Datenschutzrechtliche Bewertung des Projekts „OA-Statistik“ an der Universität Stuttgart

Die Bibliothek der Universität Stuttgart war Partner im Projekt „Infrastruktur für standardisierte Nutzungsstatistiken unter besonderer Berücksichtigung Institutioneller Repositories“ (OA-Statistik). Sie beabsichtigt, sich erneut an diesem Projekt zu beteiligen. Ziel ist es, Zugriffe der Nutzer auf Open-Access-Dokumente zu erfassen, auszuwerten und daraus Rückschlüsse auf das Angebot von Open-Access-Dokumenten zu ziehen.

Die Universitätsbibliothek betreibt einen Dokumentenserver. Mit diesem will sie erfassen, unter welcher IP-Adresse welche Dokumente von den Nutzern aufgerufen werden. Neben der IP-Adresse werden folgende Angaben erfasst und in einer Protokolldatei gespeichert: Datum, Uhrzeit, Zeitzone, Spezifikation des vom Nutzer eingesetzten Webbrowsers (User-Agent), Pfad (URL) und Name des abgerufenen Dokuments, Get-Parameter, Größe der Datei, verwendetes Protokoll, Texttyp (bspw. html oder pdf) sowie der Status (ob bspw. das Dokument erfolgreich ausgeliefert wurde). Nach dem internen Papier „AP3 Datenschutz: Anschreiben Datenschutzbeauftragte“ sowie der „Darstellung des DFG-Projekts Open-Access-Statistik unter besonderer Berücksichtigung datenschutzrechtlicher Aspekte“ wird überdies die verweisende Verknüpfung (Referrer) gespeichert. Diese Angaben werden in einer Protokolldatei gespeichert. Im nächsten Schritt wird die erfasste IP-Adresse durch einen „One-Way-Hash“ ersetzt, einer Prüfsumme, die sich aus sich heraus nicht in den Ursprungswert zurückrechnen lässt. Die so veränderten Protokolldaten sollen an einen zentralen Serviceprovider (eine „zentrale Instanz“) außerhalb der Universität Stuttgart weitergegeben, dort ausgewertet und die Ergebnisse z. T. an die Universität Stuttgart zurückgeleitet werden. Als zentrale Instanz fungiert zunächst die Universität Göttingen; später soll dies durch eine neutrale Stelle erfolgen, die von den beteiligten Institutionen unabhängig und „vertrauenswürdig“ ist. Durch das Projekt soll u. a.

- ein einheitlicher Standard zur Ermittlung von Zugriffszahlen und Statistiken für Publikationen entwickelt werden,
- ein "attraktiver Dienst für Wissenschaftler, Hochschulen und Forschungseinrichtungen sowie Förderinstitutionen geschaffen [werden], der standardisiert verlässlich aufzeigt,

wie wissenschaftliche Publikationen genutzt werden und welche Auswirkungen freie Zugänglichkeit auf die Nutzung und Sichtbarkeit hat",
- die Akzeptanz des Open-Access-Formats sowie institutioneller Dokumentenserver erhöht werden.

1. Maßgeblich für die rechtliche Bewertung ist das Telemediengesetz (TMG). Gemäß § 1 TMG findet es auf alle elektronischen Informations- und Kommunikationsdienste Anwendung, soweit sie nicht Telekommunikationsdienste (bei denen ausschließlich Signale über Telekommunikationsnetze übertragen werden, ohne dass eine inhaltliche Dienstleistung erbracht wird), telekommunikationsgestützte Dienste oder Rundfunk sind. Ein Dokumentenserver stellt ein derartiges inhaltliches Angebot im Sinne des § 1 TMG dar, weil elektronische Dokumente zum Abruf bereit gestellt werden; ein Telekommunikationsdienst, ein telekommunikationsgestützter Dienst oder Rundfunk liegt darin nicht.

2. Sowohl dynamische als auch statische IP-Adressen sind zumindest personenbeziehbar (siehe AG Mitte (Berlin) vom 27.03.07 – Az. 5 C 314/06). Diese umstrittene Ansicht wird in der Fachliteratur überwiegend vertreten.[1] Darüber hinaus wird in der Literatur empfohlen, diese Annahme vorsorglich einer Bewertung zugrunde zu legen.[2] Letztendlich kommt es auf diese Streitfrage jedoch nicht an. Denn zumindest ein Teil der erfassten IP-Adressen ist für die Universität Stuttgart ohne Weiteres personenbezogen. Ein Personenbezug ergibt sich für die Universität bei Anfragen aus dem universitätseigenen IP-Adressbereich. Da die Universität als Access-Provider selbst die IP-Adressen zuweist, verfügt sie über das nötige Zusatzwissen, um eine IP-Adresse einer natürlichen Person zuordnen zu können. Folglich erhebt die Universitätsbibliothek schon aus der Rolle der Universität als Access-Provider heraus in dem Projekt OA-Statistik personenbezogene Daten.
In Zusammenhang mit der IP-Adresse sind alle weiteren protokollierten Daten ebenfalls personenbezogen.

3. Indem die Universitätsbibliothek einen Webserver bzw. Dokumentenserver betreibt, hält sie eigene bzw. fremde Telemedien zur Nutzung bereit bzw. vermittelt den Zugang zur Nutzung dieser Telemedien. Die Universität Stuttgart ist damit Diensteanbieter i. S. d. § 2 Nr. 1 TMG. Für die Verarbeitung personenbezogener Daten gelten deshalb die §§ 11 ff. TMG. Das Landesdatenschutzgesetz gilt hingegen, soweit der Dienst im Dienst- und Arbeitsverhältnis zu ausschließlich beruflichen oder dienstlichen Zwecken bereitgestellt wird (§ 11 Abs. 1 Nr. 1 TMG, § 1 Abs. 1, 5 LDSG).

[1] So statt vieler auch Artikel-29-Datenschutzgruppe, Stellungnahme 4/2007 zum Begriff „personenbezogene Daten“, 20.06.07, S. 19 f.
[2] So auch Schmitz in: Hoeren/Sieber, Handbuch Multimedia-Recht, 19. Erg.,Teil 16.4 Rn. 51-52.

Soweit die Benutzungsordnung oder andere Regelungen der Hochschule nicht anordnen, dass der Dienst ausschließlich beruflich oder zu studienbezogenen Zwecken genutzt werden darf, eine private Nutzung insofern zumindest geduldet ist, muss unterstellt werden, dass der Dienst auch zu diesem Zweck bereit gestellt ist. Bei gemischter Nutzung (sowohl privat als auch dienstlich/studienbezogen) ist entscheidend, ob sich das Telemedienangebot getrennt betrachten lässt (Aufspaltung in private und dienstliche/studienbezogene Nutzung) oder ob eine untrennbare Mischnutzung vorliegt, die nach dem Wortlaut („ausschließlich“) für den gesamten Dienst zwingend zur Anwendung der §§ 11 TMG führt. Entscheidend dürfte sein, ob zum Beispiel getrennte Nutzerkonten für die private und dienstliche/studienbezogene Nutzung existieren oder Anfragen über unterschiedliche Eingabemasken abgewickelt werden.
Der fragliche Telemediendienst der Universitätsbibliothek unterfällt demnach den §§ 11 ff. TMG. Die Bibliothek hat keine Regelung erlassen, wonach der Dienst ausschließlich zu beruflichen oder studienbezogenen Zwecken verwendet werden darf. Auch wird bei dem Angebot nicht zwischen privater und dienstlicher/studienbezogener Nutzung unterschieden, zumal die Dokumente auch über Suchmaschinen auffindbar sind und durch jedermann abgerufen werden können. Der Telemediendienst wird also nicht im Dienst- oder Arbeitsverhältnis zu ausschließlich beruflichen oder dienstlichen bzw. studienbezogenen Zwecken bereit gestellt, sondern auch zu privaten Zwecken. Da sich die private Nutzung von der dienstlichen/studienbezogenen Nutzung nicht trennen lässt, gelten die §§ 11 ff. TMG.

4. Die Universität Stuttgart darf personenbezogene Daten für den Betrieb des Dokumentenservers nur erheben und verwenden, soweit das TMG oder eine andere Rechtsvorschrift mit ausdrücklichem Bezug auf Telemedien dies erlaubt oder der Nutzer einwilligt (§ 12 Abs. 1 TMG).

a) Verarbeitung zur Inanspruchnahme der Telemedien

Nach § 15 Abs. 1 TMG darf die Universität Stuttgart personenbezogene Daten des Nutzers erheben und verwenden, soweit es erforderlich ist, die Inanspruchnahme von Telemedien zu ermöglichen. (Eine Verarbeitung zum Zweck der Abrechnung ist von der Universitätsbibliothek nicht vorgesehen.) Nutzungsdaten sind z. B. Merkmale zur Identifikation des Nutzers, Angaben über Beginn und Ende sowie des Umfangs der jeweiligen Nutzung und Angaben über die vom Nutzer in Anspruch genommenen Telemedien (§ 15 Abs. 1 TMG). Entscheidend ist, ob diese Daten zur Diensterbringung notwendigerweise verarbeitet werden müssen.

Um Rechneranfragen an den Dokumentenserver überhaupt beantworten zu können, muss der Server die IP-Adresse des Nutzers erheben. Ebenfalls müssen der Dateipfad

(URL) und der Dateiname des abgerufenen Dokuments sowie etwaige Get-Parameter verarbeitet werden, um die Anfrage beantworten zu können, ebenso wie das eingesetzte Protokoll und der Status. Das Datum, und die Uhrzeit können erforderlich sein, um ggf. den Zeitablauf (timeout) einer Anfrage ermitteln zu können.
Informationen über die Größe der Datei sowie den Texttyp sind dagegen nicht erforderlich für die technische Abwicklung der Nutzeranfragen. Dasselbe gilt für Informationen über den vom Nutzer eingesetzten Webbrowser (User-Agent) sowie die verweisende Verknüpfung (Referrer); es handelt sich lediglich um Zusatzinformationen. Die Erhebung dieser Daten bedarf deshalb der ausdrücklichen Einwilligung des Nutzers; andernfalls ist sie unzulässig.

b) Weitere Verwendung der Daten

aa) Auch die Verwendung der zulässigerweise erhobenen Daten ist nur zulässig, soweit das TMG oder eine andere Rechsvorschrift unter Bezugnahme auf Telemedien dies gestattet oder der Nutzer einwilligt (§ 12 Abs. 1 TMG). § 15 Abs. 1 TMG gestattet eine Verwendung der Nutzerdaten, soweit es erforderlich ist, die Inanspruchnahme der Telemedien zu ermöglichen.
Wie aus den Protokolldatensätzen ersichtlich wird nach der Erhebung der IP-Adresse deren Domainname ermittelt. Aus dieser Angabe wird offenbar im Anschluss die Länderkennung ermittelt, um festzustellen, aus welchem Land der Zugriff erfolgte. Diese Schritte (Ermittlung des Domainnamens, Feststellung der Länderkennung) werden mithilfe der erhobenen IP-Adresse noch vor deren Ersetzung durch einen Hash-Wert durchgeführt. Bereits diese Schritte stellen eine Verwendung der IP-Adresse dar. Sie sind jedoch nicht erforderlich, um die Inanspruchnahme des Dokumentenservers zu ermöglichen. Nach den §§ 12 Abs. 1, 15 Abs. 1 TMG bedürfen sie deshalb der ausdrücklichen Einwilligung der Nutzer. In diesem Sinne haben auch die obersten Aufsichtsbehörden für den Datenschutz im nicht-öffentlichen Bereich am 26./27. November zur datenschutzkonformen Ausgestaltung von Analyseverfahren zur Reichweitenmessung bei Internetangeboten festgestellt: „Die Analyse des Nutzungsverhaltens unter Verwendung vollständiger IP-Adressen **(einschließlich der Geolokalisierung)** ist aufgrund der Personenbeziehbarkeit dieser Daten daher nur mit bewusster, eindeutiger Einwilligung zulässig. Liegt eine solche Einwilligung nicht vor, ist die IP-Adresse vor jeglicher Auswertung so zu kürzen, dass eine Personenbeziehbarkeit ausgeschlossen ist“ (Hervorhebung vom Verf.) .[3]

bb) Die Universitätsbibliothek darf die zulässigerweise (s. dazu oben) erhobenen Daten nur für Zwecke verwenden, für die sie erhoben wurden, also für die Abwicklung der Rechneranfragen. Für andere Zwecke dürfen die Daten nur verwendet werden, soweit

[3] http://www.lfd.m-v.de/dschutz/beschlue/Analyse.pdf

das TMG oder eine andere Rechtsvorschrift mit ausdrücklichem Bezug auf Telemedien dies erlaubt oder der Nutzer einwilligt (§ 12 Abs. 2 TMG). Die weitere Verwendung der erhobenen Nutzerdaten im Rahmen des Projekts OA-Statistik stellt eine zweckändernde Nutzung dar, weil sie über die bloße Diensterbringung hinausgeht und eine andere Zielrichtung verfolgt.

Eine zweckändernde Nutzung gestattet § 15 Abs. 3 TMG. Demnach dürfen die erhobenen Daten u. a. „zur bedarfsgerechten Gestaltung der Telemedien“ verwendet werden, indem aus den Daten Nutzungsprofile erstellt werden. Jedoch müssen die personenbezogenen Daten dabei durch Pseudonyme ersetzt werden, d. h. durch Kennzeichen, die die Identifizierung des Betroffenen ausschließen oder wesentlich erschweren (vgl. § 3 Abs. 6 a BDSG[4]). Das ist der Fall, wenn die personenbezogenen Daten so verändert werden, dass sie – ohne Kenntnis von der Zuordnungsregel – nur mit unverhältnismäßigem Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren (natürlichen) Person zugeordnet werden können. Die obersten Aufsichtsbehörden für den Datenschutz im nicht-öffentlichen Bereich haben ausdrücklich festgestellt, dass IP-Adressen keine Pseudonyme im Sinne des Telemediengesetzes sind.[5]

Der Dokumentenserver soll laufend alle Zugriffe auf die gehosteten Dokumente erfassen. Täglich sollen diese auf einen „Logfile-Server“ kopiert und in einem weiteren Schritt ebenfalls täglich in eine Datenbank kopiert werden. Es ist beabsichtigt, die erhobene IP-Adresse beim Einlesen in die Datenbank durch einen „One-Way-Hash“ zu ersetzen, einer Prüfsumme, die sich aus sich heraus nicht in den Ursprungswert zurückrechnen lässt. Eine IP-Adresse ergibt dabei stets denselben Hash-Wert. Damit die Ersetzung der IP-Adresse durch einen Hash-Wert den Anforderungen des TMG genügt, müsste der Hash-Wert ein Pseudonym i. S. d. TMG sein. Das setzt voraus, dass der Hash-Wert einen Rückschluss auf die ursprünglichen IP-Adresse ausschließt oder zumindest wesentlich erschwert.
Ein One-Way-Hash wird nicht aufgrund einer geheimen Zuordnungsregel gebildet, sondern aufgrund einer allgemein bekannten mathematischen Rechenfunktion. Damit ist es jedermann möglich, aus einer beliebigen IP-Adresse den zugehörigen Hash-Wert zu reproduzieren. Einem One-Way-Hash fehlt insofern ein wichtiges Pseudonym-Merkmal: die geheime Zuordnungsregel.
Eine solche geheime Zuordnungsregel ergibt sich auch nicht daraus, dass ein Hash-Wert aus sich heraus nicht in den Ursprungswert zurückgerechnet werden kann. Derzeit wird fast ausschließlich IPv4 verwendet, ein Internetprotokoll, bei dem IP-Adressen eine Länge von 4 Oktetts haben. Mit IPv4 können rechnerisch maximal 2^{32} (= 4.294.967.296) IP-Adressen weltweit gebildet werden. Da jede IP-Adresse stets den-

[4] Vgl. auf landesrechtlicher Ebene § 3 Abs. 7 LDSG.
[5] http://www.lfd.m-v.de/dschutz/beschlue/Analyse.pdf

selben Hash-Wert ergibt und die Rechenmethode allgemein bekannt ist, kann jede beliebige Person mit heute verfügbaren Rechnerkapazitäten sehr leicht und schnell aus allen weltweit existierenden IP-Adressen die zugehörigen Hash-Werte errechnen. Ist eine solche Zuordnung (Rainbow-Table) erst erstellt, kann ohne Weiteres aus einem Hash-Wert auf die ursprüngliche IP-Adresse geschlossen werden. Dafür ist es nicht erforderlich, die Hash-Werte aller knapp 4,3 Milliarden IP-Adressen zu berechnen. Ca. 14 % des IP-Adressbereichs scheiden für Rechneranfragen an den Dokumentenserver sogar aus, weil es sich um Sonderbereiche gemäß „RFC 3330“[6] und „RFC 5735“ handelt (z. B. IP-Adressen 10.0.0.0 bis 10.255.255.255 reserviert für Netzwerke für den privaten Gebrauch). Sofern man nicht auf bestehende Rainbow-Tables zurückgreift, würde die Berechung von 3,6 Milliarden IP-Adressen vollkommen ausreichend sein und ca. 11 Stunden dauern.
Dies führt insgesamt dazu, dass jedermann ohne unverhältnismäßigen Aufwand an Zeit, Kosten und Arbeitskraft von einem Hash-Wert auf die ursprüngliche IP-Adresse schließen kann. Die vollständige Berechung der IP-Hash-Paare der Universität Stuttgart dauert nur 3 Sekunden.
Für die rechtliche Bewertung spielt es keine Rolle, ob eine solche Reidentifikation beabsichtigt ist. Auch ist die Erstellung einer solchen Zuordnungsliste gesetzlich nicht verboten. Eine ggf. vertragliche Verpflichtung der Projektpartner, von dieser Möglichkeit keinen Gebrauch zu machen, ändert nichts an dem grundsätzlichen Personenbezug. An dieser Bewertung ändert auch § 13 Abs. 4 Nr. 6 TMG nichts, der es dem Diensteanbieter verbietet, die Nutzungsprofile mit Angaben zur Identifikation des Trägers des Pseudonyms zusammenzuführen. Denn diese Verpflichtung trifft allein den Diensteanbieter, nicht aber den zentralen Serviceprovider oder Dritte, denen eine solche Reidentifikation nicht verboten ist, so dass der Hash-Wert als Pseudonym ungeeignet ist. Damit führt die Ersetzung einer IP-Adresse durch einen One-Way-Hash nicht zu einer Pseudonymisierung, sondern lediglich zur Ersetzung eines personenbezogenen Merkmals durch ein anderes. Die von § 15 Abs. 3 TMG geforderte Pseudonymisierung der Nutzerdaten ist auf diese Weise nicht erfüllt. Unter diesen Voraussetzungen dürfen folglich keine Nutzungsprofile aus den Protokolldaten erstellt werden.

Eine Verwendung der zulässigerweise erhobenen Nutzungsdaten ist deshalb nur zulässig, wenn die IP-Adressen durch ein Merkmal ersetzt werden, das eine Bestimmung der ursprünglichen IP-Adresse wesentlich erschwert. Da sich die Größe des fraglichen IP-Adressraums nicht erweitern lässt, kann eine sinnvolle Pseudonymisierung nur dadurch geschehen, dass eine nicht allgemein bekannte, d. h. geheime Zuordnungsregel genutzt wird.

[6] http://www.zendas.de/technik/texte/rfc3330.txt

Eine andere Möglichkeit bestünde in der Anonymisierung der IP-Adressen. Die Adressen müssten in diesem Fall entweder gelöscht oder durch ein Merkmal ersetzt werden, das auch der Universität Stuttgart keinen Rückschluss mehr auf die IP-Adresse zulässt oder dies nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft möglich ist (vgl. § 3 Abs. 6 BDSG[7]).

c) Ein Personenbezug kann sich auch aus den weiteren Protokolldaten ergeben. Zu denken ist hier insbesondere an den Zeitstempel. Da Zeitstempel oft auch an anderer Stelle im Netzwerk gespeichert werden (z. B. zur Störungserkennung und –beseitigung) und dort wiederum mit anderen personenbezogenen Daten verarbeitet werden (z. B. IP-Adresse), können über den Zeitstempel die Protokolldaten für das Projekt OA-Statistik u. U. mit anderen personenbezogenen Daten verknüpft werden. Entscheidend, ob ein Personenbezug besteht, sind sämtliche an der Universität Stuttgart gespeicherten Protokolldaten, nicht nur von den Servern der Bibliothek gespeicherte Daten. Denn maßgeblich sind die Identifizierungsmöglichkeiten der verantwortlichen Stelle. Das ist die gesamte Universität Stuttgart (vgl. § 3 Abs. 7 BDSG[8]).
Da § 15 Abs. 3 TMG eine Anonymisierung durch den Diensteanbieter nicht verlangt, steht eine solche Verknüpfbarkeit mit anderen Daten der Erstellung von Nutzungsprofilen nicht entgegen. Bedeutung erlangt dies jedoch, wenn beabsichtigt wird, die Protokolldaten zu anonymisieren. In diesem Fall müsste das verbindende Merkmal (z. B. Zeitstempel) gelöscht werden, um einen Personenbezug auszuschließen.

d) Zulässigkeit von Nutzungsprofilen

Nutzungsprofile dürfen nach § 15 Abs. 3 TMG ausschließlich für drei Zwecke erstellt werden: Werbung, Marktforschung und bedarfsgerechte Gestaltung der Telemedien. In Betracht kommt hier allenfalls eine bedarfsgerechte Gestaltung der Telemedien. Darunter dürften alle Maßnahmen fallen, die aus Sicht des Nutzers der Optimierung des Diensteangebots dienen.
Es bestehen große Zweifel, ob sich die im Projektantrag angegebenen Verwendungszwecke unter die bedarfsgerechte Gestaltung der Telemedien subsumieren lassen. Ein Problem ergibt sich insbesondere daraus, dass die Verwendungszwecke nicht hinreichend klar und abschließend festgelegt sind. Die Zweckbestimmung der Datenverarbeitung gibt jedoch den Rahmen für die zulässige Datenverarbeitung vor und muss deshalb vor Durchführung des Projekts so konkret wie möglich erfolgen.
So ist in dem Projektantrag an die DFG davon die Rede, verlässliche Nutzungsdaten seien eine wichtige Dienstleistung für Autoren, die die absoluten Nutzungszahlen ihrer Publikationen und den verhältnismäßigen Stand im Wettbewerb um Aufmerksamkeit

[7] Vgl. auf landesrechtlicher Ebene § 3 Abs. 6 LDSG.
[8] Vgl. auf landesrechtlicher Ebene § 3 Abs. 3 LDSG.

interessieren würden. Eine Information der Autoren stellt jedoch an sich noch keine Maßnahme zur bedarfsgerechten Gestaltung des Dokumentenservers dar.
Weiter wird ausgeführt, dass sich aus den Nutzungszahlen ermessen lasse, welche Bedeutung bzw. welchen Einfluss („Impact“) ein Dokument besitzt. Darüber hinaus werde so die Akzeptanz von Dokumentenservern und Open Access „unterstützt“, weil nachgewiesen werden könne, dass frei zugängliche Dokumente von einem internationalen Publikum rezipiert würden. Auch dies sind noch keine Maßnahmen zur bedarfsgerechten Gestaltung des Dokumentenservers. Dasselbe gilt, wenn das Projekt mit der Entwicklung und Etablierung eines einheitlichen Standards zur Ermittlung von Zugriffszahlen und Statsistiken begründet wird.

Anders dagegen, wenn in dem Projektantrag ausgeführt wird, dass der „zentrale Serviceprovider“ Statistikdaten an die Universitätsbibiothek zurückmeldet, damit diese bei zukünftigen Suchanfragen in den Trefferlisten bspw. als Sortierfunktion angezeigt werden, um so dem Nutzer Auskunft zu geben, welche Relevanz das gefundene Dokument hat bzw. haben könnte. Diese Zusatzfunktion dient der Verbesserung des Dienstes für die Nutzer; die Erstellung von pseudonymen Nutzungsprofilen für diese Zusatzfunktion dient insofern dem Zweck der bedarfsgerechten Gestaltung der Telemedien. Die für diesen Zweck erstellten (anonymen) Statistikdaten dürfen dann auch für andere Zwecke verwendet werden (z. B. als Mehrwertdienst für Autoren). Das gilt jedoch nur, wenn dieser andere Zweck mit den dazu bereits vorliegenden Statistikdaten erfüllt werden kann. Etwas anderes gilt aber, wenn für den Mehrwertdienst dagegen andere Nutzungsprofile erstellt oder andere Auswertungen anhand der Nutzungsdaten durchgeführt werden müssen. Dies wäre nicht zulässig, da die Nutzungsprofile für andere als in § 15 Abs. 3 TMG benannte Zwecke (also insb. bedarfsgerechte Gestaltung der Telemedien) nicht verwendet werden dürfen (also z. B. wenn lediglich „ein attraktiver Dienst für Wissenschaftler, Hochschulen und Forschungseinrichtungen sowie Förderinstitutionen geschaffen“ werden soll, „der standardisiert und verlässlich aufzeigt, wie wissenschaftliche Publikationen genutzt werden und welche Auswirkungen freie Zugänglichkeit auf die Nutzung und Sichtbarkeit hat.“, DFG-Antrag S. 18).
Aus datenschutzrechtlicher Sicht ist zu fordern, dass die Zwecke, für die die Nutzungsdaten verwendet werden sollen, konkret und abschließend benannt werden. Es müsste dargelegt werden, inwiefern die Erstellung der Nutzungsprofile dem Zweck der bedarfsgerechten Gestaltung der Telemedien dient.

Darüber hinaus ist darauf zu achten, dass nur solche Daten zur Erstellung von Nutzerprofilen verwendet und deshalb in einer Protokolldatei gespeichert werden dürfen, die für die bedarfsgerechte Gestaltung der Telemedien benötigt werden. Fraglich ist insbesondere, ob Angaben wie z. B. Zeitstempel (Datum, Uhrzeit, Zeitzone) für eine solche bedarfsgerechte Ausgestaltung verarbeitet werden müssen. Hier wäre genau zu be-

gründen, inwiefern jedes einzelne der gespeicherten Merkmale für eine bedarfsgerechte Ausgestaltung benötigt wird.

In diesem Zusammenhang ist darauf hinzuweisen, dass die Universität Stuttgart durch technische und organisatorische Maßnahmen sicherstellen muss, dass die erstellten Nutzungsprofile nicht mit Angaben zur Identifikation des Trägers des Pseudonyms zusammengeführt werden können. Eine Orientierungshilfe, welche Maßnahmen in Betracht kommen, bietet § 9 Abs. 3 LDSG.[9]

e) Widerspruchsrecht

Werden Nutzungsprofile erstellt, haben die Nutzer ein Widerspruchsrecht (§ 15 Abs. 3 S. 1 TMG), wenn nicht ohnehin eine Einwilligung der Nutzer erforderlich ist (siehe dazu oben). Auf dieses Widerspruchsrecht müssen sie **zu Beginn** der Nutzung des Dokumentenservers hingewiesen werden (§ 15 Abs. 3 S. 2 TMG). In diesem Zusammenhang ist darauf hinzuweisen, dass die Nutzer gleichzeitig auch über Art, Umfang und Zweck der Erhebung und Verwendung ihrer personenbezogenen Daten in allgemein verständlicher Form unterrichtet werden müssen, sofern eine Unterrichtung nicht bereits anderweitig erfolgt ist (§ 13 Abs. 1 TMG). Der Inhalt der Unterrichtung muss für die Nutzer jederzeit abrufbar sein.
Die Möglichkeit des Widerspruchs muss im Verfahren vorgesehen sein. Widerspricht ein Nutzer, dürfen seine Daten nicht für die Erstellung eines Nutzerprofils verwendet werden. Vielmehr sind sie nach dem Ende des Nutzungsvorgangs zu löschen. Die Daten dürfen folglich keine Verwendung im Projekt OA-Statistik finden und zwar auch dann nicht, wenn sie hierfür „frühestmöglich“ pseudonymisiert werden.
Sieht das Verfahren eine unverzügliche Löschung von Daten der widersprechenden Nutzer nicht vor und können diese auch nicht in der Protokolldatei ausgefiltert werden, müssen sämtliche Protokolldaten unverzüglich gelöscht werden, wenn ein Nutzer widerspricht, weil sie dann zumindest teilweise rechtswidrig verwendet werden.
Eine mangelnde Implementation einer Widerspruchsmöglichkeit birgt also die Gefahr, dass sämtliche Daten gelöscht werden müssen.

e) Auskunftsrecht
Des Weiteren steht den Nutzern ein Auskunftsrecht über die zu ihrer Person gespeicherten Daten zu. Auf Verlangen ist ihnen deshalb mitzuteilen, welche Daten zu ihnen bzw. zu ihrem Pseudonym gespeichert sind (§ 13 Abs. 7 TMG). Auch dies muss bei der Verfahrensgestaltung berücksichtigt werden.

f) Datenweitergabe an zentralen Serviceprovider

[9] http://www.zendas.de/recht/texte/ldsg/p9.html

Die Weitergabe der Daten an einen Projektpartner, der als zentrale Speicherstelle fungiert, ist nach der bisherigen Verfahrensgestaltung nicht zulässig, da bereits die Erhebung und Verwendung der Protokolldaten an der Universität Stuttgart in derzeitiger Form nicht zulässig ist.
Unterstellt, die Daten würden bei der Universität Stuttgart ausreichend pseudonymisiert oder gar anonymisiert, dürften die Daten an einen zentralen Serviceprovider weitergegeben werden, wenn dies der bedarfsgerechten Gestaltung des Dokumentenservers der Universitätsbibliothek dient (§ 12 Abs. 3, 15 Abs. 3 TMG). Für den Serviceprovider handelt es sich in diesem Fall um anonyme Daten, so dass für ihn weder die Anforderungen der §§ 11 ff. TMG noch der Landesdatenschutzgesetze gelten; in diesem Fall liegt insbesondere keine Auftragsdatenverarbeitung vor.

5. Alternativen/Fazit

Keine besonderen Anforderungen an den gesamten Prozess bestehen dann, wenn die zulässigerweise erhobenen Daten (siehe dazu 4. a)) unverzüglich anonymisiert werden. Hierfür müssten alle Merkmale, die zu einem Personenbezug führen, gelöscht werden. Das betrifft insbesondere die IP-Adresse. Diese könnten durch Kürzung um das letzte Oktett anonymisiert werden, bevor eine weitere Verwendung stattfindet (also noch vor der Ermittlung des Domainnamens bzw. der Länderkennung).
Eine weitere Möglichkeit wäre die Verarbeitung mit Einwilligung der Betroffenen. Die Einwilligung kann elektronisch eingeholt werden (z. B. Anklicken eines Schalters). In diesem Fall darf von den Vorgaben des TMG abgewichen werden, wenn die Nutzer hierin einwilligen. So müsste nicht zwingend eine sofortige Anonymisierung oder Pseudonymisierung erfolgen, die Daten dürften auch für Mehrwertdienste genutzt werden, die selbst nicht unmittelbar eine bedarfsgerechte Gestaltung der Telemedien-dienste bezwecken, und es könnte eine Geolokalisierung durchgeführt werden.

Soll weder von einer Anonymisierung Gebrauch gemacht werden noch eine Einwilligung der Nutzer eingeholt werden, so muss das Verfahren folgendermaßen geändert werden:

- Erhebung nur der zulässigen Daten (siehe 4. a))
- Ermittlung des Länderkürzels nur anhand zumindest pseudonymisierter IP-Adresse, wenn dieses für die bedarfsgerechte Gestaltung des Dokumentenservers benötigt wird[10] (siehe 4. b) aa), bb))
- (Wirksame) Pseudonymisierung aller Nutzerdaten (insb. IP-Adressen) noch vor der weiteren Verwendung (siehe 4. b) bb))

[10] Strenger dagegen der Beschluss der obersten Aufsichtsbehörden, wonach eine Geolokalisierung ohne Einwilligung der Nutzer offenbar nur mit anonymisierten Daten für zulässig gehalten wird.

- Festlegung konkreter Verwendungszwecke, die der bedarfsgerechten Gestaltung der Telemedien dienen; keine Vewendung der Nutzungsprofile für andere Zwecke ohne Einwilligung der Nutzer (siehe 4. d))
- Einbeziehung nur solcher Daten in die Nutzerprofile/Protokolldateien, die für die bedarfsgerechte Gestaltung des Dokumentenservers benötigt werden (siehe 4. d))
- Berücksichtigung des Widerspruchsrechts der Betroffenen sowie Unterrichtung der Nutzer (siehe 4. e))

01.04.2010