

Spezifikation von Privacy Policies und Privacy Preferences

Nuh Keles

Betreuer: Mirco Stern

Universität Karlsruhe (TH)

1 Einführung

Wir befinden uns heute in einer Zeit, in der die Information und die Kommunikation von ganz essentieller Bedeutung sind. Den Vermutungen zufolge verdoppelt sich alle 20 Monate die Informationsvielfalt auf der ganzen Welt [2]. Der explosive Fortschritt im Bereich der Informationstechnologie bringt jedoch neue Probleme und Herausforderungen mit sich. Eines dieser Probleme ist die Privatheit.

1.1 Definition von Privatheit

Privatheit hat viele Definitionen. Es ist das Recht jedes Einzelnen, jeder Gruppe und jeder Institution für sich selbst zu entscheiden, wann, wie und zu welchem Zweck Informationen über sie an andere weitergegeben werden. Unsere Definition baut auf die vorherige Definition auf und verwendet eine Definition der Privatheit, als ein auf Information und Kommunikation basierendes Konzept, das heißt die Methode und der Umfang, mit dem man kontrollieren kann, welche Informationen über einen *gesammelt, gespeichert, verwendet* oder *weitergegeben* werden [5]. Somit entscheidet jeder für sich selbst, welche Informationen von wem gesammelt, wie lange aufbewahrt und für welche Zwecke sie benutzt und mitgeteilt werden.

1.2 Sicherstellung der Privatheit

Privatheit kann auf verschiedenen Wegen sichergestellt werden.

1.2.1 Keine Herausgabe privater Daten

Dies ist eine der einfachsten Möglichkeiten zur Sicherstellung der Privatheit, jedoch nicht immer möglich. Ein Verzicht auf die Herausgabe privater Daten würde beiden Parteien nicht weiterhelfen. Z.B. ein Kunde möchte ein gebrauchtes Notebook bei *ebay* kaufen. Für den Kauf sind Informationen wie Geburtsdatum, Anschrift und Kontoinformationen notwendig. In diesem Fall ist man gezwungen, seine persönlichen Informationen preiszugeben, sonst käme es nicht zu einer Verhandlung.

1.2.2 Anonyme Herausgabe privater Daten

Die nächste Alternative zur Sicherstellung der Privatheit ist die anonyme Herausgabe privater Daten. Um diese Anonymität zu bewahren tendieren viele Menschen dazu, sich mit falschen Angaben zur Person vorzustellen. Ein Beispiel hierfür wäre das Chatten im Internet (Angabe des Nicknames) oder bei der Erstellung der E-Mail Adresse (Angaben zur Person). Aber anonymer Dienstaufwurf macht nicht immer Sinn. Bei den meisten praktischen Anwendungen ist die Preisgabe der persönlichen Daten notwendig oder vielleicht nützlich. Beispiele hierfür wären Online-Bewerbung, Online-Shopping.

1.2.3 Kontrolle der Verwendung bei Herausgabe privater Daten

Die letzte Möglichkeit ist die Kontrolle der Verwendung bei Herausgabe privater Daten, welches das Hauptaugenmerk für uns bildet. Preisgabe persönlicher Informationen ist alltäglich! Z.B. Wenn man über das Internet etwas bestellt, ist man gezwungen dem Händler persönliche Daten, wie seinen Namen und Adresse mitzuteilen. Falls man beispielsweise via Kreditkarte bestellt, ist man auch noch gezwungen seine Kreditkarteninformationen weiterzugeben. Manche Händler gehen weiter und bitten um die Angabe weiterer Informationen, wie Telefonnummern oder E-Mail Adressen. Viele Händler legen so genannte Accounts an, damit man als Kunde nicht jedes Mal wieder seine Daten angeben muss, sobald man eine erneute Bestellung aufgeben möchte. Auf diese Accounts kann der Kunde über ein Benutzernamen und Passwort zugreifen. Diese Accounts können dafür benutzt werden, um das Kaufverhalten des Kunden über einen längeren Zeitraum aufzuzeichnen und zu beobachten. Doch man kann sich niemals sicher sein was mit diesen Daten geschieht.

Um die Herausgabe privater Daten kontrollieren zu können, ist zunächst eine Spezifikationssprache notwendig, um herauszustellen, wie sich Dienstgeber einerseits und Dienstnehmer andererseits die Verwendung vorstellen. Wir werden zwei von diesen genannten Spezifikationssprachen detaillierter erklären:

- P3P ist eine Policy-Beschreibungssprache und wird vom Dienstgeber benutzt
- APPEL ist eine Präferenz-Beschreibungssprache und wird vom Dienstnehmer benutzt.

Beide, sowohl P3P als auch APPEL sind zwei wichtige Spezifikationssprachen. P3P lässt zu, dass eine Webseite Ihre Datenschutzpraktiken beschreibt. Mit Hilfe von APPEL können Benutzer, ihre Präferenzen ausdrücken. Im Folgenden werden die formalen Beschreibungen von P3P Policies und APPEL Präferenzen anhand von Beispielen erklärt. APPEL hat einige Schwächen, welche näher erläutert werden. Darüber hinaus wird das Zusammenspiel zwischen P3P und APPEL aufgeführt.

Im Abschnitt 3 werden die Probleme bei der Einhaltung von Privatheit thematisiert sowie die Durchsetzung der Privatheit mit Hilfe Hippokratischer Datenbanken aufgezeigt. Zum Schluss fassen wir die wichtigsten Punkte noch einmal zusammen.

2 Spezifikationssprachen

2.1 P3P

P3P (Platform for Privacy Preferences) ist eine durch das World Wide Web-Konsortium (W3C) standardisierte Plattform zum Austausch von Datenschutzinformationen im World Wide Web. P3P ermöglicht den schnellen Zugriff auf Datenschutzinformationen einer Webseite. Mit Hilfe von P3P sollen die Benutzer einen schnellen Überblick über die Datenschutzpraxis einer Webseite bekommen

können [7]. Außerdem ermöglicht P3P den automatisierten Abgleich der Datenschutzerklärung des Anbieters mit den diesbezüglichen Präferenzen des Nutzers. Damit wird es dem Besucher einer Webseite möglich, schnell, unkompliziert und automatisch auf die jeweilige Datenschutzpraxis zu reagieren.

2.1.1 P3P Privacy Policies

Die P3P Policies erlauben einer Website ihre Datenkollektionen und Datengebrauchspraktiken zu beschreiben. Policies werden in einem XML-Format ausgedrückt. P3P Policies werden als eine Folge von STATEMENT Elemente beschrieben [3], welche die folgenden Subelemente haben:

- **CONSEQUENCE:** beschreibt den beabsichtigten Zweck für Informationssammlung in menschenlesbarem Text.
- **PURPOSE:** beschreibt für welchen Zweck die Informationen gesammelt werden. Mehrere PURPOSEs können in einem STATEMENT beschrieben werden, wenn alle dieselben Werte für RECIPIENT, RETENTION und DATA-GROUPS haben. Sonst werden Sie in verschiedenen STATEMENT-Elementen beschrieben.

Wichtige PURPOSE-Elemente sind:

- *current:* Beendigung und Unterstützung der Tätigkeit, für die Daten zur Verfügung gestellt wurden,
- *individual-decision:* Schließen auf Gewohnheiten, Interessen und andere Eigenschaften der Einzelpersonen,
- *contact:* Verbindungsherstellung mit dem Benutzer für Marketing von Dienstleistungen oder von Produkten durch einen Kommunikationskanal.

- **RECIPIENT:** beschreibt für welche Benutzer die Informationen beabsichtigt sind. Mehrere RECIPIENTs können in einem STATEMENT beschrieben werden.

Wichtige RECIPIENT-Elemente sind:

- *ours:* Dienstgeber selbst,
- *same:* Dritte, die unserer Praxis folgen,
- *unrelated:* Dritte, dessen Praxis uns unbekannt ist.

- **RETENTION:** beschreibt die Aufbewahrungsdauer der gesammelten Informationen.

Wichtige RETENTION-Elemente sind:

- *stated-purpose:* Die Daten werden so schnell wie möglich gelöscht.
- *business-practice:* langfristige Aufbewahrung aber mit aktuellem Verwendungszweck.
- *indefinitely:* unbeschränkt

- **DATA-GROUP:** beschreibt die individuellen Daten, die übertragen werden.

```

<POLICY>
  <STATEMENT>
    <PURPOSE><current/></PURPOSE>
    <RECIPIENT><ours/><same/></RECIPIENT>
    <RETENTION><stated-purpose/></RETENTION>
    <DATA-GROUP>
      <DATA ref="#user.name"/>
      <DATA ref="#user.home-info.postal/>
      <DATA ref="#dynamic.miscdata">
        <CATEGORIES><purchase/></CATEGORIES>
      </DATA>
    </DATA-GROUP>
  </STATEMENT>
  <STATEMENT>
    <PURPOSE>
      <individual-decision required="opt-in"/>
      <contact required="opt-in"/>
    </PURPOSE>
    <RECIPIENT><ours/></RECIPIENT>
    <RETENTION><business-practices/></RETENTION>
    <DATA-GROUP>
      <DATA ref="#user.home-info.online.email/>
      <DATA ref="#dynamic.miscdata">
        <CATEGORIES><purchase/></CATEGORIES>
      </DATA>
    </DATA-GROUP>
  </STATEMENT>
</POLICY>

```

Abb. 1: Beispiel einer Privacy Policy einer Buchhandlung in P3P

Z.B. eine Buchhandlung benötigt bestimmte minimale persönliche Informationen, um einen Verkauf durchzuführen. Diese Informationen schließen Name, Versandadresse und die Kreditkartennummer ein. Sie verwendet auch die Verkaufsgeschichte der Kunden, um persönliche Buchempfehlungen anzubieten. Für die Buchempfehlung benötigt die Buchhandlung zusätzlich die Emailadresse des Kunden.

Abbildung 1 zeigt, wie die Policy der Buchhandlung in der P3P Policy Sprache aussehen kann. Die Policy beinhaltet zwei STATEMENTS. Das erste STATEMENT gibt den Namen, die Anschrift und verschiedene Verkaufsdaten der Person (d.h., Buchtitel, Kreditkartennummer, etc.) für die Durchführung des gegenwärtigen Verkaufs an. Das zweite STATEMENT zeigt verschiedene Verkaufsdaten an, die für persönliche Empfehlungen verwendet werden.

Der *opt-in* Wert des *required* Attributes der Zwecke *individual-decision* und *contact* deutet an, dass die explizite Kundenzustimmung notwendig ist.

2.1.2 Erstellung einer P3P-Datenschutzerklärung

Die Datenschutzerklärungen im Volltext und im P3P-Format müssen letztlich auf einem Webserver des Dienstgebers bereitgestellt und referenziert werden.

Um dem P3P-Agent des Nutzers das Auffinden der P3P-Datenschutzerklärung zu ermöglichen, empfiehlt der P3P-Standard eine Referenzdatei im Verzeichnis *w3c* mit dem Namen *p3p.xml* zu erstellen. Diese Datei verweist auf die P3P-Datenschutzerklärung(en), die für die vorliegende Website gelten. Die Erstellung der Referenzdatei können ebenfalls Software-Tools oder spezialisierte Firmen übernehmen. Die P3P-Datenschutzerklärung(en) können an einem beliebigen Speicherort hinterlegt werden, soweit dieser in der Referenzdatei korrekt angegeben ist. Die Datenschutzerklärung als Textdokument muss innerhalb der P3P-Datenschutzerklärung verlinkt werden und sollte zusätzlich auf der Homepage der Website durch einen Link abrufbar sein.

Derzeit bieten ca. 30 % der 100 meistbesuchten Websites in den USA P3P-Datenschutzerklärungen. Die Zahlen zeigen, dass sich der P3P-Standard in der Version 1.0, der im April 2002 verabschiedet wurde, langsam aber stetig insbesondere bei kommerziellen Webanbietern durchsetzt. Sie zeigen auch, dass für P3P offensichtlich ein Bedarf besteht und die Verwendung einen Mehrwert liefert. Dies lässt auch bei europäischen Websites ein kontinuierlich wachsendes P3P-Angebot erwarten [7].

P3P ist ein technischer Standard und kann keine Garantien geben, dass Anbieter sich auch wirklich an die von ihnen veröffentlichten Praktiken halten. So wie sich die Datenschutzrechte des Internetsurfers in der Hauptsache aus dem jeweils anwendbaren Datenschutzrecht eines Staates oder vertraglichen Vereinbarungen zwischen den Parteien ergeben, so wird die tatsächliche Durchsetzung dieser Rechte vom Rechtssystem eines Staates gewährleistet [7].

2.2 APPEL

APPEL ist eine Präferenz-Beschreibungssprache und wird vom Dienstnehmer benutzt. Präferenzen werden als eine Liste von RULEs beschrieben, welche mit Hilfe eines regelbasierten Mechanismus verarbeitet werden. Die erste Regel, die zutrifft, wird angewandt. Eine Regel besteht aus zwei Teilen:

- **RULE behavior:** beschreibt die Aktion, die vorgenommen wird, wenn die RULE zutrifft.
- **RULE Body :** Hier definiert man seine Präferenzen, welche gegen P3P Policy überprüft werden.

Abbildung 2 zeigt ein Präferenzbeispiel in APPEL. Der RULESET besteht aus zwei RULEs (Regel). Die erste Regel spezifiziert die Bedingungen unter welchen der Zugang zu einer Webseite blockiert werden muss. In diesem Beispiel werden die Webseiten blockiert, dessen Zwecke *contact* (d.h. Informationen können für die Promotion eines Produktes oder des Services verwendet werden) oder *admin* (d.h. Informationen können für die technische Unterstützung der Webseite und des Systems verwendet werden) sind. Wenn die erste Regel zutrifft, wird sie angewendet und die zweite Regel nicht kontrolliert. Ansonsten

```

<appel:RULESET>
  <appel:RULE behavior="block">
    <POLICY>
      <STATEMENT>
        <PURPOSE appel:connective="or">
          <contact/>
          <admin/>
        </PURPOSE>
      </STATEMENT>
    </POLICY>
  </appel:RULE>

  <appel:RULE behavior="request"/>
  <appel:OTHERWISE/>
</appel:RULESET>

```

Abb. 2: Ein Präferenzbeispiel in APPEL

wird die zweite Regel kontrolliert. Die zweite Regel wird garantiert zutreffen, wenn die erste Regel nicht zutrifft und erlaubt den Zugriff auf die Webseite.

Somit blockiert diese Präferenz die Webseiten, dessen Zwecke *contact* oder *admin* sind und erlaubt alle andere Webseiten den Zugriff

APPEL wurde schon in *Internet Explorer 6*, in *Mozilla 1.4*, in *Netscape 7* und natürlich in neueren Versionen implementiert. *Privacy Bird* ist eine Browser Extension für Internet Explorer 5 und beinhaltet eine APPEL Engine, die die Präferenzen des Benutzers gegen die P3P Policy einer Website vergleicht.

2.2.1 Kritik an APPEL

APPEL ist kompakt, lesbar und benutzt XML für die Syntax, aber hat es auch einige Probleme.

Man kann nicht spezifizieren was akzeptabel ist: Der Benutzer können direkt spezifizieren, was in der Methode inakzeptabel ist, aber nicht was akzeptabel ist. Eine Webseite ist akzeptabel, wenn der Benutzer bereit ist, seine persönliche Daten weiter zu geben.

Beispiel: Eine Person möchte eine Präferenz beschreiben, die die Zwecke (PURPOSE) *current* oder *telemarketing* erlaubt, und somit alle andere Webseiten blockiert. Sie schreibt die in Abb.3 angezeigte *Präferenz 1*. Leider erzielt diese Präferenz nicht, was die Person beabsichtigt. Die Quelle des Problems ist, dass die Policy Sprache eine Policy erlaubt, die mehrere STATEMENTS enthält und die Regel dann zutrifft, wenn irgendwelche STATEMENTS die Regel zufrieden stellen. Z.B eine Webseite kann eine Policy haben die zwei STATEMENTS enthält; das erste STATEMENT enthält nur *current* als Zweck und das zweite STATEMENTS enthält *admin* als Zweck. Diese Präferenz wird aufgrund des ersten STATEMENTS zu der Policy der Webseite passen.

```

<appel:RULESET>
  <appel:RULE behavior="request">
    <POLICY>
      <STATEMENT>
        <PURPOSE appel:connective="or-exact">
          <current/><telemarketing/>
        </PURPOSE>
      </STATEMENT>
    </POLICY>
  </appel:RULE>

  <appel:RULE behavior="block"/>
    <appel:OTHERWISE/>
  </appel:RULE>
</appel:RULESET>

```

Abb. 3: Präferenz 1. Diese Präferenz blockiert die inakzeptablen Webseiten nicht, da andere STATEMENTS in der Policy die Präferenz verletzen können.

```

<appel:RULESET>
  <appel:RULE behavior="block">
    <POLICY>
      <STATEMENT>
        <PURPOSE appel:connective="or">
          <admin/><develop/><tailoring/>
          <pseudo-decision/>
          <individual-analysis/>
          <individual-decision/>
          <contact/> <historical/>
          <pseudo-analysis/>
          <other-purpose/>
        </PURPOSE> </STATEMENT> </POLICY>
      </appel:RULE>

  <appel:RULE behavior="request"/>
    <appel:OTHERWISE/>
  </appel:RULE>
</appel:RULESET>

```

Abb. 4: Präferenz 2

Die Person versucht dieses Problem zu lösen, indem sie akzeptable Spezifikationen in inakzeptable Spezifikationen transformiert. Sie merkt, es gibt eine feste Anzahl von vorbestimmten Zwecken in P3P. So schreibt sie die *Präferenz 2*, wie im Abb.4 angezeigt. Sie zählt die Zwecke auf, die inakzeptabel sind. Zu beachten ist, dass die Präferenz nun schwerer zu verstehen ist. Bei kleinen Präferenzen ist es vielleicht nicht so wichtig, aber bei praktischen Anwendungen führt dies zu größeren Problemen.

Einfache Präferenzen sind schwer zu definieren: Dieses Problem entsteht durch eine grundlegende Designwahl und kann nicht vollständig gelöst werden ohne eine Erneuerung des Sprachdesigns.

Wie oben genannt wird, weist APPEL einige Probleme auf. Eine Alternative zu APPEL ist XPref. XPref basiert auf einer festen Untermenge von XPath [1].

2.3 Der Abgleich von P3P und APPEL

Einen genauen Überblick über die Funktionsweise von P3P liefert Abb.5.

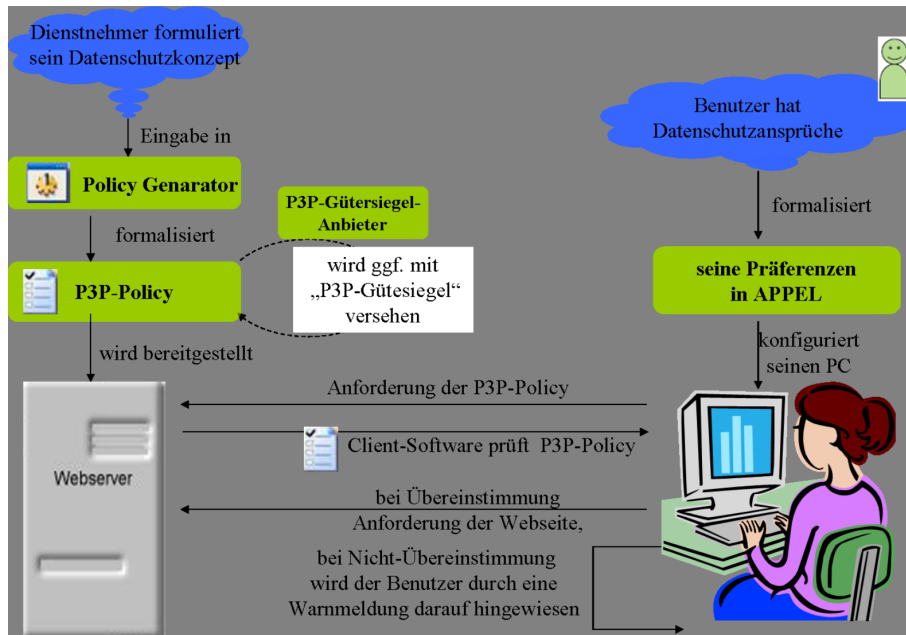


Abb. 5: Die Funktionsweise von P3P und APPEL

Der Dienstgeber hinterlegt eine Beschreibung seiner Datenverarbeitungspraktiken (Datenschutzerklärung) als normale Textversion und im P3P-Format auf dem Webservice. Der Benutzer legt einmal seine Vorstellungen vom Schutz der eigenen Daten in einem P3P-Agent, z.B. einem P3P-fähigen Browser, fest.

Vor Abruf einer Website ruft der P3P-Agent des Nutzers die Datenschutzerklärung auf dem Webserver ab und vergleicht die dort gemachten Angaben mit den Datenschutzeinstellungen des Benutzers. Liegen Abweichungen vor, wird der Benutzer durch ein Warnsignal darauf hingewiesen. Zur weiteren Information liefert P3P dem Nutzer die essentiellen Informationen der Datenschutzerklärung zusammengefasst und in natürlicher Sprache, wobei auf die Unterschiede zu den eigenen Vorstellungen besonders hingewiesen wird.

2.4 Nutzen von P3P und APPEL

Die Nutzung von P3P und APPEL trägt dazu bei, dass Benutzer Herr ihrer personenbezogenen Daten bleiben. Im Einzelnen helfen dem Nutzer folgende Funktionen, die Übersicht zu behalten

- P3P liefert eine übersichtliche Darstellung aller Inhalte einer Webseite mit ihren Datenschutzinformationen.
- Die Zusammenfassung der Datenschutzerklärung wird auch bei fremdsprachigen Websites in der Sprache des P3P-Agent angezeigt. Damit kann der Benutzer auch die Datenschutzerklärungen fremdsprachiger Websites ohne Kenntnisse der Rechtssprache in Ihren wichtigsten Aussagen verstehen.
- Die vollständige Datenschutzerklärung im normalen Textformat kann über einen Link schnell aufgerufen werden; langwieriges Suchen entfällt.
- Der jeweils verantwortliche Ansprechpartner ist leicht auffindbar. So kann der Benutzer seine Datenschutzrechte (z.B. einen Widerruf) unmittelbar geltend machen.
- Die Maschinenlesbarkeit der Datenschutzinformationen ermöglicht die Steuerung von Softwarefunktionen (z.B. die Annahme von Cookies) anhand der Datenverarbeitungspraktiken des Dienstgebers. Einmal eingestellte Datenschutzeinstellungen des Nutzers werden ohne zusätzlichen Aufwand und ohne Zeitverzug beim Surfen berücksichtigt.

2.5 Werkzeuge

Inzwischen wurden viele Werkzeuge in Zusammenhang mit Privacy Policies und Privacy Präferenzen entwickelt. Insbesondere vereinfachen die P3P Editoren und APPEL Editoren die Erstellung der Privacy Policies und Privacy Präferenzen.

P3PBuilder: P3PBuilder ist ein P3P Policy Generator. Er ist Web-basiert und einfach zu verwenden.

P3PEdit: P3PEdit ist ein kommerzieller P3P Policy Generator, der von vielen Firmen benutzt wird.

P3PDisplay, P3PDeveloper.com, Privacybot.com sind andere wichtige Policy Editoren.

P3P Validator: W3C stellt den P3P Validator Service bereit, der überprüft, ob die Webseite mit P3P kompatibel ist.

JRC APPEL Präferenz Editor: ist ein Java-basierter Editor für die Erzeugung von APPEL Präferenzen.

2.6 Probleme beim Abgleich

Realistische Policies neigen dazu, empfindliche private Informationen zu enthalten. Also auch Policies müssen wie Ressourcen geschützt werden. **Automatisierter Vertrauensaustausch** ist eine Annäherung zum Herstellen des Vertrauens zwischen Fremden durch **iterative Freigabe** der digitalen Bescheinigungen [6].

Vertrauen zwischen zwei Fremden basiert auf Eigenschaften der Parteien, die durch Freigabe der digitalen Bescheinigungen nachgewiesen werden. Eine digitale Bescheinigung ist eine nachweisbare, unfälschliche, digital unterzeichnete Erklärung über die Eigenschaften der in der Bescheinigung erwähnten Parteien, die durch einen Beglaubigungsaussteller herausgegeben wird. Jede Partei kann Zugriffskontroll-Policies definieren, um externe Zugriffe auf ihre empfindlichen Ressourcen zu kontrollieren. Vertrauen wird nicht in einem einzigen Schritt hergestellt. Stattdessen werden die digitalen Bescheinigungen Schritt für Schritt freigegeben. Weniger empfindliche Bescheinigungen werden zuerst freigegeben. Später wenn ein bestimmtes Niveau des Vertrauens hergestellt worden ist, können empfindlichere Bescheinigungen freigegeben werden

Z.B. ein Benutzer möchte einige Medikamente von einer Online-Apotheke kaufen. Die Apotheke fordert von dem Benutzer ein Rezept und eine gültige Kreditkarte. Das Rezept und die Kreditkarte enthalten sensitive private Informationen. So fordert der Benutzer von der Apotheke zunächst eine gültige Apothekenlizenz.

3 Einhaltung der Privatheit

3.1 Privatheit versus Sicherheit

Privatheit ist etwas Anderes als Sicherheit. Sicherheit hat die Aufgabe, die Verarbeitung, Speicherung und Kommunikation von Informationen so zu gestalten, dass die Vertraulichkeit, Verfügbarkeit und Integrität der Informationen und Systeme in ausreichendem Maß sichergestellt wird [8]. Es ist leichter Sicherheit herzustellen als Privatheit, da Sicherheit auf vergangenen oder gegenwärtigen Bedingungen basiert, und somit vor der Herausgabe der Daten geprüft werden können. Man spricht hier von Zugriffskontrolle. In diesen Situationen sind die Bedingungen beobachtbar.

Aber Privatheit basiert auf Bedingungen, die sich auch auf die Zukunft (d.h. nach der Herausgabe der Daten) beziehen, was wir als Verpflichtungen bezeichnen [4]. Genau hier treten die Probleme mit der Privatheit auf. Wie kann man sicher sein, dass seine Informationen z.B. nach 3 Monaten wirklich gelöscht oder nicht weitergegeben werden. Ein weiteres Problem ist, dass man nicht immer beobachten kann, ob die Bedingungen erfüllt werden oder nicht.

Möchte man beispielsweise eine Digitalkamera bei *amazon.de* erwerben, fordert *amazon.de* persönliche Informationen wie Adresse, Geburtsdatum und Kreditkartennummer. Nach der Privacy Policy von *amazon.de* werden diese persönlichen Informationen nach 3 Monaten gelöscht. Ob dies tatsächlich eingehalten wurde, kann nicht nachgeprüft werden.

3.2 Hippokratische Datenbanken

Hippokratische Datenbanken sind von der Idee des hippokratischen Eids inspiriert. Datenbanken, welche die Privatheit als einen zentralen Aspekt sehen, bezeichnet man als hippokratische Datenbanken. Hippokratische Datenbanken bieten Lösungsvorschläge zu den oben genannten Problemen.

Die folgenden 2 Eigenschaften sind grundlegend für Datenbanksysteme.

1. die Fähigkeit, persistente Daten zu managen.
2. die Fähigkeit, eines effizienten Zugriffs auf eine große Menge von Daten.

3.2.1 Die zehn Prinzipien

Nun artikulieren wir die grundlegenden Prinzipien von hippokratischen Datenbanksystemen. Diese Prinzipien definieren die Erwartung eines Benutzers von einer Datenbank, die sich als hippokratisch bezeichnet.

1. **Spezifikation der beabsichtigten Nutzung:** Der Zweck, für den Informationen gesammelt wurden, soll mit der Information verbunden werden.
2. **Zustimmung:** Der Zweck, der mit persönlichen Informationen verbunden ist, soll die Zustimmung des Informationsgebers haben.
3. **Beschränkte Kollektion:** Die Informationssammlung wird auf das Minimum begrenzt, um den spezifizierten Zweck zu erreichen.
4. **Beschränkter Gebrauch:** Die Datenbank wird nur jene Abfragen laufen lassen, die mit dem Zweck übereinstimmen, für den die Informationen gesammelt worden sind.
5. **Beschränkte Freigabe:** Persönliche Informationen werden nicht weitergegeben ohne Erlaubnis des Benutzers.
6. **Beschränkte Beibehaltung:** Persönliche Informationen werden nur solange behalten, bis die vordefinierte Dauer erreicht ist.
7. **Aktualität:** Persönliche Informationen, die in der Datenbank gespeichert sind, sollen genau sein und aktuell bleiben.
8. **Sicherheit:** Persönliche Informationen sollen durch Schutzmechanismen gegen Diebstahl und anderen Missbrauch geschützt werden. (Zur Durchsetzung dieses Prinzips werden die Sicherheitsmechanismen des Datenbanksystems herangezogen)
9. **Offenheit:** Ein Benutzer soll Zugriffsrechte auf alle Informationen haben, die über ihm gesammelt wurden.
10. **Übereinstimmung:** Ein Benutzer soll in der Lage sein, Übereinstimmung mit den oben genannten Prinzipien verifizieren zu können.

3.2.2 Die Durchsetzung der Privatheit

Abbildung 6 zeigt die Grobarchitektur von Hippokratischen Datenbanken. Nun erklären wir, wie die Hippokratischen Datenbanken Privatheit durchsetzen.

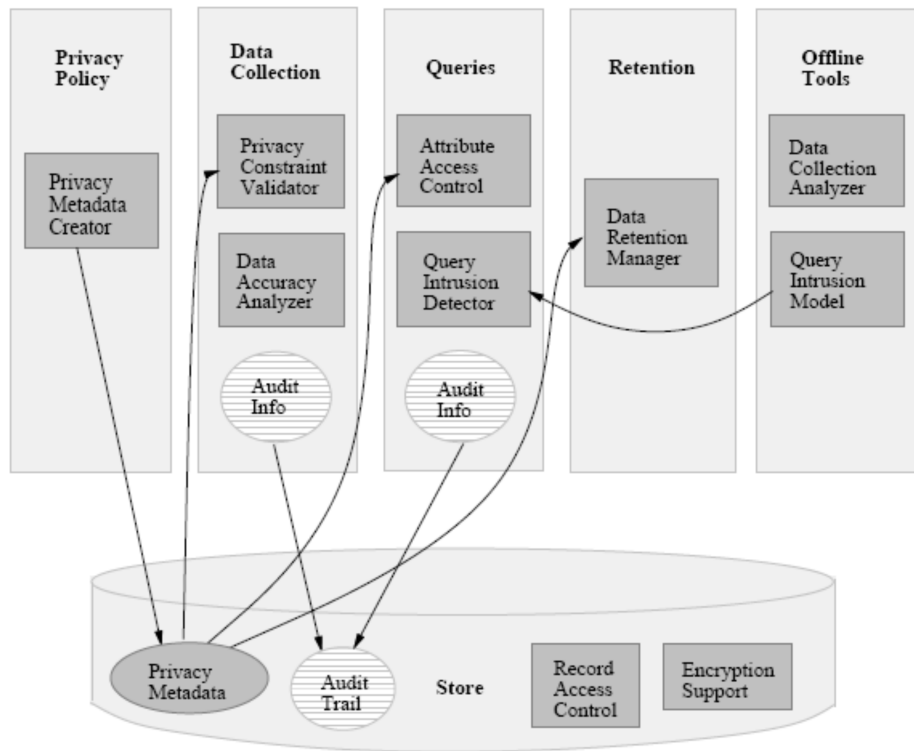


Abb. 6: Grobarchitektur [2]

1. Abfragen werden von den Datenbanken mit ihrem Zweck (purpose) verbunden. Also werden Policies mit Zugriffsrechten verknüpft. Eine Abfrage wird nur erlaubt, wenn die Gruppe der befugten Benutzer (authorized users) für diesen Zweck (purpose) in einer *Privacy Authorization Tabelle* den Benutzer einschließt, der die Abfrage ausführt. In der *Privacy Authorization Tabelle* wird definiert welcher Benutzer, mit welchem Zweck, auf welche Informationen zugreifen kann. Für jede Abfrage stellt die Komponente **Record Access Control** sicher, dass nur Einträge deren purpose-attribut (Zweckattribut) den Zweck der Abfrage einschließen, zur Abfrage sichtbar sind.

Somit wird garantiert, dass

- nur auf Daten von befugten Benutzern zugegriffen wird
- nur für vordefinierte Zwecke zugegriffen wird

Damit wird die Einhaltung schon zu einem Großteil gewährleistet.

2. Was passiert jedoch, wenn jemand Zugriffsrechte hat und versucht die Daten zu missbrauchen? Z.B. ein Kundendienstmitarbeiter einer Online-Bücherei möchte die Emailadressen aller eingetragenen Benutzer für seine eigenen Zwecke missbrauchen. Die oben beschriebene Zugriffskontrolle stoppt die Abfrage nicht, da dieser Mitarbeiter Zugriffsrechte auf die Emailadressen hat und regelmäßig auf den Emailadressen zugreift, um auf Fragen über den Auftragstatus zu antworten.

Zur Lösung dieses Problems ein **Query Intrusion Detector** prüft die Frageresultate, um die Abfragen zu entdecken, deren Zugangsmuster sich von den üblichen Zugangsmustern dieses Benutzers mit diesem Zweck(purpose) unterscheiden. Der Detektor macht dies, indem er frühere Abfragen für jeden Zweck und für jeden befugten Benutzer analysiert. In unserem Beispiel, könnte das Profil für die Abfragen, die vom Kundendienst ausgeführt werden, das folgende sein:

- die Abfragen greifen nur auf Kunden zu, deren Bestellstatus nicht-erfüllt ist oder
- die Abfragen greifen täglich auf weniger als 100 Einträge zu.

In anderen Situationen werden die Abfragen als verdächtig gekennzeichnet und je nach Einstellungen der Datenbank nicht beantwortet.

3. Ein weiterer Mechanismus ist die Erstellung der Prüflisten aller Abfragen (Logging). Im vorherigen Beispiel konnte *Query Intrusion Detector* den Missbrauch nicht verhindern. Wenn aber ein Kunde bemerkt, dass seine Emailadresse nach dem Verkauf weiterhin benutzt wird, tritt er mit der Online-Bücherei in Verbindung. Der Kunde hat damit die Möglichkeit, zu beobachten, ob seine Daten nur von berechtigten Personen und für vordefinierte Zwecke benutzt werden. Der Administrator der Datenbank würde durch Kontrolle der Prüflisten erkennen, ob ein Verletzung der Privatheit vorliegt.
4. Der **Data Collection Analyzer** prüft, ob die gesammelten Daten auch benötigt würden (Prinzip: Minimalität).
5. **Data Retention Manager** löscht die Daten automatisch, die die Bewahrungszeit erfüllt haben.

4 Zusammenfassung

Nun werden wir die wichtigsten Aspekte der Privatheit kurz zusammenfassen.

Privatheit ist das Recht jedes Einzelnen, jeder Gruppe und jeder Institution für sich selbst zu entscheiden, wann, wie und zu welchem Zweck Informationen über sie an andere weitergegeben werden. Privatheit kann auf verschiedenen Wegen sichergestellt werden. Jedoch ist unser Schwerpunkt die Kontrolle der Verwendung bei Herausgabe privater Daten. Für diese Kontrolle ist eine Spezifikationsprache notwendig. P3P ist eine Policy-Beschreibungssprache, die vom Dienstgeber und APPEL eine Präferenz-Beschreibungssprache, die vom Dienstnehmer benutzt wird. Während des Abgleichs treten einige Probleme auf, da Policies persönliche Informationen enthalten, die auch wie Ressourcen geschützt werden müssen. Z.B durch iterative Freigabe der digitalen Bescheinigungen.

Privatheit sollte man nicht mit Sicherheit verwechseln. Sicherheit kann mit Zugriffskontrollen hergestellt werden. Bei der Privatheit stellt der Zukunftsbezug (Nicht-Beobachtbarkeit) das Hauptproblem dar. Für diese Probleme bieten Hippokratische Datenbanken im Datenbankbereich Lösungsvorschläge.

Literatur

- [1] R.Agrawal, J.Kiernan, R.Srikant, Y.Xu: An XPath-based Preference Language for P3P. *Proceedings of the 12th international conference on World Wide Web 2003, Budapest, Hungary.*
- [2] R.Agrawal, J.Kiernan, R.Srikant, Y.Xu: Hippocratic Databases. *Proceedings of the 28th VLDB Conference, Hong Kong, China, 2002.*
- [3] R.Agrawal, J.Kiernan, R.Srikant, Y.Xu: Implementing P3P Using Database Technology. *Proceedings of the 19th International Conference on Data Engineering, Bangalore, India, March 2003.*
- [4] M.Hilty, D.Basin, A.Pretschner: On Obligations. *Information Security, ETH Zurich, Switzerland. ESORICS 2005, LNCS 3679, pp. 98-117, 2005.*
- [5] Dr. Robert P.Minch: Privacy Issues in Location-Aware Mobile Devices. *Proceedings of the 37th Hawaii International Conference on System Sciences, 2004.*
- [6] T.Yu, M.Winslett: A Unified Scheme for Resource Protection in Automated Trust Negotiation. *Proceedings of the 2003 IEEE Symposium on Security and Privacy (SP.03), pp. 110-112.*
- [7] Das Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein. <http://www.datenschutzzentrum.de/faq/p3p.htm#wasistp3p>
- [8] Die freie Enzyklopedie Wikipedia. <http://de.wikipedia.org>