

METIS in PArADISE

Provenance Management bei der Auswertung von Sensordatenmengen für die Entwicklung von Assistenzsystemen

Position Statement

Andreas Heuer, Universität Rostock, Institut für Informatik

andreas.heuer@uni-rostock.de

Langfassung eines fünfseitigen Workshop-Beitrags
veröffentlicht in Lecture Notes in Informatics
Band 242, BTW 2015 Workshop-Band, S. 131 ff.

Abstract: In diesem Beitrag soll ein langfristiges Forschungsvorhaben im Bereich der Informatik und Elektrotechnik an der Universität Rostock vorgestellt werden, in dem wissenschaftliche Experimente in der Informatik, der Zellbiologie und der Medizin (neurodegenerative Erkrankungen) durch effiziente Analyseverfahren auf sehr großen Mengen von Mess- oder Sensordaten unterstützt und im Sinne des Provenance Management rückverfolgbar gemacht werden sollen. Im Bereich der Informatik ist das experimentelle Anwendungsgebiet das der Erforschung und systematischen Entwicklung von Assistenzsystemen. Da in Assistenzsystemen unterstützte Personen durch eine Vielzahl von Sensoren beobachtet werden, müssen auch Privatheitsaspekte bereits während der Phase der Modellbildung berücksichtigt werden, um diese bei der konkreten Konstruktion des Assistenzsystems automatisch in den Systementwurf zu integrieren.

Die Datenbankteilaspekte dieses Forschungsgebietes werden im Beitrag näher beleuchtet: Neben der effizienten Auswertung großer Mengen von Mess- und Sensordaten sind dies das Provenance Management und die Integration von Privatheitsbedingungen. Um diese Problemkreise zu verknüpfen, treffen zwei extrem unterschiedliche Datenbankthemen aufeinander: (1) Ableitung inverser Schema- und Instanzabbildungen, die üblicherweise in der Datenbankintegration, -föderation und -evolution benötigt werden, aus dem Projekt METIS, sowie (2) Effizienz von Analyseverfahren und Integration von Privatheitsaspekten durch Anfragetransformationen für die Entwicklung von Assistenzsystemen im Projekt PArADISE. Im Beitrag werden wir den gemeinsamen Kern beider Themen in den theoretischen Grundlagen von Datenbanken identifizieren.

1 Einleitung

Wissenschaftliche Experimente, die durch Messungen oder ständige Beobachtungen laufend eine Vielzahl von Daten erheben, müssen durch effiziente Analyseverfahren auf die-

sen Mess- oder Sensordaten unterstützt werden. Die Mess- und Sensordaten sind so zu verwalten, dass sie im Sinne des Provenance Management rückverfolgbar werden. An der Universität Rostock soll ein langfristiges Forschungsvorhaben im Bereich der Informatik und Elektrotechnik etabliert werden, in dem wissenschaftliche Experimente in der Informatik, der Zellbiologie und der Medizin (neurodegenerative Erkrankungen) auf eine solche Weise unterstützt werden.

Im Bereich der Informatik ist das experimentelle Anwendungsgebiet das der Erforschung und systematischen Entwicklung von Assistenzsystemen. Da in Assistenzsystemen unterstützte Personen durch eine Vielzahl von Sensoren beobachtet werden, müssen auch Privatheitsaspekte bereits während der Phase der Modellbildung berücksichtigt werden, um diese bei der konkreten Konstruktion des Assistenzsystems automatisch in den Systementwurf zu integrieren. Somit gibt es für die Datenbankforscher unter anderem folgende Teilprobleme, die zu lösen sind und die in zwei langfristigen Projektgruppen des Datenbanklehrstuhls zusammengefasst sind:

PARADISE (Privacy AwaRe Assistive Distributed Information System Environment): In dieser Projektgruppe werden Techniken zur Auswertung von großen Mengen von Sensordaten entwickelt, die definierte Privatheitsansprüche der späteren Nutzer per Systemkonstruktion erfüllen.

METIS (Management, Evolution, Transformation und Integration von Schemata): In dieser Projektgruppe geht es unter anderem um Integrationsverfahren für heterogene Datenbanken, die per se bereits eine Rolle bei der Zusammenführung der unterschiedlichen Datenquellen des Assistenzsystems spielen. Als Seiteneffekt werden die für die Datenintegration entwickelten Techniken der *Global-as-local-view-extension* (GaLVE) auch für die nötigen inversen Abbildungen zum Provenance Management bei der Verarbeitung der Sensordaten und Ableitung von (Situations-, Aktivitäts-, Intentions-)Modellen benötigt.

Wissenschaftliche Experimente außerhalb der Informatik, etwa zur Auswertung von Sensordaten bei Patienten mit neurodegenerativen Erkrankungen [DEW⁺12] liegen nicht im Fokus dieses Beitrags.

Im Folgenden werden wir im Abschnitt 2 kurz die Architektur von Assistenzsystemen einführen, wobei wir den Schwerpunkt auf die Phase der Situations-, Aktivitäts- und Intentionserkennung legen. Danach werden wir die Erforschung und Entwicklung von Assistenzsystemen als ein wissenschaftliches Experiment ansehen, in dem Forscher eine große Anzahl von Sensordaten auswerten müssen und aus ihnen Situations-, Aktivitäts- und Intentionsmodelle entwickeln (Abschnitt 3). Zwei der Grundlagenforschungsthemen für die Datenbankforscher, Provenance Management und die Integration von Privatheitsanforderungen, werden wir in Abschnitt 4 einführen. Schließlich werden wir den State of the Art und die eigenen Vorarbeiten betrachten, wobei zu letzteren überraschenderweise auch inverse Schemaabbildungen für Datenbankintegrationsverfahren als Basis für das Provenance Management zählen. Das Positionspapier endet mit einem kurzen Fazit in Abschnitt 5.



Abbildung 1: Pyramidenarchitektur von Assistenzsystemen

2 Assistenzsysteme

Ähnlich einem menschlichen Assistenten soll ein Assistenzsystem mich unterstützen, im Hintergrund arbeiten (ambient), mich nicht stören, zum richtigen Zeitpunkt eingreifen und Hilfe anbieten (diese in üblichen Fällen auf optischem oder akustischem Wege), vertrauenswürdig und diskret sein und sich bei Bedarf abschalten lassen.

Um seine Assistenzaufgaben zu erfüllen, besteht ein Assistenzsystem üblicherweise aus fünf Schichten (siehe Abbildung 1 nach [HKHT06], auch in [HKG14]). Dabei deutet die Pyramidenform an, dass in der untersten Schicht ständig viele Daten (etwa von Sensoren) erzeugt werden, in der obersten Schicht aber nur im Bedarfsfall (also eher selten) ein akustischer oder optischer Hinweis, also eine geringe Datenmenge, ausgegeben wird.

Sensoren in der Umgebung der Person sollen Situation und Tätigkeit der Person erfassen, um ihr assistieren zu können. **Ortungskomponenten** sollen die genaue Position der Person bestimmen, etwa zur Detektion dementer Patienten mit Weglauftendenzen. Sensoren und Ortungskomponenten befinden sich in der Umgebung, in benutzten Geräten oder am Körper der Person (Armband, Brille, ...).

Damit ein Assistenzsystem seine Aufgabe erfüllen kann, müssen verschiedene (heterogene) Geräte in der Umgebung der Person **vernetzt** und zur Erreichung des Assistenzziels **spontan gekoppelt** werden.

Sensordaten müssen gefiltert, erfasst, ausgewertet, verdichtet und teilweise langfristig verwaltet werden. Aufgrund der extrem großen Datenmenge (Big Data) muss die **Verarbeitung verteilt** erfolgen: teilweise eine Filterung und Verdichtung schon im Sensor, im nächsterreichbaren Prozessor (etwa im Fernseher oder im Smart Meter in der Woh-

nung) und im Notfall über das Internet in der Cloud. Neben Daten des Assistenzsystems müssen auch fremde Daten etwa über das Internet berücksichtigt werden, beispielsweise Wartungspläne beim Auto oder die elektronische Patientenakte beim Patienten. Allgemein können hier natürlich auch die Daten sozialer Netzwerke, Kalenderdaten der Nutzer oder Wettervorhersage-Daten ausgewertet werden, falls sie für das Assistenzziel eine Rolle spielen.

Nach Auswertung der Sensordaten erfolgt die **Situations-, Handlungs- und Intentionserkennung** (siehe den folgenden Abschnitt 3 für die Entwicklung diesbezüglicher Modelle).

Schließlich erfolgt die **multimodale Interaktion** mit dem Nutzer. Üblich sind optische Signale wie eine Meldung auf dem Fernseher oder über eine Warnlampe, sowie akustische Signale wie eine Ansage über ein Radio oder Alarmtöne über einen Lautsprecher. Wenn der assistierten Person ein Signal nichts mehr nützt, kann auch eine Meldung an Angehörige, Ärzte oder Notfallzentralen ausgelöst werden. Umgekehrt kann man dem Assistenzsystem auch selbst Hinweise geben, etwa über Touchscreen, Gesten oder Bewegungen oder mittels Sprache (Kommandos).

3 Big Data Analytics für die Entwicklung von Assistenzsystemen

Aufgrund der Sensor- und Ortungsdaten sowie der weiteren über das Internet erhältlichen Daten muss das Assistenzsystem eine Situations- und Handlungserkennung vornehmen sowie eine Handlungsvorhersage (Intentionserkennung), um proaktiv eingreifen zu können.

Die Situation ist dabei die aktuelle Umgebungsinformation, die Handlung das, was die Person, der assistiert wird, gerade durchführt. Die Intentionserkennung oder Handlungsvorhersage muss voraussagen, was die die Person in Kürze tun wird.

Auch die Handlungs- und Intentionserkennung ist ein aktueller Forschungsgegenstand der Informatiker an der Universität Rostock [KNY⁺ 14]. Dabei erheben die Forscher in langen Versuchsreihen eine extrem hohe Anzahl von Sensordaten, aus denen sie mit diversen Analyseverfahren die entsprechenden Modelle ableiten. Diese Analyseverfahren sind — wenn sie ohne Datenbankunterstützung auf Dateisystemen mit Analysewerkzeugen wie R ausgeführt werden — mehrwöchige Prozesse.

Der Umfang der Daten soll hier kurz an einem Szenario abgeschätzt werden: Die Erfassung eines kurzen Handlungsablauf von 40 Minuten einer Versuchsperson durch einen Motion-Capturing-Anzug, einer EMG-Instrumentierung (Elektromyografie) mit paralleler Video-Aufzeichnung erzeugt einen Rohdatenbestand von 10 GByte. Werden für die Entwicklung eines Handlungsmodells, das den Alltag repräsentiert, Daten von vierzehn Versuchspersonen über jeweils etwa 100 Stunden aufgezeichnet, ergibt dieser Versuch Daten im Umfang von 14 Terabyte, die für Analysen verfügbar sein müssen, um entsprechende Modelle zu rechnen.

Ziel der Forscher ist neben der Modellbildungen für Handlung und Intention die Erkenntnis, wie die große Anzahl von Sensoren im Versuch für den praktischen, späteren Einsatz des Assistenzsystems drastisch eingeschränkt werden kann, ohne die Vorhersagequalität



Abbildung 2: Verdichtungsmethoden für Sensordaten

zu mindern. Für die Ableitung dieser Informationen müssen unter anderem alle Analysefunktionen invertiert werden, um die für die Modellbildung entscheidenden Anteile der Originaldaten zu finden. Letzteres ist auch ein Problem im Provenance Management, das die Experimentverläufe mit den Ergebnisableitungen begleiten soll. Eine Einschränkung sowohl der Anzahl der Sensoren als auch der Menge und Granularität der erfassten Daten ist auch aus einem anderen Grund wichtig: sie kann die Privatheitsanforderungen der Nutzer des Assistenzsystems realisieren helfen.

4 Nötige Grundlagenforschung: Privatheit und Provenance

4.1 Privatheit

Die Entwickler der Analysewerkzeuge müssen ihr Assistenzziel und die notwendigen Sensordaten zur Erreichung des Ziels und zur grundlegenden Situations-, Handlungs- und Intentionserkennung formulieren. Diese Zielformulierung wird dann in Anfragen auf Datenbanken transformiert. Weiterhin können die **Privatheitsansprüche** des Nutzers vordefiniert oder von jedem Nutzer selbst individuell verschärft werden. Auch diese Privatheitsansprüche werden in Anfragen (Sichten) auf Datenbanken umgesetzt. Durch Abgleich des Informationsbedarfs des Assistenzsystems und der Privatheitsansprüche des Nutzers kann dann die Datenbankkomponente des Assistenzsystems entscheiden, wie die Menge an Sensordaten selektiert, reduziert, komprimiert oder aggregiert werden muss, um beiden Parteien im System gerecht zu werden [Gru14]. Abbildung 2 zeigt die möglichen Datenquellen, die Verdichtungsmethoden aufgrund vom Matching von Privatheitsanforderungen und Assistenzzielen sowie die mögliche Datenverarbeitung in verschiedenen Zielsystemen (nach Grunert [Gru14]).

Ein entscheidendes Kriterium für die Vertrauenswürdigkeit eines Assistenzsystems ist noch die Frage, wie nah am Sensor die Daten bereits reduziert und verdichtet werden können: Wenn der Sensor so intelligent ist, dass er bestimmte Filtermechanismen von Datenbanksystemen beherrscht, so kann dieser bereits eine Vorfilterung vornehmen. Nur die für das Assistenzziel unabdingbaren Daten, die die Privatheit des Nutzers nicht verletzen,

können dann im Rahmen des Cloud Data Management des Anbieters der Assistenzfunktionalität entfernt und verteilt gespeichert werden.

4.2 Provenance

Herausforderungen und existierende Ansätze der **Data Provenance** diskutiert [CAB⁺14], während die Arbeiten [BSC⁺10, ZSGP11] erforderliche Provenance-Maßnahmen bei der Anfrageverarbeitung diskutieren. Die meisten Ansätze im Bereich des **Provenance Management** legen den Fokus auf die Reproduzierbarkeit der Ergebnisse aus den Rohdaten, d.h. sie liefern eine Vorwärtsverfolgung. Weitergehende Ansätze berücksichtigen zusätzlich die Nichtfälschbarkeit sowie den Urhebernachweis, führen aber zu intransparenten Prozessen, über die „unterwegs“ keine aggregierten und anonymisierten Informationen abfragbar sind.

Ein Teil der Datenbankforschung im Gebiet Provenance Management konzentriert sich auf einzelne Modelle, Operatoren und Annotationen für die Rekonstruktion der Rohdaten [CAB⁺14, Fre09]. Eine durchgängige und automatisierte Vorgehensweise fehlt jedoch. Daneben wurden Provenance-Eigenschaften charakterisiert und formalisiert, aber leider nicht für die automatische Rekonstruktion der Originaldaten rückwärts [BKT01], [GKT07] [CWW00, PIW11].

4.3 Grundlagen: Inverse Schemaabbildungen

Für Letzteres werden Charakterisierungen inverser Analyseoperatoren und inverser Schemaabbildungen benötigt. Inverse Schemaabbildungen und ihre Eigenschaften werden zwar untersucht [Fag07], derzeit aber noch nicht im Bereich Provenance angewandt, sondern eher bei der Schema-Evolution [CMDZ10].

In [FHLM96] wurde die Idee der **inversen Schemaabbildungen** bereits für die Integration heterogener Datenbanken eingeführt. Diese Idee wurde in [SBLH14] auf neuere Entwicklungen in der Theorie inverser Schemaabbildungen angepasst. Dabei wurden die klassischen Schemaabbildungen verallgemeinert auf den Fall, dass nicht alle relevanten Daten im (relationalen) Schema repräsentiert werden können, sondern zusätzlich auf Instanzebene erfasst werden müssen. Diese zusätzlichen Annotationen helfen bei der inversen Abbildung verdichteter Daten.

Im Gegensatz zur traditionellen Datenintegration sollen Anfragen im integrierten System bei diesem Ansatz nicht an das globale Schema gestellt werden, sondern weiterhin an die lokalen Schemata, wobei diese um Daten aus anderen Quellen erweitert werden. Dieser Ansatz wird als *Global-as-local-view-extension* (GaLVE) bezeichnet.

Inverse Schema-Instanz-Abbildungen werden nun auch für die Invertierung, also Rückverfolgung, von Analyseprozessen, also allgemeineren Datenbankanfragen, benötigt. Bisher wurden allerdings für die Datenbankintegration nur einfache Anfragen, bestehend aus

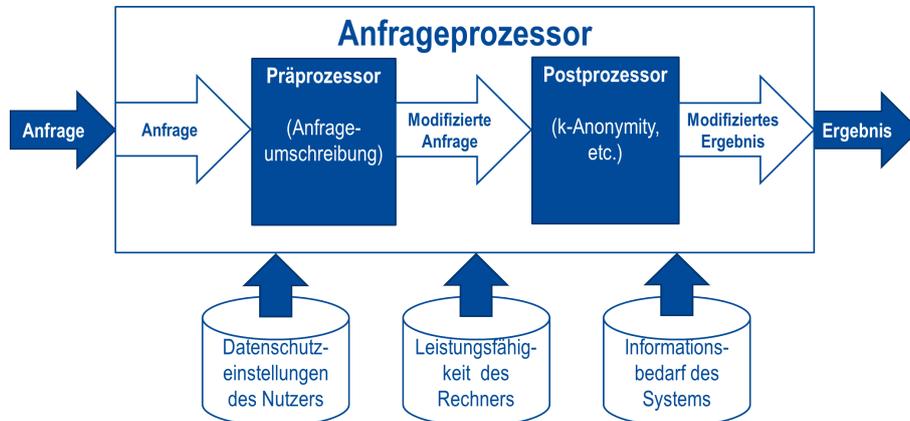


Abbildung 3: Anfragetransformationen zur Berücksichtigung von Privatheitsansprüchen des Nutzers und Informationsbedarf des Systems

Selektion, Projektion und Verbund, berücksichtigt. Analyseprozesse erfordern eine Erweiterung der bestehenden Technik auf statistische Funktionen (skalare Funktionen, Aggregatfunktionen, OLAP) und allgemeine Workflows. Hier muss ermittelt werden, welche Zusatzinformationen zur Gewährleistung der Rückverfolgbarkeit erfasst werden müssen.

4.4 METIS in PArADISE

Im Projekt **PArADISE** (Privacy AwaRe Assistive Distributed Information System Environment): arbeiten wir derzeit an Techniken zur Auswertung von großen Mengen von Sensordaten, die definierte Privatheitsansprüche der späteren Nutzer per Systemkonstruktion erfüllen.

Ein erster Prototyp ist von einer studentischen Arbeitsgruppe erstellt worden. Derzeit können Analysen zur Modellbildung auf Sensordaten in SQL-92, SQL:2003 oder iterativen Ansätzen über SQL-Anweisungen realisiert und auf die Basissysteme DB2 (zeilenorientierte oder spaltenorientiert: DB2 BLU), PostgreSQL (zeilenorientiert) sowie MonetDB (spaltenorientiert und hauptspeicheroptimiert) abgebildet werden. In nächster Zeit wird eine automatische Abbildung der R-Analysen unserer Assistenzsystem-Forscher auf diese SQL-Lösungen umgesetzt.

Die Privatheitsansprüche werden dabei durch automatische Anfragetransformationen abgebildet (nach [Gru14], siehe Abbildung 3). Sowohl die Privatheitsansprüche des Nutzers (Datenschutzeinstellungen wie k-Anonymität und schützenswerte Informationen oder Handlungen) als auch der Informationsbedarf des Assistenzsystems sollen dabei aufeinander abgeglichen werden (was auf ein *Query-Containment-Problem* [Chi09] führt). Später soll noch die Leistungsfähigkeit der berechnenden Prozessoren (am Sensor oder am Ser-

ver) berücksichtigt werden, um über Vorfilterungen und Verdichtungen etwa direkt am Sensor entscheiden zu können.

Im Projekt **METIS** (Management, Evolution, Transformation und Integration von Schemata) wurde die für die Datenintegration entwickelte Technik der *Global-as-local-view-extension* (GaLVE) umgesetzt. Diese wird derzeit für die nötigen *inversen Abbildungen* zum Provenance Management bei der Verarbeitung der Sensordaten und Ableitung von (Situations-, Aktivitäts-, Intentions-)Modellen umgesetzt. Dabei soll unter anderem rekonstruiert werden, welche der Originaldaten maßgeblichen Einfluss auf die Modelle hatten, welche Sensoren untergeordnete Bedeutung haben und in welcher Granularität die Daten für die langfristige Sicherung der Provenance gespeichert werden müssen.

Als Kombination der beiden Projekte **PARADISE** und **METIS** müssen auch die Aspekte der Privatheit und der Provenance gegeneinander abgeglichen werden, da die Forderung der vollständigen Rekonstruierbarkeit von Originaldaten den Privatheitsansprüchen der Nutzer widersprechen kann. Hier werden wir zwischen der Erforschungs- und Entwicklungsphase des Assistenzsystems (mit Testnutzergruppen und eingeschränkter Privatheit, aber voller Provenance) und der Einsatzphase des Assistenzsystems (mit realen Nutzern, voller Umsetzung der Privatheitsansprüche, aber eingeschränkter Provenance) unterscheiden. Allerdings sollen bereits in der Entwicklungsphase des Systems die späteren Privatheitsansprüche konstruktiv durch passende Anfragetransformationen in das Ziel-Assistenzsystem implantiert werden. Während es bei der Entwicklung von Handlungsmodellen für Assistenzsysteme um *Big Data* im eigentlichen Sinn geht, ist im Ziel-Assistenzsystem für eine Person (in einer Wohnung, im Auto) zwar immer noch eine große Menge von Stromdaten aus Sensoren vorhanden, die aber in einschlägigen Artikeln auch als *small data, where $n = me$* [Est14] bezeichnet wird: hier ist also die Anonymität der Person nicht mehr zu verstecken, aber die Herausgabe von schützenswerten Informationen über diese Person.

4.5 Gemeinsamer Kern: Query Containment und inverse Schema-Abbildungen

Der gemeinsame Kern beider Forschungsgebiete ist auf der theoretischen Seite einerseits das Query-Containment-Problem für den Abgleich der Privatheitsansprüche mit den Assistenzzielen und andererseits die inversen Schema-Abbildungen für die Rekonstruktion von Originaldaten. Beide Forschungsgebiete müssen hier miteinander kombiniert werden. Beide Forschungsgebiete haben aber auch ähnliche Probleme: bei einfachen Anfragen mit Konstanten-Selektion, Projektion und Verbund (konjunktiven Anfragen) gibt es Algorithmen zur Lösung der Probleme. Da im Fall der Modellbildung für Assistenzsysteme aber komplexe Analysen (mit Korrelation und Regression als Kernfunktionen) nötig sind, muss die Theorie auf diese Operationen möglichst pragmatisch ausgeweitet werden.

5 Fazit

In diesem Positionspapier sollten zwei Projekte der Rostocker Datenbankgruppe vorgestellt werden, die in Kombination die Basis für ein größeres Vorhaben bilden: die Entwicklung von Assistenzsystemen. Als Big-Data-Problem wurde dabei die sensordaten-gestützte Entwicklung von Handlungsmodellen identifiziert. Abzusichernde Aspekte sind dabei die Privatheitsansprüche des späteren Nutzers des Systems und das Provenance Management für die Forscher, um Entscheidungen für Modelle bis zu den Originaldaten hin rekonstruieren zu können. Beide Probleme sind nicht unabhängig voneinander. Der Leser sollte durch unseren Beitrag sowohl an dem großen Forschungsziel als auch an den Detail-Fragestellungen interessiert werden. Wir hoffen dadurch auf vielfältige Kontakte in den einschlägigen Forschungs-Communities.

Literatur

- [BKT01] Peter Buneman, Sanjeev Khanna und Wang Chiew Tan. Why and Where: A Characterization of Data Provenance. In Jan Van den Bussche und Victor Vianu, Hrsg., *ICDT*, Jgg. 1973 of *Lecture Notes in Computer Science*, Seiten 316–330. Springer, 2001.
- [BSC⁺10] Roger S. Barga, Yogesh L. Simmhan, Eran Chinthaka, Satya Sanket Sahoo, Jared Jackson und Nelson Araujo. Provenance for Scientific Workflows Towards Reproducible Research. *IEEE Data Eng. Bull.*, 33(3):50–58, 2010.
- [CAB⁺14] Lucian Carata, Sherif Akoush, Nikilesh Balakrishnan, Thomas Bytheway, Ripduman Sohan, Margo Seltzer und Andy Hopper. A Primer on Provenance. *Commun. ACM*, 57(5):52–60, Mai 2014.
- [Chi09] Rada Chirkova. Query Containment. In Ling Liu und M. Tamer Özsu, Hrsg., *Encyclopedia of Database Systems*, Seiten 2249–2253. Springer US, 2009.
- [CMDZ10] Carlo Curino, Hyun Jin Moon, Alin Deutsch und Carlo Zaniolo. Update Rewriting and Integrity Constraint Maintenance in a Schema Evolution Support System: PRISM++. *PVLDB*, 4(2):117–128, 2010.
- [CWW00] Yingwei Cui, Jennifer Widom und Janet L. Wiener. Tracing the lineage of view data in a warehousing environment. *ACM Trans. Database Syst.*, 25(2):179–227, 2000.
- [DEW⁺12] Martin Dyrba, Michael Ewers, Martin Wegrzyn, Ingo Kilimann, Claudia Plant, Annahita Oswald, Thomas Meindl, Michela Pievani, Arun L. W. Bokde, Andreas Fellgiebel, Massimo Filippi, Harald Hampel, Stefan Klöppel, Karlheinz Hauenstein, Thomas Kirste und Stefan J. Teipel. Combining DTI and MRI for the Automated Detection of Alzheimer’s Disease Using a Large European Multicenter Dataset. In Pew-Thian Yap, Tianming Liu, Dinggang Shen, Carl-Fredrik Westin und Li Shen, Hrsg., *Multimodal Brain Image Analysis - Second International Workshop, MBIA 2012, Held in Conjunction with MICCAI 2012, Nice, France, October 1-5, 2012. Proceedings*, Jgg. 7509 of *Lecture Notes in Computer Science*, Seiten 18–28. Springer, 2012.
- [Est14] Deborah Estrin. Small data, where n = me. *Commun. ACM*, 57(4):32–34, 2014.
- [Fag07] Ronald Fagin. Inverting schema mappings. *ACM Trans. Database Syst.*, 32(4), 2007.

- [FHLM96] Guntram Flach, Andreas Heuer, Uwe Langer und Holger Meyer. Transparente Anfragen in föderativen Datenbanksystemen. In *Proceedings zum Workshop Föderierte Datenbanken*, Seiten 45–49, 1996.
- [Fre09] Juliana Freire. Provenance Management: Challenges and Opportunities. In Johann Christoph Freytag, Thomas Ruf, Wolfgang Lehner und Gottfried Vossen, Hrsg., *BTW*, Jgg. 144 of *LNI*, Seite 4. GI, 2009.
- [GKT07] Todd J. Green, Gregory Karvounarakis und Val Tannen. Provenance semirings. In Leonid Libkin, Hrsg., *PODS*, Seiten 31–40. ACM, 2007.
- [Gru14] Hannes Grunert. Distributed Denial of Privacy. In Erhard Plödereder, Lars Grunske, Eric Schneider und Dominik Ull, Hrsg., *44. Jahrestagung der Gesellschaft für Informatik, Informatik 2014, Big Data - Komplexität meistern, 22.-26. September 2014 in Stuttgart, Deutschland*, Jgg. 232 of *LNI*, Seiten 2299–2304. GI, 2014.
- [HKG14] Andreas Heuer, Thomas Karopka und Esther Geisler. *Einführung in Ambient Assisted Living*. Lehrbuch Projekt BAAL (Weiterbildung in Ambient Assisted Living) der Universität Rostock, 2014.
- [HKHT06] Andreas Heuer, Thomas Kirste, Wolfgang Hoffmann und Dirk Timmermann. Coast: Concept for proactive assistive systems and technologies. Bericht, Fakultät für Informatik und Elektrotechnik der Universität Rostock, 2006.
- [KNY⁺14] Frank Krüger, Martin Nyolt, Kristina Yordanova, Albert Hein und Thomas Kirste. Computational State Space Models for Activity and Intention Recognition. A Feasibility Study. *PLOS ONE*, November 2014. 9(11): e109381. doi:10.1371/journal.pone.0109381.
- [PIW11] Hyunjung Park, Robert Ikeda und Jennifer Widom. RAMP: A System for Capturing and Tracing Provenance in MapReduce Workflows. *PVLDB*, 4(12):1351–1354, 2011.
- [SBLH14] Georgi Straube, Ilvio Bruder, Dortje Löper und Andreas Heuer. Data Integration in a Clinical Environment Using the Global-as-Local-View-Extension Technique. In *HIS*, Jgg. 8423 of *Lecture Notes in Computer Science*, Seiten 148–159. Springer, 2014.
- [ZSGP11] Jing Zhao, Yogesh Simmhan, Karthik Gomadam und Viktor K. Prasanna. Querying Provenance Information in Distributed Environments. *I. J. Comput. Appl.*, 18(3):196–215, 2011.