Der Landesbeauftragte für den Datenschutz Mecklenburg-Vorpommern
Der Landesbeauftragte für den Datenschutz und für das Recht auf Akteneinsicht Brandenburg

Data Warehouse und Data Mining im öffentlichen Bereich

- Datenschutzrechtliche und -technische Aspekte -

(Stand: August 2002)

Autoren: Gabriel Schulz, Andreas Waldenspuhl (Der Landesbeauftragte für den Datenschutz Mecklenburg-Vorpommern), Sven Hermerschmidt (Der Landesbeauftragte für den Datenschutz und für das Recht auf Akteneinsicht Brandenburg)

Inhalt

1. Einleitung

1.1 Ausgangslage

Mit der ständig zunehmenden Leistungsfähigkeit der in Wirtschaft und Verwaltung eingesetzten Informations- und Telekommunikationstechnik wächst die Menge der automatisiert gespeicherten *personenbezogenen Daten* unaufhaltsam. Aus allen Lebensbereichen werden Daten beispielsweise durch Nutzung von Chipkartensystemen und neuen Kommunikationsmedien preisgegeben und sowohl in Privatunternehmen als auch im Bereich der öffentlichen Verwaltung gespeichert. Datensparsamkeit und Datenvermeidung spielen noch immer und zum Teil ganz bewusst eine untergeordnete Rolle.

In Unternehmen und Behörden wächst das Interesse, das gesammelte Datenmaterial effektiver als bisher zu nutzen. Dabei wird davon ausgegangen, dass zu einem Betroffenen (Kunden) nicht durch Bewertung einzelner Daten ein aussagefähiges Gesamtbild entsteht, sondern erst durch die Analyse der Gesamtheit aller verfügbaren Daten einer Person und ihrer Beziehungen zueinander.

Das *Data Warehouse* ermöglicht diese neue Betrachtungsweise der Daten, indem in ihm alle im Unternehmen bzw. in der Behörde verfügbaren Daten nach bestimmten Kriterien sortiert gespeichert und zur Analyse und Auswertung bereitgehalten werden. Bisher unbekannte Zusammenhänge zwischen Einzeldaten sollen mit Hilfe des so genannten *Data Mining* aus der Gesamtheit des Datenbestandes erkannt werden, um aus diesen – die Aussagekraft der einzelnen Angaben meist deutlich übersteigenden – Informationen vor allem wirtschaftliche Vorteile für die speichernde Stelle erzielen zu können. Dabei ist beabsichtigt, dem Manager und künftig wohl bald auch dem Behördenleiter aus der Fülle des Datenmaterials nur die strategisch wichtigen Informationen – möglichst in ansprechender und visuell einprägsamer Form – darzustellen, ohne dass er den Ballast der täglich anfallenden operativen Daten wahrnimmt und dabei auf die ständige Mitwirkung von Informatikern und Statistikern angewiesen ist. Die sich rasant entwickelnde Leistungsfähigkeit von Hardware und Datenbanktechnologien bietet dabei immer bessere Möglichkeiten zur Analyse und zur verständlichen, grafischen Präsentation von Datenbeständen.

1.2 Ziel der Abhandlung

Auch wenn Data-Warehouse-Konzepte bisher meist nur im Bereich der Privatwirtschaft anzutreffen sind, erscheint die Untersuchung dieser neuen Technologie aus datenschutzrechtlicher und -technischer Sicht auch durch die für den öffentlichen Bereich zuständigen Datenschutzbeauftragten geboten, um für die Diskussion über Anwendung dieser Konzepte in der Verwaltung gewappnet zu sein. Künftige Anwender sollen den möglichen Nutzen für Wirtschaft und Verwaltung sorgfältig gegenüber den Gefahren für die Privatsphäre des Einzelnen abwägen können und im Ergebnis dieses Prozesses die jeweils erforderlichen und angemessenen datenschutzfreundlichen Technologien einsetzen.

Den Schwerpunkt der rechtlichen Bewertung bilden die für den öffentlichen Bereich geltenden Datenschutznormen, wobei aber versucht wird, allgemeingültige Aussagen herauszuarbeiten, welche auch auf den privatwirtschaftlichen Bereich übertragbar sind.

1.3 Gang der Darstellung

Zunächst werden die wichtigsten, mit dem Konzept des Data Warehouse zusammenhängenden Begriffe definiert und erläutert (2. Kapitel) sowie die erforderlichen Softwaregrundlagen im Überblick vorgestellt (3. Kapitel). Anschließend werden verschiedene Überlegungen dargestellt, die den Einsatz des Data Warehouse in der Verwaltung interessant erscheinen lassen

(4. Kapitel) und konkrete Realisierungsbeispiele aufgeführt (5. Kapitel). Den Hauptteil der Arbeit bilden die datenschutzrechtliche Bewertung des Data-Warehouse-Konzeptes (6. Kapitel) und die damit einhergehenden Empfehlungen für seine datenschutzgerechte Umsetzung (7. Kapitel). Am Ende werden die Ergebnisse zusammengefasst und die Konsequenzen für das weitere Vorgehen gezogen (8. Kapitel).

2. Begriffe

2.1 Operative Datenbasis

Den Bestand an Daten einer *Daten verarbeitenden Stelle*, der ständig zur unmittelbaren Aufgabenerfüllung benötigt wird, bezeichnet man als operative Datenbasis. Diese Daten resultieren aus den Einzeloperationen des Tagesgeschäfts und kennzeichnen den Ablauf oder Status einzelner Geschäfts- bzw. Verwaltungsvorgänge. Aus datenschutzrechtlicher Sicht korrespondiert die operative Verarbeitung eng mit der Aufgabe der Daten verarbeitenden Stelle und hat nicht die Gewinnung strategischer Aussagen zum Ziel.

2.2 Data Warehouse

Im Data Warehouse werden alle von einer Daten verarbeitenden Stelle jemals erhobenen und gespeicherten Daten nach einem einheitlichen System geordnet und jederzeit verfügbar zum Abruf für unterschiedliche Zwecke bereitgehalten. Die Daten kommen aus allen Bereichen der Daten verarbeitenden Stelle. So können im nicht-öffentlichen Bereich beispielsweise neben Vertriebsinformationen und Kommunikationsdaten auch die Transaktionsdaten der Zahlungsvorgänge sowie weitere Informationen über die Kauf- und Zahlgewohnheiten des einzelnen Kunden einfließen. In der öffentlichen Verwaltung wäre denkbar, dass sämtliche Vorgänge aller Organisationseinheiten einer Stelle (z. B. Ämter einer Kommune oder Abteilungen eines Ministeriums usw.) betroffenen- oder bearbeiterbezogen recherchierbar gespeichert werden.

Im Gegensatz zur operativen Datenbasis, bei der Daten i. d. R. nach Abschluss des Vorgangs schwer zugreifbar archiviert oder gelöscht werden, hält das Data Warehouse Daten über einen langen Zeitraum umfassend für Recherchezwecke vor.

2.3 Management-Informations-System

Die Idee des Data Warehouse ist nicht neu. Bereits in den 70er Jahren wurde der Versuch unternommen, Managemententscheidungen auf der Basis der gespeicherten Daten einer Daten verarbeitenden Stelle zu treffen. Die damals propagierten Management-Informations-Systeme konnten die an sie gestellten Erwartungen jedoch nie erfüllen, weil im Gegensatz zum Data-Warehouse-Konzept strategische Daten direkt aus der operativen Datenbasis gewonnen werden sollten.

2.4 Extraktionswerkzeuge

Das Zusammenführen aller operativen Daten einer Daten verarbeitende Stelle in das Data Warehouse ist keinesfalls eine triviale Aufgabe. Hierfür sind spezielle Softwarekomponenten erforderlich, die als Extraktionswerkzeuge bezeichnet werden. Sie haben u. a. die Aufgabe, Rohdatenstrukturen zu analysieren, Daten zu selektieren und für die Zusammenführung in das Data Warehouse vorzubereiten. Nach erfolgter Zusammenführung werden diese Werkzeuge für umfangreiche Prüfungen und ggf. Korrekturen genutzt. Erschwert wird das Zusammenführen der Daten insbesondere dadurch, dass in einer Daten verarbeitenden Stelle verschiedene Computersysteme und -programme eingesetzt werden, die ganz unterschiedliche Daten-

strukturen erzeugen. Aus diesem Grund enthalten auf dem Markt angebotene Data-Warehouse-Systeme bis zu 170 verschiedene Importschnittstellen.

2.5 Data Mart

Aus den o. g. Gründen ist das allumfassende, unternehmens- bzw. behördenweite Data Warehouse zur Zeit kaum realisierbar. In der Praxis erfolgt die Zusammenführung von Datenbeständen aber schon erfolgreich auf der Ebene von Abteilungen oder Geschäftsbereichen einer Daten verarbeitenden Stelle. Ein derart verkleinertes, themenspezifisches Data Warehouse wird als Data Mart (Marktplatz) bezeichnet.

2.6 Data Mining

Die Zusammenführung aller Daten im Data Warehouse bzw. im Data Mart und die danach erfolgte Trennung von der operativen Datenbasis ist Voraussetzung für die weitere Informationsgewinnung. Unter dem Begriff des Data Mining werden alle Verfahren subsumiert, mit denen die scheinbar zusammenhanglosen Daten des Data Warehouse nach bisher unbekannten, wissenswerten Zusammenhängen durchsucht werden. Dabei kommt es nicht mehr darauf an, dass der Verdacht eines Zusammenhangs durch kluge Fragestellungen an das System bestätigt werden soll. Vielmehr soll der Computer selbstständig nach unbekannten, bisher verborgenen Mustern oder Trends suchen.

3. Softwaregrundlagen

Softwarebasis des Data Warehouse ist i. d. R. ein auch als Standardsoftware verfügbares *Datenbank-Management-System* (DBMS). Die Menge der unterschiedlichen Daten, die im Data Warehouse vorgehalten werden (bei Warenwirtschaftssystemen können Größenordnungen von mehr als 20 Terabyte erreicht werden), hat auch zur Folge, dass die Zahl der Nutzer eines solchen Systems groß ist. Diese Nutzer sind in der Mehrheit jedoch keine IT-Fachleute, die detaillierte Kenntnisse über DBMS und Abfragesprachen haben. Deshalb ist der erste Schritt zur effektiveren Gestaltung der Informationsgewinnung die Bereitstellung von *grafischen Benutzeroberflächen* und leicht bedienbaren *Visualisierungswerkzeugen*. Auch Nicht-Experten werden damit in die Lage versetzt, u. a. durch intelligente Menüführung Abfragen zu formulieren und die Ergebnisse mit Hilfe grafischer Darstellungen auszuwerten.

Ein Data Warehouse soll jedoch mehr leisten als ein konventionelles DBMS. Deshalb werden *Data-Mining-Tools* angeboten, die in den Daten verborgene Regeln entdecken, bisher nicht erkannte Zusammenhänge aufdecken oder bestimmte Sachverhalte voraussagen. Es können Unregelmäßigkeiten aufgespürt und „Ausreißer“ sichtbar gemacht werden. Zur Datenanalyse werden neue Softwaretechnologien wie *neuronale Netze*, *Case Based Reasoning*, *Regelinduktion* oder *Clustering* genutzt.

Der Grund für den Einsatz solcher Technologien für das Data Warehouse ist die Lernfähigkeit der darauf basierenden Softwareprodukte. Die Software trainiert sich mit einem teilweise selbstorganisierenden Lernprozess anhand vorhandener Daten selbst. Es werden weiche Fragestellungen beispielsweise zur Generierung von Hypothesen sowie deren Validierung eingesetzt.

4. Data Warehouse in der Verwaltung

Obwohl Data-Warehouse-Konzepte zur Zeit fast nur mit privaten Unternehmen in Zusammenhang gebracht werden, stammt die Idee aus der Verwaltung. Bereits Anfang der 60er Jahre wurden Vorstellungen entwickelt, dass der Daseinsvorsorgestaat die allgemeine Bedürfnis-

befriedigung durch seine Verwaltungstätigkeit sicherstellen solle. Die einsetzende Nutzung von Computern in der Verwaltung führte zu den ersten großen Datensammlungen. Man hoffte, dass mit diesen Daten auch die Verhaltensweisen in der Gesellschaft erforscht werden können. Die zur Verfügung stehende Rechentechnik war zu der Zeit jedoch nicht annähernd in der Lage, dieses hoch gesteckte Ziel zu erreichen.

Durch die rasante Entwicklung der Informations- und Kommunikationstechnik und die zunehmende Nutzung von Data-Warehouse-Technologien durch die Privatwirtschaft kommt die Nutzung des Data-Warehouse-Konzeptes auch für den öffentlichen Bereich ins Gespräch.

So gibt es inzwischen in nahezu allen Verwaltungen auf der Ebene von Bund, Ländern und Gemeinden umfangreiche Bestrebungen zur Verwaltungsmodernisierung. Um Verwaltungsabläufe effektiver und damit letztlich kostensparender zu gestalten, ist in aller Regel eine fundierte Datenbasis erforderlich, aus der die möglichen Instrumente für eine Änderung der Abläufe entwickelt werden können. Dabei liegt es nahe, Daten, die von verschiedenen Organisationseinheiten einer Behörde zu ganz unterschiedlichen Zwecken erhoben, gespeichert und verarbeitet werden, in einem Data Warehouse zusammenzuführen. Die so zusammengeführten und im Data Warehouse gespeicherten Daten könnten dann mit Werkzeugen des Data Mining nach allen möglichen Richtungen mit dem Ziel ausgewertet werden, Effektivierungspotentiale zu entdecken, auf die bisher niemand gekommen ist. Mit der Verwaltungsmodernisierung in engem Zusammenhang steht die Einführung von Kosten-Leistungs-Rechnung in die öffentliche Verwaltung, bei der ebenfalls Data-Warehouse-Technologien genutzt werden könnten.

Unabhängig von Projekten der Verwaltungsmodernisierung ist es möglich, dass Behörden für die Erstellung von behördeninternen Statistiken auf Data-Warehouse-Systeme und –technologien zurückgreifen.

Schließlich ist auch nicht auszuschließen, dass im Sozialversicherungsbereich Data-Warehouse-Systeme eingesetzt werden, um auch hier mit Hilfe von Data Mining bisher nicht bekannte Zusammenhänge zu erkennen und zu nutzen.

In allen Fällen stellt sich aus datenschutzrechtlicher Sicht die Frage, ob ein Data Warehouse in der öffentlichen Verwaltung mit den Grundsätzen des Volkszählungsurteils des Bundesverfassungsgerichts aus dem Jahre 1983 vereinbar ist. Es wird vor allem darauf ankommen, ob die von den Datenschutzgesetzen von Bund und Ländern geforderte Zweckbindung bei der Verarbeitung – insbesondere der Nutzung – personenbezogener Daten gewährleistet werden kann. Zusätzliche Einschränkungen sind dabei bei besonders sensiblen Daten, wie beispielsweise Sozial- oder Personaldaten, gegeben.

5. Realisierungsbeispiele

Data-Warehouse-Konzepte sind keineswegs nur Zukunftsmusik. Hinter Begriffen wie *„Semantic Computing“*, *„Knowledge Processing“* oder *„Computer Animation and Simulation“* verbergen sich ganz konkrete Verfahren und am Markt bereits angebotene Produkte, die als Bausteine für ein Data Warehouse dienen. Große Softwarefirmen bieten neben kompletten Data-Warehouse-Produkten und Einzelbausteinen mit wohlklingenden Namen wie *„Business Information Warehouse“* oder *„sphinxVision“* auch bereits Schulungen zu solchen Produkten an.

Was demnächst auch in Deutschland erwartet werden kann, zeigt ein Blick in die USA. Der „gläserne Kunde“ ist beispielsweise in einer amerikanischen Warenhauskette, die bereits Supermärkte in Deutschland eröffnet hat, längst Wirklichkeit. Der Konzern weiß genau, welche Produkte in welchen Filialen zu welchem Preis angeboten werden können, weil er detailliert registriert, welcher Kunde welche Waren in seinem Einkaufswagen hat.

Inzwischen werden allerdings in vielen Ländern bei Anbietern und Nutzern von Data Warehouse auch die Gefahren dieser Technologien erkannt. Das ist vor allem darauf zurückzuführen, dass zumindest in den Industriestaaten eine zunehmende Sensibilität für den Schutz der Privatsphäre zu beobachten ist und in vielen Ländern – insbesondere in Europa, aber auch in Kanada – Data-Warehouse-Systeme an rechtliche Grenzen stoßen. Deshalb unternehmen auch bedeutende Anbieter von Data Warehouses zunehmend den Versuch, datenschutzfreundliche Technologien (*privacy enhancing technologies – PET*) in ihre Systeme zu implementieren.

Die Funktionsweise eines Data Warehouse unter Nutzung von PET könnte dabei in etwa wie folgt aussehen:

Zunächst werden eine Vielzahl von sehr detaillierten personenbezogenen Kundendaten erhoben und im Data Warehouse gespeichert. Innerhalb des Data Warehouse werden dann die identifizierenden Daten von den übrigen Daten getrennt. Somit werden die Daten innerhalb des Data Warehouse pseudonymisiert. Das Data Warehouse bietet nun die Möglichkeit, die Fülle der Daten mit Hilfe von Data Mining auszuwerten und auf unterschiedlichen Ebenen (sog. *views*) zusammenzuführen. Wichtig ist, dass bei dieser Art der Nutzung der Daten keine Zugriffe auf die identifizierenden Daten möglich sind. Die Wiederherstellung eines Personenbezuges soll also weitgehend ausgeschlossen werden.

Die eingesetzten PET konzentrieren sich dabei auf die Beschränkung des Zugangs zu den Daten, vermeiden aber nicht die Erhebung und Speicherung personenbezogener Daten. Das Unternehmen stellt für sein Data Warehouse darüber hinaus verschiedene *tools* zur Verfügung, die es ermöglichen, dass z. B. auf Wunsch des Kunden einzelne Datensätze physisch gelöscht werden können oder die betroffenen Personen einen lesenden Zugriff auf alle ihre eigenen Daten erhalten können. Ob solche *tools* in Anspruch genommen werden, hängt aber vom Käufer des Data Warehouse ab.

Obwohl die bisher genannten Beispiele aus der privaten Wirtschaft stammen, ist auch zu beobachten, dass staatliche Stellen den Nutzen dieser neuen Technologie mehr und mehr erkennen. So ist beispielsweise die Bundesanstalt für Finanzdienstleistungsaufsicht dabei, für die Börsenaufsicht Softwareagenten einzuführen, die aus täglich etwa einer halben Million Meldungen über Käufe und Verkäufe von Aktien und Optionsscheinen verdächtige Transaktionen herausfiltern sollen, um verbotene Insidergeschäfte aufzudecken.

Ein weiteres konkretes Beispiel ist der von den Betriebskrankenkassen (BKK) in Deutschland geschaffene Datenverbund BKK-InfoNet mit einer gemeinsamen Datenbank für alle BKK. Diese Datenbank dient nicht nur dazu, die üblichen Datenverarbeitungsvorgänge der Krankenkassen, insbesondere zur Abrechnung mit den Leistungserbringern und zur Verwaltung der Mitglieder, an einer Stelle zu konzentrieren und damit zu effektivieren. Vielmehr werden die zu ganz unterschiedlichen Zwecken gespeicherten – zuvor pseudonymisierten – Daten von Versicherten, Leistungserbringern und anderen Leistungsträgern zusammengeführt und Auswertungen nach verschiedenen Richtungen – etwa zum Zwecke der Qualitätssicherung – vorgenommen.

Gerade im Bereich der öffentlichen Verwaltung ist zu klären, ob das Konzept des Data Warehouse in Verbindung mit dem Einsatz datenschutzfreundlicher Technologien Grundlage für datenschutzrechtlich zulässige Verfahren zur Datenverarbeitung sein kann.

6. Datenschutzrechtliche Bewertung

6.1 Einführung

Aufgabe des Datenschutzes ist es, das Recht des Einzelnen zu schützen, grundsätzlich selbst über die Preisgabe und Verwendung seiner Daten zu bestimmen. Dieses Grundrecht auf in-

formationelle Selbstbestimmung wurde vom Bundesverfassungsgericht im sog. *Volkszählungsurteil* (BVerfGE 65, 1 ff.) aus Art. 2 Abs. 1 i. V. m. Art. 1 Abs. 1 Grundgesetz hergeleitet. Der Umsetzung dieses – mittlerweile in einigen Landesverfassungen ausdrücklich genannten – Grundrechts dienen die Datenschutzgesetze des Bundes und der Länder sowie die datenschutzrechtlichen Bestimmungen bereichsspezifischer Rechtsvorschriften.

Das *Bundesdatenschutzgesetz* (BDSG) gilt für die Erhebung, Verarbeitung und Nutzung personenbezogener Daten durch die öffentlichen Stellen des Bundes und durch die Privatwirtschaft. Anwendungsbereich der Landesdatenschutzgesetze ist der Umgang mit personenbezogenen Daten durch die öffentlichen Stellen der Länder. Grundlage der folgenden Ausführungen ist das BDSG; die Landesdatenschutzgesetze werden nur dann erwähnt, wenn sie Regelungen enthalten, die im öffentlichen Bereich zu einer abweichenden Bewertung führen.

Die bereichsspezifischen Datenschutznormen basieren auf den Vorschriften und vor allem auf den Begriffen der allgemeinen Datenschutzgesetze. Als leges speciales gehen sie diesen grundsätzlich vor (§ 1 Abs. 3 Satz 1 BDSG). Diese Abhandlung beschränkt sich allerdings auf die Darstellung der Zulässigkeit nach den allgemeinen Datenschutzgesetzen. Im Anwendungsbereich der bereichsspezifischen Gesetze wäre eine Reihe zusätzlicher Fragen zu beantworten. Dies würde den Rahmen dieses Papiers sprengen.

Dem Datenschutzrecht unterliegen nur *personenbezogene Daten*, also nach der Definition des § 3 Abs. 1 BDSG „Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person (Betroffener)“. Daten einer juristischen Person, z. B. einer Anstalt oder einer Aktiengesellschaft, fallen genauso wenig unter das Datenschutzrecht wie Daten, die keiner konkreten natürlichen Person zugeordnet werden können, also anonym sind.

Eine Voraussetzung für den erfolgreichen Einsatz des Data Warehouse ist, dass alle verfügbaren Daten der Geschäfts- und Verwaltungsvorgänge und damit auch der Kunden bzw. Betroffenen erfasst werden. Um von anonymen Datenbeständen sprechen zu können, würde es nicht ausreichen, beispielsweise auf die Speicherung der Namen von Kunden oder anderer ihrer Identifikation dienenden Schlüssel (z. B. die Kundennummer) zu verzichten. Die Menge der darüber hinaus gespeicherten Daten (Anschrift, Kommunikationsdaten, Transaktionsdaten von Bezahlvorgängen, Bankverbindungen usw.) gestattet es i. d. R. ohne weiteres, den Personenbezug mit vertretbarem Aufwand an Zeit, Kosten und Arbeitskraft herzustellen. Um die Personenbeziehbarkeit sicher auszuschließen, wäre es also erforderlich, auch auf wenigstens einen Teil solcher Daten zu verzichten.

In der Wirtschaft wurde bisher in vielen Fällen der Personenbezug aber geradezu gefordert. Strategisches Ziel der Nutzung des Data Warehouse ist oft eine zielgerichtete, kundenorientierte und damit personenbezogene Werbung. Die Analyse der Daten soll mitunter auch zeigen, welcher Kunde „wechselgefährdet“ ist oder frühzeitig als potentieller „säumiger Schuldner“ erkannt wird. Mittlerweile gibt es aber auch Ansätze, wonach die vorhandenen personenbezogenen Daten frühzeitig pseudo- oder gar anonymisiert und dennoch interessante Informationen für die Unternehmen gewonnen werden (siehe 5. Kapitel).

Insgesamt ist festzuhalten: Aus der Zielsetzung des Data-Warehouse-Konzeptes folgt, dass wenigstens in einzelnen Phasen des Umgangs mit den Daten diese zumindest personenbeziehbar und damit personenbezogene Daten sind.

Gemäß § 4 Abs. 1 BDSG sind die Erhebung, Verarbeitung und Nutzung personenbezogener Daten nur zulässig, soweit

- das BDSG es erlaubt,
- eine andere Rechtsvorschrift es zulässt oder

– der Betroffene eingewilligt hat.

Die Zulässigkeit des Umgangs mit personenbezogenen Daten nach dem BDSG und den Landesdatenschutzgesetzen wird unter 6.2, die aufgrund einer Einwilligung unter 6.3 untersucht. Dabei werden die besonderen Arten personenbezogener Daten i. S. v. § 3 Abs. 9 BDSG nicht berücksichtigt. Diese sind Angaben über die rassische und ethnische Herkunft, politische Meinungen, religiöse oder philosophische Überzeugungen, Gewerkschaftszugehörigkeit Gesundheit oder Sexualleben. Der Umgang mit ihnen ist nur unter Erfüllung weiterer Voraussetzungen zulässig (z. B. §§ 4a, 13 Abs. 2, 14 Abs. 5 und 6, 28 Abs. 6 bis 9 BDSG).

6.2 Zulässigkeit nach den Datenschutzgesetzen

6.2.1 Einführung

Im Folgenden sollen die wesentlichen Probleme, die sich aus datenschutzrechtlicher Sicht bei einem Data Warehouse ergeben können, aufgezeigt werden. Dabei werden zunächst die nicht-öffentlichen Stellen betrachtet (6.2.2), bevor unten (6.2.3) speziell und ausführlicher auf die Besonderheiten von Data Warehouse in der öffentlichen Verwaltung eingegangen wird.

6.2.2 Nicht-öffentlicher Bereich

Die Ermächtigung der datenverarbeitenden Stelle, personenbezogene Daten in einem Data Warehouse zu erheben, verarbeiten oder nutzen, kann sich – wie oben (6.1) gezeigt – unter anderem aus dem BDSG ergeben. Als Rechtsvorschrift i. S. v. § 4 Abs. 1 BDSG kommt zunächst die allgemeine Datenverarbeitungsvorschrift für den nicht-öffentlichen Bereich, § 28 BDSG, in Betracht. § 28 BDSG gestattet das Speichern, Verändern, Übermitteln oder Nutzen personenbezogener Daten für die Erfüllung eigener Geschäftszwecke unter bestimmten Voraussetzungen.

So könnte möglicherweise der *Vertragszweck* (§ 28 Abs. 1 Satz 1 Nr. 1 BDSG) den Betrieb eines Data Warehouse decken. Da die Daten jedoch langfristig gespeichert werden und gerade nicht nur als operative Datenbasis zur Durchführung einzelner Geschäftsvorgänge dienen, weicht die vom Vertrag gedeckte von der im Data Warehouse vorgesehenen Nutzung der Daten ab und entfällt somit als Rechtsgrundlage.

So schließt beispielsweise der Kunde einer Bank den Vertrag zur Abwicklung von Zahlungstransaktionen. Die Ausführung der entsprechenden Anweisung durch die Bank erfordert die Verarbeitung zahlreicher Datensätze (u. a. die Übermittlung des Zahlungsgrundes). Die Bank übermittelt diese Informationen lediglich im Auftrag ihres Kunden während der Ausführung der Transaktion (Vertragszweck). Der Zahlungsgrund ist kein Vertragsdatum der Bank und darf deshalb nach Abschluss der Transaktion von ihr für eigene Zwecke nicht länger gespeichert und ausgewertet werden.

Folglich entfällt mit Erfüllung des jeweiligen Vertragszwecks die Erlaubnis zur Datenverarbeitung nach § 28 Abs. 1 Satz 1 Nr. 1 BDSG. Das gleiche gilt für vertragsähnliche Vertrauensverhältnisse im Sinne dieser Vorschrift. § 28 Abs. 1 Satz 1 Nr. 1 BDSG würde die Speicherung und weitere Verarbeitung in einem Data Warehouse nur dann erlauben, wenn gerade dies Vertragszweck wäre. Davon ist i. d. R. nicht auszugehen.

Die Datenverarbeitung im Data Warehouse könnte weiterhin aufgrund einer *Interessenabwägung* gem. § 28 Abs. 1 Satz 1 Nr. 2 BDSG zulässig sein. Dabei ist zu beachten, dass die Vorschrift kein Auffangtatbestand zu § 28 Abs. 1 Satz 1 Nr. 1 BDSG ist, so dass nach Nr. 1 erforderliche Daten nicht nach Nr. 2 zu anderen Zwecken verarbeitet oder genutzt werden können.

Nach § 28 Abs. 1 Satz 1 Nr. 2 BDSG müsste einerseits das Speichern bestimmter Daten im Data Warehouse der Wahrung berechtigter Interessen des Betreibers dienen. Dass mit dem Betrieb eines Data Warehouse tatsächlich berechtigte Interessen eines Unternehmens gewahrt werden können (welche auch immer das sein mögen), kann nie mit Sicherheit gesagt werden, da der Erfolg der Auswertungen sich vorher überhaupt nicht einschätzen lässt. Darüber hinaus dürfen andererseits schutzwürdige Interessen des Betroffenen der Verarbeitung der Daten im Data Warehouse nicht entgegenstehen. Ob die Auswertung der Daten und die Verknüpfung mit anderen Daten sowie die Interpretation des Ergebnisses schutzwürdige Interessen tatsächlich unberührt lässt, darf wohl zu Recht bezweifelt werden. Es ist beispielsweise nicht auszuschließen, dass im Ergebnis der Auswertungen der Betroffene zu Unrecht als potentieller „säumiger Schuldner“ oder aus anderen Gründen als „unzuverlässiger Kunde“ klassifiziert wird. Bereits die Tatsache, dass er als „gläserner Kunde“ geführt wird, dürfte nicht im Interesse des Betroffenen sein.

Es bleibt festzuhalten, dass auch die Interessenabwägung gem. § 28 Abs. 1 Satz 1 Nr. 2 BDSG als Rechtsgrundlage für den Betrieb eines Data Warehouse mit personenbezogenen Daten ausscheiden dürfte.

Im Data Warehouse werden u. U. auch personenbezogene *Daten aus öffentlichen Quellen* verarbeitet. Sofern nicht das schutzwürdige Interesse des Betroffenen am Ausschluss der Verarbeitung überwiegt, könnte der Betrieb des Data Warehouse dann auf § 28 Abs. 1 Satz 1 Nr. 3 BDSG gestützt werden. Diese Vorschrift kann jedoch nicht für das gesamte Data Warehouse als Rechtsgrundlage dienen. Selbst für den aus öffentlichen Quellen stammenden Teil der Daten ist zu bezweifeln, ob die Verarbeitung im Data Warehouse zulässig ist, da durchaus schutzwürdige Interessen Betroffener überwiegen können, wenn diese Daten mit anderen verknüpft und ausgewertet werden.

6.2.3 Öffentlicher Bereich

6.2.3.1 Erheben

Nach § 13 Abs. 1 BDSG ist das Erheben personenbezogener Daten zulässig, wenn ihre Kenntnis zur Erfüllung der Aufgaben der Daten verarbeitenden Stelle erforderlich ist.

Zu untersuchen ist, ob die Erhebung personenbezogener Daten direkt für die Verwendung im Data Warehouse zulässig ist. Dazu müsste das Betreiben eines Data Warehouse selbst eine Aufgabe der Daten verarbeitenden Stelle oder für eine andere Aufgabe erforderlich sein. Bestimmungen, die einer Daten verarbeitenden Stelle den Betrieb eines Data Warehouse vorschreiben, existieren bislang weder im allgemeinen, noch im bereichspezifischen Datenschutzrecht.

Ein Data Warehouse könnte aber für eine andere Aufgabe der Daten verarbeitenden Stelle erforderlich sein. Bei Daten, die in einem Data Warehouse verarbeitet werden, steht zum Zeitpunkt ihrer Erhebung/Speicherung nicht fest, auf welche Weise sie ausgewertet, mit welchen anderen Daten sie verknüpft, welche neuen Informationen durch ihr Hinzukommen zu den anderen Daten gewonnen und wie diese Informationen – und damit auch die erhobenen Daten – genutzt werden. Zumindest für die typischen öffentlichen, dem allgemeinen (Datenschutz-)Recht unterliegenden Stellen, die Aufgaben der Eingriffs- und Leistungsverwaltung erfüllen, ist eine solche Datenerhebung nicht zulässig. Denn Daten, von denen man bei der Erhebung noch nicht weiß, wofür sie „gut“ sind, können zur Erfüllung einer konkreten Aufgabe eines staatlichen Eingriffs oder einer Leistungsgewährung nicht erforderlich sein. Dass sie dafür u. U. nützlich sein können, reicht nicht aus. Es bleibt eine unzulässige Datenerhebung auf Vorrat. Denkbar wäre eine Erhebung für bereichsspezifisch geregelte spezielle, vor allem statistische Zwecke, oder mit Einwilligung des Betroffenen (Abschnitt 6.3).

Etwas anderes könnte für Datenerhebungen zu wissenschaftlichen Zwecken gelten, da Datenumgang im Forschungsbereich durch das Datenschutzrecht privilegiert wird. Auch könnte die Sammlung und Auswertung verschiedenartiger personenbezogener Daten durch Data-Mining-Technologien für die Forschung vielversprechend sein. So schreibt § 40 Abs. 1 BDSG lediglich vor, dass für wissenschaftliche Zwecke erhobene Daten nur für solche Zwecke verarbeitet werden dürfen. Dem Wortlaut nach verlangt sie keine vorhergehende Festlegung des genauen Forschungszweckes. Die originäre Datenerhebung zu Forschungszwecken wäre aber wiederum nur aufgrund einer speziellen Rechtsvorschrift oder einer Einwilligung zulässig. Da Rechtsvorschriften, die eine Erhebung zu Forschungszwecken vorsehen, nicht ersichtlich sind, bleibt als Legitimation nur die Einwilligung der Betroffenen. Für diese gilt das unten im Punkt 6.3 Ausgeführte. Damit ist auch hier zweifelhaft, ob die Datenerhebung zum Zwecke der Auswertung in einem Data Warehouse zulässig ist.

Insgesamt ist festzuhalten, dass die Erhebung personenbezogener Daten für den Zweck der Verwendung in einem Data Warehouse i. d. R. nicht durch Vorschriften des allgemeinen Datenschutzrechts gerechtfertigt werden kann.

6.2.3.2 Verarbeiten

Gemäß § 3 Abs. 4 Satz 1 BDSG ist Verarbeiten das Speichern, Verändern, Übermitteln, Sperren und Löschen personenbezogener Daten.

Eine wichtige Form der Verarbeitung ist das *Speichern*. Das Speichern personenbezogener Daten ist nach § 14 Abs. 1 BDSG zulässig, wenn es zur Erfüllung der in der Zuständigkeit der Daten verarbeitenden Stelle liegenden Aufgaben erforderlich ist und es für die Zwecke erfolgt, für die die Daten erhoben bzw., wenn keine Erhebung vorausgegangen ist, gespeichert worden sind.

In 6.2.3.1 wurde dargelegt, dass die Erhebung von personenbezogenen Daten für die Verwendung in einem Data Warehouse unzulässig ist. Die dort angestellten Überlegungen gelten auch für eine unmittelbare Speicherung ohne vorangehende Erhebung. Somit kommt eine Speicherung nach § 14 Abs. 1 BDSG grundsätzlich nicht in Betracht.

Die Speicherung personenbezogener Daten, die für andere Zwecke erhoben oder erstmals gespeichert worden sind (Zweckänderung), richtet sich nach § 14 Abs. 2 BDSG und wird ausführlich in 6.2.3.3 behandelt.

Für das *Verändern, Übermitteln, Sperren* und *Löschen* personenbezogener Daten, welche in einem Data Warehouse gespeichert sind, ergeben sich bei der Anwendung der einschlägigen Vorschriften keine rechtlichen Besonderheiten. In technischer Hinsicht kann es insoweit Probleme geben, als das Berichtigen, Sperren oder Löschen als Rechtsfolge sich nicht nur auf sämtliche Kopien eines Datums, sondern vor allem auch auf solche Daten und Informationen auswirkt, die durch Verarbeitung und Nutzung des in Rede stehenden Datums entstanden sind. Ähnliche Schwierigkeiten können bei den Verarbeitungsformen *Anonymisieren* und *Pseudonymisieren* entstehen.

6.2.3.3 Zweckbindung

Wie in 6.2.3.1 gezeigt, werden personenbezogene Daten von öffentlichen Stellen zulässigerweise dann erhoben, wenn es zu deren Aufgabenerfüllung erforderlich ist, vgl. § 13 BDSG. Alle weiteren Datenverarbeitungsvorgänge, wie z. B. Speichern (siehe 6.2.3.2) und Nutzen, setzen voraus, dass die Daten für den Zweck verarbeitet werden, für den sie erhoben worden sind.

Bei einem Data Warehouse einer öffentlichen Stelle sind die Daten in aller Regel nicht zum Zwecke der weiteren Auswertung in einem Data Warehouse erhoben worden. Sinn des Data Warehouse ist es ja gerade, dass die Daten für Zwecke verarbeitet und vor allem genutzt werden, die mit dem ursprünglichen Erhebungszweck nichts zu tun haben. Die Daten werden für die operative Aufgabenerfüllung nicht mehr benötigt. Sieht man von dem Fall ab, dass die Daten eigens zum Zwecke der Auswertung in einem Data Warehouse erhoben worden sind, so ist in den übrigen Fällen für das Nutzen der Daten in einem Data Warehouse eine Änderung des Erhebungszwecks – also eine Durchbrechung der Zweckbindung – erforderlich.

Alle Datenschutzgesetze enthalten Ausnahmetatbestände, nach denen Daten auch zu einem anderen als dem Erhebungszweck genutzt werden dürfen. Die einzelnen Vorschriften in den Bundesländern und im BDSG ähneln sich dabei mehr oder weniger stark. Hier wird wie oben auf § 14 BDSG und zwar dessen Abs. 2 abgestellt. Mit jeder Zweckänderung ist ein neuer Eingriff in das Recht auf informationelle Selbstbestimmung verbunden. Die Ermächtigung zur Zweckänderung in § 14 Abs. 2 BDSG ist wegen der Bedeutung des Grundrechts eng auszulegen.

Die Zweckänderung könnte zulässig sein, weil eine *Rechtsvorschrift* dies vorsieht oder zwingend voraussetzt, § 14 Abs. 2 Nr. 1 BDSG. Eine Rechtsvorschrift, die das zweckverändernde Nutzen von Daten in einem Data Warehouse ausdrücklich zulässt, ist nicht ohne Weiteres erkennbar. Dasselbe gilt für eine Rechtsvorschrift, die eine solche Zweckänderung *zwingend voraussetzt*. Abgesehen von den verfassungsrechtlichen Zweifeln an dieser Formulierung, müsste es sich zumindest um eine Rechtsvorschrift handeln, die nicht ausgeführt werden kann, ohne dass die Daten zweckändernd in einem Data Warehouse ausgewertet werden. Eine solche Rechtsvorschrift ist nicht denkbar und würde auf jeden Fall gegen das vom Bundesverfassungsgericht in seinem Volkszählungsurteil geforderte Gebot der Normenklarheit verstoßen.

Eine weitere Möglichkeit der zulässigen Zweckänderung mit dem Ziel der Auswertung in einem Data Warehouse wäre die *Einwilligung* des Betroffenen, § 14 Abs. 2 Nr. 2 BDSG. Dies soll hier nicht weiter besprochen werden, sondern wird unter 6.3 gesondert behandelt.

Weiterhin kommt gemäß § 14 Abs. 2 Nr. 3 BDSG in Betracht, dass die Zweckänderung *offensichtlich im Interesse des Betroffenen* liegt und nicht davon auszugehen ist, dass er in Kenntnis der Zweckänderung seine Einwilligung verweigern würde. Diese Lösung dürfte ebenfalls ausscheiden. Die Vorschrift verlangt, dass der Betroffene klar überwiegende unmittelbare Vorteile davon hat, dass seine ursprünglich zu einem anderen Zweck erhobenen Daten in einem Data Warehouse ausgewertet werden. Das kann in der Regel nicht angenommen werden. Zwar mag es zumindest mittelbare Vorteile für die Bürger geben, wenn Daten in einem Data Warehouse ausgewertet werden, da sich möglicherweise viele Verwaltungsabläufe auch zum Nutzen der Bürger verbessern lassen könnten. Konkrete Vorteile, die aus der Nutzung gerade seiner Daten entstehen, kann der Betroffene jedoch nicht erlangen. Angesichts der Gefahren, die ein Data Warehouse unter Verwendung personenbezogener Daten für das Persönlichkeitsrecht der Betroffenen darstellt, ist vielmehr davon auszugehen, dass ein Data Warehouse deutlich mehr Nachteile als Vorteile für den Betroffenen hat. Darüber hinaus kann aus eben diesen Gründen auch nicht angenommen werden, dass ein Betroffener seine Einwilligung erteilen würde, weshalb die Voraussetzungen des § 14 Abs. 2 Nr. 3 BDSG schon deshalb nicht gegeben sind.

§ 14 Abs. 2 Nr. 4 BDSG scheidet als Legitimation für die Zweckbindung von vornherein aus, da die Nutzung in einem Data Warehouse in keinem Zusammenhang mit der Unrichtigkeit von Daten steht und der Sinn des Data Warehouse nicht in der *Korrektur unrichtiger Daten* liegt.

Stammen die Daten aus *allgemein zugänglichen Quellen*, ist unter den weiteren Voraussetzungen von § 14 Abs. 2 Nr. 5 BDSG ebenfalls eine Zweckänderung möglich. Sollen solche Daten in einem Data Warehouse genutzt werden, entstehen allerdings die gleichen Probleme, die oben unter 6.2.2 hinsichtlich § 28 Abs. 1 Satz 1 Nr. 3 BDSG beschrieben wurden. Die Nutzung wird in der Regel also spätestens an den überwiegenden schutzwürdigen Interessen der Betroffenen, die Zweckbindung einzuhalten, scheitern.

Die Zusammenhänge, die mit einem Data Warehouse untersucht und festgestellt werden sollen, dienen – abgesehen von dem Fall, dass die Daten eigens zu diesen Zwecken erhoben worden sind – weder der *Strafverfolgung* oder -*vollstreckung* noch der *Abwehr von Gefahren für die öffentliche Sicherheit*. Insofern kommen nach § 14 Abs. 2 Nr. 6 und 7 BDSG keine Zweckänderungen zugunsten der Nutzung in einem Data Warehouse in Betracht.

Ebenso wenig sind Fälle denkbar, bei denen die zweckändernde Nutzung von Daten in einem Data Warehouse zur *Abwehr schwerwiegender Beeinträchtigungen der Rechte einer anderen Person* erforderlich ist, so dass auch § 14 Abs. 2 Nr. 8 BDSG als Rechtsgrundlage ausscheidet.

Die Zweckänderung nach § 14 Abs. 2 Nr. 9 BDSG (die Daten sind für die Durchführung *wissenschaftlicher Forschung* erforderlich) ist schon deshalb zweifelhaft, weil es sich um ein bestimmtes Forschungsvorhaben handeln muss. Dies ist aufgrund der Charakteristik des Data Warehouse aber vorher gerade nicht abzusehen. Darüber hinaus wird die Zulässigkeit i. d. R. an den überwiegenden schutzwürdigen Interessen der Betroffenen scheitern, die nicht wissen können, in welcher Art und Weise ihre Daten im Data Warehouse verarbeitet werden.

Zusammenfassend lässt sich feststellen, dass die mit einer Änderung des Erhebungszwecks einhergehende Nutzung personenbezogener Daten durch öffentliche Stellen in einem Data Warehouse nach § 14 Abs. 2 BDSG im Allgemeinen nicht zulässig ist.

Darüber hinaus könnte die Nutzung personenbezogener Daten in einem Data Warehouse nach § 14 Abs. 3 BDSG bzw. den entsprechenden landesrechtlichen Vorschriften zulässig sein. Die Norm stellt klar, dass die Nutzung personenbezogener Daten zu den dort genannten Zwecken (Aufsicht, Rechnungsprüfung) keine Änderung des Erhebungszwecks darstellt, sondern von diesem Primärzweck mit umfasst ist. Zu beachten ist, dass § 14 Abs. 3 BDSG keine eigenständige Befugnis ist, personenbezogene Daten zu den dort genannten Zwecken zu verarbeiten oder zu nutzen. Die Verarbeitung bzw. Nutzung muss selbstverständlich trotzdem zur Aufgabenerfüllung i. S. v. § 14 Abs. 1 BDSG erforderlich sein.

In Betracht kommt hier eine Nutzung der Daten für Organisationsuntersuchungen gemäß § 14 Abs. 3 Satz 1 BDSG. Bei solchen Untersuchungen sind vor allem Personaldaten der Beschäftigten betroffen, deren zulässige Verarbeitung nach bereichsspezifischem Recht (z. B. § 35 DSG M-V oder § 29 BbgDSG) zu beurteilen wäre. Auf diese Thematik wird nicht weiter eingegangen.

6.3 Einwilligung

6.3.1 Allgemeine Voraussetzungen nach den Datenschutzgesetzen

Eine mögliche Legitimation zur Verarbeitung personenbezogener Daten durch öffentliche Stellen stellt die *Einwilligung* des Betroffenen dar. Grundsätzlich ist jeder Datenverarbeitungsvorgang erlaubt, soweit der Betroffene darin eingewilligt hat, vgl. nur § 4 Abs. 1 BDSG. § 4a BDSG und die entsprechenden Bestimmungen der Landesdatenschutzgesetze stellen besondere Anforderungen an die Form und den Inhalt von Einwilligungen.

Formal bedarf die Einwilligung grundsätzlich der Schriftform und muss in bestimmter Art und Weise erkennbar gemacht werden.

Inhaltlich muss vor allem sichergestellt werden, dass der Betroffene den genauen Inhalt und die Tragweite seiner Einwilligung erkennt. Dies hat bestimmte Informationspflichten der Daten verarbeitenden Stelle zur Folge. Nur eine informierte Einwilligung (*informed consent*) ist wirksam. Der Betroffene muss wissen, zu welchem Zweck welche Daten gespeichert, übermittelt oder sonst verarbeitet werden. Bei Übermittlungen muss ihm auch der Adressatenkreis bekannt sein.

Im öffentlichen Bereich besteht nur scheinbar ein gleichberechtigtes Verhältnis zwischen Einwilligung und der Erlaubnis zur Datenverarbeitung aufgrund einer Rechtsvorschrift. Hier kann nicht außer Acht gelassen werden, dass öffentliche Stellen nur ausnahmsweise auf die Einwilligung zurückgreifen können, solange es an einer ausdrücklichen bereichsspezifischen Regelung fehlt. Dies liegt vor allem daran, dass öffentliche Stellen ohnehin darauf beschränkt sind, ihren gesetzlich und verfassungsrechtlich zugewiesenen Aufgaben nachzugehen. Eine Datenverarbeitung, die außerhalb der Aufgabenerfüllung der öffentlichen Stelle stattfindet, kann deshalb grundsätzlich auch nicht durch die Einwilligung des Betroffenen legitimiert werden.

6.3.2 Einwilligung und Data Warehouse

Legt man die oben beschriebenen Anforderungen an Wirksamkeit und Zulässigkeit einer Einwilligung zugrunde, so ist fraglich, ob ein Betroffener der Verarbeitung seiner personenbezogenen Daten in einem Data Warehouse einschließlich notwendiger Zweckänderungen in diesem Sinne zustimmen kann.

Wie schon mehrfach ausgeführt, gehört das Betreiben eines Data Warehouse nicht zu den Aufgaben einer öffentlichen Stelle und ist zu deren Erfüllung nicht erforderlich. Deshalb ist es zweifelhaft, ob dies durch eine Einwilligung umgangen werden kann. Wäre dies zulässig, würde die öffentliche Stelle außerhalb der ihr zustehenden Aufgaben personenbezogene Daten verarbeiten. Die Datenverarbeitung würde auf der Basis einer Einwilligung durchgeführt, die zu erteilen der Betroffene sich angesichts des hoheitlichen Charakters staatlicher Tätigkeit möglicherweise gar gezwungen sieht.

Selbst wenn man die oben genannten Bedenken ausräumen könnte, so wäre die Einwilligung auch aus anderen Gründen in jedem Fall unwirksam. Bei einem Data Warehouse ist es nicht möglich vorherzusagen, welche Datenverarbeitungsvorgänge stattfinden, was deren Ergebnis ist, oder für welchen Zweck die im Data Warehouse ablaufenden Verarbeitungsvorgänge im Einzelnen jeweils erfolgen. Der Verarbeitungszweck ergibt sich gerade erst aus den Ergebnissen nicht im Einzelnen vorhersehbarer Verarbeitungsvorgänge. Eine informierte Einwilligung wie sie § 4a BDSG und die entsprechenden Vorschriften der Landesdatenschutzgesetze verlangen, kann daher nicht erteilt werden.

Das Betreiben eines Data Warehouse durch öffentliche Stellen aufgrund einer Einwilligung nach § 4 Abs. 2 i. V. m. § 4a BDSG ist deshalb nicht möglich. Da die Anforderungen an die Einwilligung bei den bereichsspezifischen Datenschutzvorschriften entweder in etwa denen des BDSG entsprechen oder noch höher sind, gilt dies auch dort. Diese Problematik wird daher nicht gesondert ausgeführt.

7. Datenschutzfreundliche Technologien und Data Warehouse

7.1 Einführung

Die Zulässigkeit der automatisierten Verarbeitung personenbezogener Daten, wie in einem Data Warehouse, und damit auch ihre Akzeptanz beim Betroffenen/Kunden hängt in hohem Maße davon ab, inwieweit sie nach dem Prinzip der *Datensparsamkeit* erfolgt. Es verlangt, so wenig personenbezogene Daten wie möglich zu erheben, verarbeiten und zu nutzen, vgl. § 3a BDSG. *Datenvermeidung* ist die stets anzustrebende Form der Datensparsamkeit. In diesem Fall werden bei der Nutzung von Informations- und Kommunikationssystemen keine personenbezogenen Daten erhoben, verarbeitet und genutzt.

Bereits heute ist eine Reihe von Techniken und Hilfsmittel zur Umsetzung der Philosophie der Datenvermeidung und der Datensparsamkeit verfügbar. Die Technologie, die dafür gesorgt hat, dass personenbezogene Daten gespeichert, genutzt und weitergegeben werden können, ist auch zur Wahrung der Privatheit des Einzelnen nutzbar. Diese Möglichkeiten der modernen Datenschutztechnologie, die mit dem Begriff *„Privacy enhancing technologies“* (PET) bezeichnet wird, umfasst ein ganzes System technischer Maßnahmen.

Im Abschnitt 7.2 werden die Methoden Anonymisierung, Pseudonymisierung und der *Identity Protector* vorgestellt, die wichtigsten Realisierungshilfen dazu genannt und schließlich allgemeine Empfehlungen zur Implementierung eines datenschutzfreundlichen DV-Systems gegeben. Unter 7.3 wird dann – was schon in Kapitel 5 angesprochen wurde – ausführlich untersucht, wie mit Hilfe von PET datenschutzrechtliche Probleme bei der Datenverarbeitung im Data Warehouse gelöst werden können.

7.2 Methoden und Werkzeuge

7.2.1 Anonymisierung

Anonymisierung ist eine Veränderung personenbezogener Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig hohem Aufwand einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können.

Die Qualität der Anonymisierungsprozedur hängt von verschiedenen Einflussfaktoren ab. Entscheidend hierfür sind der Zeitpunkt der Anonymisierung, die Rücknahmefestigkeit der Anonymisierungsprozedur, die Mächtigkeit der Menge, in der sich der Betroffene verbirgt, und die Möglichkeit zur Verkettung einzelner Transaktionen desselben Betroffenen.

Auch konkrete Einzelangaben in einem Datensatz / einer Transaktion (z. B. Beruf/Amt = Bundeskanzler, konkrete Einkommensangaben) sind für die Qualität der Anonymisierungsprozedur von Bedeutung und können die Mächtigkeit der Menge, in der sich der Betroffene verbirgt, verringern. Sind im Wertebereich Werte vorhanden, die die Anonymität gefährden, müssen sie mit anderen zusammengefasst werden. Ist eine solche Veränderung aus technischen oder inhaltlichen Gründen nicht möglich, kann keine Anonymität erreicht werden.

Das Ziel datenschutzfreundlicher Technologien ist es u. a., Daten schon ohne Personenbezug zu erheben oder bereits personenbezogen erhobene Daten so bald wie möglich zu anonymisieren. Ein Höchstmaß an Anonymität wird erreicht, wenn personenbezogene Daten gar nicht erst entstehen. Gelungene Beispiele hierfür sind anonyme Telefonkarten und anonyme Zahlkarten im öffentlichen Personennahverkehr. Beispiele für die Anwendung der Anonymisierung sind im Bereich der Statistik und in der Forschung zu finden.

7.2.2 Pseudonymisierung

Pseudonymisierung ist das Verändern personenbezogener Daten durch eine Zuordnungsvorschrift derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse ohne Kenntnis oder Nutzung der Zuordnungsvorschrift nicht mehr oder nur mit einem unverhältnismäßig hohem Aufwand einer natürlichen Person zugeordnet werden können.

Dazu werden beispielsweise die Identifikationsdaten durch eine Abbildungsvorschrift in ein willkürlich gewähltes Kennzeichen (das Pseudonym) überführt. Ziel eines solchen Verfahrens ist es, nur bei Bedarf und unter Einhaltung vorher definierter Rahmenbedingungen den Personenbezug wieder herstellen zu können. Die Reidentifizierung kann mitunter auch ausschließlich dem Betroffenen vorbehalten bleiben.

Das Mittel der Pseudonymisierung sollte insbesondere dort eingesetzt werden, wo Anonymisierung nicht möglich ist.

Die Qualität der Pseudonymisierungsprozedur hängt von den gleichen Einflussfaktoren ab, wie die Stärke der Anonymisierungsprozedur, nämlich vom Zeitpunkt der Pseudonymisierung, von der Rücknahmefestigkeit der Pseudonymisierungsprozedur, von der Mächtigkeit der Menge, in der sich der Betroffene verbirgt, und von der Möglichkeit zur Verkettung einzelner Transaktionen/Datensätze desselben Betroffenen. Insbesondere können Transaktionen/Datensätze, die unter demselben Pseudonym getätigt/gespeichert wurden, miteinander verkettet werden.

Unter gleichen Bedingungen ist die Anonymisierung datenschutzfreundlicher als die Pseudonymisierung. Das Pseudonym kann dazu benutzt werden, den Personenbezug wiederherzustellen. Ansonsten kann ohne Berücksichtigung der genannten Faktoren nicht pauschal beurteilt werden, ob die Anonymisierung oder die Pseudonymisierung datensparsamer ist.

Je nach Verknüpfbarkeit und Geheimnisträger des Pseudonyms kann der Personenbezug

- nur vom Betroffenen (selbstgenerierte Pseudonyme),
- nur über eine Referenzliste (Referenz-Pseudonyme) oder
- nur unter Verwendung einer sog. Einweg-Funktion mit geheimen Parametern (Einweg-Pseudonyme)

wiederhergestellt werden.

Pseudonyme ermöglichen es, den Personenbezug herzustellen, so dass die Identität der Person nur in den vorab bestimmten Einzelfällen erkennbar wird.

Pseudonyme sollen zufällig und nicht vorhersagbar gewählt werden. Die Menge der möglichen Pseudonyme soll so mächtig sein, dass bei zufälliger Auswahl nicht zweimal das gleiche Pseudonym generiert wird. Ist eine hohe Sicherheit erforderlich, muss die Menge der Pseudonymkandidaten mindestens so mächtig sein, wie der Wertebereich sicherer kryptografischer Hashfunktionen.

Pseudonyme sollten insbesondere nicht anwendungsübergreifend, sondern nur für jeweils ein Verfahren eingesetzt werden. Jede anwendungsübergreifende Benutzung eines einzigen Pseudonyms würde die Gefahr erhöhen, dass aus sämtlichen, mit dem Pseudonym verbundenen Daten ein detailliertes Personenprofil erstellt werden kann, das wiederum den Rückschluss auf eine bestimmte Person erleichtert. Aber auch innerhalb einer Anwendung ist die Verwendung nur eines einzigen Pseudonyms nicht unproblematisch.

7.2.3 Der Identity Protector

DV-Systeme, für die eine anonyme Nutzungsform nicht vollständig möglich ist, sollten derart in unterschiedliche Einzelprozesse zerlegt werden, dass unmittelbar personenbezogene Daten (Identitätsdaten) nur erhoben, verarbeitet und genutzt werden, wo dies unabdingbar nötig ist.

Durch geeignete technische Maßnahmen muss dafür Sorge getragen werden, dass die Bereiche des DV-Systems, die den vollen Personenbezug mit den Identitätsdaten benötigen, strikt von jenen getrennt werden, die mit einem Pseudonym auskommen. Das heißt, nur die tatsächlich und unmittelbar benötigten Daten stehen dem jeweiligen Prozess zur Verfügung. Eine Zusammenführung von Identitätsdaten und Pseudonymdaten ist nur unter vorab und genau definierten Umständen möglich.

Diese Aufgaben kann ein Identity Protector leisten. Er kann als Systemelement (Prozess) betrachtet werden, das den Austausch von Identitätsdaten und Pseudonymdaten zwischen den übrigen Systemelementen steuert.

Für einen Identity Protector sind verschiedene Ausprägungsformen möglich:

a) eigenständiges Element in einem DV-System,

b) eigenständiges DV-System, das unter der Kontrolle des Benutzers steht,

c) eigenständiges DV-System, das unter der Kontrolle einer Vertrauensstelle steht.

Im Falle a) sollte der Identity Protector ein – auch für den Betreiber des DV-Systems – unveränderbarer Baustein sein. Die Realisierung ließe sich als Softwarebaustein im DV-System selbst, im zugrundeliegenden Betriebssystem oder auch als Hardwarekomponente mit zugehöriger Software (z. B. als „Black-Box-Lösung“) bewerkstelligen.

Im Falle b) wäre eine Abbildung des Identity Protector z. B. in Form einer Smart-Card möglich.

Der Identity Protector hat folgende Funktionalitäten:

- kontrollierte Offenlegung und Freigabe der Identität,
- Generierung von Pseudonymen,
- Umsetzung von Pseudonymen in weitere Pseudonyme,
- Umsetzung von Identitäten in Pseudonyme (Pseudonymisierung),
- Umsetzung von Pseudonymen in Identitäten (Depseudonymisierung),
- vorbeugende Missbrauchsbekämpfung (u. a. durch die erstgenannte Funktionalität).

7.2.4 Realisierungshilfen

Als Hilfsmittel zur Umsetzung der in 7.2.1 bis 7.2.3 dargestellten Methoden kommen vor allem folgende Werkzeuge in Betracht:

- Hashfunktionen
- (blinde) digitale Signaturen
- (Signaturschlüssel-)Zertifikate
- biometrische Verfahren
- Vertrauensstellen (Trust Center)

7.2.5 Empfehlungen zur Vorgehensweise

Um die o. g. Grundsätze bei der Entwicklung oder Modifizierung von DV-Systemen in ausreichendem Maße berücksichtigen zu können, ist folgende Vorgehensweise empfehlenswert:

Zunächst müssen Daten verarbeitende Systeme und Teilsysteme einschließlich ihrer Schnittstellen definiert werden. Bei dieser Definition muss auch eine Unterscheidung derjenigen Systeme und Teilsysteme erfolgen, in denen

- ohne personenbezogene Daten gearbeitet werden kann,
- personenbezogene Daten anonymisiert werden können,
- personenbezogene Daten pseudonymisiert werden können bzw.
- der direkt herstellbare Personenbezug unvermeidlich ist.

Ist eine Anonymisierung oder eine Pseudonymisierung erforderlich, so ist für das jeweilige System/Teilsystem eine entsprechende Prozedur zu finden,

- die die personenbezogenen Daten frühestmöglich anonymisiert bzw. pseudonymisiert,
- die nicht unzulässig beeinflusst werden kann (Integrität),
- die aus dem System/Teilsystem nicht mit geringem Aufwand wieder entfernt werden kann (Rücknahmefestigkeit),
- die den Betroffenen in einer hinreichend großen Menge möglicher Betroffener verbirgt und
- die die Verkettbarkeit von Einzeldaten oder Transaktionen zu Datenspuren unterdrückt.

Stellt sich heraus, dass die vorhandenen Risiken mit dem so konstruierten System nicht hinreichend reduziert werden können, so müssen ggf. Teile des Definitionsprozesses und Teile des Gestaltungsprozesses wiederholt werden.

7.3 Anwendung datenschutzfreundlicher Technologien auf Data Warehouse

7.3.1 Grundsatz

Aus Kapitel 6 folgt, dass in einem Data Warehouse grundsätzlich keine personenbezogenen Daten gespeichert, auf sonstige Weise verarbeitet oder genutzt werden können. Damit ist ein wichtiger Teil der o. g. Zielsetzung des Data-Warehouse-Konzeptes im öffentlichen Bereich nicht umsetzbar. Das Betreiben eines Data Warehouse kann aber auch ohne die Verwendung personenbezogener Daten sinnvoll sein, beispielsweise dann, wenn es um statistische Aussagen und Gruppenverhalten geht.

Auch für ein „anonymes“ Data Warehouse bildet die vorhandene operative Datenbasis, welche u. a. alle im Tagesgeschäft anfallenden personenbezogenen Daten enthält, die Grundlage für die zu verarbeitenden Daten. Dazu müssen die Daten der operativen Datenbasis unter Einsatz der PET und der dargelegten Vorgehensweise anonymisiert im Data Warehouse gespeichert werden.

7.3.2 Form der Speicherung der operativen Datenbasis

7.3.2.1 Speicherung in einem Datensatz

Zunächst soll der einfache Fall betrachtet werden, dass die in der operativen Datenbasis gespeicherten personenbezogenen Daten von jeder Person (rechtmäßig – davon wird im Folgenden ausgegangen) in einem einzigen Datensatz gespeichert sind. Dann dürfen die Datensätze nur mit einer reduzierten Anzahl von Datenfeldern in das Data Warehouse übertragen werden. Gemäß dem Abschnitt 7.2.1 müssen alle Felder entfernt werden, die einen unmittelbaren (Name) oder einen mittelbaren Personenbezug ermöglichen. Letzteres kann sich wie oben ausgeführt aus einzelnen Werten oder aus der Kombination der in verschiedenen Datenfeldern des selben Datensatzes enthaltenen Werten ergeben.

Ist der Personenbezug nur für wenige Personen über den Wert eines bestimmten Datenfeldes möglich, so könnte auch daran gedacht werden, auf die Übertragung aller Werte für diese Personen zu verzichten, dafür aber das besagte Datenfeld mit seinen Werten für die anderen Personen in das Data Warehouse zu übernehmen. Eine andere Möglichkeit wäre, die Wertemenge so zu vergröbern, dass ein Personenbezug nicht mehr hergestellt werden kann.

7.3.2.2 Speicherung in einer Datenbank

Komplizierter wird es, wenn die personenbezogenen Daten der einzelnen Personen in verschiedenen Datensätzen in einer Datenbank gespeichert sind. Im Allgemeinen gibt es sog. Schlüssel(-Felder), über die die Datensätze, welche die Daten zur selben Person enthalten, miteinander verknüpft werden können. Bei dieser Verbindung der Datenfelder zu einer Person sind dann die im vorstehenden Absatz ausgeführten Überlegungen anzustellen. Die technische Umsetzung für die überwiegend anzutreffenden relationalen Datenbanken erfolgt in der Weise, dass über eine Datenbankabfragesprache, beispielsweise SQL, ein mehrstufiger Programmbefehl abgesetzt wird, der mit Hilfe der Schlüsselfelder das Kreuzprodukt der betreffenden Datensätze bildet (Join), davon die interessierenden Datenfelder auswählt (Select) und diese schließlich in das Data Warehouse überträgt (Export). Gegen die kurzfristige Speicherung aller Datenfelder zu einer Person im Kreuzprodukt der Datensätze während der Befehlsausführung bestehen keine datenschutzrechtlichen Bedenken.

7.3.2.3 Speicherung in verschiedenen Datenbank(system)en

Sind die Daten zu einer Person in logisch und physisch verschiedenen Datenbank(system)en gespeichert, so kann i. d. R. nicht einfach wie im o. g. Fall ein mehrstufiger Programmbefehl formuliert werden, der dann das gewünschte Ergebnis liefert. Vielmehr müssen erst die technischen Grundlagen für eine Verknüpfung der Datensätze geschaffen werden. Hierbei müssen voraussichtlich folgende Schritte durchgeführt werden:

- Zurverfügungstellung entsprechenden Speicherplatzes
- Transport der von verschiedenen Systemen und Speicherorten kommenden Datensatzkopien mit Betriebssystembefehlen zum neuen Speicherplatz
- Anpassung der unterschiedlichen Formate
- Übertragung der zuvor mittels der Dateibeschreibungen ausgewählten Felder, die weder allein noch kombiniert einen Personenbezug ermöglichen, in das Data Warehouse
- Löschung der zwischengespeicherten Daten

Probleme bereitet hier die zeitlich nicht unerhebliche Zwischenspeicherung der personenbezogen zusammengeführten Datensätze, die – wie in 6.2.3.3 dargelegt – grundsätzlich als Ver-

stoß gegen § 14 BDSG unzulässig sein dürfte. Eine Lösung könnte sein, die einzelnen Schritte vom Transport der Daten auf den neu geschaffenen Speicherplatz bis hin zur Löschung aller zwischengespeicherten Daten in einer Art „black box“ so miteinander verkettet ablaufenzu lassen, dass ein Zugriff auf die Daten weder während des normal ablaufenden Prozesses noch bei dessen (von außen erzwungener) Unterbrechung möglich ist.

7.3.2.4 „Anonymisierungs-Black-box“

Zur Anonymisierung bietet sich eine Variante des in 7.2.3 beschriebenen Identity Protector an. Kontrolliert von einer Vertrauensstelle sollte ein von ihr signierter *String* eingegeben werden können. Dieser enthält die Bezeichnung der keinen Personenbezug ermöglichenden Datenfelder, die in das Data Warehouse übertragen werden können. Der *String* ist der Daten verarbeitenden Stelle zwar bekannt – sinnvoller Weise sogar von ihr vorgeschlagen –, kann aber nicht mehr von ihr geändert, insbesondere nicht erweitert werden. Statt der Pseudonymisierungsfunktion hätte der Identity Protector eine Anonymisierungsfunktion, die zwei Operationen durchführt: kontrollierte Ausgabe nur der im String vorgegebenen Datenfelder, anschließend Löschung sämtlicher Daten. Diese beiden Operationen dürfen dabei nicht von der kritischen Prozedur der personenbezogenen Verknüpfung der verschiedenen Datensätze getrennt werden können. Die für den Identity Protector vorgesehene Funktionskombination „vorbeugende Missbrauchsbekämpfung durch kontrollierte Offenlegung und Freigabe der Identität“ muss dabei – kontrolliert, versiegelt und möglichst zertifiziert – so eingestellt sein, dass die Preisgabe der Identität überhaupt nicht möglich ist.

Von den in 7.2.3 genannten drei Ausprägungsformen des Identity Protector sollte „c) eigenständiges DV-System, das unter der Kontrolle einer Vertrauensstelle steht“ gewählt werden. „Kontrolle“ bedeutet hier nicht notwendig, dass die Vertrauensstelle den Identity Protector ständig beaufsichtigt oder gar den alleinigen Zugriff auf ihn hat. Es genügt, wenn der Identity Protector von der Vertrauensstelle so (kryptografisch) sicher „abgeschlossen“ und gegebenenfalls physisch versiegelt worden ist, dass der Daten verarbeitenden Stelle ein Zugriff auf personenbezogene Daten ohne Kenntnis des Schlüsselwortes auch durch Manipulation oder Zerstörung des Identity Protector nicht möglich ist.

8. Schlussfolgerungen

Die o. g. Betrachtungen machen deutlich, dass für das Betreiben eines Data Warehouse mit personenbezogenen Daten derzeit keine Rechtsgrundlage existiert. Im Grunde genommen ist das aber keine neue Erkenntnis.

Das Bundesverfassungsgericht hat bereits 1969 festgestellt, dass es mit der Menschenwürde unvereinbar sei, wenn der Staat das Recht für sich in Anspruch nehmen könnte, den Menschen zwangsweise in seiner ganzen Persönlichkeit zu registrieren und katalogisieren. Spätestens nach dem Volkszählungsurteil ist auch klar, dass die Sammlung personenbezogener Daten auf Vorrat zu unbestimmten oder noch nicht bestimmbaren Zwecken nicht zulässig ist. Solche Datensammlungen könnten beispielsweise Basis für die Erstellung detaillierter Persönlichkeitsprofile sein. 1988 stellte deshalb auch der Bundesgerichtshof fest, dass im staatlichen und im privaten Bereich die zwangsweise und heimliche Erstellung solcher Persönlichkeitsprofile verboten ist.

Diese höchstrichterlichen Entscheidungen beschreiben im Grunde genommen genau das Data-Warehouse-Konzept, ohne diesen Begriff selbst zu nennen. Die Entscheidungen verdeutlichen, dass solche Konzepte mit personenbezogenen Daten grundsätzlich nicht realisiert werden dürfen.

Die EG-Datenschutzrichtlinie kommt zum gleichen Ergebnis. Auch dort wird die Forderung aufgestellt, personenbezogene Daten nur mit einer festgelegten Zweckbestimmung zu verarbeiten (Art. 7), wobei nicht zwischen öffentlichen und nicht-öffentlichen Stellen unterschieden wird. Darüber hinaus wird klargestellt, dass besonders sensible Daten, die z. B. politische Meinungen, die Gesundheit oder das Sexualleben betreffen, grundsätzlich nicht verarbeitet werden dürfen (Art. 8, siehe dazu auch den letzten Absatz von 6.1). Weiterhin stellt das durch Art. 15 der Richtlinie aufgestellte – und in den Datenschutzgesetzen (z. B. § 6a BDSG) umgesetzte – Verbot automatisierter Entscheidungen, die beispielsweise die Kreditfähigkeit eines Betroffenen bewerten, eine weitere erhebliche Einschränkung für den Betrieb von Data Warehouse dar.

Jetzt die Schlussfolgerung zu ziehen, ein Data Warehouse dürfte aus datenschutzrechtlicher Sicht grundsätzlich nicht betrieben werden, wäre aber sicher voreilig. Wie insbesondere die Ausführungen im Kapitel 7 gezeigt haben, können Data-Warehouse-Systeme unter Nutzung datenschutzfreundlicher Technologien so konzipiert werden, dass sie den datenschutzrechtlichen Anforderungen genügen. Natürlich sind dann technische und organisatorische Maßnahmen erforderlich, die eine Deanonymisierung verhindern.

Das Ziel, strategisch wichtige Informationen aus einem solchen Datenbestand abzuleiten, kann trotzdem erreicht werden. Durch sinnvolle Aggregierung und Partitionierung von Einzelangaben kann die Herstellung des Personenbezugs verhindert oder wenigstens erheblich erschwert werden. Für bestimmte Kundengruppen beispielsweise könnten trotzdem wertvolle Verkaufsinformationen abgeleitet werden, wobei aus datenschutzrechtlicher Sicht bereits die Zuordnung Einzelner zu solchen Kundengruppen problematisch ist. Das gleiche gilt für die Gewinnung strategischer Erkenntnisse im Bereich der Verwaltung.

Für alle technischen und organisatorischen Maßnahmen, die den Personenbezug der gespeicherten Daten verhindern oder einschränken, gilt jedoch immer, dass sie die Funktionalität des Data Warehouse beeinträchtigen. Dennoch muss an die Anwender dieser Technologien appelliert werden, sich im Interesse des Rechts auf informationelle Selbstbestimmung zu beschränken und Data-Warehouse-Systeme nicht unter Verwendung personenbezogener Daten zu betreiben.

Literatur

Arbeitskreis Technik:	(der Datenschutzbeauftragten des Bundes und der Länder) „Datenschutzfreundliche Technologien“, Schwerin, 1998
Baeriswyl, B.:	„Data Mining und Data Warehousing: Kundendaten als Ware oder geschütztes Gut?“, RDV 2000, S. 6 ff.
Bizer, J.:	„TK-Daten im Data Warehouse“, DuD 10/1998
Büllesbach, A.:	„Datenschutz bei Data Warehouses und Data Mining“, CR 1/2000, S. 11 ff.
Knoll, U.:	„Echtes Wissen kommt aus Daten“, UNIXopen 11/1998
Management Circle GmbH:	Seminarangebot SAP Business Information Warehouse, 1999
Management Circle GmbH:	Seminarangebot Praxisleitfaden für erfolgreiche Data Warehouse-Konzepte, 1999
Möller, F.:	„Ungeschliffene Diamanten“, Datenschutz Nachrichten 3/1998
Möller, F.:	„Data Warehouse als Warnsignal an die Datenschutzbeauftragten“, DuD 10/1998
Möncke, U.:	„Data Warehouses – eine Herausforderung für den Datenschutz?“, DuD 10/1998
Schweizer, A.:	„Data Mining Data Warehousing, Datenschutzrechtliche Orientierungshilfen für Privatunternehmen“, Zürich, 1999

Abkürzungsverzeichnis

Abs.	Absatz
Art.	Artikel
BbgDSG	Brandenburgisches Datenschutzgesetz
BDSG	Bundesdatenschutzgesetz
BKK	Betriebskrankenkassen
BVerfGE	Amtliche Sammlung der Entscheidungen des Bundesverfassungsgerichts nach Band und Seite
bzw.	beziehungsweise
DBMS	Datenbank-Management-System
DSG M-V	Landesdatenschutzgesetz von Mecklenburg-Vorpommern
DV	Datenverarbeitung
EG	Europäische Gemeinschaft
ggf.	gegebenenfalls
i. d. R.	in der Regel
i. S. v.	im Sinne von
IT	Informationstechnik
i. V. m.	in Verbindung mit
o. g.	oben genannte(n)
PET	privacy enhancing technologies
sog.	so genannte(n)
u. a.	unter anderem
usw.	und so weiter
u. U.	unter Umständen
vgl.	vergleiche
z. B.	zum Beispiel