



Eidgenössischer Datenschutzbeauftragter
Préposé fédéral à la protection des données
Incaricato federale per la protezione dei dati
Swiss Federal Data Protection Commissioner

Elektronische Spuren und Datenschutz

Bern, März 2004

Inhaltsverzeichnis

1. Einleitung.....	3
2. Definitionen	4
2.1 Die verschiedenen Arten elektronischer Spuren.....	4
2.2 Anwendung des Datenschutzes auf elektronische Spuren.....	7
2.3 Transparente Aufzeichnung von elektronischen Spuren.....	7
2.3.1 Mittel gegen mangelhafte Transparenz.....	8
3. Protokolle (Logfiles).....	10
3.1 Der Inhalt	10
3.2 Lebenszyklus.....	11
3.2.1 Aufbewahrungsdauer	13
3.3 Analyse	13
3.3.1 Funktionen im Zusammenhang mit Protokollen.....	13
3.3.2 Analyseinstrumente	15
3.3.3 Analyseresultate	16
4. Schlussfolgerung	17
5. Anhang: Nützliche Adressen für die Konfiguration der Server	18

1. Einleitung

In den vergangenen zehn Jahren hat sich computerunterstütztes Arbeiten rasant entwickelt. Im Bestreben nach einer effizienteren Arbeitsweise und einer gesteigerten Produktivität ist es heute selbstverständlich, einen Grossteil der Arbeit per Computer zu erledigen.

Die Tätigkeiten auf dem Computer ziehen jedoch eine ganze Reihe von Nebeneffekten nach sich. Das vorliegende Dokument behandelt die Problematik der elektronischen Spuren, die von den verwendeten Programmen automatisch und systematisch generiert werden. Jede noch so einfache Tätigkeit – etwa das Schreiben eines Briefes oder das Kopieren eines Objektes – hinterlässt mehrere Spuren, anhand welcher es möglich ist zu rekonstruieren, **wer** wann was gemacht hat. Tatsächlich schreiben fast alle Softwares Logfiles: Bereits eine einfache Suche nach allen Dateien „*.LOG“ oder „*.TMP“ auf der Festplatte genügt, um sich ein Bild über das ganze Ausmass der „zurückgelassenen“ Spuren zu verschaffen.

Die Möglichkeit, unerlaubt auf Personendaten zuzugreifen, ist daher enorm gross. Um unerwünschtes Eindringen in die Privatsphäre zu verhindern, muss der Zugriff auf elektronische Spuren deshalb reguliert werden.

In der vorliegenden Studie werden zunächst die verschiedenen Arten elektronischer Spuren definiert. Dann wird nach einer ausgewogenen Lösung in Bezug auf Erzeugung, Existenz, Analyse, Zugriffsrechte, Speicherung und Zerstörung der elektronischen Spuren gesucht.

Grundsätzlich ist es möglich, Quantität und Qualität der anfallenden elektronischen Spuren zu kontrollieren. Dies setzt jedoch voraus, dass der Systemadministrator die verschiedenen Programme, die Spuren erzeugen, korrekt konfiguriert. Das Ziel besteht darin, über ausreichend Informationen zu verfügen, um eine Tätigkeit bei Bedarf rekonstruieren zu können, ohne jedoch allzu viele Informationen analysieren zu müssen. Die ideale Quantität der Informationen festzulegen ist nicht einfach, weshalb es wichtig ist, nach einer ausgewogenen Lösung zu suchen.

2. Definitionen

2.1 Die verschiedenen Arten elektronischer Spuren

Die vorliegende Studie befasst sich mit elektronischen Spuren, die durch die Tätigkeiten des Benutzers „zurückgelassen“ werden. Die elektronischen Spuren können sich am Benutzerplatz, auf der Netzeinrichtung und auf dem Server befinden. Selbst wenn bestimmte elektronische Spuren keine Personendaten enthalten, so ist es wichtig, kurz auf sämtliche Datentypen einzugehen, die in Verbindung mit einer identifizierten oder einer identifizierbaren Person (Pseudonym) stehen.

Zunächst ist festzuhalten, dass sich elektronische Spuren sehr schnell ansammeln, wenn deren Umfang nicht eingeschränkt wird. Eine solche uneingeschränkte Ansammlung von Personendaten verursachen offensichtlich Kosten für die Verwaltung und die Analyse.

Die wichtigsten uns bekannten Arten elektronischer Spuren sind:

- **Logfiles:** Logfiles sind die bekanntesten elektronischen Spuren. Die meisten Anwendungsprogramme schreiben ein Logfile (Protokoll), in welchem alle nötigen Informationen enthalten sind, um herauszufinden, „**wer** wann was gemacht“ hat. Logfiles können häufig konfiguriert werden, d.h. es ist möglich zu bestimmen, welche Informationen bei welcher Gelegenheit gesammelt werden sollen. Logfiles sind manchmal unerlässlich, beispielsweise wenn es darum geht, sämtliche Änderungen einer besonders schützenswerten Datenbank zu rekonstruieren. Theoretisch können Angestellte am Arbeitsplatz mittels Logfiles überwacht werden, was zu einem Interessenkonflikt zwischen dem Arbeitgeber und den Arbeitnehmern führen kann. Deshalb muss ein Kompromiss in Bezug auf Erstellung, Speicherung, Inhalt und Gebrauch von Logfiles gefunden werden. Ein solcher Kompromiss reicht aber noch nicht aus, um den Umfang der Logfiles zu kontrollieren. Dazu muss die maximale Dauer der Speicherung von Ereignissen in den Logfiles definiert werden. Generell sind zwei Ansätze möglich: Der erste Ansatz besteht darin, die **maximale Dauer** zu definieren, nach welcher das Logfile geschlossen werden soll. In diesem Fall kann oder muss das Logfile bis zum gewünschten Zeitpunkt gespeichert werden. Der zweite Ansatz besteht darin, den **maximalen Umfang** zu definieren. Dabei wird die jeweils älteste Aufzeichnung durch eine neue gelöscht. Die Kombination beider Ansätze ist ebenfalls möglich: Wenn das Logfile den maximalen Umfang erreicht – und nur dann –, werden jene Aufzeichnungen gelöscht, die die maximale Dauer überschritten haben. Logfiles können in zwei Hauptkategorien unterteilt werden: Die **User Logfiles** beinhalten die Spuren der Tätigkeiten des Benutzers. Die meisten Logfiles sind dieser Kategorie zuzuordnen; sie werden normalerweise von den Verantwortlichen für Logfiles-Analysen und dem Beauftragten für Datensicherheit benutzt (vgl. 3.3.1). In der Regel sind die Systemadministratoren für die Analysen der User Logfiles verantwortlich, was ihnen eine gewisse Macht gibt. Hingegen ist ihnen auch bewusst, dass ihre Administratorentätigkeit im Prinzip in **Administrator Logfiles** vollständig aufgezeichnet wird. Diese Logfiles müssen von einer unabhängigen Person, meistens vom Datenschutzberater, analysiert werden. In diesem Zusammenhang sind noch die **Data Protection Logfiles** zu erwähnen, die die automatisierte Bearbeitung von besonders schützenswerten Personendaten oder Persönlichkeitsprofilen protokollieren, wenn die präventiven Massnahmen den Datenschutz nicht gewährleisten können (Art. 10 VDSG). Im Rahmen der Datenbankverwaltungssysteme (DBMS) kann eine Anwendung so konfiguriert werden, dass bestimmte Tätigkeiten der Benutzerinnen und Benutzer sowie der Administratoren direkt in **Database Logtables** aufgezeichnet werden. Der Vorteil besteht in der unabhängigen Verwaltung dieser Aufzeichnungen: Bei einer Datenbank können alle

Aufzeichnungen nach einer bestimmten Zeit auf viel flexiblere und effizientere Art und Weise gelöscht werden. Dank einer Sprache der 4. Generation – wie etwa SQL – sind auch die Analysemöglichkeiten um einiges besser. Bleibt noch die Familie der **Eavesdrop / Wiretap Logfiles**, die von Programmen erzeugt werden, um die Benutzer regelrecht auszuspionieren. Die Softwares der letzten Generation machen es möglich, die Gesamtheit der ausgeführten Tätigkeiten zu überwachen. Die abgefangenen Daten werden lokal gespeichert oder diskret, beispielsweise mittels einer geheimen elektronischen Post (E-Mail) an den Angreifer übermittelt. Letzterer kann somit alle besuchten Internetseiten und den Inhalt aller eingegangenen und ausgegangenen Nachrichten einsehen wie auch die **Erfassung** (keylogger) des Benutzers abfangen, also auch die eingegebenen Passworte! Es ist sogar möglich, einen „Film“ aus einer Folge von Bildschirmkopien zu erhalten, die in regelmässigen Abständen gemacht werden. De facto ist also eine regelrechte Videoüberwachung des Bildschirms möglich. **Diese Art der Überwachung ist verboten:** Das Eindringen in die Privatsphäre ist unrechtmässig, weil es gegen den Grundsatz von Treu und Glauben verstösst und stellt eine Verletzung des Verbots der systematischen Überwachung des Verhaltens dar. In bestimmten HTML-Dokumenten kann aber auch ein unsichtbares Pixel (web bug/beacon, clear/invisible GIF) untergebracht sein, das unerkant Informationen über das angeschaute Dokument an Dritte übermittelt, einschliesslich Angaben zur Uhrzeit des Zugriffs und zum benutzten Computer und Browser. Ein im lokalen Cookie enthaltener Kennzeichner ermöglicht es dann, die verschiedenen Informationen zu verknüpfen und das Profil des Surfers zu erstellen!

- **„Temporäre“ Dateien:** In der Regel speichern die Anwendungen zusätzliche Informationen in den temporären Dateien (*.TMP), die sich auf der Festplatte des Benutzers befinden, beispielsweise im Profil. Werden Personendaten bearbeitet, sind diese ebenfalls in den temporären Dateien zu finden. Im Übrigen befindet sich der virtuelle Speicher des Betriebssystems in den Swapfiles, die sogar nach Abschalten des Computers noch „zurückgelassene“ Anwendungsdaten enthalten können. Wie bei den Internet-Browsern kann es manchmal vorkommen, dass versteckte Dateien nicht automatisch gelöscht werden. Daher ist es möglich herauszufinden, welche Internetseiten besucht wurden. Es ist offensichtlich, dass damit ein Potenzial gegeben ist, in die Privatsphäre einzudringen. Zur Kategorie der „temporären“ Dateien gehören auch die **Cookies**. Diese kleinen – mehr oder weniger temporären – Dateien werden während der Benutzung bestimmter Internetseiten auf der Festplatte gelagert. Ein Cookie kann beispielsweise den Namen des Benutzers und das Passwort enthalten, das den Zugang zu einem externen Konto ermöglicht. Dies kann zwar die Benutzung erleichtern, birgt aber auch Risiken in sich. Daher ist es wichtig, dass die Speicherung von Cookies vom Benutzer kontrolliert wird. Kommt noch hinzu, dass das Löschen der verschiedenen Spuren des Surfers nicht ausreicht, um diese ganz zu beseitigen, da Indexdateien vorhanden sind (Datei Index.dat), in denen die Spurennamen bestehen bleiben! Eine dieser Indexdateien beinhaltet ausserdem Informationen für AutoComplete-Einträge in Formulare, wahlweise auch Benutzernamen und Passwort des Webkontos!
- **Registry (.INI files):** In diesem Dateityp sind verschiedene Informationen in Bezug auf die Konfiguration oder Benutzung einer Anwendung enthalten. Früher waren diese Parameter in den Erweiterungsdateien *.INI gespeichert. Heute werden diese Parameter in einer gemeinsamen lokalen Datenbank gruppiert, die als **Registry** bezeichnet wird. Das Registry kann leider Spuren enthalten, die von den Spywares hinterlassen werden (Spyware ist in der Regel ein Programm, das Informationen ohne Wissen des Benutzers erfasst).
- **Listen der zuletzt geöffneten Dokumente:** Die meisten Anwendungen (u.a. Word, Excel, Winzip, Internet Explorer) und auch die gegenwärtigen Betriebssysteme listen die zuletzt geöffneten Dokumente auf (MRU: *most recently used*). Diese Listen können sich sowohl im Registry als auch in den Temporärdateien oder in einer Reihe von Links zu

Dateien im zugeordneten Verzeichnis befinden. Meistens kann der Benutzer die Anzahl (0-n) der in der Liste aufscheinenden Dokumente festlegen – auf jeden Fall sollte es aber möglich sein, die Liste auf einfache Weise zu löschen.

- **Zwischenablage (clipboard):** Obwohl dieser temporäre Speicher beim Abschalten des Computers automatisch entleert wird, so ist es trotzdem möglich, ihn heimlich zu lesen. Deshalb sollten alle besonders schützenswerten Inhalte unverzüglich gelöscht werden, insbesondere vor dem Surfen im Internet.
- **Papierkorb:** Dateien, die gelöscht werden, werden vom Betriebssystem oder von der Anwendung an bestimmte Orte verlegt (in den Papierkorb oder in einen Ordner der gelöschten Elemente), die eine spätere Wiederherstellung ermöglichen. Der Benutzer glaubt, die Dateien gelöscht zu haben; in Wirklichkeit sind sie jedoch noch immer verfügbar. Dies kann zwar bei einem irrtümlichen Löschen der Dateien sehr nützlich sein, doch sensibilisierte Benutzer befolgen zu Recht den guten Rat, die Dateien endgültig zu löschen und den Umweg über den Papierkorb zu vermeiden! Jeder Papierkorb kann im Übrigen geleert werden, doch handelt es dabei im Prinzip nur um ein logisches, nicht um ein physisches Löschen.
- **Logisches / physisches Löschen:** Löscht eine Benutzerin oder ein Benutzer eine Datei, die auf einem Datenträger gespeichert ist, so wird zwar die Datei-Zuordnungstabelle vernichtet, der Inhalt bleibt jedoch bestehen, bis der freigewordene Speicherplatz von neuen Dateien besetzt wird. Eine logisch gelöschte Datei kann somit mittels eines geeigneten Verfahrens (undelete) wiederhergestellt werden. Für das endgültige, physische Löschen einer Datei stehen bestimmte Dienstprogramme (u.a. wipe, shred) zur Verfügung. Eine Defragmentierung der Laufwerke hat dank der Zusammenlegung des Speicherplatzes auch eine gewisse Vernichtung der Restdateien zur Folge. Im Übrigen reicht das einfache Formatieren des Laufwerks nicht aus, um die Wiederherstellbarkeit der Dateien zu verhindern; einzig das mehrmalige Überschreiben der Sektore mit vordefinierten „binären Schemata“ geben diese Sicherheit. Die endgültige Vernichtung der gelöschten Objekte in einer Datenbank (u.a. *tables, rows*) ist nicht eindeutig geklärt.
- **Persönliche Archive:** Um die hauptsächlichsten Laufwerke zu entleeren, transferieren die Benutzerinnen und Benutzer manchmal bestimmte „alte“ Objekte in ein sekundäres Laufwerk, von dessen Standort (lokaler Datenträger oder Netz) die Sicherheit und der Schutz der archivierten Daten abhängt. In diesem Fall spricht man von einem persönlichen Datenarchiv, das de facto eine elektronische Spur darstellt, die gemäss einer angemessenen Datenschutzpolitik behandelt werden muss. Namentlich wird empfohlen, den Zugang zu solchen Archiven mit einem Passwort zu schützen oder die Dateien unter einem verschlüsselten Format zu speichern.
- **„Backups“:** Backups sind Sicherheitskopien, die sowohl Spurendateien als auch Dateien enthalten, deren Original vom Benutzer gelöscht worden sein kann. Die Sicherungsdateien (u.a. .BAK, .WBK), mittels welcher die vorletzte Version eines Dokuments gespeichert werden kann, stellen eine besondere Form von Backup dar, die nicht vergessen werden darf. Es besteht somit das Risiko, dass dort eine grosse Anzahl von Personendaten – einschliesslich elektronischer Post – gefunden werden kann. Die Notwendigkeit für eine effizientere Politik im Umgang mit Backups ist daher offensichtlich. Insbesondere gilt es stets zu bedenken, dass die in den Backups enthaltenen Personendaten nach einer festgelegten Aufbewahrungsdauer ebenfalls gelöscht werden müssen (vgl. 3.2.). Ausserdem bergen Backups ein zusätzliches Risiko: da sie sich oft auf separaten Informationsträgern befinden, müssen sie nicht nur gegen Umweltrisiken wie Wasser und Feuer geschützt werden, sondern auch gegen Deliktrisiken wie Diebstahl oder unerlaubten Zugriff seitens Dritter.

2.2 Anwendung des Datenschutzes auf elektronische Spuren

An dieser Stelle bedarf es einer Einführung in die Grundsätze des Bundesgesetzes über den Datenschutz (DSG). Jede elektronische Spur, die Personendaten beinhaltet, stellt eine Datensammlung im Sinne des DSG dar. Bei der Bearbeitung von Personendaten einschliesslich elektronischer Spuren sind folgende Punkte zu beachten:

- **Rechtmässigkeit, Treu und Glauben:** Die betroffenen Personen sind über den genauen Zweck der Datensammlung und über die Bearbeitung jeglicher elektronischer Spur in Kenntnis zu setzen. Dabei ist nicht nur der Zweck anzugeben, sondern es sind auch Angaben zu machen über den **Inhaber** einer Datensammlung (für deren Funktion vgl. 3.3.1) wie auch über Art und Verwendung der Analysen, die für jede einzelne Spur durchgeführt werden. Nur unter diesen Bedingungen kann eine tatsächliche **Transparenz** des Aufzeichnungsvorgangs erzielt werden; andernfalls muss die Datensammlung dem Eidgenössischen Datenschutzbeauftragten (EDSB) angemeldet werden!
- **Verhältnismässigkeit:** Die Aufzeichnung elektronischer Spuren muss nach dem Grundsatz der Verhältnismässigkeit erfolgen, das heisst, dass einzig jene Bereiche bearbeitet werden dürfen, die dem Zwecke dienen, der den betroffenen Personen angegeben wurde. Konkret geht es darum, möglichst wenige oder gar keine Daten zu sammeln oder zu analysieren.
- **Auskunftsrecht:** Jede Person ist berechtigt, von vom Inhaber einer Datensammlung Auskunft darüber zu verlangen, ob Daten über sie bearbeitet werden. Falls dies zutrifft, hat der Inhaber der Datensammlung der gesuchstellenden Person Einsicht in alle sie betreffenden Daten zu gewähren und ihr den Zweck der Bearbeitung sowie die Kategorien der bearbeiteten Personendaten, die an der Sammlung Beteiligten und der Datenempfänger mitzuteilen. Die Beteiligung an den Kosten beträgt maximal 300 Franken, wenn der antragsstellenden Person in den zwölf Monaten vor dem Gesuch die gewünschten Auskünfte bereits erteilt wurden oder wenn die Anfrage ein hohes Arbeitsvolumen verursacht. Der Inhaber der Datensammlung kann wegen überwiegender Interessen die Auskunft verweigern, einschränken oder aufschieben, muss jedoch den Grund dafür angeben.

2.3 Transparente Aufzeichnung von elektronischen Spuren

Alle elektronischen Spuren, die Personendaten enthalten, stellen Datensammlungen im Sinne des DSG dar. Aus diesem Grund sind die betroffenen Personen oder der EDSB über das Vorhandensein dieser Spuren zu informieren (vgl. 2.2).

In Bezug auf die Transparenz ist Ersteres vorzuziehen. Für jede Sammlung elektronischer Spuren müssen Unternehmen den Benutzer vor allem über folgende Punkte informieren:

- **Zweck** der Spurensammlung (vgl. 2.2). Diese Information ist von entscheidender Bedeutung: Laut DSG darf eine Sammlung ausschliesslich für einen bestimmten Zweck verwendet werden, und dieser muss rechtmässig sein.
- Enthaltene **Informationen** (vgl. 3.1). Besondere Aufmerksamkeit erfordern Informationen, auf Grund welcher eine Person direkt oder indirekt (beispielsweise mittels eines Pseudonyms) identifiziert werden kann.
- **Aufbewahrungsdauer** der Online- und Offline-Spuren (siehe Kapitel 3.2.1). Während dieses Zeitraums sind die betroffenen Personen berechtigt, von ihrem Auskunftsrecht Gebrauch zu machen.

- Art der **ausgeführten Analysen** (vgl. 3.3): Es geht nicht nur darum, Angaben über den Zweck zu machen, sondern auch über die Häufigkeit, den Empfänger und die Form des Ergebnisses (anonymisiert, pseudonymisiert oder personalisiert), über das Verfahren, das vorgesehen ist, um missbräuchliche Datenanalysen zu vermeiden, sowie über die vorgesehenen Massnahmen für den Fall, dass ein Missbrauch entdeckt oder eine missbräuchliche Analyse durchgeführt wird. In diesem Zusammenhang dürfen diejenigen Analysen nicht vergessen werden, die mittels einer Kombination von mehreren Spuren gemacht werden.
- **Lokalisierung** (Server? Netz? Client?): Es muss darüber informiert werden, ob es sich um eine Spur handelt, die vom Benutzer oder vom Unternehmen kontrolliert wird. Bei den Spuren, die vom Unternehmen kontrolliert werden, sind die Benutzer über die jeweilige Politik hinsichtlich der Spurenvernichtung nach einer Archivierungsperiode (vgl. 3.2.1) und der Datensicherung in Kenntnis zu setzen.
- **Zugriff** auf Spuren: Grundsätzlich sind die Systemadministratoren befugt, Spuren zu sammeln, einzusehen, zu analysieren und zu löschen. Die Spuren der Administrator-tätigkeiten werden einer neutralen und unabhängigen Person anvertraut, in der Regel dem Datenschutzberater. An dieser Stelle sei nochmals erwähnt, dass jede betroffene Person ihr Recht auf Auskunft über ihre eigenen Personendaten geltend machen kann.

Für jede einzelne Analyse ist die Aufbewahrungsdauer anzugeben. Die Vernichtung erfolgt spätestens dann, wenn die ursprünglichen Spuren vernichtet sind (vgl. 3.2.1).

Die Transparenz der Erstellung von Logfiles ist nicht nur ein Erfordernis für Unternehmen, die solche Logfiles sammeln, sondern auch für Unternehmen, die Software verkaufen. Es wäre wünschenswert, wenn die Softwares nur die absolut notwendigen – oder gar keine – Logfiles schreiben würden. Das Schreiben von Logfiles sollte als Option im Programm vorhanden sein, die – wenn nicht anders programmiert – deaktiviert bleibt. Falls die Erstellung von Logfiles systematisch erfolgt, **müssen** die Hersteller ihre Kundschaft über das Vorhandensein dieser Logfiles informieren und klar auf deren **Inhalt, Lokalisation, Persistenz** und **Zweck** hinweisen.

2.3.1 Mittel gegen mangelhafte Transparenz

Die meisten Spuren, die auf dem Client gesammelt werden, unterstehen der Kontrolle des Benutzers; sie oder er sollte sich bemühen, sie zu vernichten. Dies ist jedoch kompliziert und selbst Experten auf diesem Gebiet bekunden Schwierigkeiten, alle Spuren zu identifizieren. Doch es gibt spezifische Software für die Erkennung, Sperrung und Vernichtung dieser parasitären Spuren. Da letztere mehrere Megabytes Speicherplatz auf dem Computer beanspruchen können – was sowohl bei den Benutzern als auch beim Unternehmen meistens unerwünscht ist – wird eine regelmässige Vernichtung empfohlen. Es ist im Interesse des Unternehmens, die nötige Software zur Verfügung zu stellen, um eine Ausbreitung unkontrollierbarer Tools, die negative Auswirkungen haben können, zu verhindern (etwa der Konflikt mit einem Standardprogramm oder die Vernichtung nützlicher Spuren). Angesichts der Tatsache, dass manche Spuren nützlich und erwünscht sind, muss entschieden werden, welche Spuren mittels einer präzisen Konfiguration der betreffenden Software zu vernichten sind.

Der Selbstschutz verläuft in drei Phasen :

- **Sperre**: Am besten geschützt ist, wer Invasionen vorbeugt, indem sie abgeblockt werden, bevor sie den anvisierten Computer erreichen. Auf diese Weise können unerwünschte Spuren vermieden werden.

- **Erkennung:** Manchmal ist es nicht möglich, Invasionen abzublocken. Daher müssen die erzeugten Spuren zuerst erkannt und dann eventuell vernichtet werden. Dies kann beispielsweise bei temporären Daten aus dem Internet oder bei noch unbekanntem Viren der Fall sein.
- **Vernichtung:** Gelingt es, die Spuren zu erkennen, müssen diese erst noch vernichtet werden können. Mittels Spurenerkennungsprogrammen können diese im Prinzip vernichtet werden. Bestimmte Probleme, wie beispielsweise das logische und nicht physische Löschen der Spurendaten, bleiben aber bestehen.

Auf dem Internet ist eine grosse Anzahl lizenzfreier Software (freewares) für die Bekämpfung von Spuren zu finden. Ein guter individueller Schutz wird – abgesehen von den „üblichen“ Antivirus-Programmen und persönlicher Firewalls – gewährleistet durch einen **Trace Eraser**, einen **Cookie Manager**, eine **Antispyware** (eventuell einen Antibanner) und einen **File Wiper**. Auf den Websites <http://www.staff.uiuc.edu/~ehowes/main.htm> und <http://www.securitynews.cc> werden beispielsweise verschiedene Tools kostenlos zur Verfügung gestellt.

3. Protokolle (Logfiles)

In diesem Kapitel werden wir uns zunächst mit den verschiedenen Phasen und Aspekten im Lebenszyklus elektronischer Spuren befassen.

3.1 Der Inhalt

Bei den Protokollen gibt es zwei wesentliche Aufzeichnungsarten:

- jene, die eine Person *direkt identifizieren* lassen oder beinahe, beispielsweise mittels einer E-Mail-Adresse oder einer Benutzeridentifikation (UserID).
- jene, die eine Person *indirekt identifizieren* lassen, meistens auf Grund eines Pseudonyms wie z. Bsp. einer IP-Adresse.

Ein typisches Protokoll enthält folgende Angaben:

- **Wann:** Datum und Uhrzeit des Ereignisses
- **Wer:** In der Praxis handelt es sich oft um die Benutzeridentifikation (UserID) oder um die IP-Adresse des Computers. Mittels dieser Pseudonyme ist die formelle Identifizierung des Benutzers oder wenigstens des benutzten Computers relativ einfach. Die Verbindung zwischen Computer, Benutzeridentifikation und Benutzer beruht auf anderen Elementen (Passworte); diese sind missbrauchs anfällig.
- **Was:** Dieses Aufzeichnungsfeld ist spezifisch und ändert sich bei jeder Spur. Zu den typischen Beispielen gehören die Adresse einer besuchten Internetseite (URL), Absender und Empfänger von elektronischer Post oder ein (gelungener oder misslungener) Versuch, sich mittels eines Passwortes Zugang zu geschützten Daten zu verschaffen.
- **Bedeutung/Typ:** Nicht alle Aufzeichnungen sind gleich wichtig. Beispielsweise ist ein gelungener Versuch, ein Passwort zu ändern, weit weniger wichtig als ein Fehler während desselben Vorgangs. Wenn man um den Typ weiss, erleichtert dies die Analyse.
- **Andere Angaben:** Dabei handelt es sich um eine Vielzahl verschiedener Informationen. Darin enthalten sind Hinweise auf das benutzte Betriebssystem, die verwendete Software (z.B. welcher Browser benutzt wurde), deren Version etc.

Ein Protokoll kann in Form eines zweidimensionalen Schemas dargestellt werden, das aus verschiedenen Spalten besteht (aufgezeichneter Teil), in welchen die Angaben über jedes Ereignis auf einer Zeile festgehalten werden. Es gibt noch andere Versionen von Protokollen, die viel weniger strukturiert und daher schwieriger zu analysieren sind.

Beispiel eines Browser-Logfiles:

```
#Date: 2003-05-28 06:33:28
#Fields: date time c-ip cs-username s-ip s-port cs-method cs-uri-stem cs-uri-query sc-status cs(User-Agent)
2003-05-28 06:33:28 192.168.0.100 - 192.168.10.10 80 GET / - 401 Mozilla/4.0+(compatible);+MSIE+6.0;+Windows+NT+5.0)
2003-05-28 13:04:30 192.168.0.10 DOMAINE\Utilisateur 192.168.10.10 80 GET /ListePrix.html - 200
Mozilla/4.0+(compatible);+MSIE+5.01;+Windows+NT+5.0)
2003-05-28 13:04:35 192.168.100.100 DOMAINE\Utilisateur 192.168.20.200 80 GET /index.html - 304
Mozilla/4.0+(compatible);+MSIE+5.01;+Windows+NT+5.0)
```

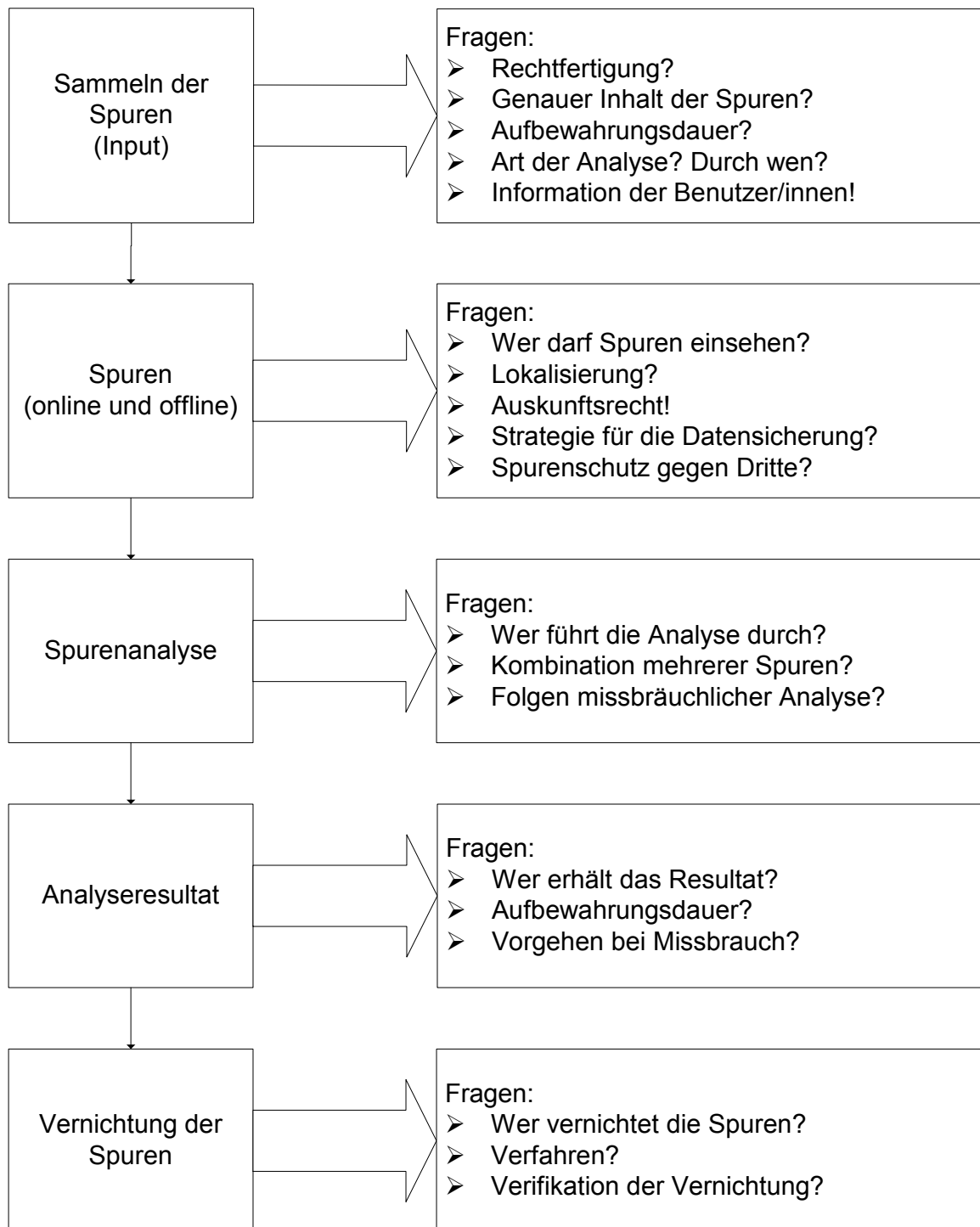
Beispiel eines E-Mail-Logfiles:

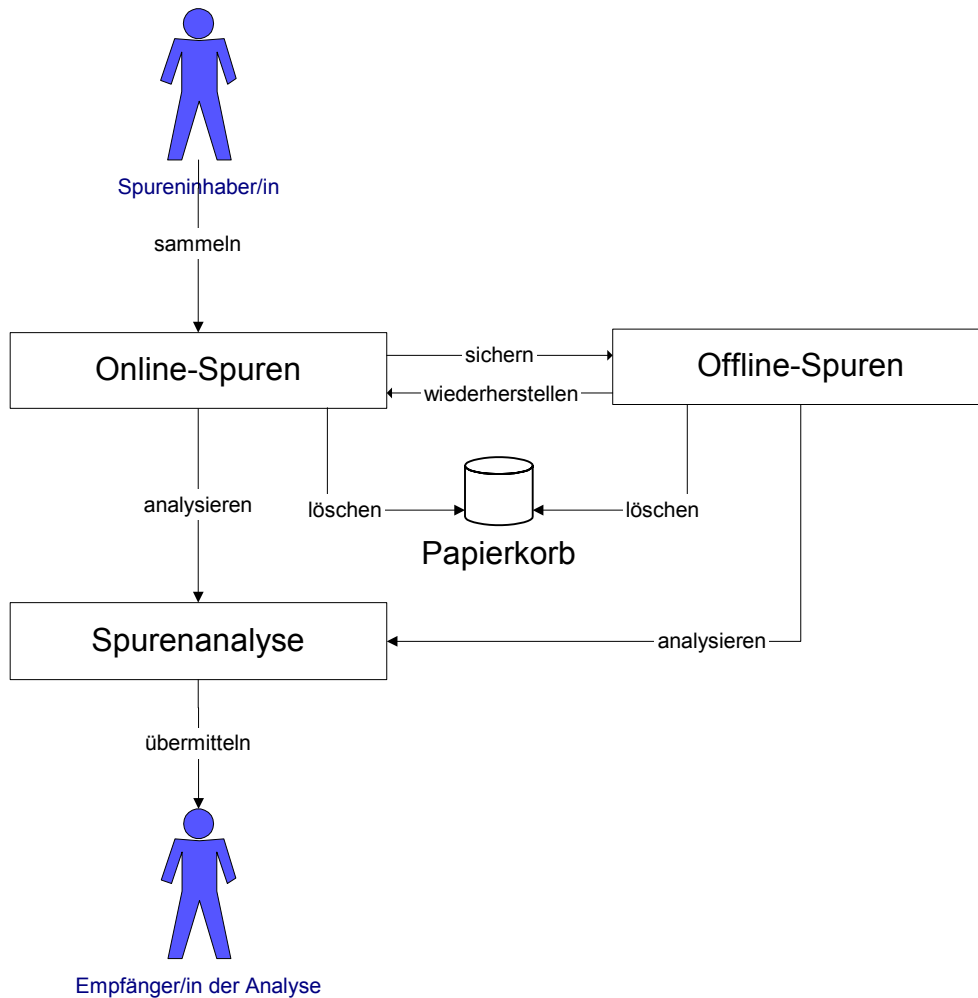
```
2003/02/24 15:52:44 192.168.44.44 mail from untel@entreprise.ch to unautre@fournisseur.ch successfully
2003/02/24 15:52:44 192.168.44.45 mail from unautre@fournisseur.ch to untel@entreprise.ch successfully
2003/02/24 16:38:12 192.168.44.44 mail from untel@entreprise.ch to unautre@fournisseur.ch successfully
2003/02/24 16:40:18 192.168.44.45 mail from unautre@fournisseur.ch to untel@entreprise.ch successfully
```

Die Protokolle können als Grundtext (Format ASCII) oder in einem besonderen und manchmal benutzerdefinierten Format gespeichert werden. In letzterem Fall wird für das Lesen der Protokolle eine Software benötigt; wenn mehrere Protokolle kombiniert werden müssen, sind diese in ein anderes Format zu konvertieren (vgl. 3.3.2).

3.2 Lebenszyklus

Der Lebenszyklus eines Protokolls kann mit folgendem Diagramm dargestellt werden:





Die gebräuchlichen Anwendungen schreiben Protokolle über die Tätigkeiten des Benutzers. Die Protokolle werden auf dem Server, dem Netzwerk und dem Client gespeichert. Da sie direkt konsultiert werden können (mit den entsprechenden Zugriffsrechten, vgl. 2.2 und 3.3.1), werden diese Datentypen als Online-Protokolle bezeichnet. Diese Protokolle dienen normalerweise der Sicherheitsanalyse und der Kontrolle über die Einhaltung der Richtlinien des Unternehmens (vgl. 3.3).

Um zu verhindern, dass der Umfang der Datenprotokolle zu gross wird (cf.2.1), muss er mittels einer geeigneten Strategie eingeschränkt werden: Nach der Aufbewahrungsfrist werden die Online-Protokolle durch Überschreiben oder regelmässiges Löschen entfernt. Ab diesem Zeitpunkt sind die Protokolle nur noch in der Datensicherung vorhanden, ein direkter Zugriff ist jedoch nicht mehr möglich. Die Protokolle können nur eingesehen werden, wenn man Zugang zum betreffenden Datenträger hat, weshalb sie als Offline-Protokolle bezeichnet werden. Die darin enthaltenen Informationen werden benötigt, um eine Situation zu rekonstruieren oder um die Ursache eines Absturzes oder die für einen Missbrauch verantwortliche Person zu ermitteln.

Nach Ablauf der Aufbewahrungsfrist (vgl. 3.2.1) müssen alle Protokolle endgültig gelöscht werden. Von nun an wird die Datensicherung erst wirklich problematisch: Führt ein Unternehmen eine vollständige Datensicherung aller Informationen eines Servers auf demselben Datenträger durch, so muss es innerhalb bestimmter Fristen den Datenträger ändern, um die überholten Daten selektiv zu löschen. Dies ist offensichtlich ein komplizierter Vorgang. Deshalb ist es wünschenswert, die Datensicherung entsprechend den unterschiedlichen Anforderungen auf verschiedenen Datenträgern durchzuführen. Infolgedessen werden getrennte Sicherungen für das Betriebssystem, die Benutzerdaten und

die Protokolle vorgenommen. Bei jeder Datensicherung kann die Aufbewahrungsdauer wie auch die Zugriffspolitik variieren.

3.2.1 Aufbewahrungsdauer

Bei der Aufbewahrungsdauer ist zu unterscheiden zwischen den Spuren, die vom Unternehmen, und denjenigen, die vom Benutzer kontrolliert werden:

- **Vom Benutzer kontrollierte Protokolle und Spuren:** Ein Clientcomputer ist normalerweise einem Benutzer zugewiesen. Diese Person kann daher einen Grossteil der Personendaten, die sich auf dem Computer befinden, selbst kontrollieren. Die Verantwortung des Unternehmens beschränkt sich darauf, den Benutzer über das Vorhandensein dieser elektronischen Spuren auf dem Clientcomputer zu informieren (vgl. 2.2, 2.3 und 3.1). Da man nicht erwarten kann, dass alle Benutzer wissen, wie sie die eigenen Spuren löschen können, sollte das Unternehmen, aus Gründen der Transparenz, diesbezügliche Anweisungen geben und die dafür notwendigen Tools zur Verfügung stellen.
- **Vom Unternehmen kontrollierte Protokolle und Spuren:** Die Benutzer haben keinen direkten Zugang zu den Spuren, die vom Unternehmen kontrolliert werden. Deshalb muss das Unternehmen es als seine Pflicht betrachten, die Benutzer u.a. über das Vorhandensein (vgl. 2.2, 2.3 und 3.1), die Aufbewahrungsdauer (online und offline) und die Vernichtungspolitik zu informieren und darlegen, wer zu welchen Daten Zugang hat und wozu. Das Unternehmen hat also für vollständige Transparenz zu sorgen. Im Übrigen müssen die Benutzer den Inhalt der sie betreffenden Daten überprüfen können. Bei den Protokollen, die vom Unternehmen kontrolliert werden, sollte die Archivierungsdauer für Online-Daten nicht mehr als 1 Monat und bei Offline-Daten nicht mehr als 6 Monate betragen. Die von der Datenschutzgesetzgebung vorgesehenen Protokolle müssen jedoch 1 Jahr lang in revisionsgerechter Form aufbewahrt werden.

3.3 Analyse

Angesichts der grossen Zahl von Protokollen, ihres Umfangs und ihrer spezifischen Eigenheiten ist eine direkte, manuelle Analyse nicht möglich. Damit die Gesamtheit der Protokolle anhand einer Kombination verschiedener Protokolle beurteilt werden kann – ohne jedoch andere potenziell wichtige Informationen aus den Augen zu verlieren –, ist es unerlässlich, spezielle Analyseanwendungen einzusetzen. Diese verhindern den direkten Zugriff der für die Analyse verantwortlichen Personen auf die Protokolle und schützen die Personendaten daher gegen jeden unnötigen direkten Zugriff. Es ist klar, dass die unterschiedlichen Speicherformate der Protokolle und insbesondere die unterschiedlichen Zielsetzungen der Analysen den Einsatz verschiedener spezifischer Tools für die Analyse von Protokollen nötig machen, wobei auch die direkte, manuelle Analyse in Ausnahmefällen nicht auszuschliessen ist.

3.3.1 Funktionen im Zusammenhang mit Protokollen

Die Protokollanalysetools geben ihren Anwendern offensichtlich viel Macht. Diese Analysetools können die Tätigkeiten des Benutzers überwachen. Deshalb dürfen sie auch nur einer sehr beschränkten Anzahl von Personen zur Verfügung gestellt werden. Damit verhindert werden kann, dass eine einzige Person zu viel Macht erhält, sollte das Recht, solche Tools zu benutzen, auf mehrere Personen verteilt sein. Die ideale Lösung bestünde darin, die Protokollanalyse ausschliesslich vom Inhaber eines Protokolls oder von der stellvertretenden Person durchführen zu lassen. Um schliesslich auch die Tätigkeiten der für die Analyse verantwortlichen Personen zu überprüfen (Ad-hoc-Funktion), werden deren Tätigkeiten ebenfalls von den Analyseprogrammen aufgezeichnet, die dann von einer neutralen Person

analysiert werden. Betrachten wir nun die verschiedenen Funktionen im Zusammenhang mit Protokollen:

- **Unternehmen:** Das Unternehmen legt Richtlinien für die Benutzung der Informatiktools fest und trifft Massnahmen, mit denen die Einhaltung der Richtlinien überprüft werden kann. Das Unternehmen ist Inhaber aller bei ihm gesammelten Protokolle, da es direkt oder indirekt deren Ziel und Inhalt bestimmt. Im Falle eines Missbrauchs muss das Unternehmen (Direktion, Personalbüro oder vorgesetzte Person) den Täter identifizieren und allfällige Sanktionen bekannt geben. Dazu muss sich das Unternehmen auf die Analysen der Spuren der unter Verdacht stehenden Person stützen können.
- **Benutzer:** Jede Person, die an einem Computer arbeitet, ist Benutzer. Die Person muss über das Vorhandensein der Spuren, die sie betreffen, informiert sein und ist befugt, die unter ihrer Kontrolle stehenden Spuren zu verwalten. Ebenso kann sie ihr Recht auf Auskunft über die sie betreffenden Spuren, die das Unternehmen sammelt, geltend machen.
- **Administrator (u.a. Systeme, Netze, Datenbanken):** Aus Gründen des Zugriffsrechts müssen praktisch alle Protokollanalysen von einem zuständigen Administrator durchgeführt werden. Diese Person besitzt offensichtlich viel Macht, weshalb ihre Tätigkeiten ebenfalls protokolliert und von einem Dritten (vom Datenschutzberater) analysiert werden müssen, damit Missbräuchen vorgebeugt werden kann.
- **Beauftragter für Datensicherheit:** Der Beauftragte für Datensicherheit hat den Auftrag, die Gesamtheit der administrativen Aufgaben auf ihre Kohärenz zu überprüfen, damit das anvisierte Schutzniveau erreicht werden kann. Bei Unklarheiten, Pannen oder Störungen kann er oder sie die Hilfe der Administratoren in Anspruch nehmen, um die Situation zu analysieren und eventuelle Schwachstellen raschmöglichst zu beseitigen.
- **Datenschutzberater/in:** Der Datenschutzberater muss zu allen Spuren der administrativen Tätigkeit Zugang haben, um jegliche missbräuchliche Überwachung von Seiten der Administratoren auszuschliessen. Sie oder er fungiert zudem als neutrale Vertrauensperson zwischen der Direktion und den Benutzern.

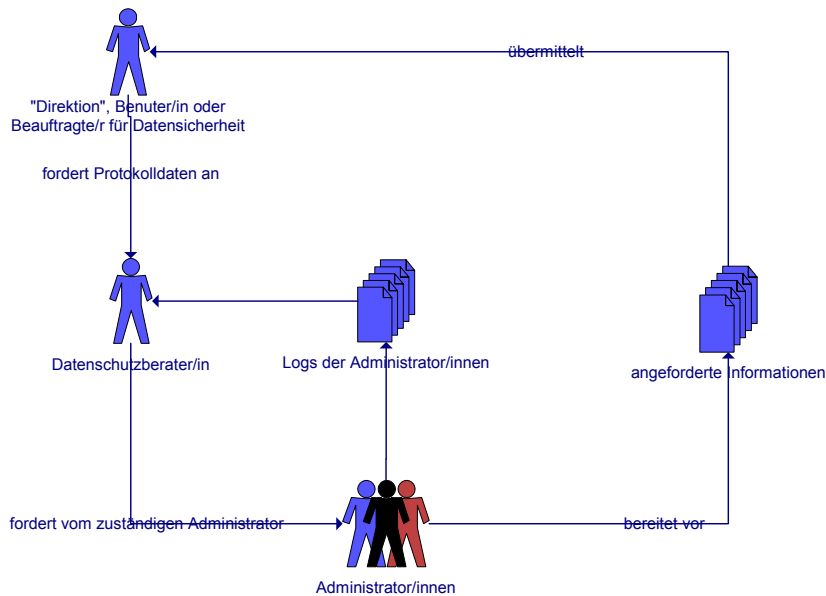
	Anonyme Analysen	Analysen unter Pseudonym ohne Missbrauch (oder Missbrauchsverdacht)	Personenbezogene Analysen ohne Missbrauch (oder Missbrauchsverdacht)	Personenbezogene Analysen der Tätigkeiten der Administrator-/innen	Personenbezogene Analysen bei Missbrauch (oder Missbrauchsverdacht)	Analysen bei Pannen
Benutzer	Z		A, Z			
Datenschutzberater	Z	Z		D, Z	Z	
Beauftragter für Datensicherheit	Z				Z	A, Z
Administrator	D, Z	D	D		D	D
Unternehmen	I, A, Z	I, A, Z	I	I, Z	I, A, Z	I

I: Inhaber/in der Datensammlung (entscheidet über das Sammeln)

A: fordert Analyse an (Anforderer)

D: Durchführer der Analyse

Z: Zugang zu den Resultaten der Analyse



3.3.2 Analyseinstrumente

Wie in Kapitel 3.1 bereits erwähnt, können die Protokolle verschiedene Formate haben. Unterschiede kann es beim Datum-Format geben (Bsp.: 29.08.2003, 08/29/03, 20030829 etc.), beim Zeit-Format (Bsp.: 23:12, 11:12PM), bei der Bezeichnung der verschiedenen Felder ("IP Address" oder einfach "IP"), bei der Kodierung der Information (ASCII, Unicode u.a.) etc. Die Liste möglicher Unterschiede ist tendenziell unendlich lang.

Die Aufgabe, die unterschiedlichen Protokollformate zu konvertieren, damit sie vergleichbar und kombinierte Analysen möglich werden, kann sich als ziemlich schwierig und entsprechend teuer erweisen.

Wie so oft greift man aus Kostengründen auf bereits existierende spezifische Tools zurück. Microsoft beispielsweise bietet «LogParser» an, der die Arbeit erheblich erleichtert, indem er für die verschiedenen Analysearten kleine Batches in Pseudo-SQL anbietet (Bsp.: die meistgelesenen Seiten des eigenen Site: >LogParser -i:FS "SELECT TOP 20 Path FROM C:\inetpub\wwwroot*.*)"¹.

LogParser unterstützt die gängigsten Formate (IISW3C, IIS, IISMSID, NCSA, ODBC, BIN, URLSCAN, HTTPERR, EVT, TEXTWORD, TEXTLINE, CSV, W3C und FS), womit bei ihm das Problem der Formatunterschiede weitgehend gelöst ist.

Sofern der Umfang der Daten es erlaubt, gibt es die Möglichkeit, dass die Protokolle, die für eine bestimmte Analyse erforderlich sind, in eine eigentliche Datenbank importiert werden. In diesem Fall lassen sich die gesuchten Informationen mittels klassischer SQL-Suche herausziehen. Der Import von Protokollen in eine Datenbank stellt wiederum das Problem der Formatumwandlung. Das "Resource Kit" von Windows 2000 enthält z. Bsp. das "Event Log Query Tool" (elogdmp.exe), mit dem lokale oder entfernte Protokolle angeschaut, gefiltert oder analysiert und in Datenbanken wie Access oder SQL-Server importiert werden können.

Mehr Informationen gibt es unter: <http://is-it-true.org/nt/nt2000/atips/atips137.shtml>

¹ Mit einer etwas elaborierteren Abfrage kann man z. Bsp. die zehn meistbesuchten Sites pro Tag erfragen: LogParser "SELECT TOP 10 TO_STRING(TO_TIMESTAMP(date, time), 'yyyy-MM-dd') AS Day, cs-uri-stem, COUNT(*) AS Total FROM C:\WINNT\system32\LogFiles\w3svc1\ex*.log GROUP BY Day, cs-uri-stem ORDER BY Total DESC" -rtp:-1

3.3.3 Analyseresultate

Da die Protokolle sehr oft sehr umfangreich und schwierig zu interpretieren sind, ist es unumgänglich, die interessierenden Fakten daraus herauszuziehen oder das ganze in eine aussagekräftige Statistik umzumünzen. Die Art und Weise, wie man Resultate präsentiert, ist also sehr wichtig. Die Resultate müssen eindeutige Angaben enthalten über den Auftraggeber, die untersuchte Periode, die Gültigkeitsdauer, die einschlägigen Protokolle, das Datum und die analysierende Person, das Ziel (vgl. 2.2 und 2.3) und die Empfänger (gegebenenfalls mit dem Vermerk "vertraulich").

Zu Protokollen von identifizierten oder identifizierbaren Personen gibt es drei mögliche Arten von Analysen:

- **Anonymisierte Analyse:** Es gibt keinerlei Hinweise oder Daten, auf Grund deren es möglich wäre zu rekonstruieren, wer was gemacht hat. Vielmehr handelt es sich hier um eine statistische Analyse, mit der das Verhalten einer Gruppe oder die Unterschiede im Verhalten verschiedener Gruppen illustriert werden können.
- **Analyse unter Pseudonym:** Dabei handelt es sich um eine Analyse, die die Aktivitäten identifizierbarer, jedoch nicht identifizierter Individuen aufzeigt. Das gewählte Pseudonym darf natürlich die Identität der wirklichen Person nicht verraten. Eine Identifizierung wäre einzig über eine Zuordnungsliste möglich; diese muss jedoch gegen jeglichen nicht autorisierten Zugriff konsequent geschützt sein.
- **Personalisierte Analyse:** Dabei handelt es sich um eine Analyse der Aktivitäten einer identifizierten Person. Eine solche Analyse ist unbedingt vertraulich zu behandeln, jeglicher Zugang für nicht befugte Dritte muss verunmöglicht werden, und die Aufbewahrungsfrist muss festgelegt sein.

4. Schlussfolgerung

Wie wir gesehen haben sind die Protokolle nur ein Teil, wenn auch der wichtigste, der elektronischen Spuren. Solche Spuren können auf der Ebene des zentralen Servers eines Unternehmens, auf der Ebene der Netzwerkinstallationen wie auch auf der Ebene der einzelnen Clients gesammelt werden. Der Trend zu immer systematischerem Sammeln von Spuren der Aktivitäten der Benutzer muss unbedingt gebrochen werden; das Ziel solchen Sammelns und die Verhältnismässigkeit müssen kritisch hinterfragt werden. Natürlich gibt es hier kein Patentrezept; wir können lediglich folgende Grundsätze des Datenschutzes empfehlen, die auf jede Form elektronischer Spuren zutreffen:

- Anwendung nur dann, wenn andere, gleich wirksame, aber weniger einschneidende Massnahmen nicht zur Verfügung stehen.
- Beachtung der geltenden gesetzlichen Bestimmungen (DSG, StG, URG, StGB etc.).
- Das Ziel der Datenbearbeitung muss vom Inhaber der Datensammlung klar vorgegeben sein.
- Vollständige Transparenz gegenüber den betroffenen Personen.
- Es werden nur diejenigen Daten gesammelt, die zur Erreichung des gesteckten Ziels unbedingt nötig sind.
- Strikte Einhaltung der Aufbewahrungsfristen.
- Klar definierte und kontrollierte Analysevorgänge.
- Vollständiger Schutz aller bearbeiteten Daten.
- Trennung der involvierten Funktionen.
- Die betroffenen Personen sind auf ihr Auskunftsrecht hinzuweisen.

Diese Vorgaben mögen formalistisch und übertrieben ehrgeizig erscheinen; es ist jedoch absolut möglich, sie in jedem spezifischen Unternehmensumfeld in die Praxis umzusetzen. Der Nutzen, den man aus ihrer Beachtung zieht, ist enorm: er drückt sich direkt aus in einer Reduktion der Kosten, einer Verbesserung des Image und einer Steigerung des Vertrauens.

5. Anhang: Nützliche Adressen für die Konfiguration der Server

Die modernen Betriebssysteme verfügen über eine eindruckliche Zahl von Parametern, die es jedem Administrator erlauben, sie seinen spezifischen Bedürfnissen entsprechend einzustellen. Die Sicherheitsbedürfnisse können sich von Unternehmen zu Unternehmen stark unterscheiden. Deshalb ist es unmöglich, eine Standardkonfiguration für den Datenschutz zu empfehlen.

Mit einer einfachen Internetrecherche findet man eine Vielzahl von nützlichen Adressen mit detaillierten Erklärungen, wie die verschiedenen Sicherheitsparameter eines Servers zu konfigurieren sind. Die Administratoren werden daraus die Konfiguration wählen können, die ihren Bedürfnissen am ehesten entspricht.

Nachfolgend eine nicht abschliessende Liste von Links auf nützliche Informationsquellen:

<http://www.intersectalliance.com/projects/Win2kConfig/index.html>

<http://nsa1.www.conxion.com/win2k/download.htm>

<http://www.microsoft.com/windows2000/technologies/security/default.asp>

<http://www.windowsecurity.com/>

<http://www.microsoft.com/downloads/details.aspx?FamilyId=9964CF42-E236-4D73-AEF4-7B4FDC0A25F6&displaylang=en>

<http://www.datenschutz-bremen.de/technik/nt/quellenhinweise.htm> (D)

<http://helpdesk.rus.uni-stuttgart.de/~rustomfi/download/hardenW2K12.pdf> (D)