

Private Daten: Herausforderungen für das Datenmanagement

Heinrich Jasper

Institut für Informatik
Technische Universität Bergakademie Freiberg
Bernhard-von-Cotta-Str. 2
D-09599 Freiberg
jasper@tu-freiberg.de

Abstract: Privatpersonen sind immer stärker in vernetzte Prozesse eingebunden, die heute schon teilweise und zukünftig ganz sicher mit digitalen Informations- und Kommunikationstechnologien unterstützt und realisiert werden. Während der Ausführung dieser Prozesse fallen Daten an, die für die Personen von langfristigen Wert sind, z. B. Zeugnisse oder Vertragsdaten. Für das Management privater Daten existiert eine Vielzahl von Herausforderungen, insbesondere auf den Gebieten der langfristigen und sicheren Speicherung dieser Daten und dem vertrauenswürdigen Austausch der Daten mit anderen Personen oder Behörden. Im Vortrag werden einige technische orientierte Problemstellungen präsentiert und erste Ideen für die Herangehensweise bei deren Lösung vorgestellt.

1 Private Daten

Die hinlänglich bekannten Szenarien einer vernetzten Welt sehen Menschen beim on-line Banking, in virtuellen Lernumgebungen, bei der on-line Immatrikulation an Hochschulen, bei der Heimarbeit, bei on-line Wahlen und vieles mehr. In diesen und weiteren Prozessen fallen Daten an, die mit der jeweils handelnden Person verbunden und auf einem oder mehreren Computersystemen gespeichert sind. Im Folgenden bezeichne ich mit „privaten Daten“ die in beliebiger digitaler Form gespeicherten Informationsobjekte einer natürlichen Person, über die diese verfügen darf. Oft hat diese Person das ausschließliche Recht an den privaten Daten.

Die Art der zu verwaltenden Daten ergibt sich aus den zu berücksichtigenden Prozessen, wie sie oben beispielhaft aufgelistet wurden. Zusätzlich dringt digitale Technologie immer weiter in private Bereiche vor, so dass dort ebenfalls private Daten entstehen und verwaltet werden müssen, etwa in der digitalen Fotografie. Es lassen sich folgende Arten privater Daten ableiten:

- Selbst erstellte Dokumente z.B. während einer Ausbildung etc. Diese sind heute meist in digitaler Form vorhanden.
- Selbst aufgenommene digital gespeicherte Fotos oder Videos.
- Zeugnisse und Urkunden (ausbildungsbezogene, digitale, signierte Daten etc.).
- Verträge und Vertragsdaten (Versicherungsunterlagen, Unterlagen zur Kontoführung etc.).

Darüber hinaus gehören zu den privaten Daten von Dritten bereitgestellte Daten, an die die Privatperson damit Rechte erhält, z. B. Familienfotos, Musikdateien etc.

2 Anforderungen an die Verwaltung privater Daten

Anforderungen an die private Datenhaltung ergeben sich aus den von der Personen gewünschten Eigenschaften und aus den Eigenschaften, die aufgrund der Einbindung der Daten in soziotechnische Prozesse notwendigerweise resultieren. So kann eine Person Interesse daran haben, digitale Fotos ihr Leben lang aufzubewahren. Ein Beispiel für einen soziotechnischen Prozess ist z. B. die Immatrikulation an einer Hochschule zu der das digital vorliegende Zeugnis der entsprechenden Hochschulreife in einer vertrauenswürdigen Form der Hochschule vorgelegt werden muss. Allgemein lassen sich die zwei wesentliche Anforderungen an private Daten herausarbeiten:

- **Langfristigkeit:** Digital gespeicherte Informationen sollen das Leben der zugehörigen Privatperson begleiten oder sogar überdauern und für die notwendigen bzw. gewünschten Operationen jederzeit zur Verfügung stehen. Dieses umfasst die jederzeitige Erschließung des Inhalts mittels der natürlichen Sensoren – also audiovisueller, gegebenenfalls taktile Zugriff.
- **Sicherheit und Schutz einschließlich Authentifikation:** Nur berechtigte Personen und Institutionen dürfen auf den privaten Daten festgelegte Operationen ausführen. Neben den juristischen Aspekten spielt hier das Vertrauen der Privatperson in die sichere Ausgestaltung der Verwaltungs- und Übermittlungsprozesse der privaten Daten eine wichtige Rolle.

Lösungsszenarien zu den Themen Langfristigkeit und Sicherheit müssen zwei wesentliche Aspekte berücksichtigen: Erstens ist aufgrund des mehrere Jahrzehnte umfassenden Nutzungszeitraums der Daten von konkreten Technologien zu abstrahieren und zweitens ist der Nutzer gemäß seiner menschlichen Fähigkeiten adäquat zu würdigen. Ersteres unterstützen z. B. virtuelle Informations- und Kommunikationssysteme, siehe z. B. das Konzept des Universal Virtual Computer (UVC) in [Lo01].

Die Würdigung der menschlichen Fähigkeiten beruht prinzipiell darauf, Privatpersonen von dem technischen Ballast, der mit der Nutzung der Informations- und Kommunikationstechnologie einhergeht, zu befreien. Das betrifft u. a. die Migration von Daten und Metadaten bei neuen Infrastrukturen („Ich kaufe einen neuen PC“) und die Selbstadministration und Selbsterklärbarkeit von gesicherten Prozessen z. B. mittels digitaler Signaturen.

3 Ontologien und Agenten

Zwei Ansätze aus Gebieten der Künstlichen Intelligenz eignen sich meines Erachtens besonders zur Unterstützung der Privatpersonen und sollen hier kurz vorgestellt werden: Ontologien und Agenten.

3.1 Ontologien

Ontologien stellen die Begriffe einer Domäne und ihre logischen Zusammenhänge formal dar [So00]. Für private Nutzer ist eine Ontologie wichtig, die die Infrastruktur zur Verwaltung privater Daten einschließlich ihrer Nutzungsmöglichkeiten erklärt. Dazu gehören technische Begriffe wie Daten, Datei, Datenbank, Suchverfahren, Änderungsoperatoren, Versenden von Daten, zertifizierte Dokumente etc. Weiterhin sind die mit der Datenbeschreibung verbundenen Informationen in der Ontologie zu berücksichtigen, also Begriffe von Datenmodellen und Metadatenmodellen, wobei letztere insbesondere dem Austausch von Daten und Metadaten zwischen Systemen dienen [St02] und damit die im nächsten Abschnitt vorgestellten Agenten unterstützen.

Bei der Verwaltung privater Daten ist die Ontologie der technischen Infrastruktur durch die private Ontologie einer Person zu ergänzen. Dieses ermöglicht es, die im privaten Bereich festgelegten Bedeutungen von Begriffen („Bild von Opas 70sten Geburtstag“: in der engeren Familie ist dadurch „Opa“ eindeutig; weitere Beispiele: „unser Haus“, „Ferien auf Norderney“, ...) über Systemgrenzen hinweg zu transferieren. Damit werden zwei Ziele erreicht: Der Inhalt privater Daten erschließt sich auch langfristig dem privaten Nutzer. Gleichzeitig kann der Inhalt von den Beteiligten an den oben aufgeführten soziotechnischen Prozessen – sowohl weiteren Privatpersonen als auch Organisationen – verstanden und zur ordnungsgemäßen Abwicklung der Prozesse genutzt werden.

Die in [St02] vorgestellten Konzepte zum Wissensmanagement lassen sich somit auf das private Wissensmanagement ausweiten. Noch zu untersuchen ist die Transparenz – hier verstanden als Unsichtbarkeit – des Wissens über die technischen Infrastrukturen für den privaten Nutzer: Wie weit können wir die Technik voranbringen, so dass sie sowohl beherrschbar aber auch einfach für den „causal user“ ist? In der „Künstlichen Intelligenz“ sind u.a. hierzu Agenten entwickelt worden.

3.2 Agenten

Ein Agent ist ein Softwaresystem, das als kontinuierlicher Prozess in einer Umgebung läuft, diese Umgebung beobachtet und nach vorgegebenen Regeln auf Ereignisse in der Umgebung selbständig reagiert [FG97]. Für das Szenario der privaten Daten sind Agenten von Interesse, die die oben aufgeführten Herausforderungen Langfristigkeit und Sicherheit unterstützen. Sie treten damit als Helfer für die Privatperson – möglichst von diesen nicht bemerkt – auf und übernehmen entsprechende Aufgaben.

Im Kontext der privaten Daten sind Agenten zu entwickeln, die die automatische Überführung aller Daten von einem System auf ein Folgesystem unterstützen. Dabei ist natürlich Unterstützung von Seiten der Hersteller dieser Systeme in vordefinierter Form notwendig. Für das genutzte System und das Folgesystem – beide verstanden als Hard-/Software-Einheiten – ist Wissen über die Datenhaltung in den Systemen sowie über die Migration der Daten zwischen den Systemen für die Agenten bereitzustellen.

Weiterhin sind Agenten notwendig, die die sichere Abwicklung der Kommunikationsprozesse unterstützen und von allen Beteiligten als solche akzeptiert sind. Beispielsweise ist bei der Immatrikulation an Hochschulen das digitale Zeugnis von der ausstellenden Schule, der beteiligten Privatperson und den annehmenden Hochschule anzuerkennen. Diesen Prozess soll nicht der private Nutzer managen, sondern von den Agenten übernommen werden.

Zu den Themen Ontologien und Agenten für die langfristige, sichere Haltung privater Daten in einer vernetzten Welt laufen erste Arbeiten an der Technischen Universität Bergakademie Freiberg. Im Rahmen dieser Arbeiten werden auch andere Projekte mit ähnlichen Zielen bei nicht-privaten, organisationsunterstützten Informationssystemen berücksichtigt, z. B. das IBM-Projekt „Autonomic Computing“ [IBM01].

4 Ausblick

Generell ist meines Erachtens die Verbreitung von digitalen Verfahren in den vorgestellten Szenarien von dem Vertrauen abhängig, dass Privatpersonen den Verfahren zu Speicherung und Übermittlung eigener digitaler und digitalisierter Assets entgegenbringen. In den obigen Ausführungen wurde deutlich, dass digitale und digitalisierte private Daten langfristig ausschließlich in einem Kontext existieren werden, der mehrerer Partner einbezieht. Damit handelt sich bei dem Gesamtszenario um Sozialsystem und für dessen Akzeptanz ist Vertrauen wichtig: „... Sozialsysteme, die durch ein hohes Vertrauenskapital gekennzeichnet sind, gründen sich auf langfristig angelegten sozialen Bindungen und wechselseitigen Verpflichtungen zwischen Individuen und kollektiven Akteuren. Loyalität und Identität bestimmen das Verhältnis zwischen individuellem Akteur und kollektiver Organisationseinheit. ...“ [Ma02]. Wie sollen in der aktuellen technischen und gesellschaftlichen Entwicklung dieser Techniken die Vertrauensfrage definiert werden: Vertrauen wir Microsoft, AOL und anderen? Wer setzt die Maßstäbe, definiert die Standards? Welchen Einfluss haben die Vertreter der Privatpersonen, z. B. gewählte Politiker, Landesvertreter oder auch Organisationen wie der Verbraucherschutz? Was können wir tun, dass es nicht auf Dauer das persönliche „Hobby“ jedes Einzelnen bleibt, die eigenen privaten Assets in dieser vernetzten Welt zu sichern?

Literaturverzeichnis

- [FG97] Franklin, S., Graesser, A.: Is it an agent, or just a program? A taxonomy for autonomous agents. In Jennings, M., Wooldridge M.: Intelligent Agents, Springer, 1997.
- [IBM01] IBM: Autonomic Computing: IBM's Perspective on the State of Information Technology. <http://www.research.ibm.com/autonomic/manifesto>
- [Lo01] Lorie, R. A.: Long term preservation of digital information. ACM/IEEE Joint Conference on Digital Libraries, JCDL 2001, Roanoke, Virginia, USA, June 24-28, 2001.
- [Ma02] Mayer, K. U.: Mißtrauen im Reformprozeß. Forschung und Lehre, 6/2002.
- [So00] Sowa, J. F.: Knowledge Representation: Logical, Philosophical, and Computational Foundations. Brooks Cole Publishing Co., Pacific Grove, CA, 2000.
- [St02] Staab, S.: Wissensmanagement mit Ontologien und Metadaten. Informatik Spektrum, Band 25, Heft 3, Juni 2002.