



# Ein Grid-basiertes, föderiertes Intrusion Detection System zur Sicherung der D-Grid Infrastruktur (GIDS)

*Datenschutzmodell für ein Grid-basiertes IDS (MS 13)*
*Meilenstein zum Abschluss des Arbeitspakets 3*

*Autoren:*
Dr. Wolfgang Hommel (Leibniz-Rechenzentrum)
Dr. Nils gentschen Felde (Ludwig-Maximilians-Universität München)
Felix von Eye (Leibniz-Rechenzentrum)
Jan Kohlrausch (DFN-CERT GmbH)
Christian Szongott (Regionales Rechenzentrum für Niedersachsen)

GEFÖRDERT VOM

# Inhaltsverzeichnis

# Kapitel 1

# Einleitung

Dieses Dokument präsentiert als Ergebnis des Arbeitspakets 3 des Projekts „Ein Grid-basiertes, föderiertes Intrusion Detection System zur Sicherung der D-Grid Infrastruktur“ (GIDS) ein Datenschutzkonzept für die Erhebung, Verarbeitung und den Transport der Daten innerhalb des Grid-basierten Intrusion Detection Systems (GIDS). GIDS (`http://www.grid-ids.de`) ist ein Teilprojekt im Rahmen des D-Grid (`http://www.d-grid.de`) und wird vom Bundesministerium für Bildung und Forschung (BMBF, `http://www.bmbf.de`) gefördert. Weitere Projektinformationen und Unterlagen können der Projekt-Webseite entnommen werden.

## 1.1 Projektbeschreibung

### 1.1.1 Problemstellung

Im Umfeld von Grids ergeben sich im Vergleich zu konventionellen vernetzten Systemen eine Reihe bisher ungelöster Probleme, die es im Falle des D-Grid zu bewältigen gilt. So begegnet man im Grid-Kontext unter anderem einem sehr dynamischen Umfeld. Die hohe Dynamik an verfügbaren Ressourcen, wie auch dynamische Benutzergruppen, dessen Benutzer in virtuellen Organisationen (VOs) zusammengefasst werden, erfordern individuelle, dynamische Nutzersichten, die sich in den Kontext einer VO einbetten lassen und deren individuelle Bedürfnisse erfüllen. Weiter ergibt sich ein Grid-typisch heterogenes Umfeld, welches ebenfalls auf mehreren Ebenen existiert. Nicht nur im Bereich der Ressourcen ist diese Heterogenität zu erkennen. Auch die eingesetzten Grid-Middlewares und vorhandenen Grid-Dienste zeigen dies. Nicht zuletzt die zum Teil bereits von den beteiligten Organisationen eingesetzten Sicherheitskomponenten und -werkzeuge zur Erkennung von Angriffen sind von unterschiedlichster Art.

Hier ist häufig keine Koppelung bestehender Komponenten möglich und der Grid-weite Austausch von Informationen bezüglich sicherheitsrelevanter Ereignisse wird nicht umgesetzt. Dies ist nicht nur auf die Heterogenität in diesem Umfeld zurückzuführen, sondern auch auf Randbedingungen, wie beispielsweise unterschiedliche Sicherheits- und Informationsverbreitungsrichtlinien („security and information sharing policies“) der beteiligten realen Organisationen. Darüber hinaus bieten Firewalls derzeit keinen umfassenden Schutz für Grids. Aufgrund fehlender Mechanismen zur dynamischen Erkennung und Freischaltung von Kommunikationsanforderungen müssen große Portbereiche zum Teil sogar ohne einschränkende Angabe von IP-Adressen permanent freigegeben werden.

Zurzeit existiert kein Gesamtkonzept für ein kooperatives, Grid-weit föderiertes Intrusion Detection System (GIDS) mit entsprechenden Reporting-Komponenten, das sich in ein Umfeld wie dem D-Grid einbettet. Daher soll ein Konzept für ein GIDS entwickelt, im D-Grid implementiert und in die Produktion überführt werden.

### 1.1.2 Ziel

Ziel dieses Projekts ist die Bereitstellung eines GIDS-Dienstes für das D-Grid. Hierbei gilt es, soweit wie möglich bestehende Ansätze zu integrieren und ein domänen- und organisa-

tionsübergreifendes Gesamtsystem zu entwickeln. Insbesondere die Fähigkeit, mit Virtuellen Organisationen (VO) umzugehen und diese auch als Kunden in Betracht zu ziehen, ist dabei von entscheidender Bedeutung. Die Grundidee ist es, Angriffe durch die kooperative Nutzung und Auswertung von lokalen Sicherheitssystemen zu erkennen. Dazu ist der Austausch von Angriffsdaten und somit deren datenschutzkonforme Aufarbeitung, auch zur Wahrung individuell bestehender Sicherheits- und Informationsverbreitungsrichtlinien, notwendig. In einem kooperativen IDS besteht die Möglichkeit, Angriffe schneller zu erkennen, als dies mit unabhängigen und nur die lokale Sicht berücksichtigenden Sicherheitssystemen möglich ist. Somit kann eine Verkürzung der Reaktionszeit der beteiligten Parteien erzielt werden. Weiterhin können Vorwarnungen an zum Zeitpunkt der Erkennung eines Angriffs noch nicht betroffenen Parteien herausgegeben sowie gegebenenfalls präventive Gegenmaßnahmen ergriffen werden.

Eine Auswertung der Daten kann sich zu großen Teilen auf bereits vorhandene Ansätze klassischer IDS stützen. Bei der Auswertung der verfügbaren Datengrundlage ist darauf zu achten, dass VO-spezifische Zugriffsrechte und Befugnisse eingehalten werden. Nach erfolgreicher Auswertung aller verfügbaren Informationen durch ein kooperatives und föderiertes GIDS, unter Beachtung individueller Sicherheits- und Datenschutz-Policies, erfolgt eine Berichterstattung über die erkannten Angriffe auf das Grid oder einzelne beteiligte Partner. Auch hier ist es von Bedeutung, dass eine VO-spezifische Sicht auf die bereitgestellten Informationen realisiert wird. Dazu ist eine Anbindung an die im D-Grid bestehenden VO Managementsysteme zu schaffen. Nach der Entwicklung einer geeigneten Architektur für ein kooperatives und föderiertes IDS in Grid-Umgebungen steht die Implementierung und Produktivführung des Systems. Es soll nach Abschluss der Projektlaufzeit ein produktives Intrusion Detection System als Grid-Dienst im D-Grid zu Verfügung stehen, das sowohl von Ressourcenanbietern als auch von Kunden (VOs, Communities etc.) genutzt werden kann.

## 1.2 Erstellung eines Datenschutzkonzeptes

Spezielle Herausforderung des Datenschutzkonzeptes ist es, den unterschiedlichen und teilweise sogar widersprüchlichen Anforderungen gerecht zu werden. Diese ergeben sich aus der Erkennungsleistung des GIDS-Sensornetzwerkes, den Anforderungen der Datenzulieferer und dem gesetzlichen Datenschutz. Ziel ist es, letztendlich ein Konzept zu erstellen, dass den bestmöglichsten Kompromiss dieser Anforderungen darstellt. Dazu werden die technischen Ansätze zur Anonymisierung und Pseudonymisierung auf ihre Eignung für das GIDS und dessen Anforderungen untersucht.

Es zeigt sich, dass es insbesondere auf dem Gebiet der Pseudonymisierung von IDS-Daten erfolgsversprechende Ansätze gibt, auf denen aufgebaut werden kann. Als spezieller Vorteil der Pseudonymisierung kann der direkte personengebundene Bezug zu Merkmalen aus den Daten entfernt werden, ohne die Korrelation und Aggregation der Daten zu beeinträchtigen. Diese ist insbesondere im Grid-Umfeld wichtig, um verteilte Angriffe zu erkennen. Weiterhin stellt sich die Wahl des Austauschformats IDMEF zur Übertragung der Daten als vorteilhaft heraus. Weil alle Daten in einem einheitlichen Format übertragen werden, müssen nicht alle proprietären Merkmale der Daten auf die Verträglichkeit mit dem Datenschutz überprüft werden. Es ist ausreichend, dies für IDMEF durchzuführen.

## 1.3 Struktur des Dokuments

Im zweiten Kapitel werden die technischen und organisatorischen Grundlagen beschrieben, soweit diese für das Verständnis des Datenschutzkonzeptes notwendig sind. Darauf folgen die Prinzipien und die Vorgehensweise, auf denen die Erstellung des Datenschutzkonzeptes basiert. Im vierten Kapitel wird dann das Datenschutzkonzept vorgestellt. In diesem Kapitel wird auf die technischen Verfahren zum Schutz und der Pseudonymisierung der Daten eingegangen. Dies bezieht sich auf die verschiedenen Datentypen einschließlich der IP-Adressen und Audit-Daten.

# Kapitel 2

# Grundlagen eines Datenschutzkonzeptes

In diesem Kapitel werden die rechtlichen, technischen und organisatorischen Grundlagen beschrieben, die zum Verständnis des GIDS-Datenschutzkonzeptes notwendig sind. Beispielsweise ist es notwendig, die Verwendung der Daten und die Rechte für deren Zugriff vorher festzulegen. Weiterhin werden Ansätze zum Schutz personenbezogener Daten zusammengefasst.

## 2.1 Rechtliche Aspekte des Datenschutzkonzeptes

Das Datenschutzkonzept als integraler Bestandteil eines funktionierenden Datenschutzmanagements findet im Gesetz lediglich einmal Erwähnung im Zusammenhang mit Datenschutzaudits in § 9a BDSG. Trotz der Erwähnung existiert keine gesetzliche Definition. In der Praxis wird als Maßstab für ein Datenschutzkonzept oft der Baustein B 1.5 aus dem BSI-Grundschutzkatalog herangezogen. In der Gesamtbetrachtung ist unter einem Datenschutzkonzept die planvolle Berücksichtigung des Datenschutzes unter Berücksichtigung der Projektgegebenheiten im Einzelfall zu verstehen.

### 2.1.1 Personenbezug

Der Datenschutz und somit das Datenschutzrecht findet nur dann Anwendung, wenn es sich bei den erhobenen und verarbeiteten Daten um personenbezogene Daten im Sinne von § 3 Abs. 1 BDSG handelt. Personenbezogene Daten sind demnach Einzelangaben über persönliche und sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person.

#### 2.1.1.1 Bestimmtheit einer Person

Die Bestimmtheit ist der einfachste Fall des Personenbezugs von Daten. Hier sind die Daten direkt mit dem Namen oder einer anderen eindeutigen Bezeichnung der Person verknüpft.

#### 2.1.1.2 Bestimmbarkeit einer Person

Aber auch bei einer nur auf Umwegen möglichen Bestimmbarkeit einer Person ist der Personenbezug anzunehmen. Hierbei kommt es darauf an, ob aufgrund der Angaben und zusätzlicher Kenntnisse, Mittel und Möglichkeiten, der Bezug zu einer natürlichen Person herstellbar ist. Allerdings muss der Bezug mit den normalerweise zur Verfügung stehenden Hilfsmitteln und ohne unverhältnismäßigen Aufwand herstellbar sein. Folglich kann sich auch aus der Zusammenschau von Daten oder unter Berücksichtigung der realistischen Heranziehung weiterer Daten die Annahme des Personenbezugs ergeben. Die rein theoretische Möglichkeit der Identifizierung reicht aber allein nicht aus. Zusätzlicher Maßstab ist die Verhältnismäßigkeit der Ermittlung der Person unter wirtschaftlichen Gesichtspunkten. Faustregel hierbei ist, dass die

Verhältnismäßigkeit nicht mehr gegeben ist, wenn der (hypothetische) Aufwand der Erhebung der Daten bei der betroffenen Person niedriger ist.

Bezüglich der Bestimmbarkeit ist in der Diskussion, inwieweit IP-Adressen als personenbezogene Daten zu behandeln sind. Dies hat vor allem deshalb Relevanz, da bei dieser Annahme auch andere, mit der IP-Adresse zusammen gespeicherte Daten Personenbezug erlangen können. Eigentlich handelt es sich bei einer IP-Adresse nur um eine maschinelle Adressierung zwischen Rechnern. Allerdings kann nicht gänzlich ausgeschlossen werden, dass hinter einer IP-Adresse eine natürliche Person als Nutzer steht, die unter Umständen (Anfrage an RIPE, Provider) mit der maschinellen Adressierung in Verbindung gebracht und somit individualisiert werden kann. Statische IP-Adressen werden schon länger wie personenbezogene Daten behandelt, da die Auflösbarkeit zu einer natürlichen Person nicht ausgeschlossen werden kann. Bei dynamischen IP-Adressen bestehen unterschiedliche Auffassungen. Das AG Berlin-Mitte (Urt. v. 27.3.2007, Az.: 5 C 314/06) bejaht den Personenbezug, da der zuständige Provider die Auflösung zur Person seines Kunden ggf. herstellen kann. Das AG München (Urt. v. 30.9.2008, Az.: 133 C 5677/08) sieht das anders. Demnach kommt es jeweils auf die tatsächliche Auflösungsmöglichkeit desjenigen an, der von der IP-Adresse Kenntnis erhält. Der Streit ist noch immer offen. Wo möglich werden in der Praxis IP-Adressen sicherheitshalber stets wie personenbezogene Daten behandelt.

#### 2.1.1.3 Anonymisierung

Nach der gesetzlichen Definition in § 3 Abs. 6 BDSG ist Anonymisieren das Verändern personenbezogener Daten derart, 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 Person zugeordnet werden können. Somit ist zu einer Anonymisierung nicht nur die Löschung von Daten geeignet. Es reicht auch eine irreversible Veränderung der Daten derart, dass der Bezug zu einer Person nicht mehr oder nur mit unvertretbarem Aufwand möglich ist, wobei diesbezüglich die gleichen Kriterien wie oben bei der Bestimmbarkeit gelten. Bei erfolgreicher Anonymisierung bedarf es zur weiteren Verarbeitung der Daten aus Sicht des Datenschutzes keiner besonderen Rechtsgrundlage mehr.

#### 2.1.1.4 Pseudonymisierung

Die Pseudonymisierung wird in § 3 Abs. 6a BDSG gesetzlich definiert als Ersetzung des Namens oder anderer Identifikationsmerkmale durch ein Kennzeichen zu dem Zweck, die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren. Dies ist beispielsweise der Fall, wenn der Name durch eine Personenkennziffer oder durch ein neutrales Synonym ersetzt wird. So kann auch eine dynamische IP-Adresse als Synonym für den entsprechenden Kunden des Zugangsproviders angesehen werden. Eine Pseudonymisierung führt aber nicht unbedingt auch zu einer Anonymisierung, da oftmals zumindest durch einen Beteiligten der Bezug zur Person wieder hergestellt werden kann. In diesem Fall handelt es sich weiter um personenbezogene Daten die dem Datenschutz unterliegen. Eine Anonymisierung durch Pseudonymisierung ist deshalb nur unter den eben genannten Voraussetzungen aus § 3 Abs. 6 BDSG anzunehmen.

### 2.1.2 Rechtsgrundlage der Datenverarbeitung

Ist der Personenbezug gegeben, bedarf es aufgrund des grundsätzlichen Verbots der Erhebung und Verarbeitung von personenbezogenen Daten stets einer Rechtsgrundlage für die Datenverarbeitung. Nach § 4 Abs. 1 BDSG kann dies ein Gesetz oder eine andere Rechtsvorschrift sein, die die Datenverarbeitung erlaubt oder anordnet. Darüber hinaus ist auch die Einwilligung des Betroffenen eine Rechtsgrundlage.

Stets gilt es aber, die Zweckbindung der jeweiligen Rechtsgrundlage zu beachten. D.h. die Daten dürfen grundsätzlich nur zu dem Zweck erhoben werden, der von der gesetzlichen Erlaubnis oder der Einwilligung umfasst ist. So dürfen z.B. zur Erkennung von Störungen erhobene IP-Adressen nicht für personengebundene Nutzungsanalysen zur Optimierung der Tarifstruktur genutzt werden.

Darüber hinaus ist stets der Grundsatz der Datensparsamkeit einzuhalten. Selbst wenn eine Rechtsgrundlage für die Verarbeitung besteht, muss sich der Umfang der tatsächlichen Erhebung und Verwendung an der Notwendigkeit orientieren.

#### 2.1.2.1 Einwilligung

Die Einwilligung ist der freiwillige Verzicht der betroffenen Person auf ihr Recht auf informationelle Selbstbestimmung. Damit der Schutz durch das Datenschutzrecht nicht durch Trivialeinwilligungen unterlaufen werden kann und die Einwilligung tatsächlich freiwillig erfolgt, stellt das Gesetz in § 4a Abs. 1 BDSG und in den jeweiligen bereichsspezifischen Gesetzen zum Datenschutz Anforderungen an Inhalt, Verfahren und Form. Sie muss sich auf eine bestimmte Datenverarbeitung beziehen, darf nicht in AGB versteckt werden und muss grundsätzlich schriftlich erfolgen. Bei Telemedien- (§ 13 Abs. 2 TMG) und Telekommunikationsdiensten (§ 94 TKG) kann die Einwilligung auch elektronisch erfolgen. Allerdings müssen Protokollierung, jederzeitige Abrufbarkeit und Widerrufbarkeit der Einwilligung gewährleistet werden.

In Bezug auf die Erhebung und Verwendung von IP-Adressen ist die Einwilligung als Rechtsgrundlage generell wenig praktikabel, da die Einwilligung bereits vor der Erhebung vorliegen muss.

#### 2.1.2.2 Gesetzliche Erlaubnisnormen

Bei den Erlaubnisnormen ist zwischen den allgemeinen und den bereichsspezifischen Gesetzen zu differenzieren. Besteht für einen bestimmten Bereich (z.B. Telekommunikation) eine anwendbare gesetzliche Erlaubnis, geht diese den allgemeinen Erlaubnisnormen aus dem BDSG oder den LDSG vor. Umgekehrt heißt dies aber auch, dass auf die allgemeine Erlaubnis nicht ohne Weiteres zurückgegriffen werden kann, da bei einer sachlich beschränkten speziellen gesetzlichen Regelung davon ausgegangen werden muss, dass die Beschränkung durch den Gesetzgeber so beabsichtigt ist.

##### 2.1.2.2.1 Spezialgesetzliche Erlaubnisnormen

Spezialgesetzliche Erlaubnisnormen finden sich in der Regel in dem Gesetz, in dem die Materie geregelt ist.

So enthält beispielsweise das Telekommunikationsgesetz (TKG) für Telekommunikationsanbieter Erlaubnisnormen zur Verarbeitung von bei der Nutzung der Dienste anfallenden Verkehrsdaten. Hierzu gehören neben IP-Adressen auch Zeitangaben, Angaben über die Anzahl der Pakete u.s.w. Praktisch relevant sind die Regelungen zur Telekommunikation vor allem deshalb, weil z.B. auch ein Arbeitgeber als Telekommunikationsdiensteanbieter angesehen wird, wenn er die private E-Mail oder Internetnutzung durch seine Mitarbeiter erlaubt.

Dann erlaubt § 97 TKG die Erhebung und Verarbeitung der Verkehrsdaten zur Ermittlung des Entgelts. Allerdings nur dann, wenn tatsächlich abgerechnet wird und nur in dem dafür erforderlichen Umfang.

§ 100 Abs. 1 TKG erlaubt die Erhebung und Verwendung der Verkehrsdaten zum Zweck der Erkennung und Beseitigung von Störungen. Die Erlaubnis enthält aber mit dem Begriff “soweit erforderlich” eine erhebliche Einschränkung. Die Daten dürfen nicht pauschal vorgehalten werden, sondern nur dann, wenn eine begründbare Erforderlichkeit dieser Daten zur Störungserkennung oder -beseitigung besteht, wobei dies bei einem positiven Befund des IDS der Fall sein dürfte. Ohne weitere Anhaltspunkte für eine konkrete Störung wird in der Rechtsprechung derzeit eine anlasslose Speicherung bis zu 7 Tagen geduldet. Sind die Daten danach nicht mehr für die Eingrenzung oder Behebung der Störung erforderlich, müssen sie gelöscht werden. Werden die Daten noch zu statistischen Zwecken benötigt, kann statt der Löschung auch eine Anonymisierung erfolgen.

#### 2.1.2.2.2 Allgemeine Erlaubnisnormen

Bestehen für einen Bereich keine spezialgesetzlichen Regelungen zum Datenschutz oder sind diese nicht anwendbar, kann ggf. auf die allgemeinen Erlaubnisnormen zurückgegriffen werden. Beispielsweise sind die Normen aus dem TKG nicht anwendbar, wenn die E-Mail- und Internetnutzung ausschließlich zu dienstlichen Zwecken erfolgen darf. Dann fehlt es an der Anbietereigenschaft des Arbeitgebers, womit das TKG auf diesen Sachverhalt insgesamt unanwendbar ist. In diesem Fall beurteilt sich der Umgang mit den Verkehrsdaten nach den allgemeinen Erlaubnisnormen.

Für nichtöffentliche Stellen z.B. aus der Privatwirtschaft enthält §28 BDSG weit gehende Erlaubnistatbestände. So dürfen Daten verarbeitet werden, wenn dies der Zweckbestimmung eines Vertragsverhältnisses dient oder soweit dies zur Wahrung berechtigter Interessen erforderlich ist. Letzterenfalls ist allerdings eine Abwägung mit den schutzwürdigen Interessen des Betroffenen erforderlich.

Für die öffentlichen Stellen auf Bundesebene enthält § 14 Abs. 1 BDSG eine allgemeine Erlaubnis, wenn die Datenverarbeitung zur Erfüllung der in der Zuständigkeit der verantwortlichen Stelle liegenden Aufgaben erforderlich ist. Wichtigste Begrenzung ist die konkrete Erforderlichkeit der Datenverarbeitung. Unter die allgemeine Erlaubnis fällt auch die Datenverarbeitung zu Forschungszwecken, wobei hinsichtlich der Anforderungen ergänzend § 40 BDSG gilt. Für die öffentlichen Stellen der Länder enthalten die jeweiligen Landesdatenschutzgesetze dem 14 Abs. 1 BDSG vergleichbare allgemeine Erlaubnisnormen, so dass sich hier keine wesentlichen inhaltlichen Abweichungen ergeben.

## 2.2 Technische Grundlagen

In diesem Abschnitt werden die technischen Grundlagen zusammengefasst, die für das Datenschutzkonzept des GIDS relevant sind. Da die Erhebung, Speicherung und der Transport von personenbezogenen Daten gesetzlich eingeschränkt sind, werden technische Methoden für die Aufbereitung der Daten benötigt. Dadurch kann eine Grundlage für deren Verarbeitung geschaffen werden.

Daten sind genau dann personenbezogen, wenn Merkmale in den Daten die Identifizierung einer natürlichen Person zulassen, auf die sich die Daten beziehen. Grundsätzlich wird zwischen bestimmten und bestimmbaren Identifikationsmerkmalen unterschieden. Diese werden in der Literatur auch häufig als primäre und sekundäre[1] Merkmale bezeichnet. Die primären ermöglichen die direkte Identifizierung der Person. Wichtigstes Attribut ist der Name der Person; weitere Attribute sind beispielsweise die Nummer des Ausweises oder die Sozialversicherungsnummer. Sekundär sind Attribute, die in Verbindung mit anderen Attributen und Hintergrundwissen eine Identifizierung ermöglichen. Durch Korrelation oder Kombination dieser Merkmale mit Hintergrundinformationen kann unter Umständen auf die Identität der Person geschlossen werden. Dies kann beispielsweise die Identifizierung des Webbrowsers in Kombination mit der Betriebssystemversion und Informationen über die installierten Browser-Plug-Ins sein.

## Anonymisierung

Der einfachste und effizienteste Weg, den Bezug in Datensätzen zu Personen zu löschen, ist deren Anonymisierung. Dies kann zuerst dadurch geschehen, indem alle Felder mit Personenbezug aus den Datensätzen gelöscht werden. Diese Methode ist speziell dann sinnvoll, wenn die Daten nur zu statistischen Zwecken verwendet werden sollen. Beispielsweise kann die Anzahl von Verbindungsversuchen oder bestimmter Netzwerkpakete gezählt werden. Dadurch lassen sich effizient neuartige Internet-Würmer entdecken, die bei ihrer Verbreitung ein erhöhtes Verkehrsaufkommen erzeugen. Neben dem Löschen der Daten können diese auch durch Verallgemeinerung anonymisiert werden. Wird beispielsweise eine IP-Adresse auf die Nummer des

[1] Sekundäre Identifikationsmerkmale werden teilweise auch als Quasi-Identifikatoren bezeichnet

Autonomen Systems (AS) abgebildet, in der die IP enthalten ist, entfällt der Personenbezug.

Ein Maß für die Güte der Anonymisierung ist die Eigenschaft der $k$-**Anonymität**. Diese besagt, dass für jeden Datensatz deren sekundären Identifikationsmerkmale auf mindestens $k-1$ weitere Datensätze mit der gleichen Wahrscheinlichkeit zutreffen. Das bedeutet also, dass die Attribute auf mindestens $k$ Individuen zutreffen. Je größer $k$ ist, umso größer ist die Anzahl der Individuen, auf die die Attribute zutreffen und umso schwieriger ist die Identifizierung der wahren Identität.

Als Nachteile der Anonymisierung der Datensätze verhindert diese eine Korrelation von Datensätzen und die Zurückverfolgung von Angriffen zu der Quelle. Die Korrelation ist beispielsweise dann entscheidend, wenn der Angriff nur aus dem Kontext von mehreren verschiedenen aber doch zusammenhängenden Informationen hervorgeht. In diesen Fällen lässt jede Information für sich noch nicht eindeutig auf einen Angriff schließen. Zumindest erschließt sich der Angriff noch nicht in seinem vollen Umfang. Dieser ergibt sich erst durch Kombination bzw. Korrelation der Daten.

## Pseudonymisierung und Reidentifizierung

Die Pseudonymisierung zielt darauf ab, die Nachteile der vollständigen Anonymisierung der Daten zu vermeiden. Die grundsätzliche Idee ist es, das personenbezogene Attribut eindeutig auf ein Pseudonym abzubilden, dessen Bezug zu der realen Person ohne die Kenntnis der Abbildungsfunktion nicht mehr nachvollziehbar ist. Da allerdings eine eindeutige Abbildung existiert, lässt sich die Pseudonymisierung wieder rückgängig machen.

Grundlage der Pseudonymisierung ist eine eindeutige Abbildung des personenbezogenen Attributes auf ein Pseudonym. Die Abbildungen unterscheiden sich in der Eigenschaft, ob die Umkehrfunktion berechenbar ist oder nicht. Umkehrbare Abbildungen sind dann vorteilhaft, wenn die Pseudonymisierung aufgehoben werden soll. Nicht berechenbare Umkehrfunktionen lassen zwar eine Korrelation der transformierten Daten zu, verhindern aber die Identifizierung der Identität für Dritte. Bei der Rücknahme der Pseudonymisierung wird zwischen den verschiedenen Rollen des Sicherheitsbeauftragten und der Rolle des Datenschützers unterschieden. Aufgrund der Pseudonymisierung hat der Sicherheitsbeauftragte nur Zugriff auf pseudonymisierte Daten; kann also keinen Bezug der IDS-Daten zu den Benutzern herstellen. Allerdings bietet sich die Möglichkeit, die Anonymisierung mit Einwilligung der betroffenen Seite oder Person rückgängig zu machen, oder wenn eine rechtliche Erlaubnis vorliegt. Ein Datenschutzbeauftragter kann diese Vorgänge beobachten, um Missbrauch zu verhindern und Empfehlungen zu geben. Dies trifft beispielsweise dann zu, wenn dies für die Abwehr eines Angriffes oder zur Warnung der Betroffenen notwendig ist.

Technisch existieren verschiedene Ansätze, die hier aufgrund der Vielzahl nicht alle aufgezählt werden können. Eine Möglichkeit ist die Verwendung von Hash-Werten. Dabei wird aus dem Datum ein Hash-Wert berechnet, wobei verschiedenen bekannte Verfahren wie beispielsweise **MD5** oder **SHA-X** zum Einsatz kommen. Eine Eigenschaft dieser Hash-Algorithmen ist die Unumkehrbarkeit. Diese bewirkt, dass aus dem Hash-Wert das eigentliche Datum nicht berechnet werden kann. Ist der Werteraum des Datums ausreichend groß, lässt sich die Pseudonymisierung also nicht umkehren.

Für IP-Adressen wurde mit Crypto-PAn in [3, 4] ein eigenständiger Ansatz entwickelt. Wichtigstes Merkmal ist die Erhaltung des Präfixes der IP-Adresse während der Transformation. Stimmen zwei IP-Adressen in den ersten $n$-Bits überein, so bleibt dies auch nach der Transformation erhalten. Dies bewirkt, dass zwei IP-Adressen aus dem gleichen Netzblock auch noch nach der Transformation in einem Netzblock liegen. Dadurch bleibt also der Zusammenhang zwischen zwei IP-Adressen gewahrt, was bei der Untersuchung und Interpretierung von Angriffen notwendig sein kann.

Für die Reidentifizierung pseudonymisierter Daten muss die entsprechende Transformation umgekehrt werden. Technisch lässt sich dies beispielsweise durch einen Krypto-Algorithmus

lösen, bei dem der Datenschutzbeauftragte den entsprechenden Schlüssel für die Rücktransformation besitzt. Weiterhin wurde von Shamir ein Ansatz vorgeschlagen, bei dem ein Datum in eine beliebige Anzahl von *Shares* aufgeteilt werden kann. Das Datum lässt sich nur durch das Zusammenführen aller Shares wieder rekonstruieren. Mittels dieses Ansatzes lassen sich also die Informationen zur Reidentifizierung auf eine beliebige Anzahl an Seiten oder Personen aufteilen. Eine Reidentifizierung ist dann nur durch Kooperation aller beteiligten Seiten möglich.

In Bezug auf die Verwendung von Pseudonyms kann zwischen verschiedenen Klassen unterschieden werden:

**Entity-basiertes Pseudonym** Dem Identifikator (Person oder beispielsweise IP-Adresse) wird ein festes Pseudonym zugeordnet.

**Rollen-basiertes oder Relationship-basiertes Pseudonym** Das Pseudonym hängt von der Rolle der Person oder dem Verhältnis der Personen zueinander ab. Je nach Rolle oder dem Verhältnis wird also ein anderes Pseudonym verwendet.

**Transaktions-basiertes Pseudonym** Das Pseudonym ist nur innerhalb einer geschlossenen Transaktion gültig. Eine Transaktion kann beispielsweise eine Sitzung eines Benutzers oder eine Transaktion an einem Web-Portal sein.

Damit ergibt sich eine Abwägung der möglichen Zuordnung eines Pseudonyms zu einer Person und der Möglichkeit, Daten zu korrelieren. Je größer der Kontext ist, in dem ein Pseudonym verwendet wird, umso größer wird die Wahrscheinlichkeit, durch Korrelation mehrerer unterschiedlicher Informationen, die entsprechende Person zuzuordnen. Allerdings schwindet mit abnehmendem Kontext die Möglichkeiten der Interpretierung und Auswertung der Daten.

### Schutz der Daten

Der Schutz der Daten ist ein wichtiger Punkt des Datenschutzkonzeptes, bei dem die Sicherheit bei der Speicherung, dem Transport sowie dem Zugriff berücksichtigt werden. Allerdings haben sich dafür Methoden etabliert, die als Standard angesehen werden können. Allgemein sind die Vertraulichkeit, Integrität und Verfügbarkeit der Daten zu gewährleisten. Weiterhin darf der Zugriff nur aus authentifizierter und autorisierter Quelle stammen.

Für alle Punkte ist ein Grundschutz zu gewährleisten. Dies beginnt mit dem Absichern des Standortes und des Systems, in denen die Daten gespeichert werden. Zu berücksichtigen sind sowohl die physikalische Sicherheit der Systeme als auch die Sicherheit gegenüber Angriffen aus dem Internet. Hierfür wird auf das Grundschutzhandbuch des BSI[2] verwiesen. Das Gleiche gilt für den Transport der Daten. Beim Zugriff ist neben der Authentifizierung die Autorisierung einzuhalten. Dies kann dadurch geschehen, dass die notwendigen Berechtigungen als Matrix der Datentypen und Rollen aufgestellt werden. Für jede Rolle wird damit festgelegt, auf welche Daten diese Zugriff hat.

## 2.3 Organisatorische Grundlagen

Bei dem Betrieb eines verteilten IDS sind verschiedene Aspekte zu berücksichtigen. Dies gilt insbesondere für den Fall, dass verschiedene, unabhängige Partner beteiligt sind. Grundsätzlich sind zu berücksichtigen:

**Sicherheitsanforderungen der beteiligten Seiten** Sowohl beim Betrieb als auch bei der Administration eines verteilten IDS werden sicherheitskritische Daten erhoben. Dies sind einerseits Informationen der Seiten über die Position der Sensoren und die Struktur des Netzwerkes. Andererseits liefern die Sensordaten Informationen über die Verwundbarkeit

[2] Es ist im Internet unter der Adresse
https://www.bsi.bund.de/cln_165/DE/Themen/ITGrundschutz/itgrundschutz_node.html verfügbar

und den Zustand der überwachten Netzwerke. Damit diese Daten nicht an Unbefugte weitergegeben werden, ist eine genaue Festlegung der Verwendungszwecke der Daten notwendig, in die die beteiligten Seiten einwilligen.

**QoS Anforderungen der beteiligten Seiten** Die rechtzeitige Erkennung und Behebung von Angriffen ist für den ordnungsgemäßen Betrieb eines Grids wichtig. Dafür bildet das IDS die wichtigste Grundlage. Um eine Verbindlichkeit beim Betrieb zu schaffen, kann die Qualität des Dienstes (QoS) festgelegt werden. Dies betrifft sowohl die Seite der Datenzulieferer als auch die Betreiber.

**Rechtliche Regelungen** Wie bereits in Abschnitt 2.1 beschrieben, unterliegt der Betrieb eines verteilten IDS verschiedenen rechtlichen Regelungen. Um Schwierigkeiten zu vermeiden, bietet es sich an, die rechtlichen Regelungen in der Kooperationsvereinbarung zu berücksichtigen. In diesem Fall kann auch aus dieser Perspektive die Berücksichtigung des geltenden Rechts belegt werden.

Für die oben genannten Punkte wird üblicher Weise ein **Non-Disclosure Agreement (NDA)** und eine **Kooperationsvereinbarung** aufgestellt. Mit der Unterzeichnung dieser Dokumente willigen die beteiligten Seiten ein, dass keine Daten weitergegeben, aufgehoben oder anderweitig verwendet werden, die dem vorher festgelegten Verwendungszweck widersprechen. Weiterhin belegt diese Regelung die Beachtung des geltenden Rechtes in Bezug auf den Datenschutz.

# Kapitel 3

# Entwurf des Datenschutzkonzeptes

Ziel des GIDS ist das zuverlässige Erkennen von Angriffen und die Grundlage für die Wiederherstellung der Systeme nach Angriffen zu schaffen. Damit einher geht die Anforderung nach aussagekräftigen Informationen. Die spezielle Herausforderung des Datenschutzkonzeptes ist es damit, einen Kompromiss zu finden, der diese Anforderungen unter Beachtung der gesetzlichen Regelungen realisiert.

Das Datenschutzkonzept basiert auf verschiedenen Ansatzpunkten:

**Rechtliche Einschränkungen** Diese regeln die Verwendung von personenbezogenen Daten und bilden den gesetzlichen Rahmen für den Betrieb des GIDS.

**Anforderungen des GIDS** In den vorherigen Arbeitspaketen wurden Anforderungen an das GIDS gestellt. Eine Anforderung an das Datenschutzkonzept ist es, diese mit den gesetzlichen Regelungen in Einklang zu bringen.

**Least Privilege Prinzip** Dieses Prinzip liefert eine wichtige Gestaltungsregel, indem die Privilegien der beteiligten Seiten auf das Notwendige eingeschränkt werden. Dadurch wird eine hohe Grundsicherheit und Fehlertoleranz erreicht.

Zuerst werden die Anforderungen an das Gids und Datenschutzkonzept aufgezählt, die in den vorherigen Arbeitspaketen erarbeitet wurden. Danach werden die Rollen im GIDS und deren Anforderungen und Abhängigkeiten aufgezählt. Den Abschluss bildet die Aufzählung der Anforderungen und Abhängigkeiten des Datenschutzkonzeptes zu der Architektur des GIDS. Die Realisierung des Konzeptes erfolgt dann auf deren Basis und der Anwendung des Least Privilege Prinzips und umfasst die folgenden Punkte:

- Die Kooperationsvereinbarung und Non-Disclosure Agreement
- Entscheidung der Daten, auf die die Pseudonymisierung angewendet wird
- Entscheidung der Daten, die an die GIDS-Datenspeicher weitergegeben werden
- Prozeduren für die Aufhebung der Pseudonymisierung

## 3.1 Anforderungen an das Datenschutzkonzept

In [7] in Kapitel 3.3 und ergänzend in 4.6 wurden eine Reihe von Kriterien für ein Grid-IDS angeführt. Diese wurden in fünf Kategorien eingeteilt. Der Übersicht halber werden im Folgenden relevante Anforderungen geordnet nach den Kategorien noch einmal wiederholt.

**Funktionale Anforderungen:**

**F01:** Unterstützung verschiedener Granularitätsstufen bei der Berichterstattung

**F02:** Berichterstattung zu qualitativ differierenden Angriffen

**F05:** Variationsmöglichkeit der Informationsquellen/Datenbasis zur Laufzeit

**F06:** Proaktive Benachrichtigung der Kunden

**F11:** Anbindung an bzw. Nutzung von bestehenden VO-Managementsystemen

**F12:** Aussagekräftige Informationsaufbereitung

**Nichtfunktionale Anforderungen:**

**N12:** Dynamik der Nutzer und VOs

**N13:** Dynamik der Ressourcen

**N15:** Unterstützung Virtueller Organisationen

**N16:** Interoperabilität der Sensoren

**Sicherheitsanforderungen:**

**Kryptographische Anforderungen:**

**S01:** Vertraulichkeit von Daten und Nachrichten

**S02:** Authentizität von Daten und Nachrichten

**S03:** Integrität von Daten und Nachrichten

**S06:** Schutz der Grid-IDS Daten

**Nutzerverwaltung:**

**S07:** Integration in PKI

**S11:** Zugriffsbeschränkung auf Informationen

**Organisatorische und Datenschutzanforderungen:**

**Organisatorische Anforderungen:**

**D01:** Etablierung einer vertrauenswürdigen Koordinationseinheit

**D03:** Gewährleistung der Autonomie beteiligter Informationsanbieter

**D05:** Juristisch verwertbare Speicherung der Daten

**Datenschutz:**

**D06:** Anonymisierungs- und/oder Pseudonymisierungsmöglichkeiten

**D07:** Durchsetzung des Datenschutzes

**D08:** Nachhalten historischer Berichte

**D09:** Archivierung von Sensordaten

**Angriffstypen und -muster:**

**E04:** Erkennung verschiedener Angriffstypen (aktiv, passiv/autonom, DoS)

**E05:** Entdecken kurzzeitig angelegter bis hin zu zeitlich lang andauernder Angriffe

Die Anforderungen an die Sicherheit und den Datenschutz unterstützen das Datenschutzkonzept direkt. So sind Sicherheitsmaßnahmen für den Transport, Speicherung und Zugriff auf die Daten gefordert, die direkter Bestandteil des Datenschutzkonzeptes sind. Im Detail sind bereits in dem Grobkonzept des GIDS in [7] Komponenten zur Verschlüsselung und Integritätssicherung der Daten vorgesehen. Der Zugriff der Daten ist dort durch eine Einbindung in bestehende Grid-Systeme zur Benutzer-Authentifizierung abgesichert. Als weitere Anforderung

ist der Schutz der GIDS-Daten gefordert, der sowohl die Ergebnisse als auch Betriebsdaten beinhalten.

Andere Anforderungen haben indirekte Konsequenzen zur Folge. So wird zum Beispiel eine aussagekräftige Informationsaufbereitung, die Erkennung verschiedener Angriffstypen und die proaktive Benachrichtigung der Kunden gefordert. Weiterhin wird eine juristisch verwertbare Speicherung der Daten gefordert, um im Fall eines Vorfalls rechtliche Schritte einleiten zu können. Um einen rechtlich konformen Umgang mit den Daten zu erreichen, kann der Personenbezug aus den Daten entfernt werden. Das vollständige Entfernen dieser Daten widerspricht allerdings der Forderung nach der Verwertbarkeit der Ergebnisse. In diesem Fall kann beispielsweise ein Vorfall nicht weiter bearbeitet werden, bei dem kompromittierte Benutzer-Accounts missbraucht wurden. Weiterhin verhindert dies die Warnung von Betroffenen, deren Identität oder Systeme bei einem Sicherheitsvorfall missbraucht worden sind. Eine vollständige Anonymisierung der Daten verursacht des Weiteren einen Konflikt mit der Forderung nach juristisch verwertbaren Beweisen für den Einbruch in ein Grid-System. Aus der hochgradigen Vernetzung der Grid-Systeme ergeben sich bedeutende Konsequenzen für die Erkennung von Vorfällen. In vielen Fällen kann ein Angriff erst durch Zusammenführung bzw. Korrelation von verschiedenen Daten der verteilten Systemen erkannt werden. Spezielles Beispiel ist der Missbrauch einer kompromittierten Benutzer-Identität, die sich nur aus den nachfolgenden Schritten des Angriffs nachvollziehen lässt; beispielsweise wenn der Angreifer die Identität für nachfolgende Angriffe missbraucht. Dies gilt in gleichem Maße für die Bewertung von Vorfällen, die Bemessung des Ausmaßes und die Reaktion auf Vorfälle. Eine selektive Anonymisierung der Daten ist zwar grundsätzlich möglich, würde aber die Erkennungsleistung und die Verwertbarkeit der Ergebnisse aus den eben genannten Gründen negativ beeinflussen. Alternative ist die Pseudonymisierung der Daten, die in den technischen Grundlagen beschrieben wurde.

Weitere Anforderungen, die Konsequenzen für das Datenschutzkonzept haben, ist die Variationsmöglichkeit der Informationsquellen/Datenbasis zur Laufzeit und die Gewährleistung der Autonomie beteiligter Informationsanbieter. Aus dem ersten Punkt folgt, dass die Typen der verarbeiteten Sensordaten nicht eingeschränkt sind. Es ist deshalb nicht möglich, a-priori alle personenbezogenen Merkmale aus den Daten zu bestimmen. Der zweite Punkt hat Konsequenzen für die Filterung der Daten und Durchführung der Pseudonymisierung, wodurch ein zentralistischer Ansatz ausscheidet.

Ein weiterer Ansatz liefert die Korrelation und Aggregation von IDS-Daten. So können Daten auf der Seite des Datenzulieferers aggregiert und korreliert werden. Ergebnis ist ein Ereignis, das alle Schritte eines Angriffs aussagekräftig wiedergibt. Da diese Informationen zur Bearbeitung und Abwehr des Angriffs und zur Wiederherstellung der Systeme zwingend benötigt werden, hat der Gesetzgeber erweiterte Möglichkeiten deren Verwendung geschaffen.

## 3.2 Rollen im GIDS und deren Anforderungen

Wie bereits vorher beschrieben, bilden die Rollen und deren Anforderungen an die Verwendung und den Schutz der Daten eine Grundlage des Datenschutzkonzeptes. Dafür wird zwischen den folgenden Rollen unterschieden:

**Benutzer des Grids** Die zentrale Anforderung der Benutzer des Grids ist der Schutz deren Daten. Dies sind einerseits die Daten mit Bezug auf deren Person - also Anmeldevorgänge beim Grid und Benutzung von Grid Systemen - und andererseits die Ergebnisse von den Grid-Jobs. Zwar unterliegen die Ergebnisse nicht immer dem Datenschutz, sie sind aber in der Regel vertraulich. Dies gilt insbesondere für medizinische Daten und Resultate von gewerblich genutzten Grids. Für das GIDS sind diese Daten kritisch, weil sie entweder personenbezogene Merkmale enthalten und dem Datenschutz unterliegen oder aus gewerblichen Gründen als absolut vertraulich gehandhabt werden müssen. Allerdings wird von den Benutzern der zuverlässige Betrieb des Grids und die transparente Behebung von Störungen erwartet. Dies erfordert die effektive Erkennung und Beseitigung von Sicherheitsproblemen.

**Datenzulieferer** sind als föderale Partner Teil der GIDS-Infrastruktur, können aber in ihrer administrativen Domäne unabhängig agieren. Sie betreiben Sensoren zur Erkennung von Angriffen und sind technisch über den GIDS-Agenten an der Architektur angeschlossen. Die Ergebnisse der eigenen Sensoren werden in einer privaten Datenbank gespeichert. Weiterhin erhalten sie über den GIDS-Bus Daten über Angriffe oder Anomalien, die von den andereren Datenzulieferer exportiert werden. Damit teilt sich die Rolle auf in:

- **Datenexport:** In dieser Rolle ist die Anforderung, vertrauliche Informationen geheim zu halten. Dies können Daten über Angriffe in der eigenen administrativen Domäne sein, oder administrative Daten zum Betrieb des GIDS.
- **Datenimport:** Die importierten Daten liefern Informationen über Angriffe der anderen am GIDS beteiligten Seiten. Diese helfen den Datenzulieferern, damit zusammenhängende Angriffe im eigenen Umfeld zu entdecken und sich davor erfolgreich zu schützen.

Daraus kann die Konsequenz abgeleitet werden, dass grundsätzlich alle Daten mit Interna von Datenzulieferer möglichst vertraulich gehandhabt werden müssen, während die Informationen über Angreifer möglichst offen gehalten werden sollten. Ein sinnvoller Kompromiss ist insbesondere dann zu finden, wenn legitime Daten im Rahmen von Angriffen missbraucht werden. Das ist beispielsweise dann der Fall, wenn der Angreifer einen Benutzer-Account erfolgreich kompromittiert hat.

**Betreiber** Das Ziel des Betreibers ist es im Rahmen der gesetzlichen Regelungen und Anforderungen der Rollen einen reibungsfreien und effizienten Betrieb des GIDS zu ermöglichen. Damit fallen die Anforderungen mit denen der anderen Rollen weitesgehend überein. Darüber hinaus besteht die Notwendigkeit eines präzisen Datenschutzkonzeptes. Dies ist insbesondere deshalb wichtig, weil der Betreiber letzendlich gegenüber allen Seiten den sicheren und gesetzeskonformen Betrieb garantieren muss.

**Analysten** Die wichtigste Anforderung der Analysten ist die effiziente Aufbereitung der Daten zur Angriffserkennung. Dies gilt insbesondere der Fähigkeit, durch Korrelation von Daten Angriffe und deren Ausmaß zu erkennen. Werden die Daten zur Vorfallsbearbeitung verwendet, ist unter Umständen eine Aufhebung der Pseudonymisierung der Daten notwendig. Dies ist beispielsweise dann notwendig, wenn ein konkreter Benutzer oder Betreiber eines Systems im GIDS Umfeld über Sicherheitsprobleme informiert werden soll.

**Weitere Rollen** Weitere potentielle Rollen sind besipielsweise Kunden, die Dienstleitungen bezüglich Angriffserkennung auslagern. Eine weitere potentielle Rolle ist ein Datenschutzbeauftragter, der die Einhaltung des Datenschutzes überwacht und Entscheidungen über die Aufhebung der Pseudonymisierung treffen kann.

## Abhängigkeiten zwischen den Rollen

Die Abhängigkeiten zwischen den Rollen ergeben sich aus der Grobarchitektur in Abb 3.1. Sie bilden die Grundlage für die Anwendung des “Least Privilege” Prinzips, um die benötigten Rechte und Daten für jede Rolle zu bestimmen.

Insgesamt lassen sich die folgenden Abhängigkeiten für die verschiedenen Rollen ableiten:

**Betreiber** haben die nachfolgenden Abhängigkeiten:

**Datenzulieferer/Provider** : Verarbeiten, Analysieren und korrelieren die Daten der Datenzulieferer/Provider. Diese enthalten potentiell Referenzen auf die Identitäten der Grid-Benutzer und weitere vertrauliche Daten. Des weiteren müssen zum Betrieb der GIDS-Infrastruktur Daten über die Datenzulieferer und deren Systeme vorgehalten werden.

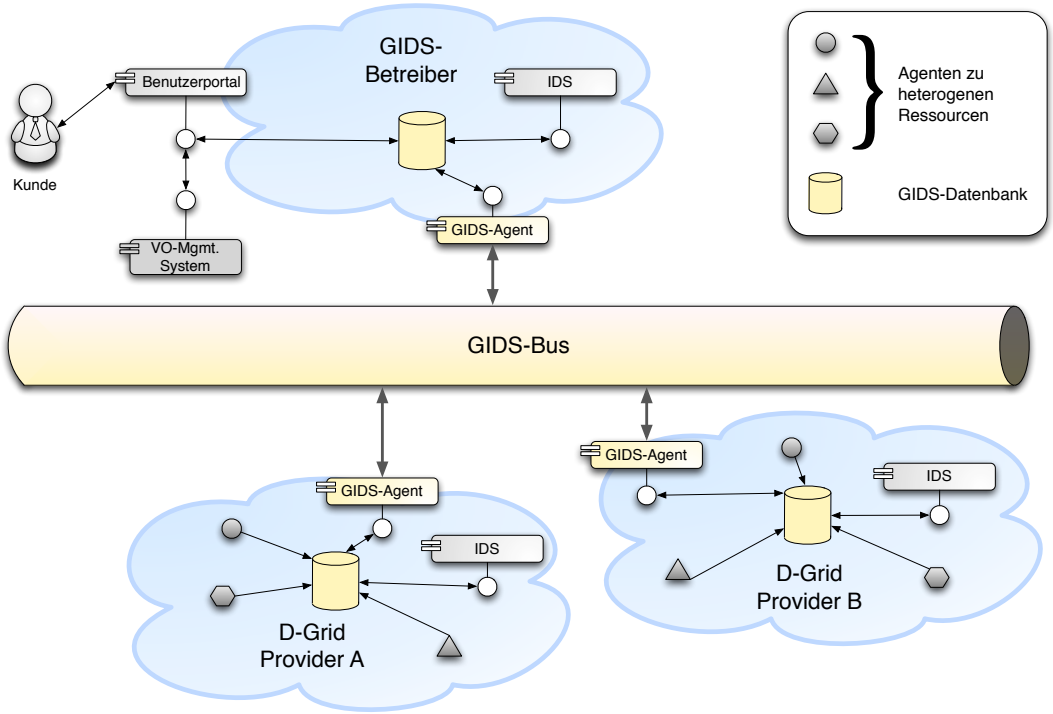


Abbildung 3.1: Grobgranulare Übersicht der Architektur des GIDS

**Kunden/Benutzer der GIDS-Infrastruktur** : Die Ergebnisse und Daten des GIDS werden Kunden und anderen Benutzer der GIDS-Infrastruktur[1] präsentiert. Deren Daten sind für den Betrieb inklusive der Authentifizierung und Autorisierung notwendig.

**Datenzulieferer/Provider** haben die nachfolgenden Abhängigkeiten:

**Benutzer der Grid-Infrastruktur** : Sie verarbeiten deren Daten und leiten diese gegebenenfalls an den GIDS-Betreiber weiter. Dies können entweder Daten aus Angriffen oder auch produktive Daten des Grid-Betriebs sein. Diese enthalten potentiell Referenzen auf die Identitäten der Grid-Benutzer und müssen dementsprechend behandelt werden.

**Analysten** haben die nachfolgenden Abhängigkeiten:

**Benutzer der Grid-Infrastruktur** : Sie analysieren deren Daten auf Angriffspuren, bewerten diese und leiten potentiell Gegenmaßnahmen zur Eindämmung der Angriffe ein. Dies können entweder Daten aus Angriffen oder auch produktive Daten des Grid-Betriebs sein. Diese enthalten potentiell Referenzen auf die Identitäten der Grid-Benutzer und müssen dementsprechend behandelt werden. Dabei ist es kritisch, eine rechtliche Grundlage für deren Tätigkeit einzurichten.

**Weitere Rollen** Im produktiven Betrieb des GIDS sind weitere Rollen denkbar. Dies kann beispielsweise ein Datenschutzbeauftragter sein, der die Einhaltung des Datenschutzes überwacht und Entscheidungen über die Aufhebung der Pseudonymisierung treffen kann. Dies führt zu der folgenden Abhängigkeit:

[1] Das können beispielsweise auch die Datenzulieferer oder Benutzer der Grid-Systeme bzw. innerhalb einer VO sein

**Datenzulieferer/Provider** : Ein Datenschutzbeauftragter hat die Übersicht über die Entscheidungsgewalt, pseudonymisierte Daten der Datenzulieferer bei Bedarf rückgängig zu machen. Dieser Fall kann beispeilsweise bei der der Reaktion auf einen Sicherheitsvorfall eintreten. In diesem Fall ist sicherzustellen, dass die Informationen über die Umkehr der Pseudonymisierung die befugten Personen und ausschließlich diese erreichen.

## 3.3 Abhängigkeiten zur Architektur

Das Datenschutzkonzept ist in mehreren Punkten von der Architektur des GIDS abhängig. Zuerst sind technische Komponenten zur Pseudonymisierung und Filterung der Daten notwendig. Sinnvoller Weise werden diese Komponenten in den administrativen Domänen der Datenzulieferer eingesetzt, wie es in der Grobarchitektur in Abb. 3.1 bereits vorgesehen ist.

Weitere Abhängigkeiten ergeben sich aus den Daten, die auf der einen Seite zum Betrieb des GIDS notwendig sind und auf der anderen Seite von dem System produziert werden. Eine wichtige Anforderung der GIDS-Architektur ist die technische Flexibilität der eingesetzten IDS-Sensorik. Dies führt dazu, dass es keine begrenzte Liste von Sensoren gibt und dass deshalb keine feste Liste der Datenarten gibt. Jedoch lassen sich verschiedenen Daten in verschiedene Klassen unterteilen, die gesondert behandelt werden müssen: Zusammenarbeit mit Ermittlungsbehörden oder

**Daten zum Betrieb des GIDS** dienen zur Kommunikation der GIDS-Komponenten untereinander, für die Erkennung von Fehlern und die Authentifizierung und Autorisierung von Bentzern. Da diese Daten beim Betreiber des GIDS erhoben werden und für den betrieb des GIDS unerlässlich sind, gibt es für deren Verarbeitung eine rechtliche Grundlage.

**Ergebnisse des GIDS** resultieren aus der Auswertung und Erhebung der IDS-Sensorik bei den Ressourcenprovidern bzw. Datenzulieferern. Im Idealfall ist kein Personenbezug durch vollständige Anonymisierung vorhanden. Jedoch kann ein Personenbezug nicht vollständig vermieden werden, wenn die Daten zur Reaktion und Aufklärung von Sicherheitsvorfällen verwendet werden. Deshalb ist die beste Alternative eine Pseudonymisierung bei der nur die Seiten diese Aufheben können, die die Daten erhoben haben. Jedoch gibt es gesetzliche Erlaubnisse für die Erhebung, Verarbeitung und Übertragung der Daten mit personenbezogenen Merkmalen (siehe Abschnitt 2.1).

**Produktive Daten der Grid-Infrastruktur** fallen beim Betrieb und der Benutzung der Grid-Infrastruktur an. Dies sind sowohl die Ergebnisse der Grid-Jobs als auch die Log-Daten der Grid-Komponenten. Zwar sind diese Daten für die Erkennung und Korrelation von Angriffen wichtig, allerdings sind sie potentiell personenbezogen und unterliegen den Datenschutzbestimmungen. Die Daten können deshalb nur unter bestimmten Voraussetzungen ausgewertet, übertragen und gespeichert werden.

# Kapitel 4

# Realisierung des Datenschutzkonzeptes

In diesem Kapitel wird auf die Realisierung des GIDS-Datenschutzkonzeptes eingegangen. Dies betrifft die technische Pseudonymisierung der Daten als auch die organisatorischen Massnahmen zur Regelung der Kooperation der am GIDS beteiligten Seiten. Grundlage dafür bilden die Konzepte zum Entwurf, die in dem vorangegangenen Kapiteln beschrieben wurden.

## 4.1 Komponenten zur technischen Durchsetzung des Datenschutzkonzeptes

Wie in Abb 4.1 gezeigt, sind bereits Komponenten zur Filterung und Pseudonymisierung der Daten im Grobkonzept vorgesehen. Die Anforderungen an diese Komponenten ergeben sich aus den Anforderungen an das GIDS und dem Datenschutzkonzept. Beispiele für letztere Anforderungen werden im Baustein "B 1.5 Datenschutz" des IT-Grundschutzes vom BSI [1] genannt. Dabei zeigt sich, dass die gesetzlichen Anforderungen wesentlichen Einfluss auf den Betrieb des GIDS haben. Dies betrifft potentiell auch IP-Adressen, die im Grid-Umfeld durch Korrelation mit weiteren Daten Bezug zu Personen haben können. Weiterhin sind alle Daten betroffen, die Anmeldevorgänge von Benutzern bei Grid-Systemen protokollieren. Dies betrifft insbesondere Audit-Daten der Grid Systeme und Anwendungen, die Unix Account- und Group-IDs beinhalten. Zwar exitieren gesetzliche Grundlagen für die Erhebung, Speicherung und den Transport dieser Daten, jedoch muss im GIDS-Umfeld im Einzelfall geklärt werden, ob diese zutreffen.

Als Konsequenz für die technischen Komponenten folgt daraus, dass diese flexibel an die jeweilige Situation angepasst werden müssen. Es muss also ohne großen Aufwand möglich sein, Daten zu anonymisieren oder zu pseudonymisieren, Daten zu filtern und zu löschen. Dadurch kann das GIDS flexibel an die jeweiligen rechtlichen Einschränkungen angepasst werden. Dies kann beispielsweise technisch durch die Vorgabe einer Policy gelöst werden, die entsprechend spezifiziert:

- Welche Daten an den GIDS-Agenten weitergegeben werden und damit übertragen werden.
- Auf welche Datentypen die jeweils passende Transformation zur Pseudonymisierung angewendet wird (z.B. IP-Adressen, Prozess-Identifikatoren, Account- oder Rechnernamen).
- Welche Daten von der Pseudonymisierung betroffen sind.
- Welche Daten gefiltert werden

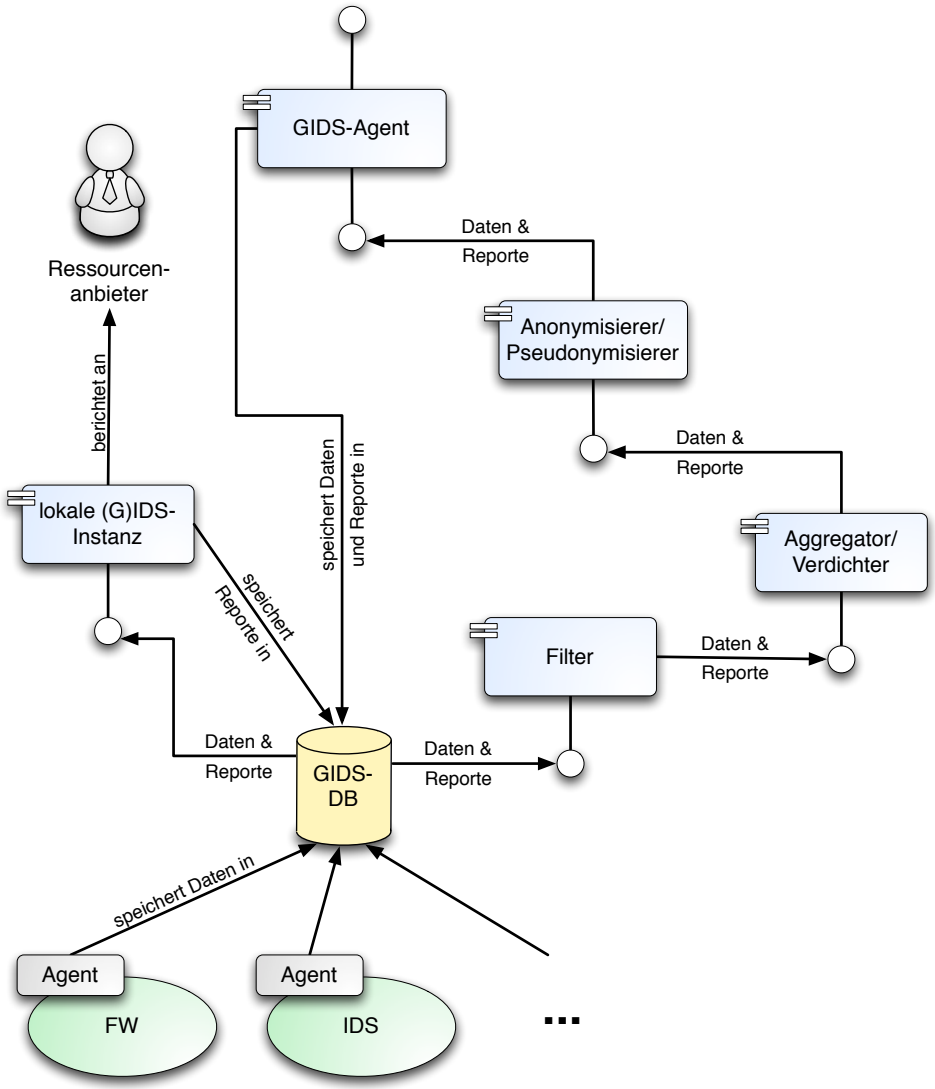


Abbildung 4.1: Architektur auf Seiten einer teilnehmenden Domäne

Zwar werden aufgrund der offenen Architektur des GIDS, die eine offene Architektur für alle IDS-Sensorik bietet, die verarbeiteten Datenarten nicht eingeschränkt. Jedoch wird mit IDMEF ein einheitliches Datenaustauschformat zur Übertragung der Daten eingeführt. Erster Vorteil ist, dass IDMEF die Syntax und Semantik der Daten exakt spezifiziert. Als zweiter Vorteil muss nur ein Datenaustauschformat berücksichtigt werden, obwohl die Daten aus verschiedenen Quellen stammen. Auf dieser Basis läßt sich leicht eine Filterung und Pseudonymisierung der Daten realisieren, indem das Grundgerüst für eine Policy auf der Basis des IDMEF erstellt wird. Inhalt der Policy ist die Vorgabe, welche Daten von welchen IDS-Sensoren exportiert oder anonymisiert werden. Des weiteren bietet es sich an, dass Datenzulieferer eine Filterung auf Basis der Klassifizierung des IDS-Alarms durchführen können. Beispielsweise können auf diese Weise alle Passwort-Rateangriffe gefiltert werden, deren Daten potentiell personenbezogene Daten beinhalten. Als weitere Beispiele kann die Filterung verwendet werden, um alle lokal aufgezeichneten Alarme zu sperren oder nur Alarme mit pseudonymisierten IP-Adressen weiterzuleiten.

Technisch kann dies über eine Spezifizierung einer Policy oder Filter für die Datentypen in den entsprechenden IDMEF-Klassen erfolgen. Beispielsweise lassen sich die folgenden Klassen dafür nutzen:

- Analyzer class
- Source address
- Classification

Da die Syntax und Semantik von IDMEF als XML-DTD spezifiziert ist, kann auch die Policy als XML-Dokument vorliegen.

## Erhebung, Transport und Speicherung der Daten

Die rechtlichen Grundlagen für die Erhebung, Transport und Speicherung der GIDS-Daten wird im Bundesdatenschutzgesetz (BDSG) festgelegt. Zwar kann die Mehrzahl der Daten durch Anonymisierung oder Pseudonymisierung so bereinigt werden, dass weder für den Betreiber noch andere dritte Seiten personenbezogene Merkmale in den Daten identifizierbar sind, jedoch ist dies nicht für jeden Anwendungszweck möglich. Dies kann beispielsweise dann vorliegen, wenn ein betroffener Benutzer zur Lösung eines Sicherheitsvorfalls benachricht werden muss. Allgemein sollten die folgenden Anforderungen berücksichtigt werden:

- **Erhebung der Daten:**
  - Wenn Daten mit personenbezogenen Merkmalen erhoben werden, muss geprüft werden ob diese anonymisiert oder pseudonymisiert werden können, ohne dass die Zielsetzungen bei der Erkennungsleistung gefährdet werden. Dies betrifft auch IP-Adressen in den Grid-Netzen. Aufgrund der Struktur von Grids, die Rechenressourcen transparent zur Verfügung stellen, ist ein Bezug von IP-Adressen der Grid-Systeme zu Personen nur in unwahrscheinlichen Ausnahmefällen zu erwarten. Allerdings gilt dies nicht für alle indirekt am Grid beteiligten Systeme. Beispiel ist ein privater Computer eines Benutzers, der sich von dort an einem Grid Portal anmeldet.
  - Eine rechtliche Grundlage für deren Erhebung der Daten muss vorliegen. Diese kann beispielsweise dann zutreffen, wenn die betroffenen Personen der Erhebung zustimmen oder kein Personenbezug vorhanden ist oder die Erhebung für den sicheren Betrieb des Grids notwendig ist.
- **Transport der Daten:**
  - Die Daten müssen sicher transportiert werden. Dies bedeutet, dass die Vertraulichkeit und Integrität der Daten und die Authentizität der Kommunikationspartner

gewährleistet sein muss. Allerdings sind diese Massnahmen im Grobkonzept in [7] bereits vorgesehen.

- Werden personenbezogene Daten zwischen unabhängigen Partnern ausgetauscht, muss dieser Austausch rechtlich legitim sein. Das gilt insbesondere dann, wenn eine Auftragsdatenverarbeitung vorliegt.

- **Speicherung der Daten:**

  - Auch bei der Speicherung der Daten existieren gesetzliche Regelungen. Hier muss beachtet werden, dass die Dauer der Speicherung begrenzt ist. Werden personenbezogene Daten gespeichert, müssen diese dann gelöscht werden, wenn sie zu den ursprünglich mit der Erhebung erzielten Zwecken nicht mehr erforderlich sind.

## 4.2 Anonymisierung und Pseudonymisierung von IP-Adressen

Für die Anonymisierung und Pseudonymisierung von IP-Adressen sind einige Ansätze vorgeschlagen worden, die sich bereits in der Praxis bewährt haben. Dabei hat die Pseudonymisierung gegenüber der Anonymisierung zwei bedeutende Vorteile. Zuerst kann die Transformation rückgängig gemacht werden. Weiterhin ist die Korrelation von IP-Adressen möglich. Allerdings muss bedacht werden, dass die pseudonymisierte Daten wie Daten mit personenbezogenen Merkmalen behandelt werden müssen, solange die Identitäten noch von einer Person nachvollzogen werden können (siehe Anschnitt 2.1). Allerdings kann mittels der Pseudonymisierung der rechtliche Rahmen für deren Verarbeitung geschaffen werden.

Mit Fan et al. [4] existiert ein Ansatz, der das Präfix der transformierten Adressen bewahrt. Als Folge davon sind zwei transformierte IP-Adressen genau dann im gleichen Sub-Netzwerk, wenn es die wahren IP-Adressen sind. Aus diesem Grund kann beispielsweise nachvollzogen werden, dass zwei angreifende IP-Adressen im gleichen Netzwerkbereich liegen, was speziell bei Denial-of-Service Angriffen wichtig ist. Mit Crypto-PAn [3] existiert eine open-source Implementierung. Allerdings hat dieser Ansatz auch Nachteile. Da die Transformation den Wertebereich von IP-Adressen beibehält, werden zusätzlich Informationen für die Kennzeichnung von transformierten IP-Adressen benötigt. Es muss also explizit zwischen pseudonymisierten und unveränderten Addressen unterschieden werden.

Die Transformation von Crypto-PAn hängt von einem privaten Schlüssel ab. Dieser kann im GIDS entweder global für alle Datenzulieferer gleich gewählt werden; oder jeder Datenzulieferer legt den Schlüssel unabhängig von den anderen frei fest. Ein globaler Schlüssel hat aber den grossen Nachteil, dass die Pseudonymisierung innerhalb der Gruppe der Datenzulieferer durch einen brute-force Angriff gebrochen werden kann.

Kann jeder Datenzulieferer den Schlüssel frei wählen, können Kollisionen der transformierten IP-Adressen auftreten. Werden alle IP-Adressen mit unterschiedlichen Schlüsseln pseudonymisiert, ist eine Korrelation nicht möglich, was zu einer wesentlichen Einschränkung der Performance des GIDS führt. Als Alternative kann die Pseudonymisierung nur für die IP-Adressen im Netzbereich der Datenzulieferer und des GIDS angewandt werden. Dadurch wird ein guter Kompromiss aus dem Schutz der IP-Adressen der Datenzulieferer und GIDS-Systeme einerseits und der Erkennungsleistung und Rückverfolgung von Angriffen andererseits gewährleistet. Ergänzend dazu können IP-Adressen bei Bedarf durch Löschen des letzten Bytes anonymisisert werden.

## Pseudonymisierung der Audit-Daten

Audit-Daten stammen aus verschiedenen Quellen, die typischer Weise auf dem lokalen System erhoben werden. Darunter fallen die Log-Daten des Systems und dort laufender Anwendungen, Log-Daten von Firewalls und anderer spezieller Programme zum erweiterten Audit des Systems (z.B. Linux Audit Daemon). In den Audit-Daten sind unter anderem Account-Namen, Rechnernamen und Prozess-Identifikatoren vorhanden, die Bezug zu Personen beinhalten können.

Grundsätzlich gilt das bereits oben beschriebene Vorgehen, die personenbezogenen Einträge, deren Voraussetzungen zur Erhebung und deren Verwendung zu identifizieren. Da alle Daten im Format IDMEF übertragen werden, reicht es aus, die kritischen Felder in diesem Format zu identifizieren.

Wie bei den IP-Adressen werden rechtliche Einschränkungen für personenbezogene Daten durch Pseudonymisierung oder Filterung berücksichtigt. Insbesondere bei Audit-Daten ist deren Interpretierung nicht ohne Kenntnis des lokalen Systems möglich. Dies betrifft beispielsweise Prozess-IDs und Referenzen auf lokale Dateien. In diesem Fall ist eine Weiterleitung durch den GIDS-Agenten nicht notwendig. Diese Daten können aber nach einem entdeckten Vorfall wertvolle Informationen über das Ausmaß und die ausgenutzte Schwachstelle beitragen. Die Vorfallsbearbeitung erfolgt aber in diesem Fall auf dem lokalen System und setzt keine a-priori Informationen in der GIDS-Datenbank voraus.

In Lee et al. [2] ist ein Ansatz zur Pseudonymisierung von Audit-Daten vorgeschlagen worden. Da dieser allerdings eine zentralen Stelle zur Berechnung der Pseudonyme voraussetzt, ist dieser Ansatz nicht direkt für das GIDS geeignet. Im Gegensatz dazu basiert der Ansatz in [6] auf der Erhebung von Audit-Daten in einer föderierten Umgebung. Aus diesem Grund kann der Ansatz ohne größere Probleme auf die GIDS-Architektur übertragen werden. Die Autoren präsentieren Lösungen für die Pseudonymisierung von Audit-Daten, die unter anderem die für das GIDS wichtigen Anforderungen erfüllen. Zuerst ermöglicht der Ansatz die Korrelation von Daten zwischen mehreren Domänen. Weiterhin wird eine Präfix erhaltende Transformation vorgeschlagen, die sich für Account- oder Rechnernamen eignet. Mathematisch basieren die Transformationen auf einem kryptographisch sicheren Hash-Verfahren, das zur Berechnung der Pseudonyme verwendet wird.

### Aufhebung der Pseudonymisierung

Die Aufhebung der Pseudonymisierung ist immer dann erforderlich, wenn während der Bearbeitung des Vorfalls die Kenntnis der entsprechenden personenbezogenen Daten notwendig ist. Beispielsweise ist dies bei der Kompromittierung eines Benutzer-Accounts oder Zertifikates notwendig. Grundvoraussetzung ist, dass die Seite, die die Daten erhoben hat, dem zustimmt. Weiterhin muss die Person, deren Identität aufgedeckt wird, über diese Schritte informiert werden. Da diese Person oder Personengruppe allerdings mit grosser Wahrscheinlichkeit von dem Vorfall betroffen ist, geschieht dies im Normalfall im Rahmen der Vorfallsbearbeitung.

## 4.3 Anwendung auf das Format IDMEF

Abbildung 4.2 zeigt den detailllierteren Aufbau einer IDMEF-Nachricht. Neben der Erweiterung durch die direkten Subklassen sind ebenfalls die Attribute der einzelnen Klassen eingezeichnet.

Datenschutzrechlich sind insbesondere die Klassen “Source” und “Target” zu beachten, weil sich hier die IP-Adressen des Angreifers und Ziels und Dateinamen, Benutzer-Identitäten und Prozess-IDs beinhaltet sein können. Dies betrifft insbesondere die Unterklassen “Node” und “User”. Allerdings ist nicht auszuschliessen, dass von einer Prozess-ID auf den entsprechenden Benutzer zurückgeschlossen werden kann. Da die Prozess-ID zusätzlich ohne die Informationen des lokalen Systems nutzlos ist, kann diese auch vollständig vor dem Transport gelöscht werden. In Dateinamen (Klasse “File”) kann der vollständige Pfad zu der Datei angegeben werden. Da dieser potentiell Rückschlüsse auf den Benutzer zulassen kann, sollte dieser ebenfalls bereinigt oder vollständig gelöscht werden. Dies hängt aber im Einzelfall von den Daten ab, die ein IDS-Sensor dort einträgt. Eine weitere kritische Klasse ist “AdditionalData”. Hier können Sensoren alle beliebige Daten ablegen, die übertragen werden können. Diese Daten können beispielsweise die rohen Daten enthalten, die vom Sensor aufgezeichnet wurden. Es können also Daten mit Personenbezug in einem unbekannten Format enthalten sein. Aus diesem Grund muss für jeden Sensor einzelnd überprüft werden, welche Daten dort abgelegt werden und ob

die Daten kritisch für den Datenschutz sind.

Die Klasse “Analyzer” beinhaltet Informationen über den IDS-Sensor, der die Angriffsdaten aufgezeichnet hat. Zwar sind in der Klasse keine Merkmale mit Personenbezug zu erwarten, jedoch enthält sie potentiell sicherheitskritische Informationen über den Sensor und die betreibende Seite. Aus diesem Grund bietet es sich an, diese Daten vor dem Transport zu bereinigen oder das IDMEF-Dokument vollständig zu sperren. Analoges gilt für die Klassen “Classification” und “Assessment”, die eine Klassifizierung und Bewertung der IDS-Meldung beinhalten. Aus der Sicht der Datenzulieferer kann die Weitergabe bestimmter Meldungen unerwünscht sein. Dies gilt für Vorfälle, in denen kritische Daten des Datenzulieferers enthalten sein können. Aus diesem Grund ist es sinnvoll, Filtermöglichkeiten auf der Basis der Werte dieser Klassen vorzusehen.

## 4.4 Kooperationsvereinbarung und NDA

Ein Kernpunkt der Architektur des GIDS ist die verteilte Infrastruktur, die eine Kooperation von verschiedenen Seiten zum Erreichen der gemeinsamen Zielsetzungen realisiert. Diese Kooperation basiert auf einer Reihe von Regelungen und Spezifizierungen, die eine geordnete Zusammenarbeit regeln und in der *Kooperationsvereinbarung* zusammengefasst werden. Dies betrifft die folgenden Punkte:

- Was sind die Regelungen der Kooperation?
- Was sind die Rechte und Pflichten des Betreibers?
- Was sind die Rechte und Pflichten der Datenzulieferer?
- Welche Daten werden erhoben und wie werden diese Daten verwendet?
- Wie dürfen diese Daten genutzt werden?

Weiterhin kann eine vertragliche Bindung vom Betreiber und den Datenzulieferern und weiteren Seiten zur Einhaltung des Datenschutzes notwendig sein. Das gilt insbesondere für eine Auftragsdatenverarbeitung bei dem produktiven Betrieb der GIDS-Infrastruktur. Neben den rechtlichen Anforderungen zur Erhebung und Verarbeitung von Daten gibt es seitens der Partner Bedürfnisse zum Schutz kritischer Daten. Beispielsweise dürfen keine sicherheitskritischen Informationen an Dritte weitergegeben werden. Um dies zu verhindern oder zumindest vertragliche Schritte treffen zu können, unterzeichnen die Partner ein *Non-Disclosure Agreement (NDA)*. Dies spezifiziert, welche Daten kritisch sind und gibt Einschränkungen für die Weitergabe und Verwendung dieser Daten an.

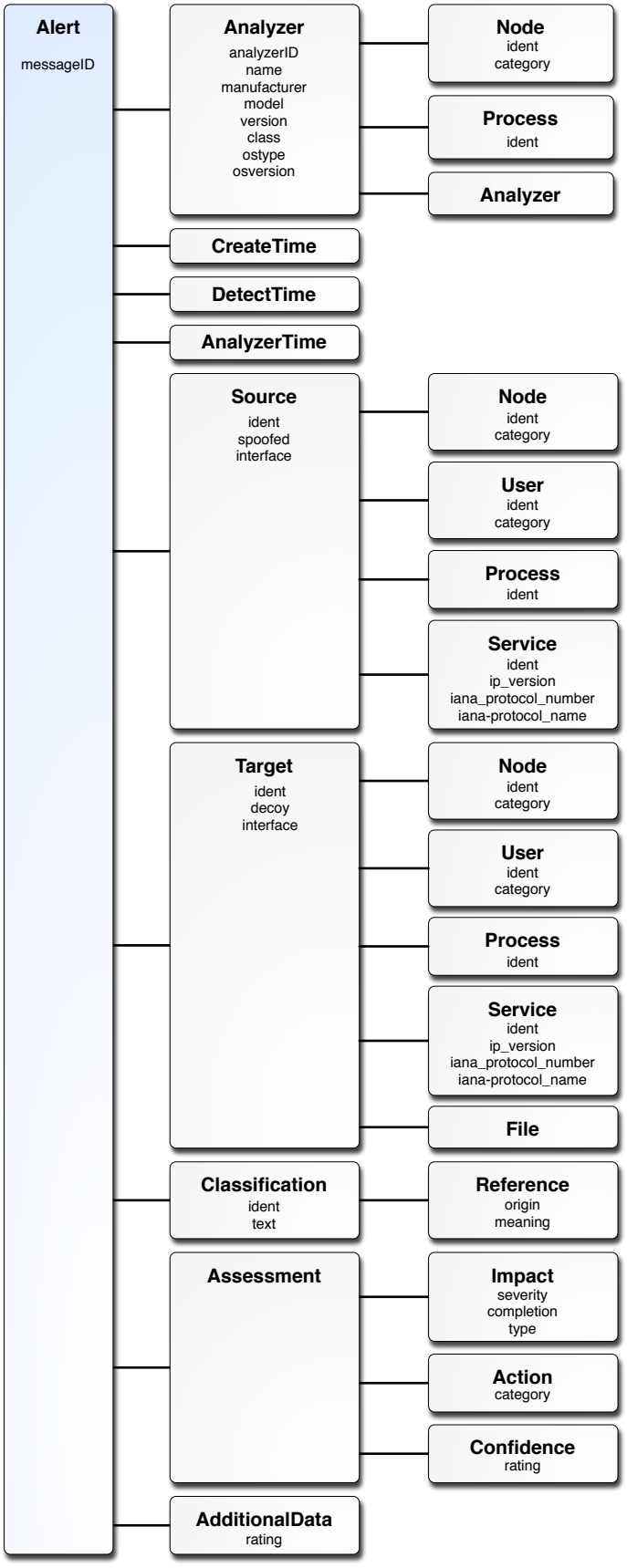


Abbildung 4.2: Das detaillierte IDMEF-Datenmodell der `Alert`-Klasse nach [5]

# Kapitel 5

# Zusammenfassung

In diesem Dokument wurde das GIDS-Datenschutzkonzept beschrieben, das den Schutz der Daten realisiert. Dabei wurden auf der einen Seite die rechtlichen Anforderungen an die Erhebung, Verarbeitung und Speicherung der Daten berücksichtigt. Auf der anderen Seite existieren Anforderungen von Seiten der am GIDS beteiligten Partner zum Schutz der Daten, die miteinbezogen werden mussten. Dies betrifft sowohl die Ergebnisse, die beim Betrieb des GIDS produziert werden als auch die organisatorischen Daten der Datenzulieferer, die für den Betrieb benötigt werden. Beide Arten sind als sicherheitskritisch einzustufen und erfordern Massnahmen zum sicheren Transport, Speicherung und Zugriff.

Im ersten Teil des Dokuments wurden die Grundlagen beschrieben, soweit diese für das Verständnis der Datenschutzkonzeptes notwendig sind. Danach wurde auf die Konzepte zur Ertellung eines Datenschutzkonzeptes eingegangen. Ziel war es, alle Anforderungen zu erfüllen, die sich aus rechtlichen, organisatorischen und technischen Aspekten ergeben. Dabei war es kritisch, einen Kompromiss zu finden, der sowohl den Schutz der Daten berücksichtigt ohne die Erkennungsleistung des GIDS grundsätzlich zu beeinträchtigen. Darauf aufbauend wurde die Realisierung des Datenschutzes beschrieben. Diese beinhaltet die technischen Massnahmen zur Pseudonymisierung der Daten sowie die Aufstellung einer Kooperationsvereinbarung und einer Non-Disclosure Vereinbarung, die eine unberechtigte Weitergabe der Daten verhindert.

# Abbildungsverzeichnis

# Literaturverzeichnis


[1] Baustein 1.5- it-grundschutzkatalog. *http://www.bfdi.bund.de/cln_134/SharedDocs/Publikationen/Orientierungshilfen/Baustein1.5.html*, 2010.

[2] Joachim Biskup and Ulrich Flegel. On pseudonymization of audit data for intrusion detection. In *International workshop on Designing privacy enhancing technologies*, pages 161–180, New York, NY, USA, 2001. Springer-Verlag New York, Inc.

[3] Crypto-Pan Homepage. *http://www.cc.gatech.edu/computing/Telecomm/projects/cryptopan/*, 2010.

[4] Jinliang Fan, Jun Xu, Mostafa H. Ammar, and Sue B. Moon. Prefix-preserving ip address anonymization: measurement-based security evaluation and a new cryptography-based scheme. *Computer Networks*, 46(2):253–272, 2004.

[5] IETF. The intrusion detection message exchange format (idmef). *http://tools.ietf.org/html/rfc4765*, 2007.

[6] Adam J. Lee. A privacy-preserving interdomain audit framework. In *Proceedings of the Workshop On Privacy In The Electronic Society*, 2006.

[7] Helmut Reiser, Nils gentschen Felde, Felix von Eye, Jan Kohlrausch, and Christian Szongott. Anforderungs- und Kritierienkatalog (MS 6). Meilensteinbericht, D-Grid, January 2010.