Modellierung und Validierung von Datenschutzanforderungen in Prozessmodellen

Sven Feja, Sören Witt, Andreas Brosche, Andreas Speck

AG Angewandte Informatik (Wirtschaftsinformatik)
Institut für Informatik - Christian-Albrechts-Universität zu Kiel
Hermann-Rodewald-Str. 3, D-24118 Kiel
(svfe|swi|abro|aspe)@informatik.uni-kiel.de

Christian Prietz

Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein (ULD)
Holstenstr. 98, D-24103 Kiel
uld37@datenschutzzentrum.de

Abstract: Die Compliance von Geschäftsprozessen gegenüber Datenschutzanforderungen ist ein grundlegendes Anliegen von Organisationen und Betroffenen. Für den gesetzeskonformen Umgang mit personenbezogenen Daten muss das Datenschutzmanagement einer Vielzahl technischer und organisatorischer Anforderungen genügen. Es ist die Aufgabe des Datenschutzmanagements, die Datenschutzkonformität in den bestehenden Geschäftsprozessen zu gewährleisten.
Dieser Beitrag stellt eine werkzeuggestützte Methode vor, Geschäftsprozesse durch Erweiterung herkömmlicher Prozessmodellierungstechniken mit Datenschutzaspekten auf grafischer Ebene zu annotieren. Diese Methode erlaubt darüber hinaus die grafische Formulierung von Regelwerken, die die Einhaltung von Datenschutzaspekten fordern. Die Prozesse können werkzeuggestützt auf die Einhaltung der Regeln geprüft werden. Das vorgestellte Verfahren vereinheitlicht und erleichtert die Planung, Verifikation und Dokumentation datenschutzkonformer Geschäftsprozesse, was beispielsweise eine Zertifizierung der Compliance erleichtert.

Einleitung

Die Einhaltung von Datenschutzanforderungen ist Pflicht und Anliegen von Organisationen wie Unternehmen und öffentlichen Einrichtungen sowie Betroffenen. Jedoch die Erfahrungen des *Unabhängigen Landeszentrums für Datenschutz Schleswig-Holstein* (ULD) zeigen, dass der Einführung eines Datenschutzmanagements in Organisationen oft zu wenig Beachtung zukommt — beispielsweise im Vergleich zur IT-Sicherheit. Gründe dafür sind unter anderem, dass Datenschutz und dessen Einhaltung als aufwändig und hinderlich empfunden werden. Dies betrifft insbesondere die Anpassung gewachsener Strukturen und Abläufe im Hinblick auf Compliance.
Dennoch muss ein funktionierendes Datenschutzmanagement für Organisationen als erstrebenswert erachtet werden. Eine Zertifizierung der Compliance kann für Unterneh-

men beispielsweise Wettbewerbsvorteile bieten, da sie die Vertrauensbasis in Geschäftsbeziehungen stärkt. Behörden profitieren unter anderem von einer erhöhten Transparenz, und allgemein erfahren Organisationen Vorteile durch erhöhte Rechtssicherheit.
Grundlage eines funktionierenden Datenschutzmanagements sind nach Bizer „definierte und abgesicherte Prozesse, die innerhalb der Organisation präventiv zur Feststellung und Minimierung von Risiken beitragen“ [Biz06]. Diese definierten und abgesicherten Prozesse steigern zum einen die Transparenz und Beherrschbarkeit der Abläufe in der Organisation, zum anderen ermöglichen sie die Integration von Datenschutzanforderungen auf der Prozessebene. Eine Voraussetzung für definierte Prozesse ist die Dokumentation der Abläufe in der Organisation. Neben reiner „Textdokumentation“, bei der sämtliche Informationen zu den Prozessen in Text- und Tabellendokumenten aufgenommen wird, findet zunehmend die Modellierung von Geschäftsprozessen Anwendung. Die dazu verwendeten Modellierungsnotationen sind bspw. die *Ereignisgesteuerte Prozesskette* (EPK) [KNS92] oder die *Unified Modeling Language* (UML). Diese Notationen werden zur Modellierung der fachlichen Abläufe genutzt, was oftmals mit Unterstützung durch Werkzeuge wie z.B. dem *ARIS Business Architect* (ABA) [IDS] geschieht. So modellierte Geschäftsprozesse — hier als *Prozessmodelle* bezeichnet — besitzen eine definierte Syntax und Semantik.
In diesem Beitrag wird eine Methode vorgestellt, wie Datenschutzaspekte in Prozessmodellen verankert und die Prozesse durch Validierung auf Compliance überprüft werden können. Die Validierung beruht auf — ähnlich wie Prozessmodelle — grafisch formulierten Regeln, die aus Datenschutzanforderungen abgeleitet werden. Diese Methode wird als *integrierte Datenschutzmodellierung* bezeichnet. Die Berücksichtigung des Datenschutzes wird hier als eine spezielle *Sicht* in einen Prozessmodellierungsworkflow integriert. Eine Sicht existiert neben weiteren und blendet für den gezeigten Aspekt irrelevante Informationen aus den Modellen aus. Datenschutzmanagement wird somit Bestandteil der Modellierung von Prozessen und nimmt keine Sonderstellung ein. Zusatzaufwand wird verringert und die Methode ist robust im Hinblick auf Veränderungen an den Prozessen. In einer weiteren Sicht könnten beispielsweise Aspekte der Webservice-Security mit eigenen Regelwerken modelliert werden, wie in [JF09] beschrieben. Der Datenschutzmodellierer sieht davon jedoch allenfalls Teile, die den Datenschutz betreffen.
Um Aspekte wie den Datenschutz in die Prozessmodelle aufzunehmen, müssen diese erweitert werden. Die Betrachtung juristischer Datenschutzanforderungen ist in bisherigen Methoden nicht berücksichtigt worden [KBK06]. Für die Entwicklung von Web-Anwendungen werden daher in [KBK06] entsprechende Erweiterung für eW3DT vorgestellt. Dabei werden Sprachkonzepte zur Abspeicherung rechtlicher Anforderungen hinzugefügt. Die Überprüfung der Korrektheit der Eintragungen obliegt aber dem Anwender. In diesem Beitrag werden hingegen die zur Einhaltung der Compliance zu erfüllenden Anforderungen mit Hilfe der grafischen Regeln formuliert, wie sie in [FF08] vorgestellt wurden. Die Prozessmodelle werden automatisiert gegen diese Regeln geprüft.
Prozessmodelle haben alleinstehend lediglich reinen Dokumentationscharakter. Ihre definierte Syntax und Semantik qualifiziert sie über die erwähnte Validierung hinaus auch als Grundlage für die Softwareentwicklung. Sie können im Rahmen der modellgetriebenen Softwareentwicklung [VS06], die die Generierung von Quellcode aus schrittweise verfeinerten Modellen verfolgt, angewendet werden. Zu nennen ist beispielsweise

die Transformation eines fachlichen Prozesmodells in einen technisch ausführbaren Prozess. Konkret wird dazu in [SI07] die Transformation einer EPK in einen BPEL-Prozess (*Business Process Execution Language* [ACD+03]) auf Basis des ABA vorgestellt. Dabei ist die BPEL eine auf Web Services basierende, automatisch ausführbare Prozessmodellierungssprache. In [LPC07] wird für die BPEL das *Privacy Inspection and Monitoring Framework* entwickelt. Grundlage dieses Frameworks sind das *Privacy Data Model* und die *Privacy Policies*. In dem Datenmodell werden Datentypen, Benutzer und Zwecke der Datenbenutzung abgelegt, die Policies (Richtlinien) beschreiben die Rechte und Pflichten bezüglich der personenbezogenen Daten. Datenmodell und Richtlinien sind auch in der hier vorgestellten integrierten Datenschutzmodellierung elementare Bestandteile, werden jedoch im Gegensatz zur textuellen Darstellung in einer grafischen Repräsentation be- und verarbeitet.
Zur automatisierten Verarbeitung der grafisch formulierten Richtlinien müssen diese in ein maschinenlesbares Format überführt werden. Sie werden dazu beispielsweise in die *Enterprise Privacy Authorization Language* (EPAL) transformiert. Diese maschinenlesbaren Richtlinien sind dann für die Generierungsprozesse im Rahmen einer modellgetriebenen Softwareentwicklung direkt einsetzbar. Aufgrund des oft als zu umfangreich empfundenen Sprachumfangs der Policy-Beschreibungsprachen existieren neben den Standards eine Vielzahl an Entwicklungen, wie sie bspw. in [NZJK06] definiert wurden.
Neben den beschriebenen technischen Aspekten hat sich in der Vergangenheit — vor allem im Bereich der Gütesiegel- und Auditarbeit des ULD — gezeigt, dass Maßnahmen und Prozesse zum Datenschutz dann am besten umgesetzt werden und die gewünschten Ergebnisse erzielen, wenn sie nicht losgelöst, sondern in das Gesamtprozessmodell der Organisation eingebunden sind. Dies gilt sowohl für die organisatorische als auch für die technische Einbindung.
Die hier vorgestellte Methode lässt sich bspw. im Rahmen des in [Dem86] vorgestellten und als *Deming-Kreis* bekannt gewordenen *Plan-Do-Check-Act-Verfahrens* (PDCA) einsetzen. In dem PDCA-Verfahren wird in der Plan-Phase zunächst das Vorgehen bzw. der Ablauf für den betrachteten Prozess festgelegt. In der sich anschließenden Do-Phase wird das geplante Vorgehen umgesetzt und in der Check-Phase ein Soll-Ist-Abgleich durchgeführt. Die Ergebnisse der Check-Phase werden in der Act-Phase korrigierend eingesetzt, um ein ggf. verfehltes Ziel aus der Plan-Phase noch zu erreichen. Um eine schrittweise Verbesserung der betrachteten Prozesse zu erzielen, wird der PDCA-Kreislauf kontinuierlich wiederholt. Für die in diesem Beitrag vorgestellte integrierte Datenschutzmodellierung kommt daher der Plan-Phase die größte Bedeutung zu, da hier die Planung und Ausgestaltung der Integration des Datenschutzes in die Prozesse und ggf. die vorherige Prozesserstellung durchgeführt wird.
Im folgenden Abschnitt werden zunächst die Grundlagen der integrierten Datenschutzmodellierung vorgestellt. Darauf folgt die Erläuterung der für das in Abschnitt 0 entwickelte Vorgehensmodell notwendigen rechtlichen als auch technischen Rahmenbedingungen.

Integrierte Datenschutzmodellierung

Die integrierte Datenschutzmodellierung ermöglicht es, neben der Modellierung der fachlichen Abläufe auch die Aspekte des Datenschutzes in die Prozessmodelle mit einzubeziehen. Grundsätzlich verfolgen auch die im vorangegangenen Abschnitt vorgestell-

ten Methoden [KBK06] bzw. [LPC07] auf Modellebene bzw. Ausführungsebene dieses Ziel. Das in diesem Beitrag vorgestellte Verfahren berücksichtigt darüber hinaus zwei grundsätzliche Prinzipien:
Alle notwendigen Aufgaben sollen auf der Prozessmodellebene realisierbar sein.
Neben der Modellierung sollen Prüftechniken die Erfüllung der gestellten Anforderungen sicher stellen
Das erste Grundprinzip benötigt ein Modellierungswerkzeug, mit dem nicht nur die fachlichen Prozesse, sondern auch die Datenschutzaspekte in Form von *Annotierungen* modellierbar sind. Dies erfordert die Erweiterung der Modellierungsnotation. Hinzu kommen grafisch formulierte Regeln, gegen die die Prozesse validiert werden. Um eine Trennung der einzelnen Modellierungsaspekte (wie Fach- und Datenschutzmodellierung) zu erreichen sowie die Übersichtlichkeit und damit Beherrschbarkeit zu wahren, erlaubt der vorgestellte Ansatz, die Modelle unter bestimmten Gesichtspunkten zu betrachten. Dies ist über eine Sichtenbildung für die entsprechenden Aspekte realisiert. Ein Überblick dazu wird in Abschnitt 0 anhand eines Beispielprozesses gegeben. Als Ergänzung zu dem Prozessmodell muss eine Beschreibung der in den Prozessen verwendeten Daten existieren. Die getrennte Beschreibung fördert neben der Übersicht über die Daten auch die Wiederverwendbarkeit der Daten in den Modellen.
Das zweite Grundprinzip ist die Unterstützung des Anwenders bei der Überprüfung der Korrektheit der Modelle, die etwa über Modellprüfverfahren erreicht werden kann. Die dafür notwendigen Regeln werden, wie im ersten Grundprinzip gefordert, grafisch formuliert. Durch einen domänenspezifischen und allgemeingültigen Regelbestand werden erste Anforderungen vordefiniert. Die Ergebnisse des Prüfverfahren werden dann im Modellierungswerkzeug angezeigt. Durch den Einsatz der Prüftechnik wird die Beherrschbarkeit von umfangreichen Prozessmodellen stark erleichtert.

2.1 Rechtlicher Hintergrund

Grundlage für die Erhebung, Verarbeitung und Nutzung personenbezogener Daten ist die in §4 BDSG geregelte Zulässigkeit. Zulässig ist die Erhebung, Verarbeitung und Nutzung hiernach, sofern sie nach dem BDSG selbst oder einer anderen Rechtsvorschrift erlaubt ist oder der Betroffene in die Verarbeitung eingewilligt hat. Eng verknüpft mit der Zulässigkeit der Datenverarbeitung ist die in §14 Abs. 1 festgeschriebene Zweckbindung. Hiernach dürfen personenbezogene Daten nur dann gespeichert, verändert oder genutzt werden, wenn dies für die Zwecke erfolgt, für die die Daten erhoben wurden. Nach der Verarbeitung sind die Daten zu löschen, wenn der für die Erhebung und Verarbeitung zugrunde liegende Zweck weggefallen ist. Die Verarbeitung personenbezogener Daten ist jedoch nicht nur mit deren Zulässigkeit verbunden. Darüber hinaus stehen dem Betroffenen sowohl in den bundes- wie auch den landesgesetzlichen Regelungen eine Vielzahl von Rechten zu, die bspw. in den §§19, 20, 34 und 35 BDSG geregelt sind.

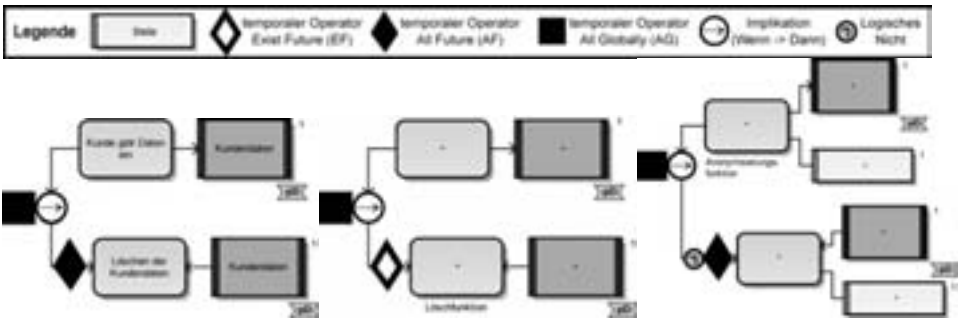
2.2 Schutzziele

Gesetzliche Bestimmungen, wie sie zuvor beschrieben wurden, unterliegen einem stetigen Wandel, die Gesetze unterscheiden sich auf Bundes- und Landesebene und es be-

steht oftmals Interpretationsbedarf. Eine direkte Anpassung der Geschäftsprozesse in einem Unternehmen an den Wortlaut des Gesetzes wird dadurch aufwändig und unterliegt einem erheblichen Wartungsaufwand. Rost und Pfitzmann haben daher, abstrahierend vom Gesetzestext, eine Menge von *Schutzzielen* definiert [RP09]. Sie beschreiben Eigenschaften eines Systems, bei deren Erfüllung angenommen werden darf, dass das System datenschutzkonform arbeitet und die gesetzlichen Vorgaben weitgehend erfüllt oder in Teilen übererfüllt. Eine Ausrichtung der Prozesse an den Schutzzielen kann daher als nachhaltiger und einfacher erachtet werden. Trotz dieser Vereinfachungen kann die Berücksichtigung des Gesetzestextes nicht vollkommen entfallen. Beispielsweise ist die Beachtung der Zulässigkeit oder die Berücksichtigung der Betroffenenrechte nicht durch die Schutzziele abgedeckt.
Die Schutzziele sollen hier nur aufgezählt werden (ihre Definitionen können [RP09] entnommen werden): Vertraulichkeit, Integrität, Verfügbarkeit, Kontingenz, (Un-)Verkettbarkeit, Transparenz. Neben diesen Schutzzielen sind Datensparsamkeit und Datentrennung wichtige Grundsätze des Datenschutzes.

2.3 Validierung der Prozessmodelle

Die Prüfung auf Einhaltung der beschriebenen Schutzziele auf der Ebene der Prozessmodellierung erfolgt durch Validierung der Modelle gegen einen Regelsatz. Die Regeln in diesem Regelsatz werden aus den Schutzzielen abgeleitet und falls notwendig aus den Gesetzen. Sie fordern beispielsweise die Berücksichtigung von Maßnahmen in einem Prozess, die der Durchsetzung eines oder mehrerer Schutzziele dienen. Bei diesen Maßnahmen kann es sich zum Beispiel um kryptografische Mechanismen handeln, die der Durchsetzung der Schutzziele Integrität, Vertraulichkeit oder Unverkettbarkeit dienen. Die Einhaltung eines Schutzziels wird jedoch nicht notwendigerweise durch genau eine Regel geprüft, es wird im Normalfall eine Vielzahl von Regeln zu beachten sein.
Beispielsweise müssen personenbezogene Daten im Prozessmodell immer mit Zulässigkeit für ihre Erhebung und Speicherung verbunden sein. Eine weitere Regel, die in nahezu jedem datenschutzkonformen System zu erfüllen ist, ist die Durchführung der Löschung von Daten, zu deren Speicherung die Grundlage entfallen ist.
Die genannten Beispiele unterscheiden sich insbesondere darin, dass die erste Regel, personenbezogene Daten mit einer Zulässigkeit zu versehen, statisch ist und somit überall im Prozess gilt. Hingegen weist die zweite Regel einen dynamischen Charakter auf, denn das Auftreten eines personenbezogenen Datums verlangt die Möglichkeit bzw. Notwendigkeit, es im Verlauf der Prozesse wieder löschen zu können oder es zu löschen. Statische Regeln werden in diesem Beitrag in boolescher Logik formuliert, dynamische Regeln bedienen sich der temporalen Logik *Computation Tree Logic* (CTL). Insbesondere für die oftmals komplizierteren CTL-Regeln wurde in [FF08] die grafische Formulierung dieser Regeln in G-CTL vorgestellt. G-CTL erlaubt es, Prozesselemente in der Regelformulierung zu verwenden, so dass eine sehr direkte Formulierung der Regeln möglich ist.



szenariospezifische Löschregel allgemeingültige Löschregel domänenspezifische Anonymisierungsregel

Abbildung **20** Beispiele für G-CTL-Regeln

Die Forderung, personenbezogene Daten wieder zu löschen, ist in der G-CTL-Regel in Abbildung 1(a) formuliert. Von oben nach unten gelesen ist sie folgendermaßen zu interpretieren: Für jedes Vorkommen der Funktion `Kunde gibt Daten ein`, welche die Ausgabedaten `Kundendaten` erzeugt, muss im gesamten Prozess mindestens eine Funktion `Löschen der Kundendaten` existieren, die genau dieses Datum `Kundendaten` wieder löscht. Die Gültigkeit der Regel im gesamten Prozess wird durch den temporalen Operator *All Globally*(AG) vor der Implikation bestimmt. Der temporale Operator *All Future*(AF) vor der Funktion `Löschen der Kundendaten` verlangt, dass diese Funktion in jedem möglichen Verlauf des Prozesses erreicht wird. Die Identität der Daten wird durch die Nummer oben rechts am Datensymbol ausgezeichnet. Durch die Identität wird das Löschen genau dieser zuvor erhobenen Daten verlangt. Bei den erzeugten und zu löschenden Daten handelt es sich um personenbezogene Daten, was durch das mit pD bezeichnete Symbol rechts unten ausgezeichnet wird. Die beschriebene Regel wird auf bestimmte Muster im Prozess angewendet, nämlich auf Funktionen mit dem Namen `Kunde gibt Daten ein`, die ein Datum `Kundendaten` erzeugen. Dieses Prinzip, Regeln auf bestimmte Muster im Prozess anzuwenden, kann zur Verallgemeinerung und Wiederverwendung genutzt werden. In Abbildung 1(b) ist die Regel zum Löschen personenbezogener Daten in einer allgemeinen Form abgebildet. Ein wesentlicher Schritt zur Verallgemeinerung besteht in der Eliminierung von konkreten Bezeichnern. Diese Regel muss gelten, wann immer durch eine Funktion personenbezogene Daten erzeugt werden. Sie fordert, dass eine beliebig bezeichnete Funktion existiert, die mit dem Attribut `Löschfunktion` ausgezeichnet ist und diese personenbezogenen Daten als Eingabe nimmt. Im Zuge der Verallgemeinerung kann hier nicht mehr gefordert werden, dass der Prozess in jedem Fall zu einer Löschung der Daten gelangt. Der Grund ist, dass beispielsweise Daten der Bestandskunden eines Webshops solange gespeichert bleiben, bis ein Kunde selbst eine Löschung veranlasst oder diese anderweitig notwendig wird. Es wird nunmehr deshalb gefordert, dass es zu einer Löschung der Daten kommen *kann*, also ein Pfad im Prozess existiert, in dem diese Funktion ausgeführt wird. Diese Änderung erfolgt durch Austausch des AF-Operators vor der Löschfunktion durch den *Exist Future* (EF) Operator. Dieser fordert lediglich die Existenz und Erreichbarkeit dieser Funktion.

Diese Verallgemeinerung ist hier also mit einer Aufweichung der konkreteren Regel verbunden, so aber auf nahezu jeden Prozess im Hinblick auf den Datenschutz anwendbar. Eine speziellere Regel, die wie jene in Abbildung 1(a) eine stärkere Forderung repräsentiert, kann dennoch widerspruchsfrei mit der weicheren Regel auf einen Prozess angewendet werden.
In [DF09] wurde eine Kategorisierung der Regeln gemäß ihres Abstraktionsgrades beschrieben. Sie werden eingeteilt in allgemeingültig, domänenspezifisch und szenariospezifisch. Allgemeingültige Regeln sind stark abstrahiert und können auf nahezu jeden Prozess angewendet werden. Ihre Erstellung verlangt keine speziellen Kenntnisse aus der fachlichen Domäne, in der die Prozesse angesiedelt sind. Domänenspezifische Regeln finden hingegen Anwendung in bestimmten Fachgebieten, ihr Abstraktionsgrad ist hoch. Sie verwenden beispielsweise fachspezifische Annotationen, um bestimmte Elemente eines Prozesses zu identifizieren. Szenariospezifische Regeln beziehen sich hingegen auf einzelne Prozesse und verwenden insbesondere konkrete Elemente der betrachteten Prozesse.
Während die Regel in Abbildung 1(a) durch ihre konkrete Formulierung als szenariospezifisch einzustufen ist, zeigt Abbildung 1(c) ein Beispiel für eine domänenspezifische Regel. Dargestellt wird eine Anonymisierungsanforderung, wie sie u.a. in der medizinischen Forschung üblich ist. Hiernach darf ein Mitarbeiter, der eine im Prozess als solche gekennzeichnete Anonymisierungsfunktion bearbeitet, mit keiner Funktion, die im zukünftigen Verlauf des Prozesses die anonymisierten Daten als Eingabe verwendet, assoziiert sein.
Es ist zu beachten, dass die anonymisierten Daten weiterhin als personenbezogen markiert bleiben. Diese Daten sind für sich genommen nicht mehr personenbezogen. Jedoch ist die Information (in Form von Mitarbeiterwissen), welche den Bezug zu einer Person herstellen kann, weiterhin im Prozess vorhanden. Somit dient diese Regel also u.a. der Umsetzung des Schutzziels Unverkettbarkeit.
Eine Überprüfung der Regeln erfolgt, wie in [FF08] beschrieben, durch einen Modellprüfer, in dem Prozess und Regeln in dessen Syntax transformiert werden.

2.4 Datenmodell der privaten Daten

Der zweite Grundpfeiler der in diesem Beitrag vorgestellten integrierten Datenschutzmodellierung ist das Datenmodell der privaten Daten. Dieses hat auf der einen Seite zur Aufgabe, die in den (Geschäfts-)Prozessmodellen verwendeten Daten in einer Übersicht darzustellen und gleichzeitig zu strukturieren. Auf der anderen Seite kann das Datenmodell aber auch zur Unterstützung der Wiederverwendung von bereits existierenden — und damit bereits als personenbezogen gekennzeichneten — Datenelementen genutzt werden.
Dabei ist in der Grundform und damit organisationsunspezifischen Form nur ein allgemeiner Datenbestand vorhanden. Dieser allgemeine Datenbestand enthält bspw. typische personenbezogene Daten wie Name, Vorname und Geburtsdatum, aber auch Daten der Kategorie *Persönlichkeitsdaten* wie Hobbies oder Angaben zum Sexualleben. Dieser allgemeine Datenbestand ist aus Erfahrungswerten des ULD entstanden und ist momentan in einer Datenbank des *ABA* abgelegt. Einen Ausschnitt der ARIS DB zeigt Abbildung 2. Die Datencluster (Rechtecke) dienen dabei der Kategorisierung, die Datenele-

mente (Ellipsen) sind den entsprechenden Kategorien zugeordnet. Das „pD“-Symbol zeigt an, dass es sich um personenbezogene bzw. personenbeziehbare Daten handelt.

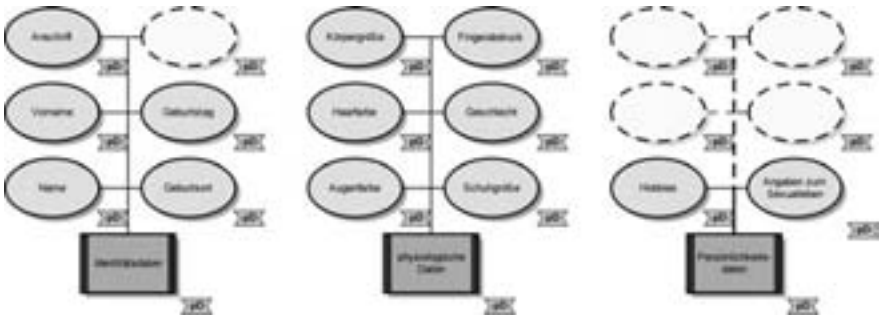


Abbildung **21**: Ausschnitt des Datenmodells der personenbezogenen Daten

Das Datenmodell ist aber nur in wenigen Fällen eine vollständige Liste und muss daher meist an die Anforderungen der entsprechenden Organisation angepasst werden. Neben der einfachen Wiederverwendbarkeit von Datenelementen, die durch die prozessmodellunabhängige Abspeicherung erreicht wird, erleichtert der allgemeine Datenbestand dem Anwender die Einordnung der in den Prozessen verwendeten Daten zu den personenbezogenen Daten.

2.5 Vorgehensmodell

Nachdem in den vorangegangenen Abschnitten die grundlegenden Komponenten der integrierten Datenschutzmodellierung das Datenmodell und die Datenschutzregeln erläutert wurden, können diese in das in Abbildung 3 dargestellte Vorgehensmodell eingebettet werden.

Abbildung **22**: Integrierte Datenschutzmodellierung

Unter Einbeziehung des Datenmodells bzw. der Datenbank der personenbezogenen Daten (DB pD) wird die fachliche Prozessmodellierung durchgeführt. D.h., soweit es möglich ist, werden die vorhanden Datenelemente auch im fachlichen Prozessmodell verwendet. Ist das Prozessmodell bereits vorhanden, können die aktuellen Bezeichnungen beibehalten werden. Es bestehen in diesem Fall zwei Möglichkeiten: Erstens kann über entsprechende Annotationen eine Zuordnung zu den in DB pD vorhanden Datenelementen durchgeführt werden, womit auch speziell auf diese Datenelemente zugeschnittene

Datenschutzregeln greifen. Zweitens ist es aber auch möglich, ein eigenes Datenmodell der personenbezogenen Daten zu erstellen.
Für den Einsatz von Prüftechniken sind die Modellelemente des Prozessmodells um die Annotationen **personenbezogene Daten** (ApD) und ggf. um die **Referenz zur DB pD** zu erweitern. ApD kann dabei die Werte `true`, `false` und `nicht gepflegt` annehmen. Der entsprechend gesetzte Wert zeigt an, ob es sich um personenbezogene Daten handelt oder nicht bzw. ob bisher keine Zuordnung statt gefunden hat. Die so annotierten Prozessmodelle können dann durch die in Abschnitt 0 beschriebene Technik automatisiert überprüft werden. Als Ergebnis wird in dem aktuell vorliegenden Prototyp eine Text-Datei angezeigt, welche im Fehlerfall die nicht erfüllten Regeln benennt. In nächsten Erweiterungen ist die Visualisierung der Fehler im Prozessmodell nach dem in [FSP09] vorgestellten Verfahren geplant.

Beispielprozess

Zur Demonstration des Einsatzes der integrierten Datenschutzmodellierung dient der in Abbildung 4 links dargestellte Geschäftsprozess aus dem Bereich des E-Commerce. Es handelt sich konkret um den Bezahlprozess eines Kunden in einem Onlineshop, der die Wahl zwischen der Einmal-Bestellung und der Bestandskunden-Bestellung hat. In diesem Beitrag soll speziell der Fall der Einmal-Bestellung betrachtet werden. Das Entscheidende an diesem Fall ist, dass die Kundendaten mit Beendigung des Geschäftsvorgangs gelöscht sein müssen. Es ist allerdings anzumerken, dass dies nur greift, wenn der Kunde während der Eingabe seiner Daten die Speicherung dieser verweigert (Opt-out-Prinzip). Aus Platzgründen wird daher beim Aufnehmen der Kundendaten (Funktion `Kunde gibt Daten ein`) die Ablehnung der Datenspeicherung angenommen. Der Onlineshop kann allerdings unter Berufung auf die Gewährleistungsfrist die Daten noch eine gewisse Zeit speichern. Im Geschäftsprozessmodell muss die Löschung aber am Ende des Geschäftsvorgangs gesichert sein.
Der vorgestellte Geschäftsprozess wurde als EPK im ABA modelliert. In Erweiterung zur fachlichen Darstellung auf der linken Seite von Abbildung 4 wird auf der rechten Seite die Datenschutz-Sicht auf den Geschäftsprozess gezeigt. Technisch gesehen wird diese Sicht im ARIS über eine Vorlage für die grafische Darstellung des Prozessmodells gelöst. Dementsprechend handelt es sich prinzipiell um den selben Prozess, nur mit unterschiedlicher Sichtweise. In der Datenschutz-Sicht gibt es zu den üblichen Prozesssymbolen noch die Annotationssymbole für die Kennzeichnung von personenbezogenen bzw. personenbeziehbaren Daten und die Vorbedinungssymbole als Annotation an die Datencluster. Die Vorbedingung ordnet dabei die rechtliche Zulässigkeit der Verarbeitung der Daten diesem Prozessschritt zu. Sowohl die Regel aus Abbildung 1(a) als auch aus Abbildung 1(b) werden erfüllt. Die Anonymisierungsregel aus Abbildung 1(c) kann in diesem Fall nicht angewandt werden.

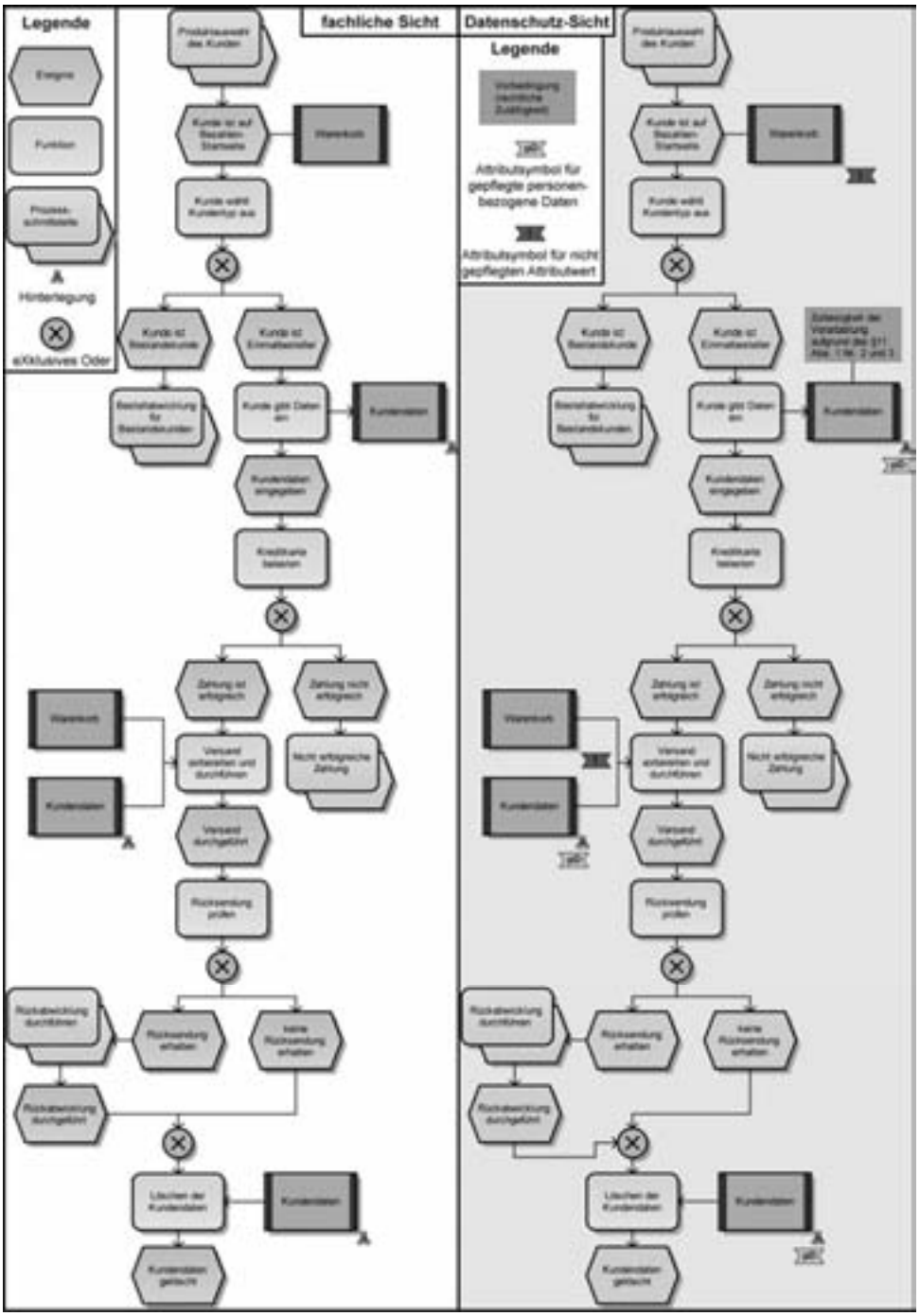


Abbildung **23**: Die fachliche und die Datenschutz-Sicht anhand des Beispielprozesses.

Zusammenfassung und Ausblick

Die in diesem Beitrag vorgestellte werkzeuggestützte, integrierte Datenschutzmodellierung ermöglicht die Einbeziehung von Datenschutzaspekten in etablierte Modellierungsverfahren für Prozesse. Insbesondere erlaubt diese Methode eine Validierung der Prozessmodelle gegen grafisch formulierte Regelwerke, die aus den Schutzzielen abgeleitet werden und bei entsprechend allgemeiner Formulierung einen hohen Grad an Wiederverwendbarkeit aufweisen. Valide Prozessmodelle stellen ferner eine wichtige Grundlage für die Verwendung im Kontext generativer Programmierung dar. Dies ist ein weiterer Mehrwert neben dem gesteigerten Vertrauen in die Compliance, deren Zertifizierung (beispielsweise durch ein Audit des ULD) so ebenfalls erheblich erleichtert werden kann. Die Datenschutzsicht nimmt keine Sonderstellung ein. Sie steht gleichberechtigt neben anderen Sichten auf die Prozessmodelle, die ebenfalls eigene Regelwerke besitzen, die im gleichen Zuge validiert werden können.
Zu den in diesem Beitrag betrachteten Datenschutzanforderungen für Datenelemente sollen in den zukünftigen Arbeiten noch Erweiterungen für die Modellierung eines Rollen- und Berechtigungskonzeptes eingeführt werden. Diese sind eine grundlegende Voraussetzung für die vollständige Erstellung bzw. Generierung von maschinenlesbaren Policies. Weiterhin ist, wie bereits erwähnt, die Visualisierung der Validierungsergebnisse innerhalb der Prozessmodelle geplant.

Literaturverzeichnis

[ACD+03] Tony Andrews, Francisco Curbera, Hitesh Dholakia, Yaron Goland, Johannes Klein, Frank Leymann, Kevin Liu, Dieter Roller, Doug Smith, Satish Thatte, Ivana Trickovic, und Sanjiva Weerawarana. Business Process Execution Language for Web Services Version 1.1. *OASIS Standard*, May 2003.

[Biz06] Johann Bizer. Datenschutz in die Prozesse. *Datenschutz und Datensicherheit - DuD*, 30(10):598, Oktober 2006.

[Dem86] William E. Deming. *Out of the crisis*. Massachusetts Institute of Technology, Cambridge, 1986.

[DF09] Jens Drawehn und Sven Feja. Anwendung von grafischen Validierungsregeln bei der Entwicklung von IT-Integrationsprozessen. In *Workshop Modellgetriebene Softwarearchitektur – Evolution, Integration und Migration (MSEIM)*, Lecture Notes in Informatics (LNI), Kaiserlautern, Germany, March 2009.

[FF08] Sven Feja und Daniel Fötsch. Model Checking with Graphical Validation Rules. In *15th IEEE International Conference on the Engineering of Computer-Based Systems (ECBS 2008), Belfast, NI, GB*, pages 117–125. IEEE Computer Society, April 2008.

[FSP09] Sven Feja, Andreas Speck, und Elke Pulvermüller. Business Process Verification. In INFORMATIK 2009, Im Focus das Leben, Beiträge zur 39. Jahrestagung der Gesellschaft für Informatik e. V. (GI)., 2009.

[IDS] IDS Scheer AG. IDS Scheer AG - Country Site DE: ARIS Software. http://www.ids-scheer.de/de/ARIS_Software_Software/7796.html. 2007-11-15.

[JF09] Meiko Jensen und Sven Feja. A Security Modeling Approach for Web-Service-based Business Processes. In 16th IEEE International Conference on the Engineering of

Computer-Based Systems (ECBS 2009), 13 - 16 April 2009, San Francisco, CA, USA. IEEE Computer Society, April 2009.
[KBK06] Ralf Knackstedt, Christian Brelage, und Noogie C. Kaufmann. Entwicklung rechtssicherer Web-Anwendungen. *Wirtschaftsinformatik*, 48(1):27–35, 2006.
[KNS92] G. Keller, M. Nüttgens, und August-Wilhelm Scheer. Semantische Prozessmodellierung auf der Grundlage Ereignisgesteuerter Prozessketten (EPK). Arbeitsbericht Heft 89, Institut für Wirtschaftsinformatik Universität Saarbrücken, 1992.
[LPC07]Yin Hua Li, Hye-Young Paik, and Jun Chen. Privacy Inspection and Monitoring Framework for Automated Business Processes. In *WISE 2007, 8th International Conference on Web Information Systems Engineering, Nancy, France, December 3-7, 2007, Proceedings*, volume 4831 of *Lecture Notes in Computer Science*, pages 603–612. Springer, 2007.
[NZJK06] Surya Nepal, John J. Zic, Frederic Jaccard, und Gregoire Kraehenbuehl. A Tag-Based Data Model for Privacy-Preserving Medical Applications. In *Current Trends in Database Technology - EDBT 2006, EDBT 2006 Workshops PhD, DataX, IIDB, IIHA, ICSNW, QLQP, PIM, PaRMA, and Reactivity on the Web, Munich, Germany, March 26-31, 2006, Revised Selected Papers*, volume 4254 of *Lecture Notes in Computer Science*, pages 433–444. Springer, 2006.
[RP09] Martin Rost und Andreas Pfitzmann. Datenschutz-Schutzziele - revisited. *Datenschutz und Datensicherheit - DuD*, 33(6):353–358, Juni 2009.
[SI07] S. Stein und K. Ivanov. EPK nach BPEL Transformation als Voraussetzung für praktische Umsetzung einer SOA. In Wolf-Gideon Bleek, Jörg Raasch, and Heinz Züllighoven, editors, *Software Engineering 2007*, volume 105 of *Lecture Notes in Informatics (LNI)*, pages 75–80, Hamburg, Germany, März 2007. Gesellschaft für Informatik (GI).
[VS06] Markus Völter und Thomas Stahl. Model-Driven Software Development : Technology, Engineering, Management. John Wiley & Sons, June 2006.