetzes (NGFN) am Universitätsklinikum Schleswig- Holstein. „Bei popgen werden gezielt Bioproben von Patienten gesammelt, die an einer bestimmten Krankheit leiden. Zur Identifizierung möglicher

5. Eine ausführliche Stellungnahme des ULD findet sich unter: https://www.datenschutzzentrum.de/bdsauditg/20081029-stellungnahme-dsag-e.html

Spender arbeitet popgen mit verschiedenen Institutionen des Gesundheitswesens zusammen. Mediziner in Krankenhäusern und niedergelassene Ärzte nehmen Kontakt zu denjenigen Patienten auf, die die zuvor festgelegten Einschlußkriterien erfüllen." (http://www.popgen.de/). Eine Beschreibung der popgen-Datenschutzmaßnahmen findet sich u.a. im Tagungsband des Workshops „Elektronische Datentreuhänderschaft – Anwendungen, Methoden, Grundlagen" auf der Jahrestagung der Gesellschaft für Informatik 2006 in Dresden, auf den bereits weiter oben hingewiesen wurde.

- Biobank der Blutspender: „Nach einer mehrjährigen Vorbereitungszeit und einem kontinuierlichem Aufbau einer Proben- und Datensammlung seit 2001 [startete] im Juni 2006 der [Blutspendedienst des Bayerischen Roten Kreuzes] mit der „Biobank der Blutspender". … Bei jeder Blutspende werden Blutproben gewonnen, die dazu dienen, das gespendete Blut auf Blutgruppe und Infektionskrankheiten zu testen. Die Proben werden bisher nach einer gesetzlich vorgeschriebenen Aufbewahrungsfrist verworfen. Mit der „Biobank der Blutspender" werden nun ausgewählte Blutproben mit dem Einverständnis der teilnehmenden Blutspender für die Wissenschaft und Forschung nutzbar gemacht. Bereits heute stehen über 100.000 Proben zur Verfügung. Damit zählt die „Biobank der Blutspender" zu den weltweit größten Biobank-Projekten." (http://www.bio-bank.de/)
- KORA-gen: „Infrastruktur zur Bereitstellung von Phänotypen, Genotypen und Bioproben für gemeinschaftliche genetisch- epidemiologische Forschungsprojekte, aufgebaut vom Helmholtz Zentrum München. Diese Infrastruktur wird von mehreren Instituten des Helmholtz Zentrum Münchens, deren kooperierenden Partnern, sowie von anderen Forschungseinrichtungen genutzt." (KORA steht für Kooperative Gesundheitsforschung in der Region Augsburg) (http://epi.helmholtz-muenchen.de/kora-gen/)
- KNHI: „Das Kompetenznetzwerk Herzinsuffizienz (KNHI) ist ein seit Juni 2003 vom Bundesministerium für Bildung und Forschung (BMBF) gefördertes interdisziplinäres Kooperationsvorhaben zwischen Wissenschaft und klinischer sowie ambulanter Versorgung. [Ein Schwerpunkt ist der] Aufbau einer zentralen Biomaterialbank zur Verarbeitung und Bereitstellung von Bioproben für das gesamte Kompetenznetz." (http://www.knhi.de/)

Die Interviews lieferten wichtige Einsichten in die innere Betriebsweise der besuchten Biobanken und halfen zu einer realistischen Sicht auf wesentliche Biobank-Prozesse (zur Definition dieses Begriffs s. 2.1.2). Ohne die Erkenntnisse aus den Interviews wäre die Modellierung der Biobank-Prozesse mit vielen Unsicherheiten behaftet geblieben. Allerdings muß kritisch angemerkt werden, daß der Aufwand für Vorbereitung, Durchführung und Nachbereitung der Interviews sehr hoch war.

Speziell für TP-3 läßt sich folgendes Resümee ziehen: Einige Interviewpartner zeigten sich von der Modellierungsmethodik, wie sie von TP-3 vorgeschlagen wurde, sehr angetan, da sie sich nicht nur als Grundlage für die datenschutzrechtliche Auditierung nutzen läßt, sondern auch als Möglichkeit begriffen wurde, die internen Biobank-Prozesse besser und gezielter zu gestalten. Andere Interviewpartner blieben skeptisch, da sie den Fokus ihrer Arbeit in anderen Problemfeldern sehen. Der Wunsch von TP-3, mit wenigstens einem Interviewpartner ein komplettes Biobank-Modell zu erstellen und damit ein im Sinne von TP-3 greifbares Interviewergebnis zu erzeugen, blieb aus Aufwandsgründen unerfüllt.

Eine ausführliche Darstellung der Interviewergebnisse findet sich im Abschlußbericht von TP-1, da TP-1 die Gesamtverantwortung für die Auswahl der Interviewpartner sowie Durchführung und Nachbereitung der Interviews trägt.

### 2.1.2 Modellierungsansatz

Eine maßgebliche Voraussetzung für die datenschutzrechtliche Auditierung einer Biobank ist ihre umfassende und vollständige Beschreibung. Im Konsens mit den Partnern im bdc\Audit-Verbundprojekt, abgesichert durch die Interviews mit den Biobank-Betreibern und in Analogie zu aktuellen betriebswirtschaftlichen Ansätzen wurde ein *prozeßorientierter* Beschreibungs- und Modellierungsansatz gewählt. Darunter wird ein Ansatz verstanden, der primär von den Prozessen ausgeht, die in einer Organisation zur Erzielung ihres Zwecks eingesetzt werden. Der Ansatz geht also *nicht* von Organisationseinheiten, Hierarchien usw. aus. Unter einem Prozeß wird eine geschlossene Handlungskette verstanden, in der mehrere Akteure, die ggf. unterschiedlichen organisatorischen Einheiten angehören, zusammenwirken, um ein bestimmtes Teilziel zu erreichen. In der Betriebswirtschaft geht man üblicherweise von sog. Geschäftsprozessen aus. In einer Biobank sind wichtige Prozesse z.B. die Einlagerung von Proben, die Vernichtung von Proben beim Widerruf des Spenders, die Verknüpfung von Proben mit klinischen Daten, die Erteilung von Auskünften zu Daten, die in einer Biobank vorhanden sind, usw. In Kap. 2.1.3 gehen wir detaillierter auf Biobank-Prozesse ein.

Bei der Beschreibung von Biobank-Prozessen stellt sich das Problem, daß ein so komplexes System wie eine Biobank nur sehr schwierig in verständlicher, vollständiger und widerspruchsfreier Art und Weise in textueller Form beschrieben werden kann. Es ist deshalb naheliegend, ein besseres Verständnis durch eine graphische Modellierung zu erlangen. So ist es beispielsweise auch in der Softwareentwicklung üblich, ein zu entwickelndes Softwaresystem im Vorfeld graphisch zu beschreiben. Auch in der Prozeßsteuerung wird die graphische Modellierung zur Visualisierung von Abläufen herangezogen. Wird sie durch textuelle Inhalte ergänzt, so ist sie ein mächtiges Werkzeug zur Beschreibung von komplexen Systemen.

Die graphische Modellierung bedient sich sog. Modellierungssprachen, welche aus abstrakten Symbolen und ihren Bezügen („Relationen") bestehen. Eine Modellierungssprache gehorcht dabei einer Grammatik, welche die Regeln für die Bildung richtiger Diagramme beschreibt. Durch die Vorgabe einer Grammatik für die Verwendung der graphischen Symbole wird der Modellierer in einem gewissen Umfang „gezwungen", sein Modell zumindest formal richtig zu gestalten.

Für die Modellierung von Biobanken für die datenschutzrechtliche Auditierung soll die Modellierungssprache

- einfach verständlich sein,
- möglichst wenig Vorkenntnisse erfordern,
- eine weitgehend vollständige Darstellung aller Zusammenhänge ermöglichen,
- mittels einer Grammatik beschrieben sein und
- durch Werkzeuge unterstützt sein.

Eine sehr weit verbreitete und akzeptierte graphische Modellierung bietet die *Unified Modeling Language* (UML). „Die *Unified Modeling Language* (UML) ist eine von der *Object Management Group* (OMG) entwickelte und standardisierte Sprache für die Modellierung von Software und anderen Syste-

men. Im Sinne einer Sprache definiert die UML dabei Bezeichner für die meisten Begriffe, die für die Modellierung wichtig sind, und legt mögliche Beziehungen zwischen diesen Begriffen fest. Die UML definiert weiter graphische Notationen für diese Begriffe und für Modelle von statischen Strukturen und von dynamischen Abläufen, die man mit diesen Begriffen formulieren kann. Die UML ist heute eine der dominierenden Sprachen für die Modellierung von betrieblichen Anwendungssystemen (Softwaresystemen)." (Wikipedia)

Die UML versteht sich nun nicht nur als eine einzelne abgeschlossene Modellierungssprache mit zugehöriger graphische Notation, sondern vielmehr als ein „Modellierungsbaukasten", der in der aktuellen Version 2.1 insgesamt dreizehn Diagrammtypen vorsieht. Diese lassen sich in Strukturdiagramme und Verhaltensdiagramme unterteilen.

Die Strukturdiagramme befassen sich mit der Gliederung von Software- und anderen Systemen in Komponenten. Sie beschreiben keine funktionalen Zusammenhänge und sind daher für die Prozeßmodellierung ungeeignet. Deshalb werden sie hier nicht weiter diskutiert.

Für die Prozeßmodellierung sieht die UML sieben Diagrammtypen vor:

- das Anwendungsfalldiagramm (auch *Use Case*- oder Anwendungsfalldiagramm genannt),
- das Aktivitätsdiagramm,
- das Sequenzdiagramm,
- das Kommunikationsdiagramm,
- das Interaktionsübersichtsdiagramm,
- das Zeitverlaufsdiagramm und
- das Zustandsdiagramm.

Für die Benutzung in der datenschutzrechtlichen Auditierung erscheinen aus Sicht von TP-3 nur zwei Diagrammtypen notwendig – die Anwendungsfalldiagramme und die Aktivitätsdiagramme.

**Anwendungsfalldiagramme** beschreiben zwar kein Verhalten und keine Abläufe, beschreiben dafür aber den Zusammenhang zwischen Prozessen („Anwendungsfällen") und den daran beteiligten Akteuren. Anwendungsfalldiagramme lassen sich also benutzen, um die Menge aller vorhandenen Prozesse aufzulisten und zu zeigen, welche Akteure („Rollen") an welchen Prozessen beteiligt sind. In der Praxis werden Anwendungsfalldiagramme oft bei der Spezifikation der Anforderungen an ein System eingesetzt. Dieser Diagrammtyp ist sehr allgemein gehalten und daher generell anwendbar.

Mit **Aktivitätsdiagrammen** lassen sich unter Verwendung weniger einfacher Diagrammelemente Prozesse verständlich darstellen. Mit Aktivitätsdiagrammen werden Prozesse in Folgen von einzelnen Aktionen aufgelöst. Eine Folge von Aktionen wird in UML auch als Kontrollfluß bezeichnet.

Zur Beschreibung von sicherheitsrelevanten Eigenschaften (im Sinne der Computer- und Netzwerksicherheit) wurde UML um einige Konstrukte erweitert, die unter dem Stichwort UMLsec zusammengefaßt werden. Für eine Vertiefung sei auf die einschlägige Literatur verwiesen.

### 2.1.3 Modellierung einer idealtypischen Biobank

In diesem Arbeitsschritt ging es darum, in Abstraktion von den Spezifika einzelner Biobanken ein Biobank-Modell zu erstellen, das alle wesentlichen Biobank-Prozesse umfaßt und dieses Modell mit Hilfe von UML-Diagrammen zu modellieren. Ein solches Modell kann einem Biobank-Betreiber als Vorbild

bei der Erstellung eines Auditierungsantrags dienen, und der versierte Auditierer wird in einem solchen Modell schnell und sicher die „Stellen" identifizieren können, an denen Datenschutz-relevante Aktionen vorgenommen werden.

Für die datenschutzrechtliche Auditierung beschreibt der Biobank-Betreiber die in der Biobank angewendeten Prozesse gegenüber einem unabhängigen Auditor. Damit ein Auditor die Einhaltung von Vorschriften zum Schutz personenbezogener Daten in einer Biobank anhand der vorgelegten Selbstbeschreibung bewerten kann, muß die Selbstbeschreibung nicht nur alle notwendigen Sachverhalte darstellen, sondern sie muß auch für den Auditor nachvollziehbar gestaltet sein. Durch TP-3 wurde in enger Zusammenarbeit mit TP-2 ein Prozeßmodell erarbeitet, das in 13 abstrakten Prozessen und 28 zugehörigen Rollen die Basis für eine Beschreibung der allermeisten Biobanken liefern kann.

Folgende Prozesse wurden identifiziert:

*Einholung Informed Consent:* Sofern eine Biobank Proben und/oder Daten selbst erhebt, holt diese vom Spender einen sog. *Informed Consent* ein. Zu diesem Prozeß gehören fünf Prozessschritte. Im ersten Schritt werden zunächst die Formblätter zur Einwilligung und Aufklärung erstellt. In einem zweiten Schritt werden die Formblätter auf rechtliche und ethisch-moralische Gesichtpunkte hin geprüft. Im dritten Schritt werden die Formblätter dem Spender im Rahmen einer Aufklärung vorgelegt. Im vierten Schritt füllt der Spender die Einwilligungserklärung aus und stimmt dieser durch seine Unterschrift bzw. die seines gesetzlichen Vertreters zu. Im fünften Schritt wird die Einwilligung in der Datenverarbeitung der Biobank vermerkt.

*Erhebung und Einlagerung von Proben und Daten:* Dieser Prozeß beschreibt in sechs Prozeßschritten die zur Erhebung und Einlagerung von Proben und Daten notwendigen Vorgänge. Den ersten Schritt bildet die Kontaktaufnahme mit dem Spender. Im zweiten Schritt werden von der Biobank für jeden Spender ein standardisiertes Entnahmeset und der Fragebogen zu den Studiendaten bereitgestellt und mit einem Erhebungspseudonym versehen. Im dritten Schritt findet der Erhebungsvorgang statt. Liegt die Einwilligung des Spenders vor, so können gemäß *Informed Consent* Proben und Daten erhoben werden. In diesem Schritt werden bereits persönliche und Studiendaten voneinander getrennt. Die persönlichen Daten wie auch der *Informed Consent* verbleiben bei der Erhebungsstelle. Die Studiendaten und die Biomaterialproben werden an die Biobank überstellt (vierter Schritt). Zum Zeitpunkt der Überstellung tragen Daten und Proben in den allermeisten Fällen zwei Identifikatoren: den Identifikator der Erhebungsstelle, mit dem ein Rückbezug zu den Daten möglich ist, die in der Erhebungsstelle verbleiben, und das o.a. Erhebungspseudonym. Im fünften Schritt werden die persönlichen Daten bei der Erhebungsstelle eingelagert. Dabei werden die Proben und die Studiendaten de-identifiziert, d.h. es wird der Identifikator der Erhebungsstelle entfernt, und das o.a. Erhebungspseudonym wird durch ein nur Biobank-intern bekanntes Vorhaltungspseudonym ersetzt. Bei Pseudonymisierung wird eine Verknüpfung zwischen dem Identifikator der Erhebungsstelle und dem Biobank-eigenem Vorhaltungspseudonym hergestellt und unter Anwendung besonderer Sicherheitsmaßnahmen abgespeichert. Im Fall der Anonymisierung wird diese Verknüpfung nicht erzeugt. Im sechsten Schritt werden die de-identifizierten Proben und Daten in die Biobank eingespeist.

*Follow-up:* Möglicherweise nimmt eine Biobank nach einer Ersterhebung daran anschließende Folgeerhebungen vor. Solche Folgeerhebungen werden als *Follow-up* bezeichnet. *Follow-up*-Prozeß gestaltet sich weitgehend analog zum Prozeß *Erhebung und Einlagerung von Proben und Daten*.

*Spenderwiderruf/Wegfall des Verwendungszwecks:* Der von einem Spender erteilte *Informed Consent* umfaßt das Recht des Spenders, seine Einwilligung zu widerrufen. Im Fall des Widerrufs der Einwilligungserklärung durch den Spender teilt dieser im ersten Schritt seinen Widerruf der Erhebungsstelle mit. Diese meldet der Biobank den Widerruf unter Angabe des entsprechenden Identifikators der Erhebungsstelle. In der Biobank werden die zugehörigen Biobank-internen Proben- und Daten-Identifikatoren des Spenders ermittelt. In einem zweiten Schritt werden die Proben und Daten vernichtet bzw. gelöscht. Dieser Vorgang wird protokolliert und dem Spender über die Erhebungsstelle kommuniziert. Im dritten Schritt werden Dritte, welche Proben oder Daten vom Spender erhalten haben, durch einen Weitergabebeauftragten benachrichtigt.

*Datenaufbereitung:* Damit erhobene Studiendaten für die Forschung nutzbar werden, ist es in der Regel notwendig, diese für den Forschungszweck aufzubereiten. Dieser Vorgang kann bereits durch die Biobank erfolgen. Im ersten Schritt des zugehörigen Biobank-Prozesses werden auf den Rohdaten Fehlerkorrekturen, statistische Analysen o.ä. Operationen durchgeführt. Im zweiten Schritt werden die Analysedaten mit den Rohdaten verknüpft und entweder als neuer Datensatz oder über die Erweiterung eines bestehenden Datensatzes in die Datenverarbeitung der Biobank eingespeist.

*Probenaufbereitung:* Für die Forschung ist es notwendig, aus den Proben verwertbare Daten zu generieren. Dies kann sowohl in einem einstufigen als auch mehrstufigen Analyseverfahren stattfinden, in dem z.B. durch DNA-Extraktion und DNA-Sequenzierung Gendaten gewonnen werden. Wir verzichten hier auf eine detaillierte Prozeßdarstellung, da in diesem Prozeß Proben in aller Regel nur mit dem Biobank-internen Vorhaltungspseudonym versehen sind und zwischen verschiedenen Labors ausgetauscht werden.

*Eigenforschung:* Eine Biobank betreibt ggf. Eigenforschung betreiben. Dazu müssen Probendaten und Studiendaten spenderweise zusammengeführt werden. Dabei werden die zusammengeführten Datensätze auf die für die Forschung notwendigen Inhalte reduziert.

*Weitergabe an Dritte:* Eine wichtige Aufgabe einer Biobank ist die Bereitstellung von Proben und Daten zur Forschung durch Dritte. Dazu muß die Biobank in einem ersten Schritt feststellen, ob für das jeweilige Forschungsvorhaben eine Weitergabe nach bestehendem *Informed Consent* zulässig ist. Im zweiten Schritt wird eine vertragliche Vereinbarung erstellt. Im dritten Schritt werden die Probendaten und die Studiendaten zusammenführt, de-identifiziert (Bildung eines sog. Herausgabefallpseudonyms) und bereitgestellt. Hierzu werden die in der vertraglichen Vereinbarung vorgesehenen Probendaten und Studiendaten auf die für das Forschungsvorhaben notwendigen Feldtypen reduziert. Ist in der vertraglichen Vereinbarung ein Rückfluß von Analyse- oder Forschungsergebnissen vorgesehen, so werden diese im vierten Schritt an die Biobank übergeben. Falls nicht nur aggregierte Daten, sondern spenderbezogene Daten an die Biobank zurückfließen, werden bei der Übergabe die Herausgabefallpseudonyme in die ursprünglichen internen Pseudonyme überführt.

*Feedback:* Feedback bezeichnet der Fall, daß ein Forschungsergebnis zur gesundheitlichen oder genetischen Konstitution eines Spenders an diesen mitgeteilt wird. (Im Fall einer Anonymisierung ist ein Feedback jedoch ausgeschlossen.) Die für ein Feedback nötigen Schritte umfassen die Information des Spenders (wenn im *Informed Consent* so vorgesehen), die Pseudonymauflösung, die Verifikation (Überprüfung, ob auch eine zweite, separat eingeholte Spenderprobe die fraglichen Merkmale aufweist) und die Mitteilung an den Spender.

*Proben-/Dateneinkauf:* Im Zuge des Proben- und Datenhandels kann es sein, daß eine Biobank Proben und Daten von anderen Biobanken erwirbt. Es gelten die Ausführungen zum Prozeß *Einlagerung von Proben und Daten* in Analogie.

*Auskunftsanfrage zur Proben- und Datenverarbeitung:* Gemäß BDSG §19 hat ein Spender Anspruch auf Auskunftserteilung zu den von ihm gespeicherten und verarbeiteten Proben und Daten. Hierzu wendet sich der Spender an die Erhebungsstelle und bittet diese um Auskunft. Die Erhebungsstelle ermittelt daraufhin die zum Spender zugehörigen Erhebungspseudonyme für Proben und Daten und leitet diese mit einem Verweis auf die gestellte Anfrage um Auskunft an die Biobank weiter. Im zweiten Schritt ermittelt die Biobank die Biobank-internen Identifikatoren für Proben und Daten und wertet die Verarbeitungsprotokolle der zugehörigen Proben und Daten aus. Insbesondere werden Weitergaben an Dritte und Einbeziehungen von Dienstleistern nachvollzogen. Im dritten Schritt werden die gewünschten Auskünfte an die Erhebungsstelle übermittelt, welche diese an den Spender in Form eines Auskunftsschreibens weiterleitet.

*Prozessmanagement – Zugriffskontrolle, Rollenvergabe:* Für die organisatorische Aufstellung einer Biobank ist es notwendig, die betrieblichen Prozeß- und Rollenstrukturen zu ermitteln und in einen Stellen- und Personalplan zu überführen. Hierbei gilt es, das eingesetzte Personal mit den notwendigen Zugriffsrechten und rechtlichen Befugnissen auszustatten. Wir verzichten hier auf eine detaillierte Darstellung.

*Monitoring, Qualitätssicherung, Zugriff durch Aufsichtsbehörde/Strafverfolgung:* Diese Managementprozesse haben alle gemein, daß Zugriff auf Proben und Daten und/oder auf Protokolldaten zur Verwendung von Proben oder Daten genommen wird. Ein solcher Zugriff bedarf jedoch vor dem Hintergrund einer datenschutzrechtlichen Auditierung besonderer Aufmerksamkeit. Im Fall der Revision zwecks Qualitätssicherung oder Monitoring beginnt der Prozeß mit dem Zugriff auf Protokolldaten. In diesem Schritt werden bei einer Biobank-internen Revision auf die für die Revision nötigen Protokolldaten zugegriffen und diese entsprechend dem Revisionsziel ausgewertet. Bei einer externen Revision werden diese Daten der externen Stelle zur Verfügung gestellt. Im zweiten Schritt wird bei Revision zwecks Qualitätssicherung oder beim Zugriff durch eine Aufsichts- oder Strafverfolgungsbehörde zusätzlich auf Proben und/oder Daten zugegriffen. Bei einer Biobank-internen Revision zwecks Qualitätssicherung werden die Proben und/oder Daten bezüglich ihrer Qualität geprüft. Bei einer externen Revision oder beim Zugriff durch eine Aufsichts- oder Strafverfolgungsbehörde werden die nötigen Proben und/oder Daten an die externe Stelle übermittelt. Der dritte Schritt behandelt die bei einem Zugriff notwendige (Re-)Identifikation der jeweiligen Spender. Hierzu werden der Biobank die Proben und/oder Daten mitgeteilt, zu denen eine Identifikation notwendig ist. Die Biobank ermittelt unter Mithilfe der Erhebungsstelle die Kontaktdaten der betroffenen Spender und teilt diese der Aufsichtsbehörde bzw. der Strafverfolgung mit. Im vierten Schritt werden im Fall des Zugriffs durch eine Aufsichts- oder Strafverfolgungsbehörde die betroffenen Spender über den Zugriff und die damit verbundene (Re-)Identifikation benachrichtigt. Diese Benachrichtigung erfolgt im Auftrag der Biobank durch die Erhebungsstelle in einem Anschreiben.

Eine ausführliche Darstellung dieser Prozesse inkl. einer tabellarischen Auflistung und der Definition der zugehörigen Rollen ist in einem internen Papier erfolgt und den Projektpartnern übergeben worden. (AG Kommunikationssysteme, Inst. für Informatik, CAU Kiel: Prozesse und Rollen in Biobanken – Basis einer prozeßorientierten Modellierung für die datenschutzrechtliche Auditierung., 21.4.2008) In

einem zweiten internen Papier, das ebenfalls an die Projektpartner übergeben wurde, werden Anleitungen zur Modellierung dieser Biobank-Prozesse mit UML und UMLsec gemacht. (AG Kommunikationssysteme, Inst. für Informatik, CAU Kiel: Prozessorientierte Biobankmodellierung in UMLsec – die Basis für die datenschutzrechtliche Auditierung. 20.6.2008.)

### 2.1.4 Nutzung der *Common Criteria* für die datenschutzrechtliche Auditierung

Die zentrale Rolle von IT-Systemen in Biobanken erlaubt es, bei ihrer datenschutzrechtlichen Auditierung auch solche Analysemethoden heranzuziehen, die aus dem Kontext der IT-Sicherheit kommen. Ein international anerkannter Standard zur Evaluation von IT-Sicherheit sind die *Common Criteria*. Die *Common Criteria* listen gemeinsame Kriterien zur Prüfung und Bewertung der Sicherheit von informationstechnischen Systemen auf. Durch das *Common Criteria Recognition Agreement* sind die CC in 24 Staaten anerkannt. Sie stellen die Weiterentwicklung und Zusammenführung der europäischen *Information Technology Security Evaluation Criteria* (ITSEC), der US-amerikanischen *Trusted Computer System Evaluation Criteria* (TCSEC) (auch: *Orange Book*) und der kanadischen *Canadian Trusted Computer Evaluation Criteria* (CTCPEC) dar.Eine Anwendung der *Common Criteria* innerhalb des Biobank-Kontextes oder der medizinischen Forschung ist allerdings noch nicht bekannt.

Die *Common Criteria* stellen eine Sprache zur Beschreibung von Sicherheitsanforderungen und Sicherheitseigenschaften einerseits und eine Methodik zur Bewertung der Einhaltung dieser Anforderungen andererseits bereit. Damit richten sich die *Common Criteria* gleichermaßen an Hersteller von IT-Systemen und -Produkten, an die Nutzern dieser Systeme und an die Zertifizierer von IT-Sicherheitstechnologien.

Die Methodik der *Common Criteria* sieht u.a. vor, daß Sicherheitsanforderungen in sog. Schutzprofilen (engl.: *Protection Profiles*) beschrieben werden. Diese beziehen sich auf eine bestimmte Gruppe von IT-Systemen bzw. -Produkten (beispielsweise Datenbanksysteme). Dabei werden systematisch die Sicherheitsumgebung, abzuwendende Bedrohungen und die hierzu notwendigen Sicherheitsanforderungen beschrieben. Der Nachweis, daß bei Erfüllung der Sicherheitsanforderungen die genannten Bedrohungen tatsächlich abgewendet werden, ist gleichfalls ein Bestandteil eines Schutzprofils.

Im Rahmen der Arbeiten von TP-3 zum Nutzen der *Common Criteria* für die datenschutzrechtliche Auditierung von Biobanken wurde ein Schutzprofil für Biobanken erarbeitet. Im Folgenden werden einige Elemente dieses Schutzprofils vorgestellt.

- Das Schutzprofil beschreibt zunächst die erforderliche **Sicherheitsumgebung**: Es werden die Annahmen zum sicheren Betrieb, zu den abzuwehrenden Bedrohungen und zu den organisatorischen Sicherheitsregeln der Biobank formuliert.
    - Durch die Darlegung von **Sicherheitsannahmen** wird die Grundlage für umfassende technisch- organisatorische Schutzmaßnahmen gelegt. Hierzu gehört u.a., daß die im Schutzprofil genannten Sicherheitsregelungen durch das Biobank-Management getragen und durchgesetzt werden.
    - Bei der Beschreibung der abzuwendenden **Bedrohungen** wird u.a. davon ausgegangen, daß Unberechtigte Zutritt, Zugang oder Zugriff erlangen wollen. Auch das Einwirken von äußerer Gewalt finden Berücksichtigung. Allgemein bedrohen alle Fehler innerhalb der Vorgänge einer Biobank und ihrer Sicherheitsvorkehrungen den Schutz und die Integrität einer Biobank. Auch das Versagen der Protokollierung wie auch das nicht Erkennen von sicherheitskriti-

schen Ereignissen stellen ernsthafte Bedrohungen dar. Neben diesen Bedrohungen während des Betriebs einer Biobank befaßt sich das Schutzprofil auch mit Bedrohungen vor der Inbetriebnahme. So stellt eine unsichere Auslieferung, Installation und Inbetriebnahme der IT-Systeme einer Biobank oder der Biobank selbst ein Sicherheitsrisiko dar, dem begegnet werden muß. Insbesondere Fehler in der Definition und Umsetzung des Rollenkonzepts können die Betriebssicherheit beeinträchtigen und sogar zu einer unerlaubten (Re-)Identifikation von genetischen Daten führen.

- Ziel der **Sicherheitsregeln** ist es, den organisatorischen Rahmen für den Betrieb der Biobank zu regeln. Insbesondere kann die IT-Sicherheit und der Datenschutz als Bestandteil der Unternehmenspolitik und damit als Managementaufgabe verstanden werden. Maßnahmen zur IT-Sicherheit und zum Datenschutz sind zu implementieren und auch bei Erweiterungen und Neuanschaffungen sowie im Betrieb und in der Forschung zu berücksichtigen. Im Einzelnen decken die Sicherheitsregelungen des Schutzprofils das gesetzeskonforme Handeln, die Nutzung und den Schutz von Daten und Ressourcen einer Biobank, den in den Ablauf integrierten Schutz von Spendern, Partnern und Mitarbeitern sowie die Nachvollziehbarkeit von Vorgängen und deren Gewährleistung ab. Insbesondere finden dabei der Umgang mit Standards, Regelwerken und deren Einhaltung sowie die Qualitätskontrolle als auch der Beratung des Managements besondere Berücksichtigung.

- Unter den **Sicherheitszielen** findet sich die zentrale Forderung nach einem Schutz gegen unbefugten Zutritt, Zugriff und Zugang – auf welche Weise auch immer dieses versucht wird. Hierzu müssen alle Systemkomponenten vor Eingriffen und Manipulationen durch materiellen Schutz selbst oder durch Dritte geschützt sein, oder zumindest in der Lage sein, derartige Eingriffe zu melden. Darüber hinaus gilt es die Vertraulichkeit und Integrität aller internen und externen Datenübermittlungen sicherzustellen. Ebenso muß es eine manipulationsresistente Protokollierung aller sicherheitsrelevanten Ereignisse geben, um den Hergang und die Folgen eines möglichen Störfalls lückenlos aufklären zu können. Während des Betriebs sind Zugangszeit und -ort stets in Bezugnahme auf den Zugreifenden zu entscheiden und für jeden Zugriff zu protokollieren. Es ist ebenfalls wichtig, daß die korrekte Funktion aller Sicherheitsfunktionen im Betrieb sichergestellt ist. Auch im Fehlerfall muß die Biobank und ihre Systeme einen sicheren Betriebszustand beibehalten und somit weitere Fehler und mögliche Verletzungen der Sicherheitsziele verhindern. Hierzu ist es notwendig, eine Archivfunktion (also Backup) zu allen Programmen, Daten und Dokumentationen zum Zweck der Wiederherstellung nach Fehlern und Systemabstürzen zu gewährleisten. Zudem muß es jederzeit möglich sein, den Sicherheitsstatus aller Systeme einer Biobank sowie der Biobank selbst abzurufen. Diese Funktionen (Backup und Abrufen der Sicherheitsstati) sowie alle anderen Funktionen zur Systemverwaltung dürfen dabei nur durch die besondere Rolle des Systemadministrators aufgerufen werden können. Um die korrekte Funktion der Sicherheitsmaßnahmen zu gewährleisten, ist es zudem notwendig, daß die Sicherheitsregeln durch das Management umgesetzt werden und die Mitarbeiter eine ausreichende Schulung zum sicheren Umgang mit den technischen Systemen einer Biobank besitzen. Insbesondere müssen die Mitarbeiter angewiesen sein, daß sie ihre Authentisierungsgeheimnisse (beispielsweise Passwörter) schützen müssen. Somit wird das Risiko einer Schutzverletzung durch Fehlbedienung oder Kompromittierung von Authentisierungsgeheimnissen erheblich vermindert.

- Schließlich umfaßt das entwickelte Biobank-Schutzprofil entsprechend der CC-Methodik eine Menge von konkreten **Sicherheitsanforderungen**:
  - Die gesetzlich vorgeschriebene Anonymisierung bzw. Pseudonymisierung von Proben und Daten kann nach *Common Criteria* durch die Klasse Privatheit (FPR) abgebildet werden. Diese Klasse enthält „Familien" (CC-Terminologie) zur Anonymität (FPR ANO) und zur Pseudonymität (FPR PSE), die es erlauben, den Sachverhalt der unerlaubten Re-Identifizierung von Proben und Daten abzubilden und diese zu untersagen.
  - Die *Common Criteria* regeln Eigenschaften zur Zugriffskontrolle in den Klassen Schutz der Benutzerdaten (FDP), Identifikation und Authentisierung (FIA) und Zugriff (FTA).
  - Ein besonderes Risiko für unerlaubte Kenntnisnahme stellt die Kommunikation im Netz dar. Dieses Risiko ist abhängig vom genutzen Kommunikationsweg und besteht grundsätzlich für jede Art von Kommunikation. Diesem Problem begegnen die *Common Criteria* mit der Klasse Vertrauenswürdiger Pfad/Kanal (FTP). Diese Klasse bildet die Forderung ab, daß bei elektronischer Kommunikation ein Mindestmaß an Sicherheit bestehen muss, so daß diese vertrauswürdig ist.
  - Um Abläufe innerhalb eines IT-Systems nachvollziehbar zu gestalten, ist es notwendig, eine Ablaufprotokollierung einzusetzen. Dabei gilt es zu beachten, daß auch die Protokollierung als auch die Einsicht und Auswertung der protokollierten Gegenstände den Ansprüchen einer datenvermeidenden und datensparsamen Verarbeitung genügen müssen. Die *Common Criteria* fassen alle protokollierungsbezogenen Eigenschaften innerhalb der Klasse Sicherheitsprotokollierung (FAU) zusammen. Mittels der Familie Generierung der Sicherheitsprotokolldaten (FAU GEN) lassen sich alle zu protokollierenden Ereignisse festlegen. Zudem wird ein automatischer Sicherheitsalarm bei Sicherheitsverstößen vorgesehen (Automatische Reaktion der Sicherheitsprotokollierung (FAU ARP) ). Eine datenvermeidende und datensparsame Protokollierung wird durch die Familien Ereignisauswahl für die Sicherheitsprotokollierung (FAU SEL) und Ereignisspeicherung der Sicherheitsprotokollierung (FAU STG) durchgesetzt. Durch die Verknüpfung mit den Aufgaben der zugriffsberechtigten Rollen gilt selbiges für die Einsicht und Auswertung von Protokolldaten, welche durch die Familien Analyse der Sicherheitsprotokollierung (FAU SAA) und Durchsicht der Sicherheitsprotokollierung (FAU SAR) geregelt sind.
  - Erst durch ein umfassendes Sicherheitsmanagement können Sicherheitsmaßnahmen ihren vollen Schutz entfallten. Das Sicherheitsmanagement regelt das Verhalten der Sicherheitsfunktionen und legt Rollen und ihre Rechte für den rollenbasierten Zugriff fest. Unmittelbar an das Sicherheitsmanagement schließt sich zudem die kryptographische Schlüsselverwaltung und der Einsatz von kryptographischen Funktionen (vgl. Kryptographische Unterstützung (FCS) ) und die Integrität und Verfügbarkeit der Sicherheitsfunktionen (vgl. Schutz der Sicherheitsfunktionen (FPT) ) an. Die *Common Criteria* erlauben es, diesen komplexen Bereich methodisch abzubilden und mit der Sicherheitsumgebung einer Biobank abzustimmen.

Zusammengefaßt:

Der Fokus der *Common Criteria* liegt klar bei der IT-Sicherheit. Mit den CC lassen sich nur schwer organisatorische Sicherheitsvorkehrungen überprüfen. Auch die Erfüllung des gesetzten datenschutzrechtlichen Rahmens läßt sich nach den *Common Criteria* nicht abschließend prüfen. Beispielsweise

haben bestimmte Anforderungen zur Erfüllung der Datentreuhänderschaft in Biobanken wie beispielsweise Spenderaufklärung und Spendereinwilligung keinen IT-Bezug und können somit gar nicht von einem Schutzprofil berücksichtigt werden. Wie (nicht nur) TP- 2 hierzu ausgeführt hat, ist eine solche Erfüllung jedoch für eine datenschutzkonforme Verarbeitung zwingend erforderlich.

Zudem bieten die *Common Criteria* keine Methodik zur graphischen Modellierung an. Die Selbstbeschreibung der IT-Sicherheit des EVG erfolgt ausschließlich textuell. Dies mag für die Vorgehensweise nach *Common Criteria* ausreichend sein. Für eine Selbstbeschreibung durch die Biobank zur datenschutzrechtlichen Auditierung ist es jedoch zweckmäßiger, eine graphische Modellierung zu verwenden.

In der datenschutzrechtlichen Auditierung einer Biobank kann die Prüfung der IT-Sicherheit durch eine Evaluation gegen ein Schutzprofil abgedeckt werden. Eine solche Evaluation kann auch vor einer datenschutzrechtlichen Auditierung stattfinden. Dies hätte zum Vorteil, daß zum Zeitpunkt der Auditierung bereits eine geprüfte IT-Sicherheit vorliegt und somit das Auditierungsverfahren zügiger und problembezogener durchgeführt werden kann. Voraussetzung dabei ist, daß das Schutzprofil nicht solche Sicherheitsannahmen voraussetzt, die eine datenschutzrechtliche Auditierung erforderlich machen.

Die Methodik der *Common Criteria* erlaubt es, die Wirksamkeit und Vertrauenswürdigkeit von IT-Sicherheit zu prüfen und zertifizieren zu lassen. Durch den engen Bezug von IT-Sicherheit und Datenschutz in IT-Systemen eignet sich diese Methodik auch zur Gewährleistung von Datenschutz in diesen Systemen. Bedingt durch die zentrale Bedeutung von IT-Systemen im Betrieb von Biobanken empfiehlt sich somit der Einsatz der *Common Criteria* begleitend zur datenschutzrechtlichen Auditierung von Biobanken.

Das im TP-3 entwickelte Schutzprofil für Biobanken kann – nach einer Aktualisierung – interessierten Biobanken zur Verfügung gestellt werden. Diese können es bei Bedarf dem Bundesamt für Sicherheit in der Informationstechnik (BSI) zur Überprüfung vorlegen. Ein vom BSI „abgesegnetes" Biobank-Schutzprofil wiederum würde es den Biobanken erlauben, ihre Anforderungen gegenüber den Herstellern von Biobank-IT-Systemen klar und präzise zu formulieren.

Weitere Details finden sich in einem internen Bericht, der den Projektpartner zur Verfügung gestellt wurde: AG Kommunikationssysteme, Inst. für Informatik, CAU Kiel: Datenschutz durch IT-Sicherheit – Bedeutung der *Common Criteria* für die datenschutzrechtliche Auditierung von Biobanken. 10.08.2008.

### 2.1.5 Bewertung von Anonymisierungs- und Pseudonymisierungsverfahren

Eine wichtige Komponente für den datenschutzgerechten Betrieb einer Biobank bilden die Verfahren für die Anonymisierung von Proben und Daten und für ihre Pseudonymisierung. Für die Bewertung von solchen Anonymisierungs-/Pseudonymisierungsverfahren ist es zum einen wichtig, Kriterien für die Bewertung dieser Verfahren zu finden, und zum anderen ist es notwendig, konkrete Verfahren für diese Bewertung transparent darlegen zu können. In einem schrittweisen Vorgehen wurden zunächst bestehende Verfahren untersucht, um dann in einem nächsten Schritt Kriterien zu entwickeln, nach denen eine Klassifizierung dieser Verfahren möglich ist.

In der Literatur gibt es bereits Ansätze zu Metriken für die Anonymisierung. So fasst Díaz[6] diese Ansätze zu zwei Gruppen von Metriken unter „effective anonymity set size" und „the degree of anonymity" zusammen. Díaz legt formal dar, wie diese zwei Arten von Metriken herangezogen werden kön-

nen, um ein Maß der Anonymität zu bestimmen. Welches Maß an Anonymität erforderlich ist, um „Privacy" zu gewährleisten, kann offensichtlich nur in Bezug auf ein konkretes Einsatzszenario bestimmt werden. Diese Frage hat TP-3 für Biobanken aufgegriffen und versucht, die Qualität von Anonymisierungverfahren auf Grundlage dieser Metriken zu bewerten. Die dabei relevanten Fragen sind: Inwieweit können die von Díaz beschriebenen Metriken in ein Auditierungverfahren einfließen und welches Maß ist für die Anonymisierung anzusetzen? Ebenfalls von Interesse ist die Übertragbarkeit auf Verfahren zur Pseudonymisierung. TP-3 hat darüber hinaus untersucht, wie sich insbesondere Verfahren für die Pseudonymisierung nicht nur aufgrund der genannten Metriken, sondern auch aufgrund ihrer „konstruktiven Eigenschaften" bewerten lassen. Z.B. verspricht ein Verfahren mit mehrstufiger Pseudonymisierung mehr Sicherheit, als eines mit einstufiger Pseudonymisierung. Ebenso spielen Faktoren wie Anzahl der beteiligten Personen, Bildungsvorschrift der Pseudonyme und möglichen Re-Identifizierungsrisiken eine Rolle.

Ein von TP-3 entwickelter erster, aber noch nicht vollständiger Ansatz sieht ein zusammengesetztes Bewertungsmaß vor, in das auch das Rollenmodell der Biobank und der zugehörige Datenfluß sowie der Anteil technischer Elemente im Anonymisierungs-/Pseudonymisierungsverfahren eingehen. Speziell für die Pseudonymisierung sollen auch die folgenden Dinge berücksichtigt werden:

- die Art des Pseudonyms (Personenpseudonym, Rollenpseudonym usw.),
- die Speicherung von *links* zwischen Namen und Pseudonymen, und
- die Bildungsvorschrift für ein Pseudonym (ableitbares Pseudonym, Zufallszahl, kryptographisch gebildete Zufallszahl)

Die entsprechenden Arbeiten bilden die Grundlage für weitere Arbeiten in der AG Kommunikationssysteme.

## 2.2 Voraussichtlicher Nutzen und Verwertbarkeit der Ergebnisse

Die von TP-3 erzielten und oben dargestellten Ergebnisse können vom ULD genutzt werden, um den Prozeß der datenschutzrechtlichen Auditierung weiterzuentwickeln. Sie können außerdem genutzt werden, um ein Verfahren für den im o.a. Gesetzentwurf der Bundesregierung vorgesehenen Datenschutzaudit zu definieren.

## 2.3 Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Die im Projekt erzielten Ergebnisse fließen ein in die Re-Auditierung des von Fa. Bayer Schering Pharma AG betriebenen GENOMatch-Systems, die noch im Jahre 2009 erfolgen soll, und die vom Leiter des TP-3 im Auftrage der Fa. Bayer Schering Pharma AG betrieben werden wird.

## 2.4 Veröffentlichungen

Die Ergebnisse wurden auf zwei Workshops dargestellt, externe Veröffentlichungen sind nicht erfolgt.

6. Díaz, Claudia: *Anonymity Metrics Revisited*. October 2005; http://drops.dagstuhl.de/opus/volltexte/2006/483/pdf/05411.DiazClaudia.ExtAbstract.483.pdf