

# Anonyme Audit-Daten im Überblick

## Architekturen für anonyme Autorisierungen und Audit-Daten

Ulrich Flegel

*Audit-Daten sind notwendig, um Missbrauch zu erkennen. Die datenschutzfreundliche Verarbeitung lässt sich durch Anonymisierung bzw. Pseudonymisierung erreichen. Anhand von Modellen werden wünschenswerte Eigenschaften der Anonymisierung erläutert. Bekannte Ansätze für die nachträgliche Anonymisierung von Audit-Daten werden kurz vorgestellt.*

*Die beschriebenen Arbeiten werden derzeit zum Teil von der Deutschen Forschungsgemeinschaft gefördert unter Bi 311/10-2.*



Dipl.-Inform.  
Ulrich Flegel

Wissenschaftlicher  
Mitarbeiter der  
Arbeitsgruppe  
Informationssysteme  
und Sicherheit (ISSI)

am Fachbereich Informatik der Universität  
Dortmund.

E-Mail: ulrich.flegel@udo.edu

### 1 Ein Besuch im Zoo

Sicherheitsmaßnahmen in der digitalen Welt sind meist bereits vorhandenen Sicherheitsmaßnahmen der realen Welt nachempfunden. Das mag daran liegen, dass Vertrauen letztlich stets in der realen Welt begründet ist und Sicherheitsmaßnahmen gerade bei fehlendem Vertrauen der Akteure notwendig sind. Wie wir in der realen Welt mit Vertrauen umgehen, lässt sich am Beispiel eines Studierenden zeigen, der den Zoo besuchen möchte. Der Zoo tritt hier als Dienst(leister) auf und bietet Studierenden kostenlosen Eintritt. Nicht-Studierende könnten versuchen, sich einen geldwerten Vorteil zu verschaffen, indem sie sich an der Zoo-Kasse als Studierende vorstellen. Der Studierendenausweis fungiert als *beglaubigte Eigenschaftsaussage*, indem er den Namen des Aussage-Subjekts der Eigenschaft *Studierender* zuordnet. Die Zoo-Kasse akzeptiert diese beglaubigte Eigenschaftsaussage, wenn die vermerkte Universität als Aussteller akzeptiert wird, das Lichtbild zur vorliegenden Person „passt“, der Studierendenausweis noch nicht abgelaufen ist und „echt“ aussieht.

Wenn die Zoo-Kasse den Studierendenausweis akzeptiert, autorisiert sie die vorliegende Person, den Zoo-Eingang zu passieren. Die Zoo-Kasse stellt eine Autorisierung in Form eines Eintrittstickets aus. Diese *Autorisierung* enthält eine dem Kunden zugeordnete Ticket-Nummer, es ist vermerkt, dass die Autorisierung zum Zoo-Eintritt berechtigt, von welcher Kasse sie ausgestellt wurde, und sie trägt Gültigkeitsinformationen wie eine Geltungsdauer sowie schwer fälschbare Echtheitsmerkmale.<sup>1</sup> Da das Ticket keine Information zur Authentisierung des Eintrittsberechtigten enthält, ist es prinzipiell übertragbar.

<sup>1</sup> Der Aufwand zur Fälschung der Echtheitsmerkmale übersteigt den Eintrittspreis.

Nach dem Zoo-Eingang springt dem Besucher ein Schild ins Auge, auf dem steht, welches Verhalten im Zoo untersagt ist. Vor allem soll man die Affen nicht ärgern, wohl weil die sich mit Bananenschalen-Geschossen rächen könnten. An kritischen Stellen (bei den Affen) kann der Zoo einen Wächter postieren, der bei entdeckten Regelverstößen sinnvoll reagiert.

### 2 Autorisierungs-Architekturen

In Autorisierungs-Architekturen werden Individuen, Rechner und andere Akteure eines verteilten IT-Systems als *Entitäten* bezeichnet. Ein *Prinzipal* ist ein Bit-String, der in seinem Anwendungsbereich eindeutig genau einer Entität als deren Surrogat zugeordnet ist. Eine Entität kann *Eigenschaften* haben, die in Sicherheitspolitiken als Entscheidungsbedingungen formuliert sind. Der Begriff *Beglaubigung* bezeichnet im Modell den Vorgang und das Resultat, wenn ein verantwortlicher Agent als *Beglaubiger* eine Aussage über Eigenschaften beglaubigt, die entitätsbezogen und nicht dienstbezogen sind. Als Beispiel für die Beglaubigung über die Eigenschaft *Studierender* trat in Abschnitt 1 der Studierendenausweis auf. Der Begriff *Autorisierung* bezeichnet im Modell den Vorgang und das Resultat, wenn ein verantwortlicher Agent als *Autorisierer* eine Aussage über dienstspezifische Erlaubnisse beglaubigt. Als Beispiel für eine Autorisierung *zoo-eintrittsberechtigt* trat in Abschnitt 1 das Eintrittsticket auf.

Im Grundmodell (s. Abb. 1) wird bei der Auswertung beglaubigter Eigenschaftsaussagen durch deren Empfänger zunächst der *verantwortliche Agent* anhand der gleichnamigen Komponente ermittelt. Der Empfänger entscheidet zunächst darüber, ob er dem Agenten hinsichtlich der Prüfung der in der *Attribut*-Komponente beschriebenen Eigenschaften und deren korrekter Zuord-

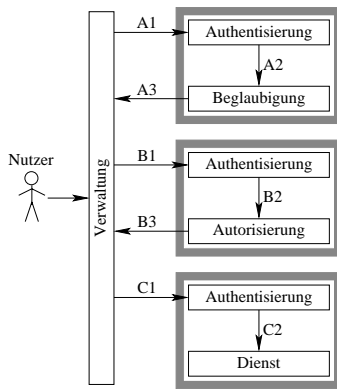


Abb. 1: Idealisierte Autorisierungs-Architektur

nung zum Prinzipal der korrekten Entität vertraut. Anschließend wird anhand der *Validitäts*-Komponente festgestellt, ob die Eigenschaftsaussage gültig ist. Dann prüft der Empfänger mittels der *Authentisierung*-Komponente, ob die vorliegende Entität der *Subjekt*-Komponente entspricht. Schließlich interpretiert der Empfänger die Attribute entsprechend seiner eigenen Politik. Die in Abb. 1 gezeigten Akteure sind die Verwaltung, ein Beglaubiger, ein Autorisierer und ein Dienst. Sie entsprechen z.B. bei *Kerberos* dem Client, dem *Authentication Server*, dem *Ticket Granting Server* und dem Dienst-Server.

### 3 Datenschutzaspekte

Audit-Daten werden auf Vorrat erhoben, gespeichert und analysiert mit dem Ziel, Missbrauch zu entdecken und zwecks Rechtsverfolgung dem Urheber zuzurechnen (Wächter bzw. Intrusion-Detection). Aufgrund der komplexen rechtlichen Situation und den datenschutzrechtlichen Einschränkungen bei der Erhebung, Speicherung und Verarbeitung personenbezogener Daten gestaltet sich der gesetzeskonforme Einsatz von Audit-Daten-gestützten Schutzmaßnahmen wie etwa Intrusion-Detection für viele Dienstanbieter diffizil [4].

Es besteht ein Spannungsfeld zwischen dem Interesse einzelner Nutzer an Datenschutz und Anonymität einerseits und der Zurechenbarkeit andererseits, um im Missbrauchsfall die Interessen anderer beteiligter Parteien schützen zu können. Wie etwa die Diskussion in [5, 6, 7] deutlich macht, kann eine für die beteiligten Parteien zufriedenstellende Lösung nicht darin bestehen, eine der beiden Anforderungen zugunsten der anderen vollständig aufzugeben. Vielmehr scheint ein fairer Ausgleich der Interessen aller beteiligter Parteien unter Berücksichti-

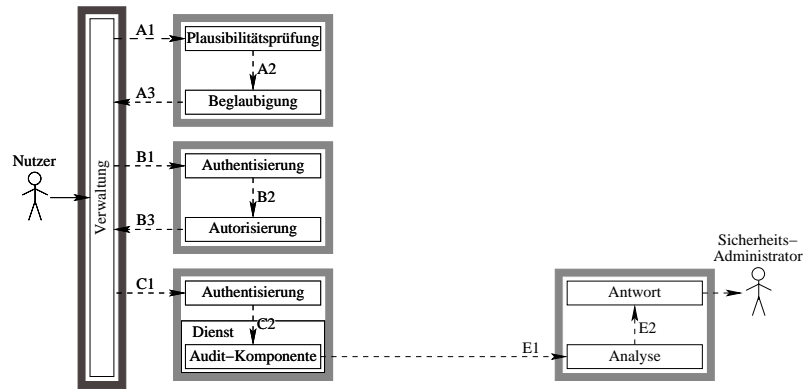


Abb. 2: Einseitig sicher: Anonymität durch die Verwaltung

gung der jeweiligen Anwendungssituation erstrebenswert (*Mehrseitige Sicherheit*).

Der in diesem Zusammenhang zentrale Begriff des *Personenbezugs* ist relativ, da die Personenbeziehbarkeit einer Information vom jeweiligen Zusatzwissen zum jeweiligen Zeitpunkt abhängt. Dementsprechend gelten die Datenschutzgesetze nur für diejenigen Datenverwender, die durch Zusatzwissen den Bezug der Daten zum Betroffenen herstellen können. Somit kann der oben beschriebene Zielkonflikt zwischen Zurechenbarkeit und Anonymität durch die Dienstnutzung unter Pseudonymen fair gelöst werden, indem über die Kontrolle von Zusatzwissen zwischen Regelfall (keine Zurechenbarkeit) und Ausnahmefall (Zurechenbarkeit möglich) unterschieden wird.

Mittels Pseudonymen werden personenbezogene Daten so verändert, dass sie ohne Kenntnis der zugehörigen Zuordnungsregel nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbar natürlichen Person zuordenbar sind, für den Ausnahmefall aber mittels der Zuordnungsregel die Identifizierung der Person ermöglichen [8]. Damit stellen Pseudonyme ein Schlüsselkonzept für den mehrseitig sicheren Umgang mit Audit-Daten dar, ja sie ermöglichen in vielen Umgebungen erst die gesetzeskonforme Erhebung und Speicherung von Audit-Daten. Die *Pseudonymisierung* bezeichnet den Vorgang der Ersetzung der Prinzipale durch Pseudonyme gemäß einer Zuordnungsregel.

### 4 Anonymisierungs-Architekturen

Die Verifizierer beglaubigter Eigenschaftsaussagen benötigen in vielen Fällen keine

Identitäten für ihre Tätigkeit. So können Eigenschaftsaussagen anonymisiert werden, indem der Subjekt-Prinzipal durch ein Pseudonym mit geeigneten Eigenschaften ersetzt wird. Das deutsche Signaturgesetz sieht bereits entsprechende Beglaubigungen vor (§7 Abs. 1-3 SigG) [8]. Entsprechend ist es etwa an der Zoo-Kasse nicht notwendig, den Namen des Studierenden zu erfahren. Wichtig ist nur, dass die Eigenschaft *Studierender* an die Person gebunden ist, die einen gültigen Studierendenausweis vorlegt, und dass dieser von einem vertrauenswürdigen Agenten ausgestellt wurde. Also könnte der Studierendenausweis anonym ausgelegt werden, indem in die Subjekt-Komponente statt des Namens die Matrikelnr. des Studierenden eingetragen würde.

Der Agent ist nun einerseits im Interesse der Zurechenbarkeit den Verwendern der Eigenschaftsaussage gegenüber zusätzlich dafür verantwortlich, dass er entsprechend seiner im Voraus festgelegten Politik zu spezifischen Zwecken gegenüber spezifischen Entitäten mittels der Zuordnungsregel Pseudonyme aufdeckt. Andererseits ist der Agent im Interesse der Anonymität den Subjekten gegenüber dafür verantwortlich, die Zuordnungsregel zu schützen und hinsichtlich der Aufdeckbarkeit und Verkettbarkeit der Pseudonyme seine dem Subjekt bekannte Politik einzuhalten.

Im Folgenden werden auf Abschnitt 2 basierend Architekturen vorgestellt, die Nutzer-Anonymität gegenüber den *Sicherheits-Administratoren* eines Dienstes herstellen, welche Beobachtungen ausschließlich auf der Basis der vom Dienst gelieferten Audit-Daten machen können (s. Abb. 2).

Die Audit-Daten werden von der *Audit-Komponente* des Dienstes erhoben und der *Audit-Analyse* der Sicherheits-Administratoren des Dienstes verfügbar gemacht (E1).

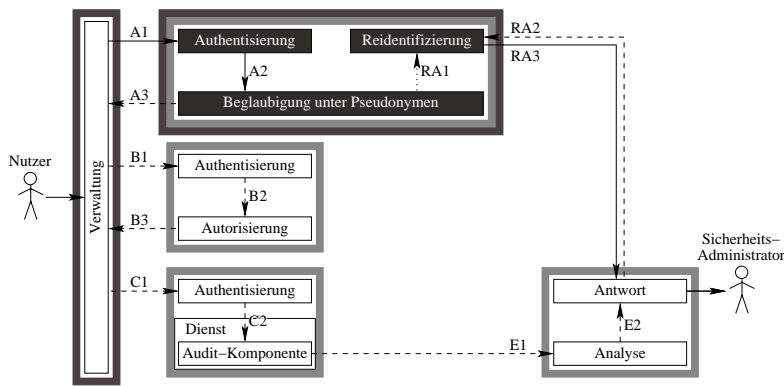


Abb. 3: Mehrseitig sicher: Anonymität durch den Beglaubiger

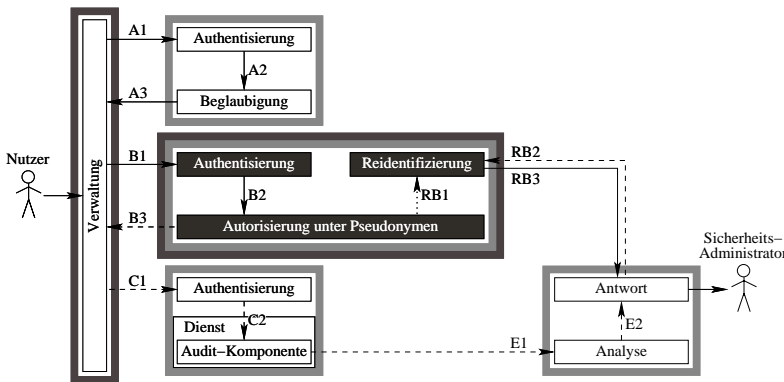


Abb. 4: Mehrseitig sicher: Anonymität durch den Autorisierer

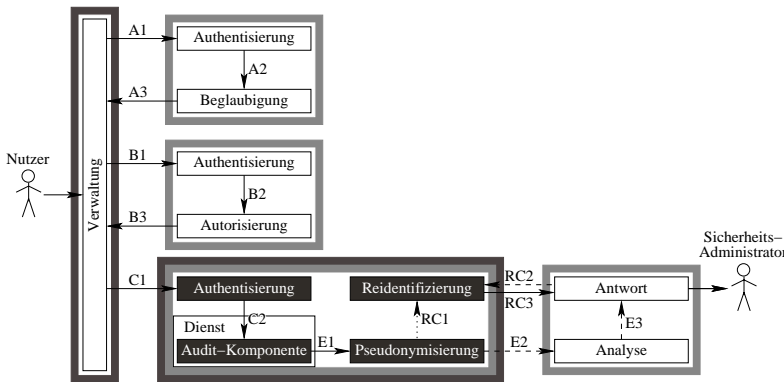


Abb. 5: Mehrseitig sicher: Anonymität durch den Dienst

Diese erzeugt entsprechend des *Analyse-Zwecks Einzelberichte* und sendet sie an die *Antwort-Einheit (E2)*, welche wiederum geeignet auf die Einzelberichte reagiert, z.B. indem sie den Sicherheits-Administrator unterrichtet und ihm Handlungsvorschläge unterbreitet. Eine konkrete Instanz dieses Szenarios wäre ein *Intrusion-Detection-System*, dessen Analyse-Zweck das Entdecken bekannter, durch die Dienstnutzer verursachter *Schutzzielverletzungen* ist.

Die Abb. 2-5 zeigen die dem Modell aus Abb. 1 entsprechenden anonymen Versio-

nen, bei denen eine Entität Anonymität der Dienst-Nutzer gegenüber den Sicherheits-Administratoren herstellt, indem sie ein Nutzer-Pseudonym erstmalig einführt.<sup>2</sup>

<sup>2</sup> Die durchgezogenen Pfeile zeigen die Flussrichtung zurechenbarer und beglaubigter bzw. nachgewiesener Aussagen über Eigenschaften an. Die gestrichelten Pfeile zeigen die Flussrichtung der anonymen und ggf. beglaubigten Aussagen über Eigenschaften an. Die gepunkteten Pfeile zeigen die Flussrichtung der Zuordnungsregel an. Jede fette graue Umrahmung schließt einen Bereich ein, in dem die Interessen einer Entität durchgesetzt werden. Die dunkelgrauen Umrahmungen stehen für das Nutzerinteresse Anonymi-

Da in mehrseitig sicheren Versionen gegenläufige Interessen mehrerer Entitäten berücksichtigt werden sollen, führt dies zum Ausschluss der Kontrolle eben dieser Entitäten über die Interessensobjekte, also die Pseudonyme in den Eigenschaftsaussagen. Entsprechend ist für mehrseitige Sicherheit die Zuordnungsregel von Agenten zu kontrollieren, denen die Interessensträger vertrauen müssen. Diese Situation ist in den Abb. 3-5 dargestellt.

Abb. 6 zeigt am Beispiel eines Dienstes mit einem Pseudonymisierer, wie der Einsatz der technischen Zweckbindung für die geordnete Aufdeckbarkeit die notwendigen Kontrollverhältnisse vereinfacht. Bei der technischen Zweckbindung der Aufdeckbarkeit wird den Pseudonymen die zweckgebunden geschützte Zuordnungsregel beigelegt (E2 unten), so dass diese nicht mehr direkt dem Reidentifizierer übermittelt werden muss (R1 oben). Die Reidentifizierung ist so unumgebar nur noch entsprechend dem Zweck der geordneten Aufdeckung möglich [1]. Demgemäß muss der Nutzer derjenigen Entität, die die Reidentifizierung kontrolliert, nicht mehr vertrauen.

Bei der Umsetzung für anonymes Audit sind die zeitnahe Pseudonym-Aufdeckbarkeit und Praktikabilität, d.h. die Unabhängigkeit von Nutzern und aufwendigen Infrastrukturen, von entscheidender Bedeutung. Eine Analyse der Architekturen hat gezeigt, dass diese Anforderungen gemeinsam nur auf Dienstebene erfüllbar sind, also bei der nachträglichen Anonymisierung von Audit-Daten. Die folgenden Abschnitte geben Beispiele ebensolcher Lösungen und die an die Lösungen zu stellenden Anforderungen.

## 5 Beispiele

Mit dem teilimplementierten Forschungs-System *Intrusion Detection and Avoidance (IDA)* wurde das Konzept der *Intrusion-Detection-Analyse* auf anonymen Audit-Daten eingeführt [9]. Das implementierte Forschungs-System *Adaptive Intrusion Detection (AID)* greift dieses mit IDA eingeführte Konzept auf, wobei sich die Architekturen von AID und IDA stark voneinander unterscheiden [10]. *Lundins Firewall*

tät und die hellgrauen Umrahmungen für das Interesse der Sicherheits-Administratoren an Zurechenbarkeit. Dunkel ausgefüllte Kästen realisieren gemeinsam mehrseitige Sicherheit. Sie befinden sich gerade in den doppelt umrahmten Bereichen, also dort, wo konfligierende Interessen durchgesetzt werden.

*Audit Anonymisierer* ist ein Forschungs-System für die Anonymisierung der Audit-Daten einer spezifischen Proxy-Firewall. Auf den anonymisierten Daten wurden Intrusion-Detection-Experimente durchgeführt [11]. Das frei verfügbare *bsmpseu* anonymisiert Solaris-BSM-Audit-Daten mittels verkettbarer Pseudonyme ohne Möglichkeit zur geordneten Aufdeckung [12]. *Jaegers Anonymisierungs-Konzept* bietet verkettbare Pseudonyme, die nicht geordnet aufgedeckt werden können. Sie können aber zur Bestätigung eines konkreten Verdachts hinsichtlich einer Identität dienen [4]. Das kommerzielle Content-Filter-System *WebWasher* kann seine Audit-Daten bzw. Berichte anonymisieren. Zur Funktionsweise der Anonymisierung ist nur bekannt, dass die organisatorische Zweckbindung bei der geordneten Aufdeckung das 4-Augen-Prinzip anwendet [13]. Der *Anonymouse Log File Anonymizer* anonymisiert Web-Server-Audit-Daten. Dabei werden nur die Top-Level-Domains der Nutzer-Adressen beibehalten. Eine geordnete Aufdeckung ist nicht möglich [14]. Das frei verfügbare und portierbare Forschungs-System *Pseudonymization with Conditional Reidentification (Pseudo/CoRe)* anonymisiert Audit-Daten im Sinne mehrseitiger Sicherheit. Dabei unterliegen die Pseudonym-Nutzungskontexte, der Pseudonym-Wechsel und die geordnete Pseudonym-Aufdeckung der technischen Zweckbindung [1, 2, 3].

## 6 Anforderungen an anonymes Audit

Das nachträgliche Pseudonymisieren von Audit-Daten erzielt eine vergleichbare Wirkung wie die Dienstenutzung mittels pseudonymer Autorisierungen [8]. Allerdings stellt die spezifische Anwendungssituation andere Anforderungen an das Konzept und die Implementierung der Pseudonym-Erzeugung. Die erste Anforderung betrifft die Performanz der Pseudonym-Erzeugung. Je nach Sorte des Dienstes, der die Audit-Daten erzeugt, kann ein extrem hohes Aufkommen zu bewältigen sein, insbesondere beim Dienst Betriebssystem, wenn es für Intrusion-Detection Systemrufe als Audit-Datensätze speichert. Die Pseudonym-Erzeugung findet idealerweise on-the-fly statt und sollte daher einen dem Datenaufkommen angemessenen Durchsatz erreichen. Im Idealfall findet die Anonymisie-

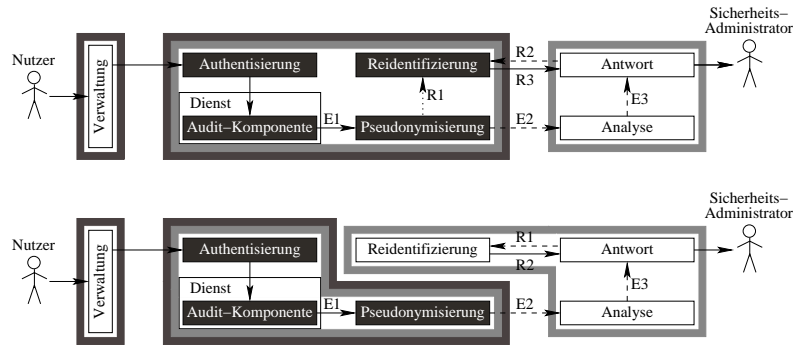


Abb. 6: Zweckbindung der geordneten Aufdeckbarkeit, organisatorisch vs. technisch

rung von Audit-Daten dort statt, wo diese Daten erhoben werden, nämlich auf dem Gerät, das die Nutzeranfragen zur Dienstleistung verarbeitet. Damit der Dienst nicht ausgebremst wird, ist es wichtig, dass die Pseudonym-Erzeugung nicht den überwiegenden Teil der Ressourcen bindet. Aufwändige Kryptoverfahren für die Pseudonym-Erzeugung scheiden daher für den dienst-lokalen on-the-fly-Einsatz aus. Die zweite Anforderung betrifft den Verwendungszweck der pseudonymisierten Audit-Daten. Erfordert der Zweck eine rasche geordnete Aufdeckbarkeit, lässt sich dies nur mittels technischer Zweckbindung erreichen.

## Fazit

Die gesetzlichen Einschränkungen bei Nutzung von Audit-Daten entfallen, wenn die Audit-Daten anonym vorliegen. Beim Dienst lässt sich dies praktikabel durch Audit-Daten-Anonymisierer erreichen. Bei dem Entwurf und bei der Auswahl von Audit-Daten-Anonymisierern ist besonders auf die notwendigen Kontrollverhältnisse zur und die Mechanismen für die Durchsetzung der Zweckbindung bei der Pseudonym-Aufdeckung zu achten.

## Literatur

[1] Ulrich Flegel/Joachim Biskup. Ausgleich von Datenschutz und Überwachung mit technischer Zweckbindung am Beispiel eines Pseudonymisierers. In S. Schubert/B. Reusch/N. Jesse (Hg.), *Informatik bewegt (Informatik 2002)*, LNI P-19, 488-494, Dortmund, Oktober 2002. Köllen Verlag.  
 [2] Ulrich Flegel. Pseudonymizing Unix log files. In G. Davida/Y. Frankel/O. Rees (Hg.), *Proceedings of the Infrastructure Security Conference (InfraSec2002)*, LNCS 2437, 162-179, Bristol, England, Oktober 2002. Springer.

[3] Ulrich Flegel. Praktikabler Datenschutz für Log-Daten. Im Tagungsband zum *10. DFN-CERT-Workshop über Sicherheit in vernetzten Systemen*, F1-20, Hamburg, Februar 2003. Books on Demand.  
 [4] Stefan Jaeger. Verbotene Protokolle. *Zeitschrift für Kommunikations- und EDV-Sicherheit (KES)*, 2000(5):6-12, 2000.  
 [5] Herbert Fiedler. Der Staat im Cyberspace. *Informatik Spektrum*, 24(5):309-314, 2001.  
 [6] Alexander Roßnagel. Freiheit im Cyberspace. *Informatik Spektrum*, 25(1):33-38, 2002.  
 [7] Herbert Fiedler. Cyber-libertär. *Informatik Spektrum*, 25(3):215-219, 2002.  
 [8] Alexander Roßnagel/Philip Scholz. Datenschutz durch Anonymität und Pseudonymität. *MultiMediaRecht (MMR)*, 2000(12):721-732, 2000.  
 [9] Simone Fischer-Hübner. *IDA (Intrusion Detection and Avoidance System): Ein einbruchsentdeckendes und einbruchvermeidendes System*. Reihe Informatik. Shaker, 1993.  
 [10] Michael Meier/Thomas Holz. Sicheres Schlüsselmanagement für verteilte Intrusion-Detection-Systeme. In P. Horster (Hg.), *Systemsicherheit*, DuD-Fachbeiträge, 275-286, Bremen, 2000. Vieweg.  
 [11] Emilie Lundin/Erland Jonsson. Anomaly-based intrusion detection: Privacy concerns and other problems. *Computer Networks*, 34(4):623-640, Oktober 2000.  
 [12] Konrad Rieck. Konzept zur datenschutzorientierten Verarbeitung von Solaris-BSM-Audit-Daten. Freie Universität Berlin, Januar 2003.  
 [13] webwasher.com AG. Den Überblick behalten, Reporting mit WebWasherEE. Januar 2003.  
 [14] Claudia Eckert/Alexander Pircher. Internet anonymity: Problems and solutions. In M. Dupuy/P. Paradinas (Hg.), *Proceedings of the IFIP TC11 16th International Conference on Information Security (IFIP/Sec'01)*, 35-50, Paris, Juni 2001. Kluwer Academic Publishers.