Verraten unsere Gene zu viel?

Ein Datenschutzkonzept für Genforschungsprojekte

von Marian Arning*, Nikolaus Forgó** und Tina Krügel***

Abstract

Die Forschung am menschlichen Gen *verspricht große medizinische Fortschritte, beinhaltet aber auch große Gefahren für die Privatsphäre der betroffenen Patienten, weil genetische Daten missbraucht werden können und für Dritte wie zum Beispiel Arbeitgeber, Versicherungen oder Strafverfolgungsbehörden einen erheblichen Wert darstellen. Um dieser Gefahr zu begegnen, ist es wichtig, bereits bei der Planung medizinischer Forschungsprojekte, aber auch bei der alltäglichen ärztlichen Arbeit, die datenschutzrechtlichen Regelungen zu beachten. In diesem Beitrag werden deshalb zunächst die datenschutzrechtlichen Grundlagen und Probleme der Genforschung dargestellt. Anschließend wird eine technische Infrastruktur für Genforschungsprojekte beschrieben, die Forschern ein optimales Arbeiten bei gleichzeitiger Einhaltung datenschutzrechtlicher Vorschriften ermöglicht. Die Kernbestandteile dieser neuartigen Infrastruktur sind die doppelte Pseudonymisierung der Gendaten, die Implementierung eines Datentreuhänders sowie die Errichtung eines projektinternen Datenschutzgremiums. Durch diese Gestaltung wird die Privatsphäre der Patienten geschützt und gleichzeitig der medizinische Fortschritt unterstützt.*

Genetic research *offers enormous chances for medical progress, but it also threatens the privacy of the concerned patients since genetic data can be abused and represent a substantial value for third parties such as employers, insurance companies or prosecution authorities. Therefore it is important to consider data protection not only in the planning stage of a medical research project but also in everyday work. The article describes the basic principles of data protection legislation and the resulting problems with regard to genetic research. Subsequently, it explains a technical infrastructure for genetic research projects which allows researchers optimal working conditions and at the same time safeguards automatic compliance with data protection regulations. The core components of that new infrastructure are the double pseudonymisation of the genetic data, a trusted third party and the establishment of a data protection board within the project. As a result, the patient's privacy is protected effectively without obstructing medical progress.*

Schlüsselworte: *Datenschutz, Genforschung, Pseudonymisierung und Anonymisierung von Gendaten, Datentreuhänder, Datenschutzgremium, internationale Genforschungsprojekte*

Keywords: *data protection, genetic research, pseudonymisation and anonymisation of genetic data, trusted third party, data custodian, transnational genetic research projects*

1 Einleitung

Humangenetische Forschung ist aus der Medizin nicht mehr wegzudenken. Sie hat ein Feld unschätzbarer Möglichkeiten eröffnet, das neue Heilungsmethoden beinhaltet und die Ursache vieler Krankheiten besser verstehen lehrt. Gleichzeitig liegt die ethische und datenschutzrechtliche Brisanz dieser Entwicklung auf der Hand: Die genetische Information eines Menschen ist einzigartig und ihrem Träger eindeutig zuzuordnen. Sie gibt Auskunft über seine ethnische Herkunft, seine Abstammung, über

*** Dipl.-Jur. Marian Arning, LL.M.**
Institut für Rechtsinformatik der Leibniz Universität Hannover, Königsworther Platz 1, 30167 Hannover · Telefon: 0511 762-8275 Fax: 0511 762-8290 · E-Mail: arning@iri.uni-hannover.de

**** Prof. Dr. jur. Nikolaus Forgó**
Institut für Rechtsinformatik der Leibniz Universität Hannover, Königsworther Platz 1, 30167 Hannover · Telefon: 0511 762-8159 Fax: 0511 762-8290 · E-Mail: forgo@iri.uni-hannover.de

***** Dr. jur. Tina Krügel, LL.M.**
Institut für Rechtsinformatik der Leibniz Universität Hannover, Königsworther Platz 1, 30167 Hannover · Telefon: 0511 762-8259 Fax: 0511 762-8290 · E-Mail: kruegel@iri.uni-hannover.de

mögliche genetische Veranlagungen oder Defekte, mit einer gewissen Wahrscheinlichkeit sogar über zukünftige Erkrankungen, möglicherweise auch über deren Heilungschancen und vieles mehr. Sie kann selbst für Blutsverwandte, die noch gar nicht geboren sein müssen, aussagekräftig sein. Schon die Offenbarung solcher Erkenntnisse gegenüber der jeweils betroffenen Person kann deren Leben elementar verändern; die Folgen wären noch gravierender, wenn solche Informationen Dritten oder gar der Öffentlichkeit bekannt würden. Viele dieser Informationen haben einen konkreten wirtschaftlichen Wert für Dritte, weil sich etwa der Eintritt eines Versicherungsfalls wegen einer genetisch (mit)bedingten Erkrankung oder der vorzeitige Verlust der Arbeitskraft mit höherer Genauigkeit als bisher vorhersagen lassen.

Werden diese hochsensiblen personenbezogenen Daten im Rahmen der diagnostischen Medizin oder medizinischen Forschung verarbeitet, ist ganz besonders die strikte Einhaltung geltender Datenschutzbestimmungen zu fordern, um Missbrauch vorzubeugen. Gleichzeitig sind die Datenschützer aufgerufen, den Wissenschaftlern und Ärzten einen gangbaren Weg zu weisen, der humangenetische Forschung im Rahmen der vom Gesetzgeber gesetzten Grenzen nicht behindert, sondern ermöglicht.

Der folgende Beitrag beleuchtet in einem ersten Schritt die datenschutzrechtlichen Besonderheiten genetischer Daten und befasst sich insbesondere mit der Frage, ob und gegebenenfalls in welcher Form humangenetische Daten anonymisiert werden können. Im zweiten Schritt stellen die Verfasser ein Datenschutzkonzept für transeuropäische Genforschungsprojekte vor.

2 Können genetische Daten anonymisiert werden?

2.1 Besonderheit gentechnischer Daten

Genetische Daten sind alle Daten über die Erbmerkmale einer Person oder über das für diese Merkmale typische Vererbungsmuster innerhalb einer miteinander verwandten Gruppe von Personen *(Artikel 29 Datenschutzgruppe 2004, 4)*. Aufgrund ihrer Aussagekraft hinsichtlich Gesundheitszustand, ethnischer Herkunft und familiärer Abstammung sind genetische Daten als besonders sensibel einzuschätzen. Durch die Verarbeitung dieser Art von Daten ist die Privatsphäre betroffener Personen besonders stark bedroht *(Schladebach 2003, 227; Weichert 2006b)*. Daten über Gesundheit, zu denen genetische Daten gehören, zählen deshalb zu den besonders schutzwürdigen Informationen.

Die Verarbeitung genetischer Daten ist datenschutzrechtlich brisant, aber in Deutschland bisher nicht spezialgesetzlich geregelt. Bislang liegt nur ein Entwurf für ein Gendiagnostikgesetz von Bündnis 90/Die Grünen vor. Die Tatsache, dass noch kein Gendiagnostikgesetz wirksam ist, stellt datenschutzrechtlich aber keine unmittelbare Gefährdungslage für die betroffenen Menschen dar. Die allgemeinen datenschutzrechtlichen Bestimmungen bieten bereits ein erhebliches Maß an Schutz vor einem unrechtmäßigen Eingriff in die Privatsphäre. So dürfen in Deutschland personenbezogene Daten gemäß Paragraf 4 Absatz 1 Bundesdatenschutzgesetz (BDSG) nur dann verarbeitet werden, wenn der betroffene Bürger in die Verarbeitung dieser Informationen einwilligt oder eine Rechtsvorschrift die Verarbeitung erlaubt. Diese Norm spiegelt die europäische Rechtslage wider. Ferner ist die Verarbeitung nur unter den sehr engen Voraussetzungen des Paragrafen 28 Absatz 6 ff. BDSG (Erheben, Verarbeiten und Nutzen von besonderen Arten personenbezogener Daten für eigene Geschäftszwecke) und des Paragrafen 40 BDSG (Verarbeitung und Nutzung durch Forschungseinrichtungen) zulässig.

2.2 Einwilligungserklärung betroffener Personen zur Verwendung genetischer Daten

Die wichtigste Ausnahme zum Verarbeitungsverbot von sensiblen Daten stellt die ausdrückliche Einwilligung der betroffenen Person in die jeweilige Datenverarbeitung dar (Paragraf 28 Absatz 6 in Verbindung mit Paragraf 4a Absatz 3 BDSG). So wird vor allem im Bereich der diagnostischen Medizin die Datenverarbeitung häufig durch die Einwilligung des Patienten erlaubt. Es deutet sich an, dass die Akzeptanz der Genforschung im medizinisch-therapeutischen Bereich in der Bevölkerung zunimmt. Nach einer 2001 durchgeführten Forsa-Umfrage sprachen sich fast drei Viertel der Befragten für die Genforschung im diagnostisch-therapeutischen Bereich aus[1]. Während nach einer Emnid-Umfrage im Jahr 1996 nur 53 Prozent der Befragten davon überzeugt waren, dass Gentechnik entscheidende Vorteile bei der Behandlung von Krebs bringe, waren es im Jahr 2002 bereits 67 Prozent, die dieser Einschätzung zustimmten *(Anonymus 2002)*.

Bei nicht medizin-diagnostisch motivierten Genanalysen, wie etwa dem sogenannten Vaterschaftstest, aber auch bei DNA-Analysen zur Überführung von Straftätern, sieht das Meinungsbild in der Bevölkerung ähnlich aus: Nach einer Forsa-Umfrage aus dem Jahr 2005 ist immerhin eine deutliche Mehrheit der Deutschen für die Ausweitung von DNA-Analysen zur Überführung von Straftätern[2].

Im Gegensatz zum medizinisch-therapeutischen Bereich liegt die Einwilligung der betroffenen Person hier allerdings nur selten vor, so dass die Erlaubnis zur Datenverarbeitung gesetzlich vorgesehen sein muss. Für den sogenannten Vaterschaftstest gibt es bisher keine gesetzliche Grundlage. Nach geltendem Recht darf ein solcher Test mithin nur mit dem Einverständnis des Kindes beziehungsweise seines Erziehungsberechtigten erfolgen. Kürzlich wurde die Notwendigkeit einer gesetzlichen Grundlage jedoch durch das Bundesverfassungsgericht festgestellt *(Bundesverfassungsgericht 2007)*. Der deutsche Gesetzgeber ist angehalten,

bis Ende März 2008 eine entsprechende Rechtsvorschrift zu erlassen.

Dennoch bleibt es schwierig vorherzusagen, wie sich die Bereitschaft betroffener Patienten, in die Verarbeitung ihrer personenbezogenen Daten zu Therapie- und Forschungszwecken einzuwilligen, im Bereich der Genforschung entwickeln wird und auch im Einzelfall auswirkt. Sicher ist hingegen, dass eine durch Einwilligung oder Gesetz zulässige Datenverarbeitung kein „Freibrief" für einen beliebigen Umgang mit diesen Daten ist. Auch und gerade dann gelten zentrale Anforderungen an den Datenschutz, so:

- der Grundsatz der Datenvermeidung und der Datensparsamkeit (Paragraf 3a BDSG), der es erforderlich macht, nicht mehr benötigte Daten zu löschen,
- der Zweckbindungsgrundsatz (u.a. Paragraf 14 BDSG), nach dem nur solche Daten zu erheben sind, die für die vorgesehenen Zwecke erforderlich sind und
- das Prinzip der Datensicherheit (Paragraf 9 BDSG), das durch technische und organisatorische Maßnahmen die betroffene Person vor der unzulässigen Weitergabe von Daten an Dritte schützt.

Soweit feststellbar, werden diese Anforderungen im klinischen Alltag nicht immer erfüllt. Dies dürfte sich aus einer Gemengelage von Gründen erklären lassen: Zu nennen sind die relativ hohe Komplexität des Datenschutzrechts und seine Unbekanntheit in der Bevölkerung sowie unter Ärzten und Forschern, die mitunter mit einer fehlenden Sensibilisierung des befassten Personals einhergeht. Die hohe Komplexität des Datenschutzrechts impliziert hohe technisch-administrative Anforderungen, denen sich die Verantwortlichen wegen des erheblichen Aufwands in der Regel nur unzureichend stellen können. Auch ist im Bereich der humangenetischen Forschung eine wirksame Einwilligung des Patienten weder unproblematisch noch ausreichend: Der Forscher braucht nämlich für jede einzelne medizinische Untersuchung der Daten eine konkrete Erlaubnis des betroffenen Bürgers, sodass eine Einwilligung für jede einzelne Datenverarbeitung einzuholen ist. Dies ist in der Praxis faktisch unmöglich.

Wird die Einwilligungserklärung des Betroffenen deshalb sehr weit gefasst, um möglichst viele Datenverarbeitungsvorgänge möglich zu machen und dadurch etwa auch sich erst im Projektverlauf ergebende neue Forschungsmethoden zu erfassen, von denen der Bürger zum Zeitpunkt seiner Einwilligungserklärung keine Kenntnis hatte, ist deren rechtliche Zulässigkeit zweifelhaft. Fasst man die Einwilligungserklärung hingegen enger, ist das Gegenteil der Fall. Neue, sich im Zuge des Projekts ergebende Forschungsmethoden wären von der Einwilligung nicht erfasst. Entsprechend bedürfte es noch nach Jahren neuer Einwilligungen von den jeweils betroffenen Patienten. Der damit verbundene organisatorische Aufwand und die sich möglicherweise ergebenden rein praktischen Probleme – etwa ob der Patient überhaupt noch einwilligungsfähig ist – liegen auf der Hand. Eine enge Auslegung der Einwilligungserklärung bleibt im Hinblick auf den medizinischen Fortschritt kontraproduktiv und würde auch Heilungschancen der Patienten aufs Spiel setzen. Umgekehrt wäre die Privatsphäre des Bürgers durch eine weite Auslegung erheblich gefährdet. Welcher Lösungsweg ist also einzuschlagen?

2.3 Anonymisierung genetischer Daten

Von dem Verbot der Verarbeitung personenbezogener Daten gibt es neben der Einwilligung des Patienten weitere Ausnahmen für die Forschung mit genetischen Daten, so auf Grundlage von

- Paragraf 28 Absatz 6 BDSG (Erheben, Verarbeiten und Nutzen von besonderen Arten personenbezogener Daten für eigene Geschäftszwecke),
- Paragraf 28 Absatz 7 BDSG (Erheben von besonderen Arten personenbezogener Daten zum Zwecke der Gesundheitsvorsorge und der medizinischen Diagnostik etc.) und
- Paragraf 40 BDSG (Verarbeitung und Nutzung personenbezogener Daten durch Forschungseinrichtungen).

Kann auf diese Ausnahmen an dieser Stelle nicht in extenso eingegangen werden, so bleibt doch festzuhalten: Den besten Schutz für die Privatsphäre des betroffenen Bürgers bietet die Anonymisierung seiner Daten. Im Entwurf von Bündnis 90/Die Grünen zum Gendiagnostikgesetz ist dies entsprechend vorgesehen, wenn es in Paragraf 28 Absatz 1 Satz 1 heißt: „Personenbezogene genetische Proben und Daten sind zu anonymisieren, soweit und sobald dies nach dem Forschungszweck möglich ist."

Anonymisieren ist gemäß Paragraf 3 Absatz 6 BDSG das Verändern personenbezogener Daten, so dass Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können. Aus den verarbeiteten anonymen Daten kann die betroffene Person also durch die datenverarbeitende Stelle nicht mehr identifiziert werden. Kann ein Datensatz nicht mehr einer Person zugewiesen werden, so geht von dessen Verarbeitung auch keine Gefahr für grundrechtliche und wirtschaftliche Interessen des Betroffenen aus. Gemäß Paragraf 1 Absatz 1 BDSG, in dem Zweck und Anwendungsbereich des BDSG geregelt ist, sind in solch einem Fall die Datenschutzgesetze auch nicht anwendbar.

Entsprechend bietet die Verarbeitung anonymer Daten den weitaus besten Schutz vor unzulässigen Eingriffen in die Privatsphäre der betroffenen Menschen. Der Anonymisierung ist deshalb gegenüber den möglicherweise einschlägigen Ausnahmen zum Verarbeitungsverbot nach Paragraf 28 Absatz 6, 7 oder Paragraf 40 BDSG Vorrang einzuräumen. Der Schutz der Privatsphäre wird sichergestellt, die medizinische Forschung nicht unnötig behindert: Anonymisierung bietet sich demnach als Lösungsweg an.

Bei der Planung eines Projekts muss sorgfältig erwogen werden, ob eine anonymisierte Datenverarbeitung möglich ist. In solch einem Fall ist es – aus rein rechtlicher Sicht – nicht mehr erforderlich, eine Einwilligung der betroffenen Person einzuholen, weil der Anwendungsbereich der Datenschutzgesetze nicht gegeben ist. Der Verarbeitung anonymer Daten sind in diesem Fall keine Grenzen gesetzt. Sie können aus datenschutzrechtlicher Sicht mangels Personenbezugs beliebig gesammelt, gespeichert und veröffentlicht werden *(Weichert 2006a, 14)*. Darüber hinaus erfüllt der Forscher seine Pflicht gemäß Paragraf 3a und Paragraf 40 Absatz 2 Satz 1 BDSG, wonach genetische Daten, sobald es der Forschungszweck zulässt, zu anonymisieren sind.

2.4 Faktische Anonymisierung von genetischen Daten

Bevor das Datenschutzkonzept für transeuropäische Genforschungsprojekte auf Grundlage dieses Lösungswegs vorgestellt werden kann, stellt sich zunächst die Frage, inwieweit sich genetische Daten überhaupt anonymisieren lassen. Außerdem gilt es in diesem Zusammenhang zu bedenken, dass für die medizinische Forschung anonymisierte Daten in vielen Fällen wenig hilfreich sind. Um den Krankheitsverlauf bei einem Patienten zu verfolgen, ihm erforschte Heilbehandlungen zugute kommen zu lassen und auch seine Reaktion auf die Behandlung beurteilen zu können, muss der Patient identifizierbar bleiben. Aus diesem Grund wird im Forschungsbereich, soweit der direkte Personenbezug für die jeweilige Datenverarbeitung nicht unerlässlich ist, vornehmlich mit Pseudonymen gearbeitet. Bietet der Einsatz von Pseudonymen für Forschung wie auch Therapie erhebliche Vorteile, birgt dieses Verfahren aber immer noch Risiken für den Persönlichkeitsschutz des Patienten.

Pseudonymisieren ist gemäß Paragraf 3 Absatz 6a BDSG das Ersetzen des Namens und anderer Identifikationsmerkmale durch ein Kennzeichen zu dem Zweck, die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren. Wird die Zuordnung eines bestimmten Pseudonyms zu einer bestimmten Person in einer Tabelle festgehalten, so kann die betroffene Person auch nach der Pseudonymisierung durch die Stelle identifiziert werden, die Zugriff auf diese Tabelle hat. Problematisch ist jedoch bereits, ob genetische Daten aufgrund ihrer bereits beschriebenen Besonderheiten wie zum Beispiel ihrer Einzigartigkeit überhaupt anonymisiert werden können.

Man stelle sich vor, dass eine für die Identifizierung einer Person ausreichend große Gensequenz ohne jeden weiteren Personenbezug (wie z.B. des Namens des betroffenen Bürgers) im Rahmen einer Studie über das HI-Virus im Internet veröffentlicht wird. Ist die genetische Information dieser Person in anderem Zusammenhang bereits als Referenzdatensatz gespeichert, sei es im Rahmen eines flächendeckenden Speicheltests oder als Voraussetzung für eine Lebensversicherung mit hoher Deckungssumme, wäre für alle Personen, die Zugriff auf diese Datenbanken haben, nunmehr eine Identifizierung der betroffenen Person und seiner HIV-Erkrankung durch ein Matchingverfahren möglich.

Mag dieses Szenario nicht unmittelbar bevorstehen, so ist die Verwendung beziehungsweise Einforderung genetischer Untersuchungen für Arbeits- und Versicherungsverhältnisse nicht mehr nur als Schreckgespenst zu betrachten. In den USA und England ist das in bestimmten Bereichen bereits Realität *(Weichert 2002, 134)*. Die Einzigartigkeit von genetischen Daten bringt das Problem mit sich, dass trotz umfassender Anonymisierung mit entsprechendem Zusatzwissen grundsätzlich ein Rückschluss auf die jeweilige Person möglich bleibt *(Weichert 2002, 134)*. Ist dies der Fall, stellt sich die Frage, ob genetische Daten überhaupt im Sinne des Datenschutzrechts anonymisiert werden können oder immer als personenbezogene Daten einzustufen sind. Genau genommen stellt sich diese Problematik auch nicht nur bei genetischen Daten. Es wird vielmehr bereits seit Anfang der 80er Jahre darauf hingewiesen, dass sich aufgrund des steten Zuwachses der Rechnerkapazitäten und Entwicklung neuer Forschungsmethoden wie zum Beispiel des Data Minings vormals anonyme Daten einer bestimmten Person zuordnen lassen, weshalb sie fortan als personenbezogen zu qualifizieren sind. Personenbezogene Daten lassen sich folglich nicht mit absoluter Sicherheit für alle Zeiten anonymisieren, da stets ein latentes Risiko der Deanonymisierung besteht *(Brennecke 1980, 159; Burkert 1980, 143; Gebhardt 1995, 108)*.

Wie bereits angemerkt, definiert das BDSG in Paragraf 3 Absatz 6 das Anonymisieren von personenbezogenen Daten als das Verändern personenbezogener Daten, so dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können. Entscheidend für die Beurteilung der Frage, ob Gendaten überhaupt anonymisiert werden können, ist also, ob die für die Datenverarbeitung verantwortliche Stelle die betroffene Person identifizieren kann, beispielsweise anhand von Informationen, die sie tatsächlich besitzt (z.B. den Namen der Person oder im Fall der Pseudonymisierung die Zuordnungstabelle) oder anhand von Informationen, auf die sie legalerweise Zugriff haben könnte (z.B. auf Referenzdatenbanken).

Für eine behandelnde Klinik ist dies immer der Fall. Hier werden bereits aus organisatorischen Gründen auch personenbezogene Daten, etwa Name und Anschrift des Patienten, gespeichert. Die im Rahmen eines nationalen beziehungsweise transnationalen Genforschungsprojekts für die Datenverarbeitung verantwortliche Stelle verfügt in der Regel aber nicht über die Namen der Probanden oder einen legalen Zugang zu Pseudonymisierungstabellen oder Referenzdatenbanken, mit Hilfe derer

sie im Wege eines Matchingverfahrens die Identität der betroffene Person feststellen könnte. Folglich würde es sich bei den Daten für sie um anonymisierte Daten handeln, die nicht in den Anwendungsbereich der Datenschutzgesetze fielen. Die verantwortliche Stelle könnte mit den von ihr zu verarbeitenden genetischen Daten nach Belieben verfahren.

Dennoch bleibt auch in einem solchen Fall eine mögliche Gefahr bestehen. Durch eine Veröffentlichung, zum Beispiel im Internet, könnten Dritte auf diese Daten zugreifen und den Personenbezug wiederherstellen, weil sie über personenbezogene Referenzdatensätze verfügen und ein Interesse an den damit verbundenen Informationen haben. Strafverfolgungsbehörden oder Versicherungen, die teilweise über Gendatenbanken verfügen, könnten beispielsweise ein großes Interesse daran haben zu erfahren, ob eine bestimmte Person, über die sie einen personenbezogenen Referenzdatensatz besitzen, eine bestimmte Krankheit hat. Der Bürger würde dadurch in seiner Freiheit verletzt, selbst zu entscheiden, wann er wem welche Daten zugänglich macht, so dass der Sinn des Datenschutzrechts unterlaufen wäre.

Besteht also die Gefahr, dass ein Dritter die von der verantwortlichen Stelle zu verarbeitenden Daten einsehen und die betroffene Person identifizieren kann, so kann hieraus ein erheblicher Eingriff in die Privatsphäre dieser Person folgen. Wenn ein Dritter auf legalem Weg Zugang zu Wissen hat, mit dessen Hilfe eine Identifizierung der Person möglich ist, handelt es sich bei den genetischen Daten für eine verantwortliche Stelle dennoch um personenbezogene Daten, obwohl diese selbst die betroffene Person gar nicht identifizieren kann. Mehr noch, da die verantwortliche Stelle in der Praxis gar nicht wissen kann, für welche der von ihr verwendeten Gendatensätze überhaupt ein personenbezogener Referenzdatensatz existiert, müsste sie immer alle Gendatensätze als personenbezogene Daten betrachten, um einer etwaigen Verantwortlichkeit zu entgehen.

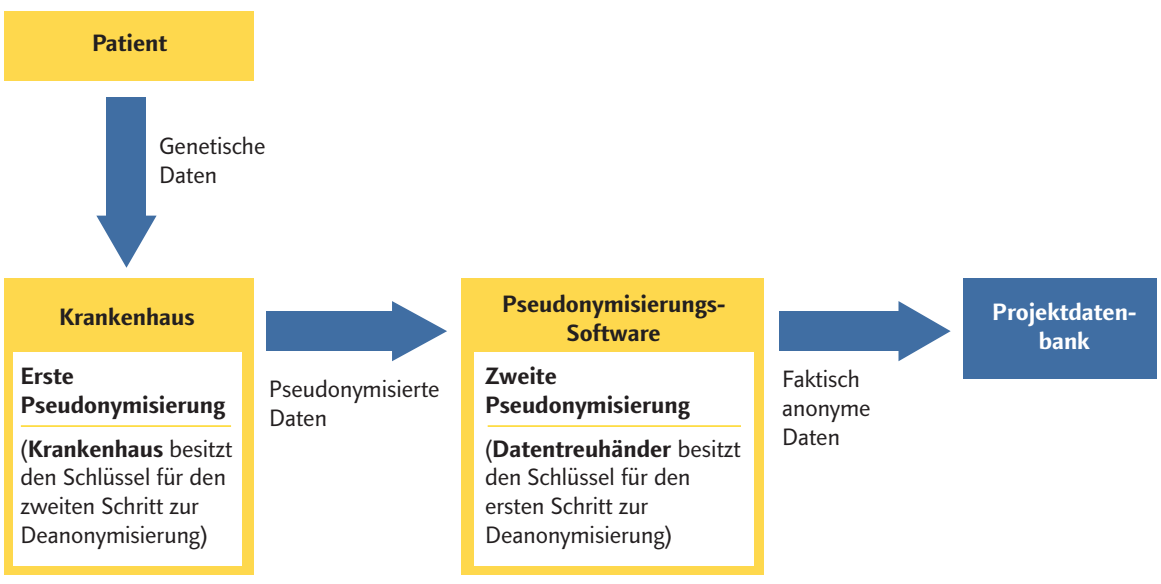
Daraus folgt, dass mögliches Zusatzwissen Dritter der verantwortlichen Stelle zugerechnet werden muss (für biometrische Daten vgl. *Hornung 2004, 430; Schaar 2005, 75)*. Folglich braucht die verantwortliche Stelle deshalb für jede Verarbeitung der anonymisierten Daten eine Erlaubnis durch eine Rechtsvorschrift oder eine Einwilligung des betroffenen Patienten.

Wäre dadurch eine betroffene Person wirksam in ihrer Privatsphäre und ihrem Recht auf informationelle Selbstbestimmung geschützt, würde auf der anderen Seite diese Auslegung die medizinische Forschung, wie oben bereits ausgeführt, erneut stark beeinträchtigen. Zugleich bleibt die Wirksamkeit einer allumfassenden Einwilligung der betroffenen Person, die auch zukünftige, zum Erklärungszeitpunkt nicht bekannte Verarbeitungsvorgänge einschließt, rechtlich zweifelhaft (mehr zu Anonymisierung von genetischen Daten und Zurechnung von Wissen Dritter: *Arning et al. 2006a, 700*).

Vor diesem Hintergrund ist festzuhalten, dass der hier dargelegte Lösungsweg einer Anonymisierung genetischer Daten restriktiv aufgefasst werden muss. Es besteht nämlich für die

Abbildung 1

Die zweifache Pseudonymisierung



Quelle: Arning, Forgó, Krügel 2007

Privatsphäre des betroffenen Bürgers überhaupt keine Gefährdung, wenn die verantwortliche Stelle erstens legalerweise nicht auf das Zusatzwissen Dritter zugreifen kann und umgekehrt die Dritten nicht auf die Daten der verantwortlichen Stelle zugreifen können. Eine Identifizierung der betroffenen Person wäre dann nicht oder nur mit unverhältnismäßigem Aufwand zu realisieren.

Folglich muss die Zurechnung des Zusatzwissens Dritter von der konkreten Art der jeweiligen Datenverarbeitung abhängig gemacht werden. Besteht also die Gefahr, dass ein Dritter die von der verantwortlichen Stelle zu verarbeitenden Daten einsehen (z.B. infolge einer Veröffentlichung oder Übermittlung dieser Daten) und die betroffene Person identifizieren kann, so müssen die Datenschutzgesetze den Betroffenen wirksam in seiner Privatsphäre schützen. Daraus folgt, dass bei Verarbeitungsvorgängen, bei denen diese Gefahr für die Privatsphäre des Betroffenen besteht, also insbesondere im Fall der Übermittlung und Veröffentlichung von Daten, das Zusatzwissen Dritter der verantwortlichen Stelle zugerechnet werden muss *(Schaar 2005, 75)*. Diese braucht also für jede Übermittlung oder Veröffentlichung der anonymisierten Daten eine Erlaubnis (durch eine Rechtsvorschrift oder Einwilligung).

Eine in diesem Sinne verstandene Lösung bietet einen ausreichenden Schutz des verfassungsrechtlich verankerten Rechts des betroffenen Bürgers auf informationelle Selbstbestimmung, das aus Artikel 2 Absatz 1 Grundgesetz (Recht auf freie Entfaltung der Persönlichkeit) in Verbindung mit Artikel 1 Absatz 1 Grundgesetz (Unantastbarkeit der Menschenwürde) abgeleitet wird, ohne die medizinische Forschung zu behindern.

3 Entwurf eines Datenschutzkonzepts für Genforschungsprojekte

3.1 Projektinternes Datenschutzgremium

Eine solche Lösung in die Verfahrensabläufe eines europäischen Genforschungsprojektes zu integrieren und dieses damit datenschutzrechtskonform auszugestalten, erfordert die Entwicklung eines komplexen Datenfluss- und Datenschutzmodells sowie eine kontinuierliche projektinterne Prüfung und Überwachung der Abläufe, um die Einhaltung geltender Datenschutzbestimmungen zu gewährleisten. Es ist daher zunächst entscheidend, das Projektkonsortium in die Lage zu versetzen, die Verfahrensabläufe zu kontrollieren und zu auditieren beziehungsweise über die entsprechenden Prozesse Rückmeldung zu geben. Ist diese Möglichkeit nicht gegeben, bliebe die Erarbeitung von Datenschutzrichtlinien, von technischen Sicherungsmaßnahmen und klar definierten Zuständigkeiten innerhalb des Projektes nutzlos und wäre den Aufwand nicht wert. Erforderlich ist daher die Schaffung eines unabhängigen Datenschutzgremiums, dessen Rechtsform es befähigt,

- die Projektbeteiligten durch Verträge zur Einhaltung der Datenschutzbestimmungen zu verpflichten und
- bei Verstoß gegen diese Bestimmungen Sanktionen zu erlassen.

Das Datenschutzgremium wird bindende Verträge mit den Projektbeteiligten schließen, die die Einhaltung des im Folgenden beschriebenen internen Datenschutzkonzepts rechtlich absichern. Zugleich ist es extern die für die gesamte Datenverarbeitung im Projekt verantwortliche Stelle und damit Ansprechpartner bei Verletzungen der Privatsphäre von Patienten.

3.2 Datenschutz-Sicherheitsnetz

3.2.1 Zweifache Pseudonymisierung

Der zuvor vorgestellte Lösungsweg im Hinblick auf faktisch anonyme genetische Daten ist der Ausgangspunkt des hier vorgeschlagenen Datenschutzmodells. Ziel eines auf intereuropäischen Datenaustausch ausgelegten Forschungsprojekts muss es sein, grundsätzlich so wenig personenbezogene Daten wie möglich zu verarbeiten. Sowohl organisatorisch, vertraglich als auch technisch muss sichergestellt werden, dass die verarbeiteten Patientendaten projektintern als faktisch anonym im zuvor beschriebenen Sinne bewertet werden können. Hierzu ist die Implementierung einer technischen Sicherungsstruktur erforderlich, die sich
a) auf den Grundsatz der zweifachen Pseudonymisierung stützt und
b) einen Datentreuhänder in die Architektur einbindet.

Wie in Abbildung 1 zu sehen, ist der Datenfluss folgendermaßen vorgesehen: Die genetischen Daten eines Patienten werden von dem jeweils behandelnden Arzt erhoben und wie gewöhnlich vor Ort im Krankenhaus oder in angeschlossenen Laboren analysiert, um sie sodann krankenhausintern zu speichern. Das Krankenhaus und angeschlossene Labore usw. sind grundsätzlich verpflichtet, mit pseudonymisierten Daten zu arbeiten, wenn der Personenbezug für die Analysen und Tests nicht erforderlich ist. Bis zu diesem Punkt obliegt der Schutz der Patientendaten ausschließlich dem jeweiligen Krankenhaus vor Ort.

Erklärt sich nunmehr ein Patient bereit, an einer Studie, die einen internationalen Datenaustausch beinhaltet, teilzunehmen, werden seine Daten unter einem vom Krankenhaus vergebenen Pseudonym (erste Pseudonymisierung) an das jeweilige Projekt übermittelt. Die Verpflichtung, die Daten ausschließlich pseudonymisiert an das Forschungsprojekt zu übermitteln, ergibt sich jetzt über die gesetzlichen Vorschriften hinaus aus der vertraglichen Verpflichtung gegenüber dem Projekt, die das jeweilige Krankenhaus mit dem Datenschutzgremium des Forschungskonsortiums eingehen muss. Ohne einen solchen Vertrag dürfte das Krankenhaus am Datenaustausch überhaupt nicht teilnehmen.

Der Schlüssel, mit dem die Daten decodiert werden, und der somit die Identifizierung des jeweiligen Patienten ermöglicht, wird ausschließlich vom Krankenhaus und dort zumeist von dem behandelnden Arzt verwaltet. Hat dieser aufgrund des Behandlungsvertrags mit dem Patienten ohnehin Zugang zu den personenbezogenen Daten, kann die erste Pseudonymisierung deshalb nur mithilfe des behandelnden Arztes aufgehoben werden.

Die Praxis zeigt jedoch, dass die Pseudonymisierungsverfahren in den unterschiedlichen Krankenhäusern europaweit vielfältig und qualitativ sehr unterschiedlich sind. Die Skala reicht von einfachen Namenskürzeln, von denen überaus leicht auf den jeweiligen Patienten zurückgeschlossen werden kann, bis zu professionellen technischen Pseudonymisierungsverfahren. Dies macht eine zweite Pseudonymisierung nach dem neuesten Stand der Technik notwendig. Erst hierdurch kann ein einheitlich hoher Datenschutzstandard in einem auf intereuropäischen Datenaustausch ausgelegten Genforschungsprojekt gewährleistet werden.

Während der Übermittlung der vom Krankenhaus erstmals pseudonymisierten Daten werden die Daten folglich mit Hilfe eines technischen Verfahrens ein weiteres Mal pseudonymisiert. Die Einbindung eines Datentreuhänders für diese Aufgabe, der etwa eine Pseudonymisierungssoftware zur Verfügung stellt, garantiert eine qualitativ hochwertige zweite Pseudonymisierung. Erst hiernach werden die nunmehr sicher faktisch anonymen genetischen Daten in das Projekt eingespeist. Es muss sichergestellt sein, dass die faktisch anonymen Daten (soweit sie in einer physisch vorhandenen Datenbank gespeichert werden sollen) und der Schlüssel für die zweite Pseudonymisierung in unterschiedlichen Datenbanken gespeichert werden. Die am Projekt beteiligten Forscher dürfen nur auf die faktisch anonymen Daten Zugriff haben, nicht aber auf die Schlüssel.

Deshalb sollte die zweite Pseudonymisierung von einem spezialisierten IT-Sicherheitsunternehmen vorgenommen werden, das idealerweise im Projektverbund auch als Datentreuhänder fungiert und die Schlüssel zu dieser Pseudonymisierung hält *(Arning et al. 2006b)*. Sollte es als Folge der Forschungsergebnisse erforderlich werden, einen der Patienten zu kontaktieren, ist für eine Identifizierung zunächst die Kooperation mit dem

Abbildung 2

Das Datenschutz-Sicherheitsnetz für intereuropäische Forschungsvorhaben

1.	2.	3.
(Doppelte) Pseudonymisierung der genetischen Daten unter Verantwortung des Datenschutzgremiums	**Erstes Fallback-Szenario: Einwilligung von jedem Patienten (Hauptfallback-Szenario)**	**Zweites Fallback-Szenario: Ausnahmen für die genetische Forschung im nationalstaalichen Recht**
Die genetischen Daten werden (doppelt) pseudonymisiert und sind projektintern als anonym zu qualifizieren. – Keine personenbezogenen Daten gemäß Paragraf 3 Absatz 1 Bundesdatenschutzgesetz (BDSG) – Genetische Daten können projektintern genutzt werden – Das Projekt hält die datenschutzrechtlichen Regelungen „automatisch“ ein	Jeder Patient, der am Projekt teilnimmt, oder sein Vertreter muss seine ausdrückliche Einwilligung geben. – Datenverarbeitung ist rechtmäßig (Paragraf 28 Absatz 6 in Verbindung mit Patagraf 4a Absatz 3 BDSG) – Einwilligung ist notwendig aus ethischer Sicht **→ „Autonomie benötigt Einwilligung“**	Artikel 8 Absatz 4 der Datenschutzrichtlinie 95/46/EG ermächtigt die Mitgliedstaaten, Ausnahmen aufgrund wichtigen öffentlichen Interesses einzuführen. – Mitgliedstaaten können die Verarbeitung von genetischen Daten für wissenschaftliche Zwecke erlauben – Nationalstaatliches Recht muss analysiert werden, wenn Stufe 1 und 2 versagen. In Deutschland: Paragraf 40 BDSG

Quelle: Arning, Forgó, Krügel 2007

Datentreuhänder erforderlich. Nur dieser verfügt über den Schlüssel für die zweite Pseudonymisierung und weiß zudem auch, von welchem der teilnehmenden Krankenhäuser der betreffende Patient behandelt wird. Das behandelnde Krankenhaus muss über den Datentreuhänder bei einer erforderlichen Identifizierung eines Patienten eingebunden werden. Das Prinzip der zweifachen Pseudonymisierung gewährleistet, dass ein Rückschluss auf den betroffenen Patienten für beide „schlüsselhaltenden" Stellen nur mit Hilfe der jeweils anderen Stelle möglich ist.

3.2.2 Einwilligung (erstes Fallback-Szenario)

Das hier vorgeschlagene Datenschutzkonzept ist gleich einem Sicherheitsnetz aufgebaut (Abbildung 2). Das Prinzip der zweifachen Pseudonymisierung ist geeignet, die gesamte Datenverarbeitung innerhalb des Projektes anonymisiert erfolgen zu lassen. Gleichwohl kann es keine Garantie dafür geben, dass ein Forscher nicht doch auf die Pseudonymisierungsschlüssel Zugriff hat. Ebenso kann es sich ergeben, dass ein Projektteilnehmer faktisch anonyme Daten vereinbarungswidrig in einer Fachzeitschrift veröffentlicht. Des Weiteren lässt sich nicht ausschließen, dass ein Mitglied des Forschungskonsortiums die Daten an eine Forschungseinrichtung außerhalb des Projektes übermittelt. In solchen Fällen können die Daten nicht mehr als anonym bewertet werden, mit der Folge, dass die Datenschutzrichtlinie in diesen Fällen Anwendung findet und damit eine Rechtsgrundlage für diese Datenverarbeitung erforderlich ist.

Trotz aller bezüglich der Reichweite oben dargestellten Schwierigkeiten, die mit einer Einwilligung eines betroffenen Patienten in die Datenverarbeitung einhergehen, wird in einem zweiten Schritt des Datenschutzkonzepts eine solche verlangt. Obwohl die Verarbeitung der Daten grundsätzlich anonym erfolgt und eine Einwilligung folglich aus rechtlicher Sicht gar nicht erforderlich wäre, hat dies verschiedene Vorteile:

- Durch die Einwilligung wird der Patient in das Projekt einbezogen. Dies sorgt für Transparenz und schafft Vertrauen, Faktoren, die aus ethischen Gründen bei Forschungsprojekten dieser Art erforderlich sind.
- Überdies schafft sie für die beteiligten Forscher eine rechtssichere Situation.

Da es auf den Willen, die betroffene Person zu identifizieren nicht ankommt, sondern ausschließlich auf die faktische Möglichkeit der Identifizierung, mag es Situationen geben, in denen es für den Forscher selbst schwierig zu bewerten ist, ob er die Möglichkeit hätte, den betreffenden Patienten zu identifizieren. Sollte dies der Fall sein, wäre seine Datenverarbeitung rechtswidrig. Das Sicherungssystem der zweifachen Pseudonymisierung würde nicht greifen. Auch wenn dies kaum vorkommen wird, bedürfte es für solche Fälle einer Legitimation der Datenverarbeitung. Die Einwilligung des Patienten hierzu ist deshalb als Fallback-Lösung immer geeignet.

3.2.3 Ausnahmen vom Verbot der Datenverarbeitung in den nationalen Gesetzen (zweites Fallback-Szenario)

Für den unwahrscheinlichen Fall, dass einmal sowohl das Prinzip der zweifachen Pseudonymisierung versagen sollte und auch die von dem betroffenen Patienten erteilte Einwilligung für die fragliche Datenverarbeitung nicht greift, ist in einem dritten Schritt das jeweils anwendbare nationale Recht zu analysieren. Die Datenschutzrichtlinie gibt in Artikel 8 Absatz 4 DSRL den Mitgliedstaaten die Möglichkeit, Ausnahmen vom Verbot der Verarbeitung sensibler Daten zu schaffen. Der deutsche Gesetzgeber hat hiervon durch die Einführung des Forschungsprivilegs in Paragraf 40 BDSG Gebrauch gemacht, so dass eine Verarbeitung von sensiblen Daten im Rahmen deutscher Forschungsarbeit durch den Gesetzgeber legitimiert ist.

4 Fazit

Die Forschung an menschlichen Genen birgt enorme Chancen für den medizinischen Fortschritt, aber auch große Risiken für die Privatsphäre der betroffenen Personen. Es ist die Aufgabe der Politik und der Datenschützer, Modelle für Genforschung zu entwickeln, die den medizinischen Fortschritt unterstützen und zugleich die Privatsphäre der betroffenen Personen wirksam schützen. Nur wenn diese beiden Faktoren gewährleistet werden können, wird die Genforschung für therapeutische Zwecke von der Bevölkerung nachhaltig akzeptiert.

Das vorgestellte Datenschutzkonzept für (transeuropäische) Genforschungsprojekte bietet bei konsequenter Umsetzung einen größtmöglichen Schutz für die informationelle Selbstbestimmung der beteiligten Patienten und eröffnet den Forschern gleichzeitig weitgehend unbeschränkte Möglichkeiten, vorhandene Daten innerhalb des Projektes zu verarbeiten. Es empfiehlt die Einbindung eines mehrstufigen Sicherheitsnetzes. Unter Einrichtung eines projektinternen Datenschutzgremiums und der Einbindung eines Datentreuhänders stützt es sich in erster Linie auf die ausschließliche Verarbeitung faktisch anonymer genetischer Daten. Die weitgehende Einbindung der Patienten wird durch die obligatorischen Einwilligungen gewährleistet, in deren Rahmen die Patienten umfassend informiert werden und die es ihnen ermöglicht, selbst über die Verarbeitung der sie betreffenden Gendaten zu entscheiden. ◆

Fußnoten

Web-Quellen
(letzter Zugriff im September 2007)

[1] http://www.medport.de/nw_read.php/22215
[2] http://www.n-tv.de/302324.html

Literatur

Anonymus (2002): Vertrauen in die Gentechnik. Deutsches Ärzteblatt, Jg. 99, Heft 19, Seite A-1262
Antonow K (2006): Der rechtliche Rahmen der Zulässigkeit für Biodatenbanken zu Forschungszwecken, Baden-Baden: Nomos
Arning M, Forgó N, Krügel T (2006): Datenschutzrechtliche Aspekte der Forschung mit genetischen Daten. Datenschutz und Datensicherheit, Heft 6, 700–705
Arning M, Forgó N, Krügel T (2006): Datenschutzrechtliche Aspekte bei der Forschung mit menschlichen Genen. In: Hochberger C, Likowski R (Hrsg.): Informatik 2006 – Informatik für Menschen, Band 1. Bonn: GI-Edition, 702–708
Artikel 29 Datenschutzgruppe (2004): Arbeitspapier über genetische Daten;
http://ec.europa.eu/justice_home/fsj/privacy/docs/wdocs/2004/wp91_de.pdf
Brennecke R (1980): Kriterien zur Operationalisierung der faktischen Anonymisierung. In: Kaase M, Krupp, H, Pflanz M et al. (Hrsg.): Datenzugang und Datenschutz. Königstein: Athenäum, 158–176
Bundesverfassungsgericht (2007): BVerfG, 1 BvR 421/05 vom 13.2.2007;
www.bverfg.de/entscheidungen/rs20070213_1bvr042105.htm
Burkert H (1980): Das Problem des Zusatzwissens. In: Kaase M, Krupp H, Pflanz M et al. (Hrsg.): Datenzugang und Datenschutz. Königstein: Athenäum, 143–148
Gebhardt U (1995): Anonymisierung als Weg aus der Mitbestimmung bei elektronischer Datenverarbeitung gemäß § 87 I Nr. 6 BetrVG? Neue Zeitschrift für Arbeitsrecht, Heft 3, 103–110
Hornung G (2005): Die digitale Identität. Baden-Baden: Nomos
Hornung G (2004): Der Personenbezug biometrischer Daten. Datenschutz und Datensicherheit, Heft 7, 429–432
Schaar P (2005): Datenschutzrechtliche Schranken der Genanalysen. In: Ronellenfitsch M, Kartmann N (Hrsg.): Genanalysen und Datenschutz. Wiesbaden;
http://www.datenschutz.hessen.de/Forum/Forum2004.pdf
Schladebach M (2003): Genetische Daten im Datenschutzrecht. Computer und Recht, Heft 3, 225–233
Weichert, T (2006): Rechtsquellen und Grundbegriffe des allgemeinen Datenschutzes. In: Kilian W, Heussen B (Hrsg.): Computerrechts-Handbuch. München: C.H. Beck
Weichert T (2006): Besonderer Datenschutz. In: Kilian W, Heussen B (Hrsg.): Computerrechts-Handbuch. München: C.H. Beck
Weichert T (2002): Der Schutz genetischer Informationen. Datenschutz und Datensicherheit, Heft 3, 133–145

Die Autoren

Dipl.-Jur. Marian Arning, LL.M.
Studium der Rechtswissenschaften an der Universität Hannover. Anschließend Mas-terstudiengang „Rechtsinformatik" an den Universitäten Hannover und Leuven (Belgien). Seit 2001 am Institut für Rechtsinformatik der Universität Hannover tätig, seit 2006 als wissenschaftlicher Mitarbeiter. Forschungsschwerpunkt: Einsatz von Informations- und Kommunikationstechnologie im Gesundheitswesen. Seit 2005 Lehrbeauftragter für Informations- und Datenschutzrecht an der Fachhochschule Hannover.

Prof. Dr. jur. Nikolaus Forgó
Studium der Rechtswissenschaften an der Universität Wien und der Université Paris II, Promotion an der Universität Wien. Seit 2002 Universitätsprofessor für Rechtsinformatik und IT-Recht in Hannover, seit 2007 Leiter des Instituts für Rechtsinformatik der Universität Hannover sowie seit 1998 Leiter des Universitätslehrgangs für Informationsrecht und Rechtsinformation an der Universität Wien. Lehrtätigkeiten an der Universität Hannover, der Universität Wien, der WU-Wien, der Donau Universität Krems und der Bucerius Law School in Hamburg.

Dr. jur. Tina Krügel, LL.M.
Studium der Rechtswissenschaften an der Universität Hannover. Anschließend Masterstudiengang „Rechtsinformatik" an den Universitäten Hannover und Oslo (Norwegen). Von 2003 bis 2004 Tätigkeit als Rechtsanwältin in einer hannoverschen Anwaltskanzlei. Seit 2004 arbeitet sie als selbstständige Rechtsanwältin mit Schwerpunkt Rechtsinformatik und als wissenschaftliche Mitarbeiterin am Institut für Rechtsinformatik der Universität Hannover.

Die Autoren sind am EU-Forschungsprojekt „Advancing Clinico-Genomic Trials on Cancer" (ACGT) beteiligt und hier für rechtliche, insbesondere datenschutzrechtliche, Belange verantwortlich (www.eu-acgt.org). Das Ziel des Projekts sind die Verbesserung und Entwicklung von Behandlungs- und Heilungsmöglichkeiten bestimmter Krebsarten durch den Aufbau einer neuartigen Grid-Infrastruktur zum Austausch und zur Erforschung genetischer Daten.