

Dr. Bertram Ludäscher

School of Information Sciences
University of Illinois at Urbana-Champaign (UIUC)
501 E. Daniel Street, MC-493
Champaign, IL 61820-6211

Phone (530) 554-1800
Fax (217) 244-3302
ludaesch@illinois.edu

Research Interests

Scientific data & workflow engineering; data integration, knowledge representation & reasoning; conceptual foundations; data science tools. Current research topics include:

- Modeling, design, and analysis of scientific workflows.
- Optimization of data-intensive and compute-intensive workflows.
- Conceptual foundations of provenance in databases and workflows.
- Optimization of provenance capture and querying.
- Principles and practice of data curation.
- Provenance and reproducibility.
- Knowledge-based data integration, knowledge representation and reasoning.
- Reasoning about taxonomies and biodiversity informatics applications.

Education

- 1998 **Ph.D.** (*Dr. rer. nat.*) Computer Science, [University of Freiburg](#), Germany
WOLFGANG-GENTNER Prize (DEM 10,000) for thesis on *Integration of Active and Deductive Database Rules*
Advisor: Prof. Dr. Georg Lausen, Databases and Information Systems ([DBIS](#)), University of Freiburg, Germany
- 1992 **M.S.** (*Dipl.-Inform.*) Computer Science, Karlsruhe Institute of Technology ([KIT](#)), Germany
Advisor: Prof. Dr. Peter H. Schmitt, Inst. for Logic, Complexity & Deduction Systems, KIT, Germany

Appointments

- 2015–present **Director**, Center for Informatics Research in Science & Scholarship (CIRSS), iSchool, UIUC
- 2015–present **Professor** (affiliate), Department of Computer Science, UIUC
- 2014–present **Professor** (affiliate), Computing & Data Sciences, National Center for Supercomputing Applications (NCSA)
- 2014–present **Professor**, School of Information Sciences (the iSchool @ Illinois; formerly: GSLIS), UIUC
- 2009–2014 **Professor**, Department of Computer Science & Genome Center, University of California, Davis
- 2004–2009 **Associate Professor**, Department of Computer Science & Genome Center, University of California, Davis
- 2008–2010 **Research Fellow**, Graduate School of Library & Information Science (GSLIS), UIUC
- 2004–2008 **Research Fellow**, San Diego Supercomputer Center, UC San Diego
- 2002–2004 **Adj. Asst. Professor**, Dept. of Computer Science & Engineering, UC San Diego
- 2004 **Associate Research Scientist**, San Diego Supercomputer Center, UC San Diego
- 2001–2004 **Lab Director, Knowledge-Based Inf. Systems**, San Diego Supercomputer Center, UCSD
- 2000–2004 **Assistant Research Scientist**, San Diego Supercomputer Center, UC San Diego
- 2000–2001 **Lecturer**, Dept. of Computer Science & Engineering, UC San Diego
- 1998–2000 **Visiting Project Scientist**, San Diego Supercomputer Center, UC San Diego
- 1998 **Research Scientist C1** (Asst. Prof.), Databases & Information Systems, U Freiburg, Germany
- 1993–1998 **Research Scientist BAT IIa**, Databases & Information Systems, U Freiburg, Germany

Summary (Reporting Period: May 2017 – April 2018)

Research and Development

I joined the iSchool as a tenured professor in August 2014, and since 2015 I have been appointed Director of the Center for Informatics Research in Science and Scholarship (CIRSS). I also hold faculty appointments (i) with the National Center for Supercomputing Applications (NCSA) in the Computing and Data Sciences group, where I spend a good amount of my time, due to large ongoing projects (see below), and (ii) with the Department of Computer Science, where I am a member of the Data and Information Systems Laboratory (DAIS).

At NCSA I lead a large collaborative project **CC*DNI DIBBS: Merging Science and Cyberinfrastructure Pathways: The Whole Tale**. This \$4,986,951 five year project was [awarded](#) in 2016 and includes collaborators from the University of Chicago, University of Texas at Austin, The University of Notre Dame, and the University of California, Santa Barbara. Given the prominence that NSF assigns to certain projects, including to [Whole Tale](#), this is not a “regular” NSF award, but instead set up as a **cooperative agreement** between NSF and UIUC. As a side-effect, there is considerable additional administrative effort that goes into running this project. In November 2017, the Whole Tale team underwent the scheduled **18 month review**, organized as a “reverse site visit” (RSV) at NSF. Before, during, and after this 18 Month RSV, considerable effort was spent on planning and documentation of project activities. The 18 Month RSV was **successful** and NSF has since released the next increment of funding. In Spring 2019 we will have our second (and final) 36 Month RSV. Whole Tale development is completely open (github.com/whole-tale).

A second large collaborative NSF project is **NSF-SMA 1637155: RIDIR: Collaborative Research: Developing and Deploying SKOPE – A resource for Synthesizing Knowledge of Past Environments**, which was [awarded](#) in 2016, with the majority funding (\$884,627) going to Illinois, and smaller collaborative awards to ASU and WSU (the collaborative total is \$1,339,658). The [SKOPE project](#) is developing and deploying an online resource for paleoenvironmental data and models. A practical (i.e., project management) **challenge** that this project has been dealing with was triggered by the move of the CyberGIS group under Shaowen Wang (SKOPE co-PI) away from NCSA and to the Department of Geography. As a result, the CyberGIS resource ROGER is no longer available to the SKOPE project. Similarly, software development efforts have been delayed for a number of reasons (e.g., lack of developer time available from the CyberGIS team). This has required additional PI effort to supervise the project and get it back on track. As in the case of Whole Tale, the SKOPE software development happens in the open (github.com/openskope).

The Whole Tale and SKOPE projects, as well as the (soon to be ending) [Kurator project](#)¹ ([awarded](#) in 2014) all entail a significant amount of **project management**, e.g., to run weekly meetings, supervise developers, postdocs, and students, or to write regular reports, etc. The nature of the research coming out of these projects is inherently application oriented [39], [46].

In addition, I maintain a strong research interest in conceptual foundations, and technical/engineering topics. For example, I continued work on the foundations of database provenance, with a focus on **why-not provenance**, where the basic problem is to explain the *absence* of certain answers in the result of a query. I had earlier pioneered a game-theoretic approach [162] with my students at UC Davis, and since my move to Illinois established a fruitful collaboration with colleagues at IIT Chicago. This collaboration led to the first practical implementation of **provenance games** and a string of publications and new results [177, 180, 182, 183, 208]. I have also recently published a **survey** on some of the practical and theoretical foundations of provenance in workflows and databases [61].

In another line of foundational research, I continue to work on the theory and applications of **taxonomy alignment**. The Euler/X system I had originally developed with my former students at UC Davis [167, 202, 38, 40] continues to be in active use, and research and development continues in various directions to consolidate, apply, and extend the tools and methods, despite the fact that the original NSF funding has long ended. Real-world taxonomy alignment problems solved with Euler/X include [201, 38, 40, 42, 207] and cover *weevils*, *grasses*, *birds*, and *primates*, to name a few illustrious examples. An ambitious proposal submitted to the NSF [206], unfortunately wasn’t funded.²

A bit closer from home for a typical iSchool audience and use cases, PhD student Jessica Cheng has demonstrated

¹In Fall 2014, the NSF-ABI proposal **Kurator: A Provenance-enabled Workflow Platform and Toolkit to Curate Biodiversity Data** was awarded (\$748,931). This is a collaboration with Prof. James Hanken, Curator in Herpetology, and Director of the Museum of Comparative Zoology at Harvard (collaborative total: \$1.6M). In Kurator, we develop data curation technologies for natural history collections, combining workflow-based approaches and script-based approaches that we aim to “workflow-enable” via methods developed in the YesWorkflow initiative.

²Success rates for many NSF programs are notoriously low and apparently continue to fall.

how Euler/X can be applied when it is necessary to “**agree to disagree**”, using a *geo-taxonomy* example presented at ASIS&T’17 [184]. Led by iSchool alumna Dr. Andrea Thomer we also published a paper [45] which applied Euler/X to produce and visualize multiple possible worlds that capture the relation between different versions of a (subset of) the database schema of Specify 6. The latter is a popular information systems in use by natural history museums.

We are not only exploring new directions for applying Euler/X to non-biological use cases, but also developing new functionality for the Euler/X system and beyond. For example SPIN intern and CS undergraduate student Sahil Gupta has developed a novel [Possible Worlds Explorer](#) that can be used to explore multiple possible worlds (answer sets of logic rules with negation in the body and disjunction/choice in the head). This student is also currently finishing a first reimplementation of Euler/X called [LeanEuler](#).

Another important area of my research is **data curation technologies**, specifically scientific workflows and other related (script-based) approaches to automating data curation workflows. Driven by practical needs in data curation and workflows, I started a new grass-roots project *YesWorkflow* (YW) [37] that is already seeing early adoption by DataONE community members doing climate modeling, archaeologists [39], bioinformaticians, and HPC users.³ The YesWorkflow project is an example how with little funding and/or leveraging existing projects, one can incubate grass-roots projects that have the potential to lead to new research projects (and funding!) down the road. A critical issue here for the iSchool is to further the culture and expand the informatics research capabilities through IS, CS, and informatics students and postdocs. Recent publications related to YesWorkflow include [179] and [212].

Last not least, after a number of prior attempts, our IMLS National Forum proposal ([Limited Access/Use TDM](#)) was finally awarded. The goal of **IMLS National Forum: Data Mining Research Using In-copyright and Limited-access Text Datasets** is to shape a research and implementation agenda for researchers and libraries content providers. This project is co-led by Beth Sandore Namachchivaya (now at the University of Waterloo), Megan Senseney (iSchool), and myself. We held a successful meeting in Chicago in early April of 2018 and are now continuing a number of activities started during the forum meeting.

³Liudmila Mainzer, a senior research scientist at NCSA and the HPCBio group has begun exploring the use of YW for scripts managing HPC jobs on Blue Waters.

Teaching

The first course I taught at the iSchool was *LIS452LE: Foundations of Information Processing in LIS*. Both the new MS/IM program at the iSchool, as well as the MCS/DS program at the CS department, created a demand for a new course on data wrangling / data cleaning. In response to this need, I have been **developing a new course IS 532: Theory & Practice of Data Cleaning** and its Coursera-based companion *CS 513* at the CS department. The course has been evolving over the last few years, as no standard curriculum or syllabus for this course has been readily available. In the current semester (Spring 2018), some new materials based on R and Python were created “on the fly”, but additional efforts are needed over the summer to consolidate and further improve these course materials.

Courses taught

- **LIS 452 Foundations of Information Processing in LIS** (Python-oriented; based on Vetle’s course)
- **IS 532 (formerly: LIS 590): Theory & Practice of Data Cleaning**
- **CS 513 (formerly CS 598): Theory & Practice of Data Cleaning** (Courserabased, for MCSDS)
- IS 592 Independent Study
- IS 599 Thesis Research
- CS 397 Independent Study
- CS 499 Senior Thesis

Other teaching

- Digital Humanities at Oxford Summer School (2016): Humanities Data: A HandsOn Approach; Hands on Data Quality, Provenance, Reproducibility
- 31st Brazilian Symposium on Databases (SBBDB): Tutorial: Provenance in Databases and Scientific Workflows
- DataONE Webinar: Provenance and DataONE: Facilitating Reproducible Science
- iDigBio Webinar: Towards userdefinable, semiautomated workflows for curating biodiversity data

Postdoc mentoring

- Dr. Yang Cao (DataONE project)
- Dr. Qian Zhang (Kurator project)

Independent studies at UIUC

- Cheng, Jessica (IS)
- Gupta, Sahil (CS)
- Jiang, Xiaoliang (IS)
- Lan, Li (IS)
- Senseney, Megan (IS)
- Shrivastava, Pratik (IS)
- Wang, Qiwen (IS)

NCSA SPIN interns mentored

- **Wang, Qiwen (2017, Regular Path Query to SQL compiler)**
- **Gupta, Sahil (2017, Possible Worlds Explorer for Answer Set Programming)**
- **Laskowski, Aubrey (2017, Counting Monotone Boolean Functions)**
- Nath, Nitesh (2016, Parallel Computation of Dedekind numbers)
- Huq, Solaman (2015, Finding Minimum Inconsistent Subset)

DataONE interns mentored

- Lee, Seokki (upcoming 2018, CS PhD student at IIT, Chicago)
- **Lyu, Hui (2017, MS LIS graduate at UIUC, now CS PhD student at UPenn)**
- **Hoang, Linh (2017, LIS PhD student at UIUC)**
- Vu, Duc (2016, ECE PhD student at UIUC)
- Yoo, Michael (2015, CE student at UIUC)
- Song, Tianhong (2014, CS PhD student at UC Davis)

Whole Tale interns mentored

- NunezCorrales, Santiago (upcoming 2018, I3 PhD student at UIUC)
- Cheng, Jessica (upcoming 2018, LIS PhD student at UIUC)
- Lan, Li (upcoming 2018, MS/IM student at UIUC)
- **Shrivastava, Pratik (2017, MS/IM student at UIUC)**

ICES evaluations Most recent evaluations (see separate documents for additional evaluations and details):

- **IS 590 (Theory & Practise of Data Cleaning), Fall 2017** (19 out of 21 students, i.e., 90.5% reporting):
 - Rate the instructor's overall teaching effectiveness: 4.58 (5)
 - Rate the overall quality of this course: 4.47 (5)
 - How much have you learned in this course? 4.58 (5)

Service

Since 2015 I am the Director of the Center for Informatics Research in Science and Scholarship (CIRSS). CIRSS aims to foster research collaborations among iSchool faculty and researchers in a number of informatics research areas. In addition to organizing the well-established CIRSS seminar, colleagues who want to brainstorm new research directions or submit research proposals are encouraged to team up with CIRSS colleagues. We also offer some support for proposal preparation and budget planning etc. My service to the iSchool and beyond during the report period includes:

iSchool

- Member, Doctorial Studies Committee
- Member, Admissions Committee

Campus

- Member, iDSI Steering Committee (Illinois Data Science Initiative)
- Member, NCSA Search Committee (Deputy Director)

Other Professional

- PeerJ Editorial Board Member
- NSF panelist (National Science Foundation)
- ISF reviewer (Israel Science Foundation)
- IPAW PC member
- SSDBM PC member
- VLDB PC member (demo track)
- ESWC PC member
- Reviewer, Distributed and Parallel Databases

Research Projects & Funding

Active Awards

- IMLS National Forum: **Data Mining Research Using In-copyright and Limited-access Text Datasets: Shaping a Research and Implementation Agenda for Researchers, Libraries Content Providers**, \$99,536. 7/1/17–6/30/18. **PI**; co-PIs: Beth Sandore Namachchivaya (University of Waterloo) and Megan Senseney (Illinois), co-I Eleanor Dickson (Illinois).
- NSF-SMA 1637155: **RIDIR: Collaborative Research: Developing and Deploying SKOPE—A resource for Synthesizing Knowledge of Past Environments**, \$884,627, 9/15/16–2/29/2020. **PI**; collaboration with ASU (\$200,842) and Washington State University (\$254,189); collaborative total: \$1,339,658.
- NSF-ACI 1541450: **CC*DNI DIBBS: Merging Science and Cyberinfrastructure Pathways: The Whole Tale**, \$4,986,951, 3/1/16–2/28/21. **PI**; collaboration with University of Chicago; University of Texas, Austin; University of California, Santa Barbara; University of Notre Dame.
- NSF Collaborative Research: ABI Development: **Kurator: A Provenance-enabled Workflow Platform and Toolkit to Curate Biodiversity Data**, \$748,931, 09/1/14-08/31/18 (NCE). **PI**; collaboration with Museum of Comparative Zoology, Harvard University.

Completed (at UIUC)

- NSF-SMA 1439603: **BCC: Collaborative Research: Designing SKOPE: Synthesized Knowledge of Past Environments**, **PI**, 9/1/14–2/28/17, \$133,877; collaboration with ASU and Washington State University.
- NSF-DBI 1147273 **Collaborative Research: ABI Development: Exploring Taxon Concepts (ETC) through Analyzing Fine-Grained Semantic Markup of Descriptive Literature**, \$392,114, 09/01/12–08/31/17 (NCE) **PI**; collaboration with University of Arizona.
- NSF- IIS 1118088: **A Logic-Based, Provenance-Aware System for Merging Scientific Data under Context and Classification Constraints**, \$479,194, 10/1/11-9/30/16, **PI**.

Completed (pre-UIUC)

- NSF/DBI: **Filtered-Push: Continuous Quality Control for Distributed Collections and Other Species-Occurrence Data**, \$414,822, 07/01/10–06/30/2015. **PI** for UCD (collaborative **PI**: J. Hanken, Harvard)
- NSF/SDCI Improvement: **Development of Kepler/CORE – A Comprehensive, Open, Robust, and Extensible Scientific Workflow Infrastructure**, \$1,700,000, 10/01/2007 – 9/30/2010 , **PI**. (co-PIs: I. Altintas (SDSC subcontract), S. Bowers, M. Jones (UCSB subcontract), T. McPhillips).
- NSF/SEI(BIO)+II: **A Collaborative Scientific Workflow Environment for Accelerating Genome-Scale Biological Research**, \$600,139, 7/1/2006 – 6/30/2009, **PI**. (co-PIs: S. Bowers, P. Farnham.)
- NSF/IIS: **Collaborative: Core Database Technologies to Enable the Integration of AToL Information (pPOD)**, \$462,000, 9/2006 – 8/2010, **PI**. (co-PIs at UCD S. Bowers, T. McPhillips; **PI** of collaborative: Val Tannen, U Penn.)
- NSF/CEO-P: **COMET: Coast-to-Mountain Environmental Transect**, \$2,158,580, 10/1/2006 – 9/30/2009, **co-PI**. Since 9/2008: **PI** (until then, **PI**: M. Gertz).
- NSF/CEO-P: **Collaborative: Management and Analysis of Environmental Observatory Data using the KEPLER Scientific Workflow System (REAP)**, \$299,041 (collaborative total: \$2.7M), 10/1/2006 – 9/30/2010, **PI** for UCD. (collaborative **PI**: M. Jones, UCSB.)

- DOE/SciDAC: **DE-FC02-07ER25811: Scientific Data Management Center for Enabling Technologies (CET)**, a.k.a. **Scientific Data Management Center (SDM-2)** \$965,000, 11/15/2006 – 11/14/2011, **PI** for UCD. (co-PI of SDSC subcontract: I. Altintas; PI of SDM-2: A. Shoshani.)
- DOE/SciDAC **Scientific Data Management Center (SDM-1)**, \$927,921, 8/15/2001 – 8/14/2006, **PI** for SDSC. (PI of SDM: A. Shoshani, LBL.)
- DOE **Center for Plasma Edge Simulation (CPES)**, 1/1/2006–12/31/2008, \$243,181 subcontract from LBL, **PI** for UCD. (PI at LBL: A. Shoshani.)
- NSF/ITR: **Science Environment for Ecological Knowledge (SEEK)**, \$2,485,683 (collaborative total: \$12.25M), 10/1/2002 – 9/30/2008, **PI** for UCSD/SDSC, UC Davis. Transfer of project and remaining funds (\$1,678,100) to UC Davis Genome Center, November 1, 2004.
- NSF/ITR: **Collaborative Research: GEON: A Research Project to Create Cyberinfrastructure for the Geosciences (GEON)**, \$5,627,953 (collaborative total: \$11M), 10/1/2002 – 9/30/2007, **co-PI** for SDSC (PI: C. Baru); main project remained at SDSC; from 2005 – 2007: \$272,000 subcontract from SDSC.

Publications

Many of the publications below are available online, see in particular my

- [Google Scholar citation page](#),
- [DBLP page](#), and
- [Illinois Experts page](#).

Some of the more recent publications are also directly hyperlinked below and have a one sentence synopsis.

Journal Articles and Communications

- [1] **B. Ludäscher**, R. Himmeröder, G. Lausen, W. May, and C. Schleppehorst. Managing Semistructured Data with FLORID: A Deductive Object-Oriented Perspective. *Information Systems*, 23(8):589–613, 1998. Pergamon (Elsevier) JOURNAL ARTICLE
- [2] **B. Ludäscher** and G. Lausen. Handling Termination in a Logical Language for Active Rules. *Informatica*, 9(1):65–84, 1998. special issue: *7th Advanced Course on Artificial Intelligence (ACAI)*, Lithuanian Academy of Sciences, Vilnius JOURNAL ARTICLE
- [3] J. Flum, M. Kubierschky, and **B. Ludäscher**. Games and Total Datalog⁺ Queries. *Theoretical Computer Science*, 239(2):257–276, 2000. Elsevier JOURNAL ARTICLE
- [4] R. Moore, C. Baru, A. Rajasekar, **B. Ludäscher**, R. Marciano, M. Wan, W. Schroeder, and A. Gupta. Collection-Based Persistent Digital Archives (Part I&II). *D-Lib Magazine*, 6(3&4), March 2000. INVITED ARTICLE
- [5] A. Paepcke, R. Brandriff, G. Janec, R. Larson, **B. Ludäscher**, S. Melnik, and S. Raghavan. Search Middleware and the Simple Digital Library Interoperability Protocol. *D-Lib Magazine*, 6(3), March 2000. INVITED ARTICLE
- [6] **B. Ludäscher**, R. Marciano, and R. Moore. Preservation of Digital Data with Self-Validating, Self-Instantiating Knowledge-Based Archives. *ACM SIGMOD Record*, 30(3):54–63, 2001. JOURNAL ARTICLE
- [7] G. A. P. C. Burns, K. Stephan, **B. Ludäscher**, A. Gupta, and R. Kötter. Towards a Federated Neuroscientific Knowledge Management System Using Brain Atlases. *Neurocomputing*, 38–40(1–4):1633–1641, 2001. Elsevier JOURNAL ARTICLE

- [8] M. E. Martone, A. Gupta, M. Wong, X. Qian, G. Sosinsky, **B. Ludäscher**, and M. H. Ellisman. A Cell-Centered Database for Electron Tomographic Data. *Journal of Structural Biology*, 138(1–2):145–155, 2002. [JOURNAL ARTICLE](#)
- [9] W. May and **B. Ludäscher**. Understanding the Global Semantics of Referential Actions Using Logic Rules. *Transactions on Database Systems (TODS)*, 27(4):343–397, 2002. [JOURNAL ARTICLE](#)
- [10] O. Boucelma, S. Castano, C. A. Goble, V. Josifovski, Z. Lacroix, and **B. Ludäscher**. Scientific Data Integration – Report on the EDBT’02 Panel. *ACM SIGMOD Record*, 31(4):107–112, 2002. [INVITED ARTICLE](#)
- [11] A. Gupta, **B. Ludäscher**, J. S. Grethe, and M. E. Martone. Towards a Formalization of Disease-Specific Ontologies for Neuroinformatics. *Neural Networks, Special Issue on Neuroinformatics*, 16:1277–1292, 2003. [JOURNAL ARTICLE](#)
- [12] W. K. Michener, J. H. Beach, M. B. Jones, **B. Ludäscher**, D. D. Pennington, R. S. Pereira, A. Rajasekar, and M. Schildhauer. A Knowledge Environment for the Biodiversity and Ecological Sciences. *Journal of Intelligent Information Systems*, 29(1), pp. 111-126. 2005. [JOURNAL ARTICLE](#)
- [13] S. A. Klasky, M. Beck, V. Bhat, E. Feibush, **B. Ludäscher**, M. Parashar, A. Shoshani, D. Silver, and M. Vouk. Data Management on the Fusion Computational Pipeline. *Journal of Physics: Conference Series*, 16:510–520, 2005. [JOURNAL ARTICLE](#)
- [14] **B. Ludäscher**, I. Altintas, C. Berkley, D. Higgins, E. Jaeger, M. Jones, E. A. Lee, J. Tao, and Y. Zhao. [Scientific Workflow Management and the Kepler System](#). *Concurrency and Computation: Practice & Experience*, 18(10):1039–1065, 2006. [Introduction and overview on scientific workflows and the Kepler system](#). [JOURNAL ARTICLE](#)
- [15] I. Altintas, O. Barney, Z. Cheng, T. Critchlow, **B. Ludäscher**, S. Parker, A. Shoshani, and M. Vouk. Accelerating the Scientific Exploration Process with Scientific Workflows. *Journal of Physics: Conference Series*, 46:468–478, 2006. [JOURNAL ARTICLE](#)
- [16] A. Shoshani, I. Altintas, A. Choudhary, T. Critchlow, C. Kamath, **B. Ludäscher**, J. Nieplocha, S. Parker, R. Ross, N. Samatova, M. Vouk, SDM Center Technologies for Accelerating Scientific Discoveries, *Journal of Physics: Conference Series* 78), doi:10.1088/1742-6596/78/1/012068, 2007 [JOURNAL ARTICLE](#)
- [17] A. Deutsch, **B. Ludäscher**, A. Nash. Rewriting Queries Using Views with Access Patterns Under Integrity Constraints. *Theoretical Computer Science*, 371(3), pp.200–226, Elsevier. 2007 [JOURNAL ARTICLE](#)
- [18] **B. Ludäscher**, N. Podhorszki, I. Altintas, S. Bowers, T. McPhillips, From Computation Models to Models of Provenance: The RWS Approach, *Concurrency and Computation: Practice & Experience*, special issue on the First Provenance Challenge, 2007. [JOURNAL ARTICLE](#)
- [19] S. Bowers, T. McPhillips, **B. Ludäscher**. Provenance in Collection-Oriented Scientific Workflows, *Concurrency and Computation: Practice & Experience*, special issue on the First Provenance Challenge, 2007. [JOURNAL ARTICLE](#)
- [20] D. Thau, **B. Ludäscher**, Reasoning about Taxonomies in First-Order Logic, *Ecological Informatics*. Vol. 2(3), pp. 195–209, October, 2007. [JOURNAL ARTICLE](#)
- [21] D. A. Batchelor, M. Beck, A. Becoulet, R. V. Budny, C. S. Chang, P. H. Diamond, J. Q. Dong, G. Y. Fu, A. Fukuyama, T. S. Hahm, D. E. Keyes, Y. Kishimoto, S. Klasky, L. L. Lao, K. Li, Z. Lin, **B. Ludäscher**, J. Manickam, N. Nakajima, T. Ozeki, N. Podhorszki, W. M. Tang, M. A. Vouk, R. E. Waltz, S. J. Wang, H. R. Wilson, X. Q. Xu, M. Yagi, and F. Zonca. Simulation of Fusion Plasmas: Current Status and Future Direction. *Plasma Science and Technology*, 9(3):312–387, 2007.

- [22] J. Cummings, A. Pankin, N. Podhorszki, G. Park, S. Ku, R. Barreto, S. Klasky, C. S. Chang, H. Strauss, L. Sugiyama, P. Snyder, D. Pearlstein, **B. Ludäscher**, G. Bateman, A. Kritz, and the CPES Team, Plasma Edge Kinetic-MHD Modeling in Tokamaks using Kepler Workflow for Code Coupling, Data Management and Visualization, *Commun. Comput. Phys.*, 4, pp. 675–702, 2008. JOURNAL ARTICLE
- [23] S. Davidson, S. Cohen-Boulakia, A. Eyal, **B. Ludäscher**, T. McPhillips, S. Bowers, M. Anand, J. Freire: Provenance in Scientific Workflow Systems. *IEEE Data Engineering Bulletin*. 30(4): 44-50 (2007) INVITED ARTICLE
- [24] T. M. McPhillips, S. Bowers, D. Zinn, and **B. Ludäscher**. [Scientific Workflow Design for Mere Mortals](#). *Future Generation Computer Systems (FGCS)*, 25(5):541–551, 2009. *Identifies desiderata of a more user-friendly scientific workflow design paradigm; shows how COMAD satisfies many desiderata.* JOURNAL ARTICLE.
- [25] J. Cheney, P. Buneman, and **B. Ludäscher**. Report on the Principles of Provenance Workshop. *SIGMOD Record*, 37(1):62–65, 2008. INVITED ARTICLE
- [26] S. Bowers, T. M. McPhillips, and **B. Ludäscher**. [Provenance in collection-oriented scientific workflows](#). *Concurrency and Computation: Practice and Experience*, 20(5):519–529, 2008. *Describes how the Collection-Oriented Model of provenance in Kepler is used to solve the First Provenance Challenge.* JOURNAL ARTICLE.
- [27] L. Moreau, **B. Ludäscher**, I. Altintas, R. S. Barga, S. Bowers, S. P. Callahan, G. C. Jr., B. Clifford, S. Cohen, S. C. Boulakia, S. B. Davidson, E. Deelman, L. A. Digiampietri, I. T. Foster, J. Freire, J. Frew, J. Futrelle, T. Gibson, Y. Gil, C. A. Goble, J. Golbeck, P. T. Groth, D. A. Holland, S. Jiang, J. Kim, D. Koop, A. Krenek, T. M. McPhillips, G. Mehta, S. Miles, D. Metzger, S. Munroe, J. Myers, B. Plale, N. Podhorszki, V. Ratnakar, E. Santos, C. E. Scheidegger, K. Schuchardt, M. I. Seltzer, Y. L. Simmhan, C. T. Silva, P. Slaughter, E. G. Stephan, R. Stevens, D. Turi, H. T. Vo, M. Wilde, J. Zhao, and Y. Zhao. [Special Issue: The First Provenance Challenge](#). *Concurrency and Computation: Practice and Experience*, 20(5):409–418, 2008. *Describes the goals and outcomes of a community workshop on the First Provenance Challenge, bringing together teams to implement provenance capture and querying across different workflow systems.* JOURNAL ARTICLE.
- [28] D. Thau, S. Bowers, and **B. Ludäscher**. [Merging Taxonomies under RCC-5 Algebraic Articulations](#). *Journal of Computing Science and Engineering (JCSE)*, 3(2):109–126, 2009. *Describes a dataset merge algorithm for taxonomically annotated data.* JOURNAL ARTICLE.
- [29] D. Zinn, S. Bowers, S. Köhler, and **B. Ludäscher**. [Parallelizing XML data-streaming workflows via MapReduce](#). *Journal of Computer and System Sciences (JCSS)*, 76(6):447–463, 2010. *Describes a novel technique for efficiently executing XML-based streaming workflows via Map-Reduce.* JOURNAL ARTICLE.
- [30] A. L. Hartman, S. Riddle, T. M. McPhillips, **B. Ludäscher**, and J. A. Eisen. [WATERS: a Workflow for the Alignment, Taxonomy, and Ecology of Ribosomal Sequences](#). *BMC Bioinformatics*, 11:317, 2010. *The paper describes a complex, real-world bioinformatics workflow and its implementation in Kepler/COMAD.* JOURNAL ARTICLE.
- [31] I. Altintas, M. K. Anand, T. N. Vuong, S. Bowers, **B. Ludäscher**, and P. M. Sloat. [A Data Model for Analyzing User Collaborations in Workflow-Driven eScience](#). *International Journal of Computers and Their Applications (IJCA)*, 18(3):160–179, December 2011. Special Issue on Scientific Workflows, Provenance and Their Applications *The paper describes a data model for collaborative provenance and describes a real-world use case from the CAMERA Metagenomics project.* JOURNAL ARTICLE.
- [32] T. Stropp, T. M. McPhillips, **B. Ludäscher**, and M. Bieda. [Workflows for microarray data processing in the Kepler environment](#). *BMC Bioinformatics*, 13:102, 2012. *Describes a new suite of Kepler workflows for microarray data processing.* JOURNAL ARTICLE.
- [33] V. Cuevas-Vicentín, S. C. Dey, S. Köhler, S. Riddle, and **B. Ludäscher**. [Scientific Workflows and Provenance: Introduction and Research Opportunities](#). *Datenbank-Spektrum*, 12(3):193–203, 2012. *Survey article on research questions in scientific workflows and provenance.* JOURNAL ARTICLE.

- [34] Y. Tsai, S. E. McPhillips, A. González, T. M. McPhillips, D. Zinn, A. E. Cohen, M. D. Feese, D. Bushnell, T. Tiefenbrunn, C. Stout, **B. Ludäscher**, B. Hedman, K. O. Hodgson, and S. M. Soltis. [AutoDrug: fully automated macromolecular crystallography workflows for fragment-based drug discovery](#). *Acta Crystallographica Section D: Biological Crystallography*, 69(5):796–803, 2013. *Describes a real-world application of the Rest-Flow workflow system, developed by McPhillips.* [JOURNAL ARTICLE](#)
- [35] T. Song, S. Köhler, **B. Ludäscher**, J. Hanken, M. Kelly, D. Lowery, J. A. Macklin, P. J. Morris, and R. A. Morris. [Towards Automated Design, Analysis and Optimization of Declarative Curation Workflows](#). *International Journal of Digital Curation (IJDC)*, 9(2):111–122, 2014. *Presents a declarative workflow approach to semi-automated curation workflow design.* [JOURNAL ARTICLE](#)
- [36] V. Cuevas-Vicenttín, P. Kianmajd, **B. Ludäscher**, P. Missier, F. Chirigati, Y. Wei, D. Koop, and S. Dey. [The PBase Scientific Workflow Provenance Repository](#). *International Journal of Digital Curation (IJDC)*, 9(2):28–38, 2014. *Describes an implementation of ProvONE.* [JOURNAL ARTICLE](#)
- [37] T. McPhillips, T. Song, T. Kolisnik, S. Aulenbach, K. Belhajjame, K. Bocinsky, Y. Cao, J. Cheney, F. Chirigati, S. Dey, J. Freire, C. Jones, J. Hanken, K. W. Kintigh, T. A. Kohler, D. Koop, J. A. Macklin, P. Missier, M. Schildhauer, C. Schwalm, Y. Wei, M. Bieda, and **B. Ludäscher**. [YesWorkflow: A User-Oriented, Language-Independent Tool for Recovering Workflow Information from Scripts](#). *International Journal of Digital Curation (IJDC)*, 10(1):298–313, 2015. *Describes the YesWorkflow approach to model scripts as workflows.* [JOURNAL ARTICLE](#)
- [38] N. M. Franz, M. Chen, S. Yu, P. Kianmajd, S. Bowers, and **B. Ludäscher**. [Reasoning over Taxonomic Change: Exploring Alignments for the *Perellesschus* Use Case](#). *PLoS ONE*, 10(2):1–34, February 2015. *Gives an overview of the Euler/X approach to taxonomy alignment using a real-world weevils use case.* [JOURNAL ARTICLE](#)
- [39] K. W. Kintigh, J. H. Altschul, A. P. Kinzig, W. F. Limp, W. K. Michener, J. A. Sabloff, E. J. Hackett, T. A. Kohler, **B. Ludäscher**, and C. A. Lynch. [Cultural Dynamics, Deep Time, and Data](#). *Advances in Archaeological Practice*, 3(1):1–15, 2015. *Recommendations for cyberinfrastructure tool building, resulting from a series of workshops.* [JOURNAL ARTICLE](#)
- [40] N. Franz, M. Chen, P. Kianmajd, S. Yu, S. Bowers, A. S. Weakley, and **B. Ludäscher**. [Names Are Not Good Enough: Reasoning over Taxonomic Change in the *Andropogon* Complex](#). *Semantic Web*, 7(6):645–667, 2016. *Uses a real-world use case to show how taxonomic names have changed over 126 years and how Euler/X can be used to “sort things out”.* [JOURNAL ARTICLE](#)
- [41] H. Cui, D. Xu, S. S. Chong, M. Ramirez, T. Rodenhausen, J. A. Macklin, **B. Ludäscher**, R. A. Morris, E. M. Soto, and N. M. Koch. [Introducing Explorer of Taxon Concepts with a Case Study on Spider Measurement Matrix Building](#). *BMC Bioinformatics*, 17:471:1–471:22, 2016. *Describes the outcomes of the NSF ETC project.* [JOURNAL ARTICLE](#)
- [42] N. M. Franz, N. M. Pier, D. M. Reeder, M. Chen, S. Yu, P. Kianmajd, S. Bowers, and **B. Ludäscher**. [Two Influential Primate Classifications Logically Aligned](#). *Systematic Biology*, 65(4):561–582, July 2016. *Another real-world application of Euler/X for aligning taxonomies.* [JOURNAL ARTICLE](#)
- [43] A. Sarkar, S. Köhler, **B. Ludäscher**, and M. Bishop. [Insider Attack Identification and Prevention in Collection-Oriented Dataflow-Based Processes](#). *IEEE Systems Journal*, 11(2):522–533, 2017. *Shows how logic rules in answer set programming can be used to model and analyze privacy attacks.* [JOURNAL ARTICLE](#)
- [44] Q. Zhang, H. Imker, C. Li, **B. Ludäscher**, and M. Senseney. [Using a Computational Study of Hydrodynamics in the Wax Lake Delta to Examine Data Sharing Principles](#). *International Journal of Digital Curation*, 11(2):138–155, July 2017. *Demonstrates the divergence of theory and practice when it comes to data sharing, using a real-world use case.* [JOURNAL ARTICLE](#)

- [45] A. Thomer, Yi-Yun Cheng, J. Schneider, M. Twidale, and **B. Ludäscher**. [Logic-Based Schema Alignment for Natural History Museum Databases](#). *Knowledge Organization*, 44(7):545–558, November 2017. *Uses Euler/X to produce “possible worlds”, relating two versions of a subset of the Specify 6 data model.* **JOURNAL ARTICLE**
- [46] A. Brinckman, K. Chard, N. Gaffney, M. Hategan, M. B. Jones, K. Kowalik, S. Kulasekaran, **B. Ludäscher**, B. D. Mecum, J. Nabrzyski, V. Stodden, I. J. Taylor, M. J. Turk, and K. Turner. [Computing Environments for Reproducibility: Capturing the “Whole Tale”](#). *Future Generation Computer Systems*, February 2018. *Provides an overview of the motivation behind and technical architecture of the Whole-Tale project.* **JOURNAL ARTICLE**

Books and Book Chapters

- [47] **B. Ludäscher**. *Integration of Active and Deductive Database Rules*. DISDBIS 45, infix-Verlag, Sankt Augustin, 1998. Ph.D. thesis, Institut für Informatik, Universität Freiburg, ISBN 3-89601-445-5 **BOOK**
- [48] G. Lausen, **B. Ludäscher**, and W. May. On Logical Foundations of Active Rules. In J. Chomicki and G. Saake, editors, *Logics for Databases and Information Systems*, chapter 12, pp. 389–422. Kluwer Academic Publishers, 1998. part of the *ACM Anthology* **BOOK CHAPTER**
- [49] J. Lu, **B. Ludäscher**, J. Schü, and V. Subrahmanian. Maintaining Well-Founded Views in Constraint Databases. In T. Martin and F. A. Fontana, editors, *Logic Programming and Soft Computing*, volume 3 of *Uncertainty Theory in AI*, pp. 129–144. Wiley, 1998. **BOOK CHAPTER**
- [50] G. Lausen, **B. Ludäscher**, and W. May. On Active Deductive Databases: The Statalog Approach. In B. Freitag, H. Decker, M. Kifer, and A. Voronkov, editors, *Transactions and Change in Logic Databases*, number 1472 in LNCS, pp. 69–106. Springer, 1998. **BOOK CHAPTER**
- [51] M. E. Martone, A. Gupta, **B. Ludäscher**, I. Zaslavsky, and M. H. Ellisman. Federation of Brain Data through Knowledge-Guided Mediation. In R. Kötter, editor, *Neuroscience Databases – A Practical Guide*, pp. 275–292. Kluwer Academic Publishers, 2002. **BOOK CHAPTER**
- [52] **B. Ludäscher**, A. Gupta, and M. E. Martone. A Model-Based Mediator System for Scientific Data Management. In T. Critchlow and Z. Lacroix, editors, *Bioinformatics: Managing Scientific Data*. Morgan Kaufmann, 2003. **BOOK CHAPTER**
- [53] **B. Ludäscher** and L. Raschid, editors. *2nd Intl. Workshop on Data Integration in the Life Sciences (DILS)*, volume 3615 of *Lecture Notes in Bioinformatics (LNBI)*, San Diego, California, July 2005. Springer. **EDITED VOLUME**
- [54] **B. Ludäscher** and C. A. Goble. Guest Editors’ Introduction to the Special Section on Scientific Workflows. *SIGMOD Record*, 34(3), 2005. **EDITED VOLUME**
- [55] **B. Ludäscher**, K. Lin, S. Bowers, E. Jaeger-Frank, B. Brodaric, and C. Baru. Managing Scientific Data: From Data Integration to Scientific Workflows. *Geoinformatics: Data to Knowledge, Geological Society of America Special Paper 397*, pp. 109–129, 2006. **BOOK CHAPTER**
- [56] **B. Ludäscher** and N. Mamoulis, editors. *20th International Conference on Scientific and Statistical Database Management (SSDBM)*, volume 5069 of *Lecture Notes in Computer Science (LNCS)*, Hong Kong, China, July 2008. Springer. **EDITED VOLUME.**
- [57] **B. Ludäscher**, I. Altintas, S. Bowers, J. Cummings, T. Critchlow, E. Deelman, D. D. Roure, J. Freire, C. Goble, M. Jones, S. Klasky, T. McPhillips, N. Podhorszki, C. Silva, I. Taylor, and M. Vouk. [Scientific Process Automation and Workflow Management](#). In A. Shoshani and D. Rotem, editors, *Scientific Data Management: Challenges, Existing Technology, and Deployment*, Computational Science Series, chapter 13. Chapman & Hall/CRC, 2009. *Introduction and overview to scientific workflows.* **BOOK CHAPTER.**

- [58] M. Gertz and **B. Ludäscher**, editors. *Scientific and Statistical Database Management, 22nd International Conference, SSDBM 2010, Heidelberg, Germany, June 30 - July 2, 2010. Proceedings*, volume 6187 of LNCS. Springer, 2010. *Program co-chair and proceedings editor*. **BOOK.**
- [59] A. Chapman, B. Ludäscher, and A. Schreiber, editors. *6th Workshop on the Theory and Practice of Provenance (TaPP)*, Cologne, Germany, June 12–13, 2014. USENIX Association.
- [60] **B. Ludäscher** and B. Plale, editors. *Provenance and Annotation of Data and Processes: 5th International Provenance and Annotation Workshop (IPAW). Revised Selected Papers*, volume 8628 of LNCS, Cologne, Germany, June 9–13, 2014. Springer.
- [61] **B. Ludäscher**. *A Brief Tour Through Provenance in Scientific Workflows and Databases*. (IDEALS version). In V. Lemieux, editor, *Building Trust in Information*, Springer Proceedings in Business and Economics, pp. 103–126. Springer, 2016. *Survey paper with a focus on the relation of workflow vs database provenance, and “worked examples” in the appendix*. **BOOK CHAPTER**

Refereed Conference and Workshop Proceedings

- [62] J. Posegga and **B. Ludäscher**. Towards First-order Deduction Based on Shannon Graphs. In *16th German Conference on Artificial Intelligence (GWAI), Bonn*, LNAI 671, pp. 67–76. Springer, 1992. **RESEARCH ARTICLE**
- [63] G. Lausen and **B. Ludäscher**. Updates by Reasoning about States. In *2nd Intl. East-West Database Workshop (EWDW)*, Workshops in Computing, pp. 17–30, Klagenfurt, Austria, 1994. Springer. **RESEARCH ARTICLE**
- [64] **B. Ludäscher**, U. Hamann, and G. Lausen. A Logical Framework for Active Rules. In *7th Intl. Conference on Management of Data (COMAD)*, pp. 221–238, Pune, India, 1995. Tata-McGraw Hill **RESEARCH ARTICLE**
- [65] **B. Ludäscher**, W. May, and G. Lausen. Nested Transactions in a Logical Language for Active Rules. In *Intl. Workshop on Logic in Databases (LID)*, LNCS 1154, pp. 197–222, San Miniato, Italy, 1996. Springer **RESEARCH ARTICLE**
- [66] **B. Ludäscher**, W. May, and J. Reinert. Towards a Logical Semantics for Referential Actions in SQL. In *6th Intl. Workshop on Foundations of Models and Languages for Data and Objects (FMLDO)*, Dagstuhl, Germany, 1996. **RESEARCH ARTICLE**
- [67] J. Flum, M. Kubierschky, and **B. Ludäscher**. Total and Partial Well-Founded Datalog Coincide. In *6th Intl. Conference on Database Theory (ICDT)*, LNCS 1186, pp. 113–124, Delphi, Greece, January 1997. Springer **RESEARCH ARTICLE**
- [68] R. Himmeröder, G. Lausen, **B. Ludäscher**, and C. Schleppehorst. On a Declarative Semantics for Web Queries. In *5th Intl. Conference on Deductive and Object-Oriented Databases (DOOD)*, LNCS 1341, pp. 386–398, Montreux, Switzerland, 1997. Springer. **RESEARCH ARTICLE**
- [69] **B. Ludäscher**, W. May, and G. Lausen. Referential Actions as Logical Rules. In *16th ACM Symposium on Principles of Database Systems (PODS)*, pp. 217–227, Tucson, Arizona, 1997. **RESEARCH ARTICLE**
- [70] W. May, **B. Ludäscher**, and G. Lausen. Well-Founded Semantics for Deductive Object-Oriented Database Languages. In *5th Intl. Conference on Deductive and Object-Oriented Databases (DOOD)*, LNCS 1341, pp. 320–336, Montreux, Switzerland, 1997. Springer **RESEARCH ARTICLE**
- [71] C. Baru, **B. Ludäscher**, Y. Papakonstantinou, P. Velikhov, and V. Vianu. Features and Requirements for an XML View Definition Language: Lessons from XML Information Mediation. In *W3C Query Languages Workshop (QL’98)*, Boston, Massachusetts, 1998. **INVITED ARTICLE**
- [72] R. Himmeröder, P.-T. Kandzia, **B. Ludäscher**, W. May, and G. Lausen. Search, Analysis, and Integration of Web Documents: A Case Study with FLORID. In *Intl. Workshop on Deductive Databases and Logic Programming (DDL’98)*, pp. 47–57, Manchester, UK, 1998. **RESEARCH ARTICLE**

- [73] R. Himmeröder, G. Lausen, **B. Ludäscher**, and C. Schleppehorst. FLORID: A DOOD-System for Querying the Web. In *Exhibition Program at Intl. Conf. on Extending Database Technology (EDBT)*, Valencia, Spain, 1998. DEMO PAPER
- [74] **B. Ludäscher**, R. Himmeröder, and W. May. Techniques and Rule Patterns for Declaratively Querying Web Data with FLORID. In *Deklarative KI-Methoden zur Implementierung und Nutzung von Systemen in Netzen (Workshop KI-98)*, Bremen, Germany, 1998. RESEARCH ARTICLE
- [75] **B. Ludäscher** and W. May. Referential Actions: From Logical Semantics to Implementation. In *Intl. Conference on Extending Database Technology (EDBT)*, LNCS 1377, pp. 404–418, Valencia, Spain, 1998. Springer RESEARCH ARTICLE
- [76] C. Baru, V. Chu, A. Gupta, **B. Ludäscher**, R. Marciano, Y. Papakonstantinou, and P. Velikhov. XML-Based Information Mediation for Digital Libraries. In *ACM Conf. on Digital Libraries (DL)*, pp. 214–215, Berkeley, CA, 1999. exhibition program DEMO PAPER
- [77] C. Baru, A. Gupta, **B. Ludäscher**, R. Marciano, Y. Papakonstantinou, P. Velikhov, and V. Chu. XML-Based Information Mediation with MIX. In *ACM Intl. Conference on Management of Data (SIGMOD)*, pp. 597–599, Philadelphia, PA, 1999. exhibition program DEMO PAPER
- [78] **B. Ludäscher** and A. Gupta. Modeling Interactive Web Sources for Information Mediation. In *Intl. Workshop on the World-Wide Web and Conceptual Modeling (WWWCM'99)*, LNCS 1727, pp. 225–238, Paris, France, 1999. Springer RESEARCH ARTICLE
- [79] **B. Ludäscher**, Y. Papakonstantinou, P. Velikhov, and V. Vianu. View Definition and DTD Inference for XML. In *Post-ICDT Workshop on Query Processing for Semistructured Data and Non-Standard Data Formats*, Jerusalem, 1999. RESEARCH ARTICLE
- [80] **B. Ludäscher**, Y. Papakonstantinou, and P. Velikhov. A Framework for Navigation-Driven Lazy Mediators. In *ACM SIGMOD Workshop on the Web and Databases (WebDB)*, pp. 85–90, Philadelphia, 1999. RESEARCH ARTICLE
- [81] W. May, R. Himmeröder, G. Lausen, and **B. Ludäscher**. A Unified Framework for Wrapping, Mediating and Restructuring Information from the Web. In *Intl. Workshop on the World-Wide Web and Conceptual Modeling (WWWCM)*, LNCS 1727, Paris, France, 1999. Springer RESEARCH ARTICLE
- [82] G. A. P. C. Burns, K. Stephan, **B. Ludäscher**, A. Gupta, and R. Kötter. Towards a Federated Neuroscientific Knowledge Management System Using Brain Atlases. In *Computational Neuroscience Meeting (CNS)*, Brugge, Belgium, 2000. RESEARCH ARTICLE
- [83] A. Gupta, **B. Ludäscher**, and M. E. Martone. Knowledge-Based Integration of Neuroscience Data Sources. In *12th Intl. Conference on Scientific and Statistical Database Management (SSDBM)*, pp. 39–52, Berlin, Germany, July 2000. IEEE Computer Society RESEARCH ARTICLE
- [84] **B. Ludäscher**, A. Gupta, and M. E. Martone. Model-Based Information Integration in a Neuroscience Mediator System. In *26th Conference on Very Large Data Bases (VLDB)*, pp. 639–642, Cairo, Egypt, 2000. Morgan Kaufmann. demonstration session DEMO PAPER
- [85] **B. Ludäscher**, Y. Papakonstantinou, and P. Velikhov. Navigation-Driven Evaluation of Virtual Mediated Views. In *Intl. Conference on Extending Database Technology (EDBT)*, LNCS 1777, pp. 150–165, Konstanz, Germany, 2000. Springer. RESEARCH ARTICLE
- [86] K. D. Munroe, **B. Ludäscher**, and Y. Papakonstantinou. BBQ: Blended Browsing and Querying of XML in a Lazy Mediator System. In *Intl. Conference on Extending Database Technology (EDBT)*, Konstanz, Germany, 2000. exhibition program DEMO PAPER

- [87] I. Zaslavsky, **B. Ludäscher**, A. Gupta, and R. Marciano. Accuracy Mediation within a Spatial Wrapper-Mediator System. In *1st Intl. Conf. on Geographic Information Science (GIScience)*, Savannah, Georgia, October 2000. ACM. **RESEARCH ABSTRACT**
- [88] A. Gupta, **B. Ludäscher**, and M. E. Martone. An Extensible Model-Based Mediator System with Domain Maps. In *17th Intl. Conf. on Data Engineering (ICDE)*, Heidelberg, Germany, 2001. exhibition program **DEMO PAPER**
- [89] **B. Ludäscher**, A. Gupta, and M. E. Martone. Model-Based Mediation with Domain Maps. In *17th Intl. Conf. on Data Engineering (ICDE)*, Heidelberg, Germany, 2001. IEEE Computer Society **RESEARCH ARTICLE**
- [90] **B. Ludäscher**, R. Marciano, and R. Moore. Towards Self-Validating Knowledge-Based Archives. In *11th Intl. Workshop on Research Issues in Data Engineering (RIDE)*, Heidelberg, Germany, 2001. IEEE Computer Society **RESEARCH ARTICLE**
- [91] I. Zaslavsky, A. Gupta, **B. Ludäscher**, and S. Tambawala. Query Evaluation and Presentation Planning with a Spatial Mediator: Extending XML-Based Mediation to Heterogeneous Sources of GIS and Imagery Data. In *5th Intl. Workshop on Query Processing and Multimedia Issues in Distributed Systems (QPMIDS)*, München, Germany, 2001. **RESEARCH ABSTRACT**
- [92] A. Gupta, **B. Ludäscher**, M. E. Martone, X. Qian, E. Ross, J. Tran, and I. Zaslavsky. KIND-2: A System for Managing Alternate Models in Model-based Mediation. In *19th British National Conference on Databases (BNCOD)*, LNCS 2405, Sheffield, UK, July 2002. Springer. **DEMO PAPER**
- [93] A. Gupta, **B. Ludäscher**, and R. Moore. Ontology Services for Curriculum Development in NSDL. In *Joint Conference on Digital Libraries (JCDL)*, Portland, Oregon, July 2002. ACM & IEEE Computer Society **RESEARCH ARTICLE**
- [94] X. Qian, **B. Ludäscher**, M. E. Martone, and A. Gupta. Navigating Virtual Information Sources with Know-ME. In *EDBT*, number 2287 in LNCS, Prague, Czech Republic, March 2002. **DEMO PAPER**
- [95] A. Rajasekar, R. Moore, **B. Ludäscher**, and I. Zaslavsky. The Grid Adventures: SDSC's Storage Resource Broker and Web Services in Digital Library Applications. In *Digital Libraries: Advanced Methods and Technologies, Digital Collections, 4th All-Russian Scientific Conference (RCDL)*, Dubna, Russia, 2002. **RESEARCH ARTICLE**
- [96] I. Zaslavsky, **B. Ludäscher**, A. Gupta, and J. Tran. Feature Interpretation in Vector Data: Reconciling Spatial and Semantic Integrity Constraints. In *Proc. 2nd Intl. Geographic Information Science Conference (GIScience)*, pp. 215–218, Boulder, USA, September 2002. **RESEARCH ABSTRACT**
- [97] A. Gupta, **B. Ludäscher**, and M. E. Martone. Registering Scientific Information Sources for Semantic Mediation. In *21st Intl. Conference on Conceptual Modeling (ER)*, LNCS 2503, pp. 182–198, Tampere, Finland, October 2002. **RESEARCH ARTICLE**
- [98] **B. Ludäscher**, I. Altintas, and A. Gupta. Time to Leave the Trees: From Syntactic to Conceptual Querying of XML. In *Intl. Workshop on XML-Based Data Management (XMLDM), in conjunction with Intl. Conf. on Extending Database Technology (EDBT)*, LNCS 2490, pp. 148–168, Prague, March 2002. Springer. **RESEARCH ARTICLE**
- [99] **B. Ludäscher**, P. Mukhopadhyay, and Y. Papakonstantinou. A Transducer-Based XML Query Processor. In *28th Conference on Very Large Data Bases (VLDB)*, pp. 227–238, Hong Kong, 2002. **RESEARCH ARTICLE**
- [100] R. Marciano, **B. Ludäscher**, R. Moore, I. Zaslavsky, and K. Pezzoli. Multi-Level Information Modeling and Preservation of eGOV Data. In *Intl. Conference on eGovernment (EGOV)*, LNCS 2456, pp. 93–100, Aix-en-Provence, France, September 2002. **RESEARCH ARTICLE**

- [101] I. Altintas, S. Bhagwanani, D. Buttler, S. Chandra, Z. Cheng, M. Coleman, T. Critchlow, A. Gupta, W. Han, L. Liu, **B. Ludäscher**, C. Pu, R. Moore, A. Shoshani, and M. Vouk. A Modeling and Execution Environment for Distributed Scientific Workflows. In *15th Intl. Conference on Scientific and Statistical Database Management (SSDBM)*, Boston, Massachusetts, 2003. DEMO PAPER
- [102] S. Bowers and **B. Ludäscher**. Towards a Generic Framework for Semantic Registration of Scientific Data. In *Semantic Web Technologies for Searching and Retrieving Scientific Data (SCISW)*, Sanibel Island, Florida, 2003. RESEARCH ABSTRACT
- [103] M. Chagoyen, M. Kurul, P. De-Alarcón, S. Santini, **B. Ludäscher**, J. Carazo, and A. Gupta. A Semantic Mediation Approach for Problems in Computational Molecular Biology. In *IEEE Computer Society Bioinformatics Conference (CSB2003)*, August 2003. RESEARCH ABSTRACT
- [104] A. Gupta, **B. Ludäscher**, and M. Martone. BIRN-M: A Semantic Mediator for Solving Real-World Neuroscience Problems. In *ACM Intl. Conference on Management of Data (SIGMOD)*, 2003. DEMO PAPER
- [105] K. Lin and **B. Ludäscher**. A System for Semantic Integration of Geologic Maps via Ontologies. In *Semantic Web Technologies for Searching and Retrieving Scientific Data (SCISW)*, Sanibel Island, Florida, 2003. RESEARCH ABSTRACT
- [106] **B. Ludäscher**, I. Altintas, and A. Gupta. Compiling Abstract Scientific Workflows into Web Service Workflows. In *15th Intl. Conference on Scientific and Statistical Database Management (SSDBM)*, Boston, Massachusetts, 2003. RESEARCH ABSTRACT
- [107] S. Bowers and **B. Ludäscher**. An Ontology Driven Framework for Data Transformation in Scientific Workflows. In *International Workshop on Data Integration in the Life Sciences (DILS)*, LNCS 2994, Leipzig, Germany, March 2004. RESEARCH ARTICLE.
- [108] S. Bowers, K. Lin, and **B. Ludäscher**. On Integrating Scientific Resources through Semantic Registration. In *16th Intl. Conference on Scientific and Statistical Database Management (SSDBM)*, Santorini Island, Greece, 2004. RESEARCH ARTICLE.
- [109] I. Altintas, C. Berkley, E. Jaeger, M. Jones, **B. Ludäscher**, and S. Mock. Kepler: Towards a Grid-Enabled System for Scientific Workflows. In *Workshop on Workflow in Grid Systems, Global-Grid Forum (GGF10)*, Berlin, Germany, March 2004. RESEARCH ABSTRACT.
- [110] I. Altintas, E. Jaeger, C. Berkley, M. Jones, **B. Ludäscher**, and S. Mock. Kepler: An Extensible System for Design and Execution of Scientific Workflows. In *16th Intl. Conference on Scientific and Statistical Database Management (SSDBM)*, Santorini, Greece, 2004. DEMO PAPER.
- [111] I. Altintas, E. Jaeger, K. Lin, **B. Ludäscher**, and A. Memon. A Web Service Composition and Deployment Framework for Scientific Workflows. In *2nd Intl. Conference on Web Services (ICWS)*, San Diego, California, July 2004. RESEARCH ABSTRACT.
- [112] A. Nash and **B. Ludäscher**. Processing Unions of Conjunctive Queries with Negation under Limited Access Patterns. In *Intl. Conference on Extending Database Technology (EDBT)*, 2004. RESEARCH ARTICLE
- [113] **B. Ludäscher** and A. Nash. Web Service Composition Through Declarative Queries: The Case of Conjunctive Queries with Union and Negation. In *20th Intl. Conf. on Data Engineering (ICDE)*, 2004. RESEARCH ABSTRACT
- [114] A. Nash and **B. Ludäscher**. Processing First-Order Queries with Limited Access Patterns. In *ACM Symposium on Principles of Database Systems (PODS)*, Paris, France, June 2004. RESEARCH ARTICLE.
- [115] A. Deutsch, **B. Ludäscher**, and A. Nash. Rewriting Queries Using Views with Access Patterns Under Integrity Constraints. In *10th Intl. Conference on Database Theory (ICDT)*, number 3363 in LNCS, Edinburgh, Scotland, January 2005. Springer. RESEARCH ARTICLE

- [116] S. Bowers and **B. Ludäscher**. Actor-Oriented Design of Scientific Workflows. In *24th Intl. Conference on Conceptual Modeling (ER)*, LNCS, Klagenfurt, Austria, October 2005. Springer. **RESEARCH ARTICLE**
- [117] S. Bowers and **B. Ludäscher**. Towards Automatic Generation of Semantic Types in Scientific Workflows. In *Intl. Workshop on Scalable Semantic Web Knowledge Base Systems (SSWS)*, LNCS, New York, November 2005. Springer. **RESEARCH ARTICLE**
- [118] I. Altintas, A. Birnbaum, K. Baldridge, W. Sudholt, M. Miller, C. Amoreira, Y. Potier, and **B. Ludäscher**. A Framework for the Design and Reuse of Grid Workflows. In *Intl. Workshop on Scientific Applications on Grid Computing (SAG'04)*, LNCS 3458. Springer, 2005. **RESEARCH ARTICLE**
- [119] W. Michener, J. Beach, S. Bowers, L. Downey, M. Jones, **B. Ludäscher**, D. Pennington, A. Rajasekar, S. Romanello, M. Schildhauer, D. Vieglais, and J. Zhang. Data Integration and Workflow Solutions for Ecology. In *2nd Intl. Workshop on Data Integration in the Life Sciences (DILS)*, volume 3615 of *LNBI*, San Diego, California, July 2005. Springer. **RESEARCH ARTICLE**
- [120] S. Romanello, J. Beach, S. Bowers, M. Jones, **B. Ludäscher**, W. Michener, D. Pennington, A. Rajasekar, and M. Schildhauer. Creating and Providing Data Management Services for the Biological and Ecological Sciences: Science Environment for Ecological Knowledge. In *17th Intl. Conference on Scientific and Statistical Database Management (SSDBM)*, Santa Barbara, California, June 2005. **RESEARCH ARTICLE**
- [121] C. Berkley, S. Bowers, M. Jones, **B. Ludäscher**, M. Schildhauer, and J. Tao. Incorporating Semantics in Scientific Workflow Authoring. In *17th Intl. Conference on Scientific and Statistical Database Management (SSDBM)*, Santa Barbara, California, June 2005. **RESEARCH ARTICLE**
- [122] E. Jaeger, I. Altintas, J. Zhang, **B. Ludäscher**, D. Pennington, and W. Michener. A Scientific Workflow Approach to Distributed Geospatial Data Processing using Web Services. In *17th Intl. Conference on Scientific and Statistical Database Management (SSDBM)*, Santa Barbara, California, June 2005. **RESEARCH ARTICLE**
- [123] S. Bowers, T. McPhillips, **B. Ludäscher**, S. Cohen, and S. B. Davidson. A Model for User-Oriented Data Provenance in Pipelined Scientific Workflows. In *Intl. Provenance and Annotation Workshop (IPAW)*, Chicago, May 2006. **RESEARCH ARTICLE**
- [124] S. Bowers, **B. Ludäscher**, A. H. Ngu, and T. Critchlow. Enabling Scientific Workflow Reuse through Structured Composition of Dataflow and Control-Flow. In *IEEE Workshop on Workflow and Data Flow for Scientific Applications (SciFlow)*, Atlanta, GA, April 2006. in conjunction with ICDE. **RESEARCH ARTICLE**
- [125] S. A. Klasky, **B. Ludäscher**, and M. Parashar. The Center for Plasma Edge Simulation Workflow Requirements. In *IEEE Workshop on Workflow and Data Flow for Scientific Applications (SciFlow)*, Atlanta, GA, April 2006. in conjunction with ICDE. **RESEARCH ARTICLE**
- [126] S. Bowers and **B. Ludäscher**. A Calculus for Propagating Semantic Annotations through Scientific Workflow Queries. In *Query Languages and Query Processing (QLQP): 11th Intl. Workshop on Foundations of Models and Languages for Data and Objects*, LNCS, Munich, Germany, March 2006. **RESEARCH ARTICLE**
- [127] U. Nambiar, **B. Ludäscher**, K. Lin, and C. Baru. [The GEON Portal: Accelerating Knowledge Discovery in the Geosciences](#). In *Intl. Workshop on Web Information and Data Management (WIDM)*, pp. 83–90, Arlington, Virginia, USA, November 2006. ACM. *Describes the GEON Portal for integrating multi-disciplinary geoscience data and tools. Semantic integration is facilitated using ontologies and spatiotemporal metadata.* **RESEARCH ARTICLE.**
- [128] S. Bowers, T. McPhillips, M. Wu, **B. Ludäscher**. Project Histories: Managing Data Provenance Across Collection-Oriented Scientific Workflow Runs, *4th Intl. Workshop on Data Integration in the Life Sciences (DILS)*, University of Pennsylvania, Philadelphia, June 27–29, 2007. **RESEARCH ARTICLE**

- [129] N. Podhorszki, **B. Ludäscher**, S. Klasky. Workflow Automation for Processing Plasma Fusion Simulation Data, *2nd Workshop on Workflows in Support of Large-Scale Science (WORKS)*, Monterey Bay California, June 25, 2007. **RESEARCH ARTICLE**
- [130] M.A. Vouk, I. Altintas, R. Barreto, J. Blondin, Z. Cheng, T. Critchlow, A. Khan, S. Klasky, J. Ligon, **B. Ludäscher**, P. A. Mouallem, S. Parker, N. Podhorszki, A. Shoshani, C. Silva, Automation of Network-Based Scientific Workflows, in: *Grid-based Problem Solving Environments, Intl. Federation for Information Processing (IFIP)*, Volume 239, eds. P.W. Gaffney, J.C.T. Pool JCT, (Boston: Springer), pp 35-61, 2007 **RESEARCH ARTICLE**
- [131] S. Bowers, T. McPhillips, S. Riddle, M. Anand, and **B. Ludäscher**, Kepler/pPOD: Scientific Workflow and Provenance Support for Assembling the Tree of Life, *Intl. Provenance and Annotation Workshop (IPAW)*, Springer LNCS, Salt Lake City, June 2008 **RESEARCH ARTICLE**
- [132] D. Thau, S. Bowers, and **B. Ludäscher**. Merging Taxonomies under RCC-5 Algebraic Articulations. In *CIKM Workshop on Ontologies and Information Systems for Semantic Web (ONISW)*, Napa Valley, California, 2008. **RESEARCH ARTICLE**
- [133] **B. Ludäscher**, S. Bowers, and T. M. McPhillips. [Scientific Workflows](#). In *Encyclopedia of Database Systems*, pp. 2507–2511. Springer, 2009. *This is the entry on “scientific workflow” in Springer’s comprehensive Encyclopedia of Database Systems.* **RESEARCH ARTICLE.**
- [134] **B. Ludäscher**, M. Weske, T. M. McPhillips, and S. Bowers. [Scientific Workflows: Business as Usual?](#) In U. Dayal, J. Eder, J. Koehler, and H. A. Reijers, editors, *BPM*, volume 5701 of *Lecture Notes in Computer Science*, pp. 31–47. Springer, 2009. *Compares and analyses differences between business workflows and scientific workflows.* **RESEARCH ARTICLE.**
- [135] M. K. Anand, S. Bowers, T. M. McPhillips, and **B. Ludäscher**. [Exploring Scientific Workflow Provenance Using Hybrid Queries over Nested Data and Lineage Graphs](#). In *Intl. Conf. on Scientific and Statistical Database Management (SSDBM)*, LNCS 5566, pp. 237–254. Springer, 2009. *Presents a new hybrid model of provenance with a horizontal (time) and vertical (data) dimension.* **RESEARCH ARTICLE.**
- [136] M. K. Anand, S. Bowers, T. M. McPhillips, and **B. Ludäscher**. [Efficient provenance storage over nested data collections](#). In *12th Intl. Conf. on Extending Database Technology (EDBT)*, pp. 958–969, Saint Petersburg, Russia, March 2009. *Storing provenance efficiently is an important challenge in scientific workflows. The paper presents and evaluates an efficient new technique for provenance storage.* **RESEARCH ARTICLE.**
- [137] D. Zinn, S. Bowers, T. M. McPhillips, and **B. Ludäscher**. [Scientific workflow design with data assembly lines](#). In *Workshop on Workflows in Support of Large-Scale Scienc (WORKS)*, Portland, OR, USA, November 2009. *Formalization of VDAL ,a variant / generalization of the COMAD workflow model of computation.* **RESEARCH ARTICLE.**
- [138] D. Zinn, S. Bowers, T. M. McPhillips, and **B. Ludäscher**. [X-CSR: Dataflow Optimization for Distributed XML Process Pipelines](#). In *Intl. Conf. on Data Engineering (ICDE)*, pp. 577–580, Shanghai, China, March 2009. **RESEARCH ARTICLE.**
- [139] M. K. Anand, S. Bowers, and **B. Ludäscher**. [A navigation model for exploring scientific workflow provenance graphs](#). In *Workshop on Workflows in Support of Large-Scale Scienc (WORKS)*, Portland, OR, USA, November 2009. *Introduces a new navigation model (e.g. expand and collapse operators) for workflow provenance graphs and different architectures for implementing the model.* **RESEARCH ARTICLE.**
- [140] P. Missier, **B. Ludäscher**, S. Bowers, S. Dey, A. Sarkar, B. Shrestha, I. Altintas, M. K. Anand, and C. Goble. [Linking multiple workflow provenance traces for interoperable collaborative science](#). In *The 5th Workshop on Workflows in Support of Large-Scale Science (WORKS)*. IEEE, November 2010. *Describes the ideas behind the Data-Tree of Life summer of code project that built an interoperable provenance tool.* **RESEARCH ARTICLE.**

- [141] M. K. Anand, S. Bowers, and **B. Ludäscher**. [Techniques for efficiently querying scientific workflow provenance graphs](#). In I. Manolescu, S. Spaccapietra, J. Teubner, M. Kitsuregawa, A. Léger, F. Naumann, A. Ailamaki, and F. Özcan, editors, *EDBT*, volume 426 of *ACM International Conference Proceeding Series*, pp. 287–298. ACM, 2010. *Develops new techniques for querying large provenance graphs and analyses their performance*.
RESEARCH ARTICLE.
- [142] M. K. Anand, S. Bowers, and **B. Ludäscher**. [Provenance browser: Displaying and querying scientific workflow provenance graphs](#). In *Intl. Conf. on Data Engineering (ICDE)*, pp. 1201–1204, Long Beach, CA, USA, March 2010. *This short paper describes the implementation of the Kepler provenance browser. At the time of writing, few if any other scientific workflow systems had provenance recording, efficient storage, querying and browsing implemented*.
DEMO PAPER.
- [143] M. K. Anand, S. Bowers, I. Altintas, and **B. Ludäscher**. [Approaches for Exploring and Querying Scientific Workflow Provenance Graphs](#). In *Intl. Provenance and Annotation Workshop (IPAW)*, pp. 17–26, 2010. *Introduced scientific workflow provenance model and describes provenance browser architecture*.
RESEARCH ARTICLE.
- [144] D. Zinn, S. Bowers, and **B. Ludäscher**. [XML-based computation for scientific workflows](#). In *Intl. Conf. on Data Engineering (ICDE)*, pp. 812–815, Long Beach, CA, USA, March 2010. *Describes a static analysis technique for XML-based scientific workflows that can be used to improve workflow design and execution*.
RESEARCH ARTICLE.
- [145] D. Zinn, Q. Hart, T. M. McPhillips, **B. Ludäscher**, Y. Simmhan, M. Giakkoupis, and V. K. Prasanna. [Towards Reliable, Performant Workflows for Streaming-Applications on Cloud Platforms](#). In *CCGRID*, pp. 235–244. IEEE, 2011. *Develops a framework and describes an implementation for scientific workflows execution in cloud environments*.
RESEARCH ARTICLE.
- [146] L. Dou, D. Zinn, T. M. McPhillips, S. Köhler, S. Riddle, S. Bowers, and **B. Ludäscher**. [Scientific Workflow Design 2.0: Demonstrating Streaming Data Collections in Kepler](#). In S. Abiteboul, K. Böhm, C. Koch, and K.-L. Tan, editors, *ICDE*, pp. 1296–1299. IEEE Computer Society, 2011. *Demonstration at ICDE showing novel collection-oriented workflow design and streaming workflow capabilities*.
RESEARCH ARTICLE.
- [147] S. C. Dey, D. Zinn, and **B. Ludäscher**. [ProPub: Towards a Declarative Approach for Publishing Customized, Policy-Aware Provenance](#). In *Intl. Conf. on Scientific and Statistical Database Management (SSDBM)*, pp. 225–243, Portland, OR, USA, 2011. *User-defined provenance publication requests and privacy concerns mark an inherent trade-off that is reconciled in a declarative framework*.
RESEARCH ARTICLE.
- [148] S. Köhler, S. Riddle, D. Zinn, T. M. McPhillips, and **B. Ludäscher**. [Improving Workflow Fault Tolerance through Provenance-Based Recovery](#). In *Intl. Conf. on Scientific and Statistical Database Management (SSDBM)*, pp. 207–224, Portland, OR, USA, 2011. *Generalizes the Rescue-DAG approach of CONDOR/DAGMan and is to-date the only crash-recovery method available in Kepler*.
RESEARCH ARTICLE.
- [149] S. Dey, D. Zinn, and **B. Ludäscher**. [Reconciling Provenance Policy Conflicts by Inventing Anonymous Nodes](#). In *4th International Workshop on Resource Discovery (RED)*, LNCS 7117, pp. 172–185, Heraklion, Greece, 2011. Springer. *Improved and extended version of [150]*.
RESEARCH ARTICLE.
- [150] S. Dey, D. Zinn, and **B. Ludäscher**. [Repairing Provenance Policy Violations by Inventing Non-Functional Nodes](#). In M.-E. Vidal, E. Ruckhaus, and Z. Lacroix, editors, *4th International Workshop on Resource Discovery (RED)*, volume 737 of *CEUR Workshop Proceedings*, Heraklion, Greece, May 2011. In conjunction with 8th Extended Semantic Web Conferences (ESWC) *First approach to reconcile user requests and provenance policies in a rule-based logic framework*.
WORKSHOP.
- [151] T. Critchlow, I. Altintas, G. Chin, D. Crawl, H. Iyer, A. Khan, S. Klasky, S. Koehler, **B. Ludäscher**, P. Mouallem, M. Nagappan, N. Podhorszki, A. Shoshani, C. Silva, R. Tchoua, and M. Vouk. [Working with Workflows: Highlights from 5 Years Building Scientific Workflows](#). In *SciDAC: Scientific Discovery through*

Advanced Computing, July 2011. *High-level summary of workflow technology development and application in the SDM Center Project.* RESEARCH ABSTRACT.

- [152] S. Köhler, **B. Ludäscher**, and Y. Smaragdakis. **Declarative Datalog Debugging for Mere Mortals**. In *Datalog 2.0: Datalog in Academia and Industry, Second Intl. Workshop*, LNCS 7494, pp. 111–122, Vienna, Austria, September 2012. Springer. *Develops a set of program rewritings for declarative Datalog debugging using an underlying provenance model.* RESEARCH ARTICLE.
- [153] D. Zinn, T. J. Green, and **B. Ludäscher**. **Win-move is coordination-free (sometimes)**. In *Intl. Conf. on Database Theory (ICDT)*, pp. 99–113, Berlin, Germany, 2012. *Sheds new light on the CALM conjecture by Hellerstein and the related notion of “coordination-free” computation. Shows that non-stratified computations may also be evaluated in a coordination-free manner.* RESEARCH ARTICLE.
- [154] S. Bowers, T. M. McPhillips, and **B. Ludäscher**. **Declarative Rules for Inferring Fine-Grained Data Provenance from Scientific Workflow Execution Traces**. In *Intl. Provenance and Annotation Workshop (IPAW)*, pp. 82–96, Santa Barbara, CA, USA, June 2012. *Presents a logic rule-based approach for enhancing and deriving new provenance information.* RESEARCH ARTICLE.
- [155] M. K. Anand, S. Bowers, and **B. Ludäscher**. **Database Support for Exploring Scientific Workflow Provenance Graphs**. In *Intl. Conf. on Scientific and Statistical Database Management (SSDBM)*, LNCS 7338, pp. 343–360, Chania, Crete, 2012. *Provides a rich provenance graph querying and navigation model, and demonstrates that the resulting language can be implemented efficiently using a standard relational database technology.* RESEARCH ARTICLE.
- [156] S. Köhler, S. Gulati, G. Cao, Q. Hart, and **B. Ludäscher**. **Sliding Window Calculations on Streaming Data using the Kepler Scientific Workflow System**. *Procedia CS: Kepler Workshop, Intl. Conf. on Computational Science (ICCS)*, 9:1639–1646, 2012. *Shows how to improve scientific workflow design for sliding window aggregates over data streams in Kepler.* WORKSHOP.
- [157] S. Dey, S. Köhler, S. Bowers, and **B. Ludäscher**. **Datalog as a Lingua Franca for Provenance Querying and Reasoning**. In *Workshop on the Theory and Practice of Provenance (TaPP)*, Boston, MA, USA, June 2012. *Shows how Datalog can be used to query and reason about provenance.* RESEARCH ARTICLE.
- [158] S. Köhler, P. Seitzer, M. T. Facciotti, and **B. Ludäscher**. **Improved Motif Detection in Large Sequence Sets with Random Sampling in a Kepler workflow**. *Procedia CS: Kepler Workshop, Intl. Conf. on Computational Science (ICCS)*, 9:1999, 2012. *Demonstrates how a bioinformatics workflow (MotifCatcher) can be scaled up via MapReduce in Kepler.* RESEARCH ABSTRACT.
- [159] L. Dou, G. Cao, P. J. Morris, R. A. Morris, **B. Ludäscher**, J. A. Macklin, and J. Hanken. **Kurator: A Kepler Package for Data Curation Workflows**. *Procedia CS: Kepler Workshop, Intl. Conf. on Computational Science (ICCS)*, 9:1614–1619, 2012. video at <http://youtu.be/DEkPbvLsud0> *Demonstrates how data curation pipelines can be implemented using Kepler and “humans in the loop” via Google spreadsheets and other cloud services.* WORKSHOP.
- [160] P. Missier, **B. Ludäscher**, S. Dey, M. Wang, T. McPhillips, S. Bowers, M. Agun, and I. Altintas. **Golden Trail: Retrieving the Data History that Matters from a Comprehensive Provenance Repository**. *International Journal of Digital Curation (IJDC), papers from the 7th Intl. Digital Curation Conference (IDCC, Bristol, UK, December 2011)*, 7(1):139–150, 2012. *Describes the model, architecture, and prototypical implementation of a workflow provenance repository, developed in the context of the DataONE Provenance Working Group.* RESEARCH ARTICLE.
- [161] R. Littauer, K. Ram, **B. Ludäscher**, W. Michener, and R. Koskela. **Trends in Use of Scientific Workflows: Insights from a Public Repository and Recommendations for Best Practice**. *International Journal of Digital Curation (IJDC), papers from the 7th Intl. Digital Curation Conference (IDCC, Bristol, UK, December 2011)*, 7(2):92–100, 2012. *Summary of a simple workflow repository study conducted by the first author as part of a DataONE summer project.* RESEARCH ARTICLE.

- [162] S. Köhler, **B. Ludäscher**, and D. Zinn. [First-Order Provenance Games](#). In *In Search of Elegance in the Theory and Practice of Computation. Essays Dedicated to Peter Buneman*, volume 8000 of *LNCS*, pp. 382–399. Springer, 2013. Preprint arxiv.org/abs/1309.2655 RESEARCH ARTICLE
- [163] P. Missier, S. C. Dey, K. Belhajjame, V. Cuevas-Vicenttín, and **B. Ludäscher**. [D-PROV: Extending the PROV Provenance Model with Workflow Structure](#). In *5th USENIX Workshop on Theory and Practice of Provenance (TaPP)*, Lombard, IL, 2013. RESEARCH ARTICLE
- [164] S. C. Dey, S. Riddle, and **B. Ludäscher**. [Provenance Analyzer: Exploring Provenance Semantics with Logic Rules](#). In *5th USENIX Workshop on Theory and Practice of Provenance (TaPP)*, Lombard, IL, 2013. RESEARCH ARTICLE
- [165] S. Dey, V. Cuevas-Vicenttín, S. Köhler, E. Gribkoff, M. Wang, and **B. Ludäscher**. [On implementing provenance-aware regular path queries with relational query engines](#). In *Proceedings of the Joint EDBT/ICDT 2013 Workshops*, pp. 214–223, Genoa, Italy, 2013. ACM. RESEARCH ARTICLE
- [166] A. Sarkar, S. Kohler, S. Riddle, **B. Ludäscher**, and M. Bishop. [Insider Attack Identification and Prevention Using a Declarative Approach](#). In *Security and Privacy Workshops (SPW), 2014 IEEE*, pp. 265–276. IEEE, 2014. RESEARCH ARTICLE
- [167] M. Chen, S. Yu, N. Franz, S. Bowers, and **B. Ludäscher**. [A Hybrid Diagnosis Approach Combining Black-Box and White-Box Reasoning](#). In *Rules on the Web (RuleML). From Theory to Applications*, volume 8620 of *LNCS*, pp. 127–141, Prague, Czech Republic, August 2014. Springer. RESEARCH ARTICLE
- [168] M. Chen, S. Yu, P. Kianmajd, N. Franz, S. Bowers, and **B. Ludäscher**. [Provenance for Explaining Taxonomy Alignments](#). In *Provenance and Annotation of Data and Processes (IPAW)*, volume 8628 of *LNCS*, pp. 258–260. Springer, Cologne, Germany, 2014. POSTER ABSTRACT
- [169] T. Song, S. Dey, S. Bowers, and **B. Ludäscher**. [Improving Workflow Design Using Abstract Provenance Graphs](#). In *Provenance and Annotation of Data and Processes (IPAW)*, volume 8628 of *LNCS*, pp. 226–228. Springer, Cologne, Germany, 2014. POSTER ABSTRACT
- [170] V. Cuevas-Vicenttín, **B. Ludäscher**, and P. Missier. [Provenance-Based Searching and Ranking for Scientific Workflows](#). In *Provenance and Annotation of Data and Processes (IPAW)*, volume 8628 of *LNCS*, pp. 209–214. Springer, Cologne, Germany, 2014. DEMO PAPER
- [171] S. C. Dey, K. Belhajjame, D. Koop, T. Song, P. Missier, and **B. Ludäscher**. [UP & DOWN: Improving Provenance Precision by Combining Workflow- and Trace-Level Information](#). In *6th Workshop on the Theory and Practice of Provenance (TaPP)*, Cologne, Germany, June 12–13, 2014. RESEARCH ARTICLE
- [172] S. Riddle, S. Köhler, and **B. Ludäscher**. [Towards Constraint Provenance Games](#). In *6th Workshop on the Theory and Practice of Provenance (TaPP)*, Cologne, Germany, June 12–13, 2014. RESEARCH ARTICLE
- [173] S. C. Dey, S. Köhler, S. Bowers, and **B. Ludäscher**. [Computing Location-Based Lineage from Workflow Specifications to Optimize Provenance Queries](#). In *IPAW*, volume 8628 of *Lecture Notes in Computer Science*, pp. 180–193. Springer, 2014. *Describes a new approach for efficient provenance queries, exploiting the fact that provenance graphs often instantiate their underlying workflow graphs.* RESEARCH ARTICLE
- [174] V. Cuevas-Vicenttín, P. Kianmajd, **B. Ludäscher**, P. Missier, F. S. Chirigati, Y. Wei, D. Koop, and S. C. Dey. [Provenance Storage, Querying, and Visualization in PBase](#). In *IPAW*, volume 8628 of *Lecture Notes in Computer Science*, pp. 239–241. Springer, 2014. *Demonstrates a PBase, a prototypical provenance store and system that can visualize workflows and execution traces; visualize reachability relations within these traces; issue SPARQL queries; and visualize query results.* DEMO PAPER

- [175] T. Dang, N. Franz, **B. Ludäscher**, and A. G. Forbes. [ProvenanceMatrix: A Visualization Tool for Multi-Taxonomy Alignments](#). In *VOILA@ISWC*, volume 1456 of *CEUR Workshop Proceedings*, page 13. CEUR-WS.org, 2015. *Proposes a new visualization approach for taxonomy alignment matrices*.
RESEARCH ARTICLE
- [176] S. C. Dey, K. Belhajjame, D. Koop, M. Raul, and **B. Ludäscher**. [Linking Prospective and Retrospective Provenance in Scripts](#). In *TaPP*. USENIX Association, 2015. *Shows how noWorkflow provenance traces can be mapped to user-specified YesWorkflow models*.
RESEARCH ARTICLE
- [177] B. Glavic, S. Köhler, S. Riddle, and **B. Ludäscher**. [Towards Constraint-Based Explanations for Answers and Non-Answers](#). In *TaPP*. USENIX Association, 2015. *Demonstrates how an ontology can provide justifications for answers and non-answers*.
RESEARCH ARTICLE
- [178] T. M. McPhillips, S. Bowers, K. Belhajjame, and **B. Ludäscher**. [Retrospective Provenance Without a Runtime Provenance Recorder](#). In *TaPP*. USENIX Association, 2015. *Presents an approach to embed retrospective provenance in a prospective provenance model without the need for a dedicated provenance recorder*.
RESEARCH ARTICLE
- [179] J. F. Pimentel, S. C. Dey, T. M. McPhillips, K. Belhajjame, D. Koop, L. Murta, V. Braganholo, and **B. Ludäscher**. [Yin & Yang: Demonstrating Complementary Provenance from noWorkflow & YesWorkflow](#). In *IPAW*, volume 9672 of *Lecture Notes in Computer Science*, pp. 161–165. Springer, 2016. *Describes a system prototype that has been developed to link Python-level runtime observables from noWorkflow with conceptual-level workflow models from YesWorkflow*.
DEMO PAPER
- [180] S. Lee, S. Köhler, **B. Ludäscher**, and B. Glavic. [Implementing Unified Why- and Why-Not Provenance Through Games](#). In *IPAW*, volume 9672 of *Lecture Notes in Computer Science*, pp. 209–213. Springer, 2016. *Description and initial demo of the PUG system, unifying why and why-not provenance*.
DEMO PAPER
- [181] Y. Cao, C. Jones, V. Cuevas-Vicenttín, M. B. Jones, **B. Ludäscher**, T. M. McPhillips, P. Missier, C. R. Schwalm, P. Slaughter, D. Vieglais, L. Walker, and Y. Wei. [DataONE: A Data Federation with Provenance Support](#). In *IPAW*, volume 9672 of *Lecture Notes in Computer Science*, pp. 230–234. Springer, 2016. *Describes the provenance capabilities in DataONE*.
DEMO PAPER
- [182] S. Lee, S. Köhler, **B. Ludäscher**, and B. Glavic. [A SQL-Middleware Unifying Why and Why-Not Provenance for First-Order Queries](#). In *ICDE*, pp. 485–496. IEEE Computer Society, 2017. *Describes how a unified approach for why and why-not provenance can be implemented on top of SQL middleware system*.
RESEARCH ARTICLE
- [183] S. Lee, X. Niu, **B. Ludäscher**, and B. Glavic. [Integrating Approximate Summarization with Provenance Capture](#). In *TaPP*. USENIX Association, 2017. *Presents an approach and implementation for computing summarized representations of why and why-not provenance*.
RESEARCH ARTICLE
- [184] Y.-Y. Cheng, N. Franz, J. Schneider, S. Yu, T. Rodenhause, and **B. Ludäscher**. [Agreeing to disagree: Reconciling conflicting taxonomic views using a logic-based approach](#). *Proceedings of the Association for Information Science and Technology (ASIS&T)*, 54(1):46–56, 2017. *Demonstrates the use of Euler/X for taxonomy alignment, using a “geo-taxonomies” example*.
RESEARCH ARTICLE
- [185] M. Senseney, E. Dickson, B. Namachchivaya, and **B. Ludäscher**. [Data Mining Research with In-copyright and Use-limited Text Datasets: Preliminary Findings from a Systematic Literature Review and Stakeholder Interviews](#). In *13th International Digital Curation Conference (IDCC)*, Barcelona, Spain, February 2018. *Initial findings from our IMLS National Forum project on data mining research with use-limited texts*.
RESEARCH ARTICLE
- [186] M. R. Gryk and **B. Ludäscher**. [Semantic Mediation to Improve Reproducibility for Biomolecular NMR Analysis](#). In *iConference*, volume 10766 of *Lecture Notes in Computer Science*, pp. 620–625. Springer, 2018. *Employs ideas of semantic mediation with ontologies to use cases from the NMRbox project to improve computational reproducibility*.
RESEARCH ARTICLE

Patents

- [187] Data Source Integration System and Method. U.S. National Patent Application, SD2001-037-1MI, **B. Ludäscher**, co-inventor, 2001.
- [188] Collection-Based Persistent Digital Archives. U.S. National Patent Application, UCSD Case No.2001-086, **B. Ludäscher**, co-inventor, 2001.

Other Conference and Workshop Proceedings

- [189] G. Lausen and **B. Ludäscher**. Database Programming with States. In *5th ERCIM Workshop Parallelism and Non-Determinism for Data Intensive Applications*, France, November 1993. INRIA. **RESEARCH ABSTRACT**
- [190] U. Hamann and **B. Ludäscher**. Zustandsorientierte Beschreibung von Updates in deduktiven Datenbanken. In *7. GI-Workshop: Grundlagen von Datenbanken*. Universität Hildesheim, 1995. ISSN 0941-3014 **RESEARCH ARTICLE**
- [191] R. Himmeröder and **B. Ludäscher**. Querying the Web with FLORID. In *10. GI-Workshop: Grundlagen von Datenbanken*, pp. 47–51, Konstanz, Germany, 1998. ISSN 1430-3558 **RESEARCH ARTICLE**
- [192] K. Lindquist, F. Vernon, T. Hansen, A. Rajasekar, **B. Ludäscher**, J. Orcutt, J. Berger, H.-W. Braun, and Y. Bock. Generalizing Seismic Processing Systems to Diverse Signal Domains. In *14th Annual IRIS Workshop*, Waikoloa, Hawaii, June 2002. IRIS Consortium. **WORKSHOP**
- [193] H.-W. Braun, T. Hansen, K. Lindquist, **B. Ludäscher**, J. Orcutt, A. Rajasekar, and F. Vernon. Distributed Data Management Architecture for Embedded Computing. In *6th Workshop on High Performance Embedded Computing (HPEC)*, MIT Lincoln Laboratory, September 2002. **WORKSHOP**
- [194] L. Dou, J. Hanken, **B. Ludäscher**, J. A. Macklin, T. M. McPhillips, P. J. Morris, R. A. Morris, and Z. Wang. [Building Specimen-Data Curation Pipelines using Kepler Workflow Technology in a Filtered-Push Network](#). In *Society for the Preservation of Natural History Collections (SPNHC), Annual Meeting*, San Francisco, May 2011. <http://youtu.be/DEkPbvLsud0> *Demonstrates how data curation pipelines can be implemented using Kepler and “humans in the loop” via Google spreadsheets and other cloud services.* **DEMO PAPER.**
- [195] L. Dou, J. Hanken, **B. Ludäscher**, J. A. Macklin, T. McPhillips, P. J. Morris, R. A. Morris, Z. Wang, and S. Koehler. [Automating Data Curation with Kepler Workflows](#). In *7th Intl. Digital Curation Conference (IDCC)*, Bristol, UK, December 2011. *Demonstration of how data curation pipelines can be implemented using Kepler with “humans in the loop”.* **POSTER.**
- [196] P. Morris, J. Macklin, L. Dou, J. Hanken, M. Kelly, D. Lowery, **B. Ludäscher**, T. McPhillips, and R. Morris. [Connecting Expert Knowledge to Specimens: Remote Annotation of Data in a Filtered-Push Network](#). In *Society for the Preservation of Natural History Collections (SPNHC), Annual Meeting*, Yale University, June 2012. *Abstract and System Demonstration of the Filtered-Push System.* **DEMO PAPER.**
- [197] J. Macklin, L. Dou, J. Hanken, M. Kelly, D. Lowery, **B. Ludäscher**, P. Morris, and R. Morris. [Planning an ApplePie Network based on Filtered-Push Technology to Connect Botanical Collections](#). In *Society for the Preservation of Natural History Collections (SPNHC), Annual Meeting*, Yale University, June 2012. *Abstract of an instance of the Filtered-Push Network, focusing on Botanical Collections.* **RESEARCH ABSTRACT.**
- [198] J. Cheney, A. Finkelstein, **B. Ludäscher**, and S. Vansummeren. [Principles of Provenance \(Dagstuhl Seminar 12091\)](#). *Dagstuhl Reports*, 2(2):84–113, 2012. *Summaries of talks, discussions, and open problems from the Dagstuhl Seminar on Principles of Provenance.* **RESEARCH REPORT.**
- [199] **B. Ludäscher**, V. Cuevas-Vicentín, P. Missier, S. Dey, P. Kianmajd, Y. Wei, D. Koop, F. Chirigati, I. Altintas, K. Belhajjame, et al. [Facilitating Scientific Research through Workflows and Provenance on the DataONE Cyberinfrastructure](#). *AGU Fall Meeting Abstracts*, 1:07, 2013. **POSTER ABSTRACT**

- [200] **B. Ludäscher**, V. Cuevas-Vicentín, P. Missier, S. Dey, P. Kianmajd, Y. Wei, D. Koop, F. Chirigati, I. Altintas, K. Belhajjame, et al. [Scientific Workflows + Provenance = Better \(Meta-\)Data Management](#). *AGU Fall Meeting Abstracts*, 1:07, 2013. POSTER ABSTRACT
- [201] N. M. Franz, N. M. Pier, D. M. Reeder, M. Chen, S. Yu, P. Kianmajd, S. Bowers, and **B. Ludäscher**. [Taxonomic Provenance: Two Influential Primate Classifications Logically Aligned](#). *arXiv preprint arXiv:1412.1025*, 2014. PREPRINT
- [202] M. Chen, S. Yu, N. Franz, S. Bowers, and **B. Ludäscher**. [Euler/X: A Toolkit for Logic-Based Taxonomy Integration](#). *CoRR*, abs/1402.1992, 2014. *System description*. TECHNICAL REPORT
- [203] M. Jones, P. Slaughter, L. Walker, C. Jones, P. Missier, **B. Ludäscher**, Y. Cao, T. McPhillips, and M. Schildhauer. [Lowering the Barrier to Reproducible Research by Publishing Provenance from Common Analytical Tools](#). In *AGU Fall Meeting Abstracts*, 2015. RESEARCH ABSTRACT
- [204] **B. Ludäscher**, K. Chard, N. Gaffney, M. B. Jones, J. Nabrzyski, V. Stodden, and M. Turk. [Capturing the "Whole Tale" of Computational Research: Reproducibility in Computing Environments](#). *CoRR*, abs/1610.09958, 2016. RESEARCH ABSTRACT
- [205] V. Stodden, **B. Ludäscher**, K. Bocinsky, K. Kintigh, T. Kohler, T. McPhillips, and J. Rush. [Querying Provenance Information: Basic Notions and an Example from Paleoclimate Reconstruction](#). In *AGU Fall Meeting Abstracts*, 2016. RESEARCH ABSTRACT
- [206] N. Franz, E. Gilbert, **B. Ludäscher**, and A. Weakley. [Controlling the taxonomic variable: Taxonomic concept resolution for a southeastern United States herbarium portal](#). *Research Ideas and Outcomes*, 2:e10610, 2016. *Research proposal (not funded)*. TECHNICAL REPORT
- [207] N. Franz, L. Musher, J. Brown, S. Yu, and **B. Ludäscher**. [Verbalizing phylogenomic conflict: Representation of node congruence across competing reconstructions of the neoavian explosion](#). *bioRxiv*, page 233973, 2017. *Full-fledged "birds" use case for Euler/X*. PREPRINT
- [208] S. Lee, S. Köhler, **B. Ludäscher**, and B. Glavic. [Efficiently Computing Provenance Graphs for Queries with Negation](#). *CoRR*, abs/1701.05699, 2017. *First practical implementation for why-not provenance over first-order queries, based on a Datalog-Neg rewriting*. TECHNICAL REPORT
- [209] **B. Ludäscher**. [From Provenance Standards and Tools to Queries and Actionable Provenance](#). In *AGU Fall Meeting Abstracts*, 2017. RESEARCH ABSTRACT
- [210] J. Wiczorek, P. J. Morris, J. Hanken, D. Lowery, **B. Ludäscher**, J. Macklin, T. McPhillips, R. Morris, and Q. Zhang. [Darwin Cloud: Mapping real-world data to Darwin Core](#). *Biodiversity Information Science and Standards (TDWG)*, 1:e20486, 2017. RESEARCH ABSTRACT
- [211] P. J. Morris, J. Hanken, D. Lowery, **B. Ludäscher**, J. Macklin, T. McPhillips, R. Morris, J. Wiczorek, and Q. Zhang. [Fitness-for-Use-Framework-aware Data Quality workflows in Kurator](#). *Biodiversity Information Science and Standards (TDWG)*, 1:e20379, 2017. RESEARCH ABSTRACT
- [212] Q. Zhang, P. J. Morris, T. McPhillips, J. Hanken, D. Lowery, **B. Ludäscher**, J. Macklin, R. Morris, and J. Wiczorek. [Using YesWorkflow hybrid queries to reveal data lineage from data curation activities](#). *Biodiversity Information Science and Standards (TDWG)*, 1:e20380, 2017. RESEARCH ABSTRACT

Abstracts of Non-Refereed Conference Proceedings

- [213] A. Gupta, **B. Ludäscher**, and M. E. Martone. [Development of a 3D Neuron-Centered Database/UCSD Neuron-Centered Database](#). In *Human Brain Project/Neuroinformatics Conference*, NIH, 2000. RESEARCH ABSTRACT

- [214] M. Martone, **B. Ludäscher**, and A. Gupta. Databases of Neuronal and Subneuronal Tomographic Structure and Molecular Distribution. In *Intl. Conf. on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS)*, Las Vegas, 2000. **RESEARCH ABSTRACT**
- [215] M. Martone, A. Gupta, **B. Ludäscher**, I. Zaslavsky, and M. Ellisman. From Molecules to Brains: Navigation Through Multiresolution Data Using Knowledge-Based Mediation. In *HBP/NIH Workshop on Web Accessible Databases and Analytical Tools for Neuroscience Research*, New Orleans, November 2000. **RESEARCH ABSTRACT**
- [216] M. E. Martone, A. Gupta, **B. Ludäscher**, I. Zaslavsky, and M. H. Ellisman. Federation of Distributed Brain Data Using Model-Based Mediation. In *Annual Meeting of the Society for Neurosciences*, 2001. **RESEARCH ABSTRACT**
- [217] M. E. Martone, S. Peltier, S. Lamont, S. K. A. Gupta, **B. Ludäscher**, T. Molina, and M. H. Ellisman. Increasing Access to Tomographic Resources: Web-Based Telemicroscopy and Database. In *Proceedings of the Microscopy Society of America*, 2001. **RESEARCH ABSTRACT**
- [218] M. E. Martone, A. Gupta, **B. Ludäscher**, I. Zaslavsky, X. Qian, M. Wong, S. Zhang, J. Tran, and M. Ellisman. Data Integration of Multi-Resolution Brain Data: Tools for Spatial and Ontological Mapping. In *Human Brain Project meeting*, NIH, 2002. **RESEARCH ABSTRACT**
- [219] B. Brodaric, **B. Ludäscher**, and K. Lin. The Role of XML in Mediated Data Integration Systems with Examples from Geological (Map) Data Interoperability. In *Geological Society of America (GSA) Annual Meeting*, volume 35(6), November 2003. **RESEARCH ABSTRACT**
- [220] A. Gupta and **B. Ludäscher**. The Many Faces of Process Interaction Graphs: A Data Management Perspective. In *Workshop on Data Management for Molecular and Cell Biology, to appear in OMICS – A Journal of Integrative Biology*, NIH Campus, Bethesda, MD, 2003. **RESEARCH ABSTRACT**
- [221] K. Lin, **B. Ludäscher**, B. Brodaric, D. Seber, C. Baru, and K. A. Sinha. Semantic Mediation Services in Geologic Data Integration: A Case Study from the GEON Grid. In *Geological Society of America (GSA) Annual Meeting*, volume 35(6), November 2003. **RESEARCH ABSTRACT**
- [222] L. Peterson, E. Yin, D. Nelson, I. Altintas, **B. Ludäscher**, T. Critchlow, A. J. Wyrobek, and M. A. Coleman. Mining the Frequency Distribution of Transcription Factor Binding Sites of Ionizing Radiation Responsive Genes. In *New Horizons in Genomics, DOE/SC-0071*, Santa Fe, New Mexico., March 30–April 1 2003. **RESEARCH ABSTRACT**
- [223] K. A. Sinha, C. Smyth, A. Zendel, B. Brodaric, C. Barnes, A. Snoke, **B. Ludäscher**, D. Seber, and C. Baru. Igneous Rocks, Terranes and GEON. In *Geological Society of America (GSA) Annual Meeting*, volume 35(6), November 2003. **RESEARCH ABSTRACT**

Other Presentations / Works

- [224] **B. Ludäscher**. *CNF-Prolog: A Meta-Interpreter for Chan's Constructive Negation*. Studienarbeit (jointly with M. Vorbeck), Universität Karlsruhe, Fakultät für Informatik, 1990. (in German).
- [225] **B. Ludäscher**. *Ein Deduktionssystem für Prädikatenlogik erster Stufe basierend auf Shannon Graphen (A First-Order Deduction System based on Shannon Graphs)*. Diplomarbeit, Universität Karlsruhe, Fakultät für Informatik, 1992. (in German). **MASTER'S THESIS**
- [226] G. Lausen and **B. Ludäscher**. Updates by Reasoning about States. Technical Report 59, Universität Freiburg, Institut für Informatik, September 1994. **TECHNICAL REPORT**
- [227] **B. Ludäscher**, U. Hamann, and G. Lausen. Reconciling Active and Deductive Databases by States. Technical Report 70, Institut für Informatik, Universität Freiburg, 1995. **TECHNICAL REPORT**

- [228] G. Lausen and **B. Ludäscher** (co-PI). Integrierte Verarbeitung heterogener Information im World-Wide Web. Funded Proposal to the German Science Foundation, 1998.
- [229] **B. Ludäscher**, G. Yang, and M. Kifer. FLORA: The Secret of Object-Oriented Logic Programming. Technical report, State University of New York, Stony Brook, June 1999. see also www.cs.sunysb.edu/~sbprolog/flora **TECHNICAL REPORT**
- [230] R. Moore, C. Baru, A. Gupta, **B. Ludäscher**, R. Marciano, and A. Rajasekar. Collection-Based Long-Term Preservation. Technical report, San Diego Supercomputer Center, June 1999. submitted to National Archives and Records Administration. **TECHNICAL REPORT**
- [231] **B. Ludäscher**, A. Gupta, M. E. Martone, and M. Ellisman. The Brain Data Grid Architecture. In *Global Grid Forum 1 (GGF1)*, Amsterdam, 2001. **RESEARCH ABSTRACT**
- [232] **B. Ludäscher**. From Database Federation to Model-Based Mediation: Databases Meets Knowledge Representation. Colloquium, Dept. of Computer Science and Engineering, U.C. San Diego, October 2001.
- [233] R. Marciano, **B. Ludäscher**, and R. Moore. The Senate Legislative Activities Collection (SLA): A Case Study (Infrastructure Research to Support Preservation Strategies). Technical report, San Diego Supercomputer Center, SDSC TR-2001-5, 2001. **TECHNICAL REPORT**
- [234] R. Moore and **B. Ludäscher** (co-PI). DOE/SciDAC Scientific Data Management Center: Development of Logic-Based Data Federation Software (Grant # DE-FC02-01ER25486), August 2001. <http://sdm.lbl.gov/sdmcenter/>. **RESEARCH GRANT**
- [235] C. Baru, M. J. Bailey, **B. Ludäscher** (co-PI), and P. Papadopoulos. NSF ITR (large): GEON: A Research Project to Create Cyberinfrastructure for the Geosciences, October 2002. <https://www.fastlane.nsf.gov/servlet/showaward?award=0225673> **RESEARCH GRANT**
- [236] **B. Ludäscher** (PI), A. Rajasekar, J. Goguen, D. Stockwell, and V. Vianu. NSF ITR (large): Enabling the Science Environment for Ecological Knowledge (SEEK), October 2002. <https://www.fastlane.nsf.gov/servlet/showaward?award=0225674> **RESEARCH GRANT**
- [237] Y. Papakonstantinou, A. Deutsch, V. Vianu, and **B. Ludäscher** (co-PI). NSF ITR (small): Querying Sequentially Accessed XML Data, August 2003. <https://www.fastlane.nsf.gov/servlet/showaward?award=0313384> **RESEARCH GRANT**
- [238] M. Ellisman, M. Martone, A. Gupta, P. Papadopoulos, C. Baru, I. Zaslavsky, **B. Ludäscher** (co-I), A. Rajasekar, and J. Grethe. Biomedical Informatics Research Network Coordinating Center (BIRN-CC) Renewal Proposal, UCSD 2003-9262, 2003. **RESEARCH GRANT**
- [239] J. A. Orcutt, F. L. Vernon, A. Rajasekar, H.-W. Braun, and **B. Ludäscher** (co-PI). NSF ITR (medium): Real-Time Data Aware System for Earth, Oceanographic, and Environmental Applications (ROADNet), September 2003. <https://www.fastlane.nsf.gov/servlet/showaward?award=0325963> **RESEARCH GRANT**
- [240] **B. Ludäscher**, I. Altintas, A. Gupta, M. E. Martone, and X. Qian. Data Integration and Mediation. Tutorial at the NPACI All Hands Meeting, San Diego, 2002. **TUTORIAL**
- [241] **B. Ludäscher**, I. Altintas, and A. Gupta. Proposal for an Executable Workflow (EWF) Language. Technical report, SDSC/SciDAC, 2002. **TECHNICAL REPORT**
- [242] **B. Ludäscher**, I. Altintas, and A. Gupta. A Semantic Mediation Approach for Scientific Workflows. Technical report, SDSC/SciDAC, 2002. **TECHNICAL REPORT**
- [243] **B. Ludäscher**. Query Containment, Minimization, and Semantic Optimization of Conjunctive Queries. Technical Report BIRN-DI-TN-2003-01, San Diego Supercomputer Center, 2003. **TECHNICAL REPORT**

- [244] **B. Ludäscher**, I. Altintas, A. Gupta, and K. Lin. Scientific Data Integration and Mediation. Tutorial at the NPACI All Hands Meeting, San Diego, 2003. **TUTORIAL**
- [245] **B. Ludäscher** and I. Altintas. On Providing Declarative Design and Programming Constructs for Scientific Workflows based on Process Networks, Technical report, San Diego Supercomputer Center, 2003. **TECHNICAL REPORT**
- [246] L. Kalinichenko, D. Atkins, E. Fox, D. Fulker, Y. Ioannidis, S. König, **B. Ludäscher** M. Marlino, P. Savino, N. Seshagiri, T. R. Sumner, A. Ushakov, and M. Wright. Digital Libraries in Education. Analytical Survey. UNESCO Institute for Information Technologies in Education, 2003. **SURVEY**
- [247] **B. Ludäscher**. Kepler: Scientific Workflows Based on Dataflow Process Networks. e-Science Workflow Services Meeting, National e-Science Center, Edinburgh, Scotland, December 2003. **RESEARCH ABSTRACT**
- [248] **B. Ludäscher**. Panel on Semantic Technologies Panel, Partnerships in Innovation – Serving a Networked Nation, National Archives and Records Administration (NARA), College Park, MD, November 16th, 2004.
- [249] **B. Ludäscher**. Managing Scientific Data: From Data Integration to Scientific Workflows (and back). Center for Hybrid and Embedded Software Systems (Chess), UC Berkeley, September 2004.
- [250] **B. Ludäscher**. Scientific Data and Workflow Engineering - Preliminary Notes from the Cyberinfrastructure Trenches. National Science Foundation (NSF), Arlington, November 2004.
- [251] **B. Ludäscher**. Semantic Technologies: Towards Making a Difference in Scientific Data Management. Opportunities for Distributed Science: A DOE National Collaboratories Program Workshop, NCAR, Boulder, December 2004.
- [252] **B. Ludäscher**, I. Altintas, and R. Moore. Scientific Process Automation (SPA), Report to the US Department of Energy SciDAC Program, Grant # DE-FC02-01ER25486. San Diego Supercomputer Center and UC Davis Genome Center, April 2005.
- [253] **B. Ludäscher**. Semantic Extensions for Scientific Workflows on the Grid. Intl. Symposium on Grid Computing, Academia Sinica, Taipei, April 2005.
- [254] M. Jones, **B. Ludäscher**, and A. Rajasekar. The SEEK EcoGrid: A Data Grid System for Ecology. Intl. Symposium on Grid Computing, Academia Sinica, Taipei, April 2005.
- [255] S. Bowers and **B. Ludäscher**. A Semantic Type System and Propagation Mechanism for Scientific Workflows. 3rd Annual GEON All-Hands Meeting, San Diego, California, May 2005.
- [256] **B. Ludäscher**. Managing Scientific Data: From Data Integration to Scientific Workflows. Göttinger Informatik Kolloquium, Georg-August-Universität, Göttingen, Germany, July 2005.
- [257] **B. Ludäscher**. Kepler Project Overview and Status. Sixth Biennial Ptolemy Miniconference: Featuring the Kepler Project, UC Berkeley, California, May 2005.
- [258] T. Wong and **B. Ludäscher**. Managing Data-Intensive Scientific Workflows with KEPLER and Storage Resource Broker (SRB). In *Student Workshop on Computing (SWOC)*. Department of Computer Science, UC Davis, October 2005.