BETH A. PLALE

Michael A and Laurie Burns McRobbie Bicentennial Professor of Computer Engineering
Chair of the Department of Intelligent Systems Engineering
Executive Director, Pervasive Technology Institute
Indiana University Bloomington
www.linkedin.com/in/bethplale
ORCID ID: 0000-0003-2164-8132

EXECUTIVE SUMMARY

Dr. Beth Plale is the Michael A and Laurie Burns Professor of Computer Engineering at Indiana University Bloomington, where she is also the Chairperson of the Department of Intelligent Systems Engineering and the Executive Director of the Pervasive Technology Institute. As a computational scientist Plale's research interests are in software, hardware, and governance infrastructure for AI, open science, provenance & reproducibility, AI ethics, and data accountability. She is a co-PI of the US National Science Foundation funded Intelligent Cyberinfrastructure for Computation in the Environment (ICICLE) AI Institute and served at the US National Science Foundation (NSF) 2017 -2022. Plale is a founder of the international Research Data Alliance (RDA) and the IU Center of Excellence for Women and Technology (CEWIT). Plale received the distinguished US Department of Energy (DOE) Early Career award and is a senior member of ACM and IEEE professional societies.

EDUCATION AND APPOINTMENTS

EDUCATION

Georgia Institute of Technology	College of Computing	Postdoctoral Fellow	1999-2001
State Univ of New York (SUNY) Binghamton	Computer Science	PhD	1998

Dissertation: Software Approach to Hazard Detection using On-line Analysis of Safety Constraints
Chairs: Sudhir Aggarwal, SUNY Binghamton and Karsten Schwan, Georgia Tech

Temple University	Computer & Info Science	M.S.	1992

Thesis: Object Oriented System for Image Rendering
Advisor: Frank Friedman

University of LaVerne	Business and Public Management	M.B.A.	1986
University of Southern Mississippi	Computer Science, Math Minor	BSc	1984

APPOINTMENTS

2023	Chair, Dept of Intelligent Systems Engineering, IUB
2021 -	Executive Director, Pervasive Technology Institute, IU

2020 -	Michael A and Laurie Burns McRobbie Bicentennial Professor of Computer Engineering, IUB
2018 –	Professor, Dept. of Intelligent Systems Engineering, IUB
2017 – 2020	Science Advisor for Public Access, National Science Foundation (NSF)
2015 – 2017	Affiliated Faculty, Ostrom Workshop on Political Theory and Policy Analysis
2011 – 2018	Professor, Dept. of Informatics and Computing, IU
2014 – 2017	Science Director, Pervasive Technology Institute (PTI), IU
2019 –	Director, Data to Insight Center (D2I), IU
2008 – 2017	
2011 – 2017	Director, HathiTrust Research Center (HTRC)
2011 – 2014	Managing Director, Pervasive Technology Institute, IU
2007 – 2009	Associate Dean of Research (ADR), School of Informatics and Computing, IUB
2006 – 2009	Director, Center for Data and Search Informatics, Indiana University
2006 – 2011	Associate Professor, Dept. of Informatics and Computing, IUB
2002 – 2007	Founder and mentor, Women in Computing (WIC), School of Informatics, IUB
2001 – 2006	Assistant Professor, Dept. of Computer Science, IUB
1998 – 2001	Postdoctoral Fellow, Georgia Institute of Technology, with: Karsten Schwan
1996 – 1997	Graduate Research Assistant, College of Computing, Georgia Institute of Technology
1994 – 1996	Adjunct Instructor, Georgia Perimeter College, Atlanta, Georgia
1991 – 1994	Graduate Research Assistant, Dept. Computer Science, SUNY Binghamton
1993 – 1994	Adjunct Instructor, Continuing Ed, SUNY Binghamton
1989 – 1991	Teaching Assistant, Dept. Computer and Information Science, Temple University
1988 – 1989	Lead Software Engineer, GTE Federal Systems, Westlake Village, California
1986 – 1987	Software Engineer, GTE Federal Systems, Westlake Village, California
1984 – 1986	Software Developer, Vitro Corporation, Oxnard, California

PROFESSIONAL MEMBERSHIIPS

Senior Member, Institute of Electrical and Electronics Engineers (IEEE)
IEEE Women in Engineering (IEEE WIE) member
Senior Member, Association of Computing Machinery (ACM)
Member, American Association for the Advancement of Science

AWARDS / DISTINCTIONS _____

Fred B Luddy Chair Faculty Fellow 2024-2025
Michael A and Laurie Burns McRobbie Bicentennial Chair of Computer Engineering 2020-
DOE Early Career Awardee 2004
Visiting Scholar, Notre Dame 2009
Visiting Scholar, University of Edinburgh, Scotland 2002

RESEARCH and SOCIETAL IMPACT _____

IU experts discuss White House AI guidelines, express concerns. Interview in piece by Isabella Vesperini, Nov 28, 2023, WTIU/WFIU News, https://indianapublicmedia.org/news/iu-experts-discuss-white-house-ai-guidelines-express-concerns.php

Artificial Intelligence in era of ChatGPT. Interview on Indiana Newsdesk, WTIU/WFIU News, Jun 2, 2023. https://www.youtube.com/watch?v=4A3zf-qKCRY&t=436s

Artificial Intelligence: Technology of the Future. Interview as part of Noon Edition WFIU, Jan 2023. https://indianapublicmedia.org/noonedition/artificial-intelligence-technology-of-the-future.php

Policy recommendations to ensure that research software is openly accessible and reusable. McKiernan EC, Barba L, Bourne PE, Carter C, Chandler Z, Choudhury S, et al. (2023), PLoS Biol 21(7): e3002204. https://doi.org/10.1371/journal.pbio.3002204

Panel Member: Building Responsible AI Review Processes for the NAIRR, Meeting #6 of the Office of Science and Technology Policy (OSTP) National Artificial Intelligence Research Resource, Apr 2022

Working Group Recommendations for an Indiana University Research Data Commons (RDC), Simon K, Dallis-Comentale D, Fang S, Huang K, Link M, Maupome G, Menczer F, Neal A, Perry B, Plale B, Raymond A, & Wiehe S (2022). https://hdl.handle.net/1805/30725

Subject Expert: NIH ODSS Innovation Lab: A Data Ecosystems Approach to Ethical AI for Biomedical and Behavioral Research, Feb 2022

Site Visit Chair: Science Foundation Ireland (SFI) $100M INSIGHT project. 2021

NASEM Panel Participant: Accelerating Scientific Discovery through Intelligent and Automated Research Workflows, Board on Research Data and Information, The National Academies of Sciences, Engineering, and Medicine, Mar 2020

Interview: Open data for better science, Weijie Zhao et al, National Science Review, Volume 5, Issue 4, July 2018, pp 593–597, https://doi.org/10.1093/nsr/nwy059, Published:07 Jun 2018

Interview: Data science perspectives, State University of New York Binghamton, Apr 2019, https://www.binghamton.edu/news/story/1780/points-of-view

Interview: On the work at Hathi Trust Research Center, Big data, Aggregations and Text Mining, Open University, UK, Dec 2013, http://kmi.open.ac.uk/news/article/18573

Board Leadership

External Advisory Board	Nat'l Center for Atmospheric Research (NCAR) Computational & Information Systems Laboratory (CISL)	2024 -
External Advisory Board	Indiana University Libraries Audiovisual Metadata Platform (AMP)	2022 - 2023
External Advisory Board Chair	NSF Institute for Harnessing Data and Model Revolution in the Polar Regions (iHARP), UMBC	2022 -
Advisory Committee	NASA Earth Science Advisory Committee (ESAC). A NASA FACA committee	2021 -
External Advisory Board	Qualitative Data Repository (QDR) Technical Advisory Board	2021 -
Board of Directors	DONA Foundation, a non-profit organization in Geneva, Switzerland on internet governance	2021 -
Editorial Board	Big Earth Data, Taylor and Francis Online	2017 - 2023
External Advisory Board	HathiTrust Research Center (HTRC)	2017 - 2021
External Advisory Board	NSF Whole Tale project, University of Illinois	2016—2017
External Advisory Board	Stanford University Center for Expanded Data Annotation and Retrieval (CEDAR)	2015 – 2017
Working Group	NOAA Data Archive and Access Requirements Working Group (subcommittee of NOAA Science Advisory Board)	2014 - 2018
Inaugural Chair	Research Data Alliance (RDA) Technical Advisory Board	2013 – 2015
Executive Steering Committee	Pacific Rim Applications and Grid Middleware Assembly (PRAGMA)	2020 – 2015 – 2017
Steering Committee	Open Grid Forum	2005 – 2007

GRANTS, CONTRACTS, and GIFTS (65) _____

	AWARD	*FUNDER*
65	RCN:CIP:Midwest Research Computing and Data Consortium, National Science Foundation, $999,998, Plale co-PI, 10/2022 – 9/2026	NSF #2227672
64	Cybershuttle: An end-to-end Cyberinfrastructure Continuum to accelerate Discovery in Science and Engineering, National Science Foundation (NSF), $2,161,272, Plale co-PI, 09/2022 – 08/2026	NSF #2209872
63	AI Institute: ICICLE: Intelligent Cyberinfrastructure with Computational Learning in the Environment, National Science Foundation (NSF), $20,000,000, Plale co-PI and IU PI (IU portion $1,999,000), 11/2021 – 10/2026	NSF #2112606
62	Data Capsule Appliance for Research Analysis of Restricted and Sensitive Data in Academic Libraries, Institute of Museum and Library Services (IMLS), $320,546, 06/2017 – 05/19	IMLS
61	SEADTrain: Cloud-hosted Hands-on Training Environment for Data Science where Data Frequently Restricted in Some Form, Microsoft Research, valued at $60,000, 04/2017 – 03/2018	Microsoft
60	CC* Storage: Robust Persistent Identification of Data (RPID), Plale (PI) with Tufts and CNRI, National Science Foundation, $199,047, 04/2017 – 03/2018. NSF Award 1659310.	NSF #1659310
59	Hardware Acquisition HathiTrust Research Center, Plale PI, HathiTrust Board of Governors, $102,000, 07-2016 – 06/2017	HathiTrust
58	Bringing Visibility to Food Security Data Results: Harvests of PRAGMA and RDA, Plale PI with National Institute of Advanced Industrial Science and Technology (AIST) Japan, MacArthur Foundation through Rensselaer Polytechnic University, $40,000, 01/2016 – 11/2016	MacArthur Foundation
57	Workshop on Data Quality in an Era of Big Data, Plale PI, National Science Foundation through CCC award to Midwest Big Data Hub, $30,000, 06/2016 – 10/2016	NSF through CCC
56	Worksets and Data Capsules: Laying foundations for secure computation with copyrighted data in HathiTrust, Plale co-PI and IU PI. Andrew W. Mellon Foundation, $1,170,000, 01/2016 – 12/2017	Andrew W. Mellon Foundation
55	HathiTrust Research Center, Plale PI with UIUC, HathiTrust, $1,000,000, 07/2014 – 06/2018	HathiTrust
54	BD Hubs: Midwest: SEEDCorn: Sustainable Enabling Environment for Data Collaboration, Ed Seidel PI (UIUC), Plale co-PI. National Science Foundation, $1,250,000, 10/2015 – 09/2018	NSF #1550320
53	Hazards SEES: Understanding cross-scale interactions of trade and policy to improve resilience to drought risk in Zambia, Kelly Caylor (Princeton) PI, IU portion Plale co-PI with Tom Evans (PI), National Science Foundation, $1,862,385, 09/2014 – 08/2019.	NSF #1534544
52	WSC-Category 2 Collaborative: Impacts of Agricultural Decision Making and Adaptive Management on Food Security, Kelly Caylor (Princeton) PI. IU portion Plale co-PI with Tom Evans (PI) and Shahzeen Attari, $1,862,385, National Science Foundation, 09/2014 – 08/2019	NSF #1360463
51	Research Data Alliance (RDA) Data Share, Plale PI/PD, with Rensselaer Polytechnic Institute, Alfred P. Sloan Foundation Grant G-2014-13746, $748,000, 01/2015 - 12/2017	Alfred P. Sloan Foundation
50	Socio-Eco-Informatics: Enhancing Predictive Capability of Social Ecological Systems Research, Plale with David Leake and Xiaozhong Liu, Faculty Research Support Program, Indiana University, $73,317, 03/2014 - 02/2016	Indiana University

49	Building the Research Data Alliance Community Through U.S. and International Engagement (RDA2), Plale co-PI, with Fran Berman (PI) Rensselaer Polytechnic Institute, $6,022,637, National Science Foundation, 10/2013 - 09/2018	NSF #1349002
48	DataNet: Sustainable Environments Actionable Data (SEAD), National Science Foundation, Plale co-PI and IU PI, with UMich (PI), UIUC, National Science Foundation, total $8,000,000, 10/2011 – 09/2016 (IU portion $2.4M)	NSF #0940824
47	SAVI: PRAGMA - Enabling Scientific Expeditions and Infrastructure Experimentation for Pacific Rim Institutions and Researchers, IU Lead, with UC San Diego (PI), UWisconsin, UFlorida, $5,693,064, National Science Foundation, 10/2012 - 09/2017	NSF #1234983
46	Workset Creation for Scholarly Analysis: Prototyping Project, Plale co-PI, with J. Stephen Downie (PI) UIUC, $436,525, Andrew W. Mellon Foundation, 07/2013 - 09/2015	Andrew W. Mellon Foundation
45	Software Sustainability: an SI2 PI Workshop, Plale PI, with Doug Thain (Notre Dame) and Matt Jones (UC Santa Barbara), $72,478, National Science Foundation, 01/2014 – 12/2014	NSF #1419131
44	CLIR/DLF Data Curation Fellowship, Council on Library and Information Resources, $77,400, 07/2012 – 06/2014	CLIR
43	Building Trident Community, Beth Plale and Stacy Kowalczyk, $50,000, Microsoft, 07/12 – 06/13	Microsoft
42	A Data Consortium: Coming Together Around Data, $98,204, National Science Foundation, 08/12 – 07/13	NSF #1238168
41	SI2 SSE: Pipeline Framework for Ensemble Runs on Clouds, Plale PI with U. Miami, National Science Foundation, $492,588, IU: $292,863, 04/2012 – 03/2014	NSF #1148359 (IU)
40	Coming Together Around Data, A PI Project Meeting for DataNet and INTEROP, National Science Foundation, Plale PI, Scott Jensen co-PI, $86,571, 08/2011 – 08/2012	NSF #1152946
39	Data Capsule for Non-Consumptive Research, Plale PI, with Atul Prakash, U Michigan, Alfred P. Sloan Foundation, $606,161, 07/2011 – 12/2014	Alfred P. Sloan Foundation
38	HathiTrust Research Center: Computational Research on the HathiTrust Repository, Plale PI with Scott Poole, UIUC and others. HathiTrust Consortium, $0*, 07/01/2011 – 06/30/2014 [*Funding of $5M was conditional on outcome of class action lawsuit by Author's Guild and publishers against Google; suit was eventually thrown out of court.]	HathiTrust
37	Microsoft Exploratory Research in Workflow and Provenance, Microsoft Research, gift through IU Foundation, $125,000, 12/2010	Microsoft
36	EAGER: In-situ Archiving of Digital Scientific Data, Plale (PI) with Elinor Ostrom and Tom Evans, National Science Foundation, $204,991, 10/2010 – 09/2012	NSF #1058452
35	Instant Karma: Applying a Proven Provenance Tool to NASA AMSR-E Data Production Stream, Plale (co-PI, IU lead) with U of Alabama Huntsville, NASA ACCESS 2009, ($260,066 to IU), 10/2009 – 09/2011	NASA
34	netKarma: GENI Provenance Registry, Plale (PI) with Christopher Small of Global Research Network Operations Center (GRNOC), Global Environments for Networked Innovations (GENI) program, NSF through BBN Corp., approx. $484,486, 09/2009 – 08/2012 [Plale .60 portion]	BBN
33	NextGen Weather Observation Networks, Plale (co-PI) with Oklahoma University, Federal Aviation Administration (FAA) through subcontract from Oklahoma University, $247,500 total ($90,000 to IU), 05/2009 – 03/2010	FAA
32	Towards Broad and Deep Community Deployment of LEAD, a Conference Grant, Plale (PI) with University of Alabama and University of Illinois Urbana Champaign, National Science Foundation, $172,664 ($94,664 to IU), 02/2009 – 09/2009	NSF

31	Pervasive Technology Institute, with Brad Wheeler, Craig Stewart, Geoffrey Fox and Fred Cate; Lilly Endowment, $15,000,000, 01/2009 – 12/2014 [Data To Insight Center portion $4.2M.]	Lilly Endowment
30	Cyberinfrastructure Software: Availability, Persistence, and Support Workshop, Plale (co-PI) with Brad Wheeler, Geoffrey Brown, Stacy Kowalczyk, and C. Stewart, National Science Foundation. $131,691, 06/2008 – 05/2011 [Workshop]	NSF
29	III:Small: Assisted Emulation for Digital Preservation, PI: Geoffrey Brown, National Science Foundation, $181,969 [minor role at project end]	NSF #1016967
28	LEAD-in-a-Box, Plale (PI) with Oklahoma University and University of North Carolina, Microsoft Corporation gift, $100,000 ($50,000 to IU), 01/2010 – 05/2011	Microsoft
27	ETF Grid Infrastructure Group: Providing System Management and Integration for the Teragrid, Plale (co-PI) with Craig Stewart, University of Chicago/National Science Foundation, $2,519,000, 08/2005 – 07/2011 [Co-PI 2009-2011, minor role]	NSF
26	Future Grid: Experimental High-Performance Grid Testbed, Plale (Sr. Investigator) with Geoffrey Fox (PI), Craig Stewart, et al., National Science Foundation, $10,000,000, 10/09 - 09/13 [Minor role.]	NSF
25	Knowledge Discovery through Provenance Collection, Representation, and Use in the Lilly Science Grid (LSG), Plale (PI) with Dennis Groth and University of Manchester, UK, Eli Lilly Corp., $220,000 ($120,000 to IU), 03/2008 – 08/2009 [Plale .75 portion]	Eli Lilly Corp.
24	SDCI Data: New Toolkit for Provenance Collection, Publishing, and Use, Plale (PI) with David Leake and Yogesh Simmhan. National Science Foundation, $437,954, 09/2007 -08/2009 [Plale .66 portion]	NSF #0721674
23	SDCI NMI: Improvement: Open Grid Computing Environments Software for Science Gateways Plale (Sr. Investigator) with Geoffrey Fox, Dennis Gannon, Marlon Pierce, and Nancy Wilkins-Diehr. National Science Foundation, NSF 07-503, $1,698,347, 07/2007 - 06/2010 [Minor role]	NSF #0721656
22	CSR-CSI: An Adaptive Programming Framework for Data and Event Driven Computation, Plale (PI) with Dennis Gannon, National Science Foundation, $300,000, 08/2007 - 07/2009 [Plale .50 portion]	NSF #0720580
21	Remote Digital Signatures and Signature Logging: A Proposal to Develop a Tool to Increase the Productivity of Physicians, Plale (PI), Midwest Proton Radiotherapy Institute, $8,953, 03/2007-08/2008	Midwest Proton Radiotherapy Institute
20	Center for Research on Multicore Computing (CRMC), with Geoffrey Fox and Dennis Gannon, Microsoft, $749,996, 07/2006 - 06/2009 [Minor role]	
19	Visual Search Tools using Existing Toolkits in LEAD, Plale (PI) with Sangmi Pallickara, National Science Foundation Research Experiences for Undergraduates (REU) associated with NSF ITR LEAD, $14,000, 06/2006 – 05/2007	NSF
18	Chemical Informatics Cyberinfrastructure, Plale with Gary Wiggins, Geoffrey Fox (PI), Mookie Baik, Dennis Gannon, Randy Bramley, John Huffman, and Marlon Pierce, National Institutes of Health, $731,750, 09/2005 - 07/2007 [minor role]	NIH
17	Science of Search: Data, Analytics, and Architectures Center (DSAAC), Plale (PI) with Dennis Gannon, National Science Foundation, $10,000, 08/2006 - 07/2007	NSF #0630322
16	2005-2006 Distinguished Lecture Series, Plale with Kay Connelly and Butler University, DePauw University, Rose-Hulman Institute of Technology, Computing Research Association CRA-W, 2005 – 2006, $10,000. [Seminar series]	Computing Research Association
15	MRI: Acquisition of a High-Speed, High Capacity Storage System to Support Scientific Computing: the Data Capacitor, Plale (co-PI) with Craig Stewart,	NSF #0521433

	Randy Bramley, Tom Hacker, and Cathy Pilachowski, National Science Foundation, 09/2005 – 08/2008, $1,999,981. [Equipment grant]	
14	DIALOGUE - Data Integration Applications: Linking Organisations to Gain Understanding and Experience", Plale (co-PI) with Malcolm Atkinson, University of Edinburgh, Joel Statz, Ohio State, Peter Brezeny, University of Vienna. Engineering and Physical Sciences Research Council (EPSRC), United Kingdom, 01/2005 – 01/2007. [Travel grant, workshop series]	EPSRC, United Kingdom
13	Expanding Science by Unified Access to Radar Data and Algorithms, Plale (PI), Microsoft Research Equipment Grant, $35,000 equip + $53,000 gift, 09/2004 – 01/2007	Microsoft
12	DOE Early Career: Time-based Data Streams: Fundamental Concepts for Data Resource for Streams, Plale (PI), Department of Energy, 09/2004 - 08/2007, $298,343	DOE
11	An Itanium Environment for Grid-Based Data Mining, Plale (co-PI) with Dennis Groth, Hewlett Packard Philanthropy and Education Equipment Grant, 09/2004 – 09/2005, $60,000. [Equipment grant]	Hewlett Packard
10	ITR Collaborative Research: Linked Environments for Atmospheric Discovery (LEAD), with Dennis Gannon, Oklahoma University, University of Alabama Huntsville, UCAR, University of Illinois Urbana Champaign, Howard University, Colorado State University, and Millersville University. National Science Foundation, 10/2003 - 9/2009, total approx. $11,000,000 ($2,234,101 IU). [Plale portion .50]	NSF #0331480 (IU)
9	Large-scale Streaming Data in Scientific Portals, Plale (PI), National Science Foundation through NCSA/University of Illinois, Scientific Portals Expedition, 10/2003 - 3/2005, $200,000	NSF passthrough to Alliance
8	Middleware for Grid Portal Development, Plale with Dennis Gannon, Marlon Pierce (PI), and Geoffrey Fox, University of Texas Austin, Argonne National Labs, National Science Foundation, 09/2003 - 08/2006, $868,803 to IU. [IU funds roughly evenly distributed]	NSF #0330613
7	A Research Infrastructure for Collaborative, High-Performance Grid Applications, Plale (co-PI) with David Wise (PI), Andrew Lumsdaine, Geoffrey Fox, Randy Bramley, NSF, 09/2002 - 08/2007, $1,311,875 plus IU match of $803,709. [Equipment grant]	NSF #0202048
6	Crane-IU-Purdue Knowledge Projection for Fleet Maintenance, Plale with Randy Bramley, Geoffrey Fox, David Leake, and Rik McMullen (PI), Crane Naval Warfare Center, 6/15/02 - 9/27/05, $500,000 [funds roughly evenly distributed]	Naval Warfare Center
5	Middleware Technology to Support Science Portals: A Gateway to the Grid, Plale (co-PI) with Dennis Gannon, Department of Energy, 01/2001 – 07/2007, $1,045,141. [Plale .33 portion]	DOE
4	ITR/SY Collaborative Research: A Unified Relational Approach to Grid Information Services, Plale (co-PI) with Northwestern University, National Science Foundation, 09/2001 - 08/2005, IU: $234,702	NSF #0128390
3	Dynamic Querying of Large-scale Streaming Data, Plale (PI), National Science Foundation through NCSA/University of Illinois, $62,000, 10/2001 – 09/2002	NSF
2	POWRE: Applying Database Techniques to Management of Large Data Flows in Interactive Scientific Simulations, Plale (PI), National Science Foundation, $75,005, 1999 - 2002	NSF #9973834
1	Profiling the Performance of SMP Servers, Plale with Karsten Schwan, Intel Corporation, $32,000, 1998. [equipment grant]	Intel

PEER REVIEWED JOURNAL / CONFERENCE PROCEEDINGS and BOOK CHAPTERS (162) _____

	YEAR	REFERENCE
162	2024	Withana, S and B. Plale, Patra ModelCards: AI/ML Accountability in Edge-Cloud Continuum, 2024 IEEE 20th Int'l Conf on eScience (eScience), Osaka, Japan Sep 2024
161		Panda, D. K., V. Chaudhary, E. Fosler-Lussier, R. Machiraju, A. Majumdar, B. Plale, R. Ramnath, P. Sadayappan, N. Savardekar, and K. Tomko. 2024. " Creating intelligent cyberinfrastructure for democratizing AI." *AI Magazine* 45: 22–28. https://doi.org/10.1002/aaai.12166
160	2023	Plale, B., Khan, S., Morales, A.. Democratization of AI: Challenges of AI Cyberinfrastructure and Software Research, *Proc 2023 IEEE 19th Int'l Conf on eScience (eScience)*, IEEE CS Press, 2023. 10.1109/e-Science58273.2023.10254950
159		Withana, S. and B. Plale, Edge AI Distributed Framework, *Proc* 2023 IEEE 19th Int'l Conf on eScience (eScience), IEEE CS Press, 2023 10.1109/e-Science58273.2023.10254827
158		Khan, S., Morales, Alfonso, Plale, Beth. Democratization is a process, not a destination: operationalizing ethics and democratization in a cyberinfrastructure for AI project, *Int'l Conf on AI for People*, 24-26 November 2023, Bologna, Italy
157		Plale B., Malakar P, D'Souza M, Kapoor HK, Simmhan Y, Altinas I, Swaminathan M. (2023), CCGrid2023: A Holistic Approach to Inclusion and Belonging, *IEEE/ACM 23rd Int'l Symp on Cluster, Cloud, and Internet Computing (CCGrid 2023)*, pp. 684-685, DOI: 10.1109/CCGrid57682.2023.00070
156		Lou Y, Plale B (2023) Knowledge Enhanced Digital Objects: a Data Lake Approach, Research Poster. *Proc. IEEE/ACM 23rd Int'l Symp on Cluster, Cloud, and Internet Computing (CCGrid 2023),* DOI: 10.1109/CCGridW59191.2023.00064
155		McKiernan EC, Barba L, Bourne PE, Carter C, Chandler Z, Choudhury S, et al. (2023), Policy recommendations to ensure that research software is openly accessible and reusable. *PLoS Biol* 21(7): e3002204. https://doi.org/10.1371/journal.pbio.3002204
154		Marru, S. Pierce, M. and Plale, B. et al., Cybershuttle: An End-to-End Cyberinfrastructure Continuum to Accelerate Discovery in Science and Engineering, *Proc. Of Practice and Experience in Advanced Research Computing 2023* (PEARC 2023), ACM, pp. 26-34 10.1145/3569951.3593602
153		Islam, S., James Beach, Elizabeth R. Ellwood, Jose Fortes, Larry Lannom, Gil Nelson, Beth Plale (2023), Assessing the FAIR Digital Object Framework for Global Biodiversity Research, *Research Ideas and Outcomes*, RIO, DOI: 10.3897/rio.9.e108808
152	2022	Cutcher-Gershenfeld, J., et al. (2022). Start with the Minimum, *Stanford Social Innovation Review*, Vol. 20, No. 2, Stanford Center on Philanthropy and Civil Society, Spring 2022
151	2021	Plale, B., Tanu Malik, Line Pouchard (2021). Reproducibility Practice in High Performance Computing: Community Survey Results, *in Computing in Science and Engineering*, 23(05), pp. 55-60. DOI: 10.1109/MCSE.2021.3096678
150		Withana, S., Kshitij Mehta, Matthew Wolf, Beth Plale (2021). Towards System for Knowledge Representation of Campaign Experimentation, Proc. *IEEE 17th Int'l Conf on eScience (eScience)*, IEEE CS Press, pp. 257-258, doi: 10.1109/eScience51609.2021.00049
149		Withana, S., Kshitij Mehta, Matthew Wolf, Beth Plale (2021). Campaign Knowledge Network: Building Knowledge for Campaign Efficiency. *arXiv preprint server,* arXiv:2112.03435
148		Kouper, I., Scheidt, Lois A. and Plale, Beth A. (2021). Fostering Interdisciplinary Data Cultures through Early Career Development: The RDA/US Data Share Fellowship. *Data Science Journal*, 20(1), DOI: 10.5334/dsj-2021-002
147	2020	Peng, Z. and Beth Plale (2020). Reliable Access to Massive Restricted Texts: Experience-based Evaluation, *Concurrency and Computation: Practice and Experience*, Wiley, 32:e5225. DOI: 10.1002/cpe.5255

146   2019   Juanillas, V., Alexis Dereeper, Nicolas Beaume, Gaetan Droc, Joshua Dizon, John Robert Mendoza, Jon Peter Perdon, Locedie Mansueto, Lindsay Triplett, Jillian Lang, Gabriel Zhou, Kunalan Ratharanjan, Beth Plale, Jason Haga, Jan E Leach, Manuel Ruiz, Michael Thomson, Nickolai Alexandrov, Pierre Larmande, Tobias Kretzschmar, Ramil P Mauleon (2019), Rice Galaxy: an open resource for plant science, *GigaScience*, 8(5), DOI: 10.1093/gigascience/giz156

145   Plale, B., Eleanor Dickson, Inna Kouper, Samitha Liayanage, Yu Ma, Robert McDonald, John Walsh and Sachith Withana (2019), Safe Open Science for Restricted Data, *in Data and Information Management*, De Gruyter, Vol. 3(1), 50-60. DOI: [10.2478/dim-2019-0005](10.2478/dim-2019-0005)

144   Gil, Y., Suzanne A. Pierce, Hassan Babaie, Arindam Banerjee, Kirk Borne, Gary Bust, Michelle Cheatham, Imme Ebert-Uphoff, Carla Gomes, Mary Hill, John Horel, Leslie Hsu, Jim Kinter, Craig Knoblock, David Krum, Vipin Kumar, Pierre Lermusiaux, Yan Liu, Chris North, Victor Pankratius, Shanan Peters, Beth Plale, Allen Pope, Sai Ravela, Juan Restrepo, Aaron Ridley, Hanan Samet, Shashi Shekhar (2019). Intelligent Systems for Geosciences: An Essential Research Agenda, *Communications of the ACM*, 62(1), pp. 76-84. DOI: [10.1145/3192335](10.1145/3192335)

143   2018   Suriarachchi, I., Sachith Withana, and Beth Plale (2018), Big Provenance Stream Processing for Data Intensive Computations, *Proc. 13th IEEE Int'l Conf on e-Science*, IEEE Computer Society

142   Withana, S., Inna Kouper, Beth Plale (2018). Data Capsule for Restricted Data in Libraries, Extended Abstract. Workshop on Cyberinfrastructure and Machine Learning for Digital Libraries and Archives, *Proc. Joint Conference on Digital Libraries,* ACM

141   2017   Elag, M., Praveen Kumar, Luigi Marini, James D. Myers, Margaret Hedstrom, and Beth Plale (2017). Identification and characterization of information-networks in long-tail data collections, *Environmental Modelling & Software*, Elsevier, vol. 94, pp. 100-111. DOI:10.1016/j.envsoft.2017.03.032

140   Plale, B.,. and Inna Kouper (2017). The Centrality of Data: Data Lifecycle and Data Pipelines, In *Data Analytics in Intelligent Transportation Systems, M.A. Chowdhury, A. Apon, and K. Dey (Eds.)* Elsevier Inc. Cambridge, MA. DOI:10.1016/B978-0-12-809715-1.00004-3

139   Ruan, G., Paul Hanson, Hilary A. Dugan, Beth Plale (2017). Mining Lake time series using symbolic representation, *Ecological Informatics*, Elsevier, vol. 39, pp. 10-22. DOI:10.1016/j.econif.2017.03.001

138   Murdock, J., Jacob Jett, Tim Cole, Yu Ma, J. Stephen Downie, Beth Plale (2017). Towards Publishing Secure Capsule-based Analysis (short paper), *Proc. 17th ACM-IEEE-CS Joint Conference on Digital Libraries*, ACM

137   2016   Stephen Downie, J., Mike Furlough, Robert McDonald, Beth Namachchivaya, Beth A. Plale, John Unsworth (2016). The HathiTrust Research Center: Exploring the Full-Text Frontier, *Educause Review*, 51(3).

136   Zeng, J. and Beth Plale (2016). Argus: A Multi-tenancy NoSQL store with workload-award resource reservation, *Parallel Computing*, Elsevier, vol. 58, pp. 76-89. DOI:10.1016/j.parco.2016.06.003

135   Kowalczyk, S., Y. Sun, Z. Peng, B. Plale, C. Willis, J. Zeng, M. Pathirage, S. Liyanage, A. Todd, G. Ruan (2016). Big Data at Scale for Digital Humanities: An Architecture for the HathiTrust Research Center. In *Big Data: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications*, IGI Global. DOI: 10.4018/978-1-4666-4699-5.ch011

134   Plale, B., Inna Kouper, Allison Goodwell, and Isuru Suriarachchi (2016). Trust Threads: Minimal Provenance for Data Publishing and Reuse, *Big Data is Not a Monolith: Policies, Practices and Problems, Cassidy R. Sugimoto, Hamid Ekbia, and Michael Mattioli, Eds.*, MIT Press

133   Chen, P., Tom Evans, Michael Frisby, Eduardo Izquierdo, and Beth Plale (2016), A Hybrid Approach to Population Construction for Agricultural Agent-Based Simulation, *Proc. 11th IEEE Int'l Conf. on e-Science*, IEEE Computer Society, Baltimore, MD

132		Suriarachchi, I., and Beth Plale (2016), Crossing Analytics Systems: A Case for Integrated Provenance in Data Lakes, *Proc. IEEE 12th Int'l Conference on e-Science*, IEEE Computer Society
131		Ruan, G., and Beth Plale (2016). Horme: Random Access Big Data Analytics, *Proc. IEEE Conf on Cluster Computing*, IEEE Computer Society, Taipei, Taiwan, Sep 2016
130		Suriarachchi, I., and Beth Plale (2016). Provenance as Essential Infrastructure for Data Lakes, Short Paper. *Proc. 6th Int'l Provenance and Annotation Workshop (IPAW)*
129		Chen, P., Tom Evans, and Beth Plale (2016), Analysis of Memory Constrained Live Provenance, *Proc. 6th Int'l Provenance and Annotation Workshop (IPAW)*
128		Pathirage, M., Julian Hyde, Yi Pan, and Beth Plale (2016). SamzaSQL: Scalable Fast Data Management with Streaming SQL, IEEE Int'l Workshop on High-Performance Big Data Computing, *Proc. IEEE Int'l Parallel and Distributed Processing Symp Workshops (IPDPSW 2016)*
127		Zeng, J. and Beth Plale (2016). KVLight: A Lightweight Key-Value Store for Distributed Access in Cloud, *Proc. 16th IEEE/ACM Int'l Symp on Cluster, Cloud and Grid Computing (CCGrid),* Cartagena, Columbia, May 2016 (20% acceptance)
126	2015	Provenance and Annotation of Data and Processes: 5th International Provenance and Annotation Workshop, IPAW 2014 (2015). Cologne, Germany, June 9-13, 2014. Revised selected papers, Bertram Ludaescher and Beth Plale, Eds, Lecture Notes in Computer Science 8628, Springer, 2015
125		Suriarachchi, I., Q. Zhou and B. Plale (2015). Komadu: A Capture and Visualization System for Scientific Data Provenance. *Journal of Open Research Software* 3(1):e4, DOI: http://dx.doi.org/10.5334/jors.bq
124		Zeng, J., and Beth Plale (2015). Workload-aware Resource Reservation for Multi-Tenant NoSQL, *Proc. IEEE Cluster*, pp. 32-41, DOI:10.1109/CLUSTER.2015.14. Best paper candidate.
123		Zeng, J., Beth Plale (2015). Short paper: Towards Building a Lightweight Key-Value Store on Parallel File System, *Proc. 17th IEEE Int'l Conference on Cluster (CLUSTER)*
122		Peng Chen, P., and Beth Plale (2015). ProvErr: System Level Statistical Fault Diagnosis using Dependency Model, *Proc. 15th IEEE/ACM Int'l Symp on Cluster, Cloud and Grid Computing*, IEEE, pp. 525-534 10.1109/CCGrid.2015.86 (25.7% acceptance)
121	2014	Cheah You-Wei., and Beth Plale (2014). Provenance Quality Assessment Methodology and Framework, *Journal of Data and Information Quality, Special issue on Provenance, Data and Information Quality*, ACM, Vol 5(3).
120		Plale, B., (2014). Synthesis of Working Group and Interest Group Activity One Year into the Research Data Alliance, *D-Lib Magazine* (2014) DOI 10.1045/january2014-plale
119		Luo, Y., Beth Plale, Zhenhua Guo, Wilfred W. Li, Judy Qiu, Yiming Sun (2014). Hierarchical MapReduce: Towards Simplified Cross-Domain Processing, *Concurrency and Computation: Practice and Experience*, Wiley, Vol 26(4), pp. 878-893
118		Chakraborty, A., Milinda Pathirage, Isuru Suriarachchi, Kavitha Chandrasekar, Craig Mattocks, Beth Plale (2014). Executing Storm Surge Ensembles on PAAS Cloud, in *Cloud Computing for Data-Intensive Applications*, X. Li and J. Qiu, Eds. Springer, pp. 257-276.
117		Ruan, G., Hui Zhang, Beth Plale (2014). Parallel and Quantitative Sequential Pattern Mining for Large-Scale Interval-based Temporal Data, *Proc. 5th SC Workshop on Big Data Analytics: Challenges and Opportunities*
116		Zhou, Q. Devarshi Ghoshal, Beth Plale (2014). Study of Usefulness of Middleware-Only Provenance, *Proc. IEEE 10th Int'l Conf on e-Science*, IEEE, pp. 215-222, 10.1109/eScience.2014.49
115		Peng, Z., Miao Chen, Stacy Kowalczyk, Beth Plale (2014). Short Paper: Author Gender Metadata Augmentation of HathiTrust Digital Library, Connecting Collections, Cultures, and Communities, *77th Ann. Meeting of American Society for Information Science and Technology (ASIS&T)*

114		Ruan, G., Hui Zhang, Beth Plale (2014). Parallel and Quantitative Sequential Pattern Mining for Large-scale Interval-based Temporal Data, Workshop in Advances in Software and Hardware for Big Data to Knowledge Discovery, *Proc. IEEE Int'l Conf on Big Data*, IEEE, pp. 32-39
113		Zeng, J. and Beth Plale (2014). Multi-tenant Fair Share in No-SQL Data Stores, *Proc. IEEE Cluster* (23.5% acceptance)
112		Zeng, J., Guangchen Ruan, Alexander Crowell, Atul Prakash, Beth Plale (2014). Cloud Computing Data Capsules for Non-consumptive Use of Texts, 5th Workshop on Scientific Cloud Computing, *Proc. ACM High Performance Distributed Computing (HPDC)*. Best paper runner up.
111		Ruan, G., H. Zhang, Eric Wernert, Beth Plale (2014). TextRWeb: Large-Scale Text Analytics with R on the Web. *Proc. XSEDE '14*. ACM.
110		Ghoshal, D., Arun Chauhan, Beth Plale (2014). Regenerating and Quantifying Quality of Benchmarking Data using Static and Dynamic Provenance, *Proc. Int'l Provenance and Annotation Workshop (IPAW 2014)*, Lecture Notes in Computer Science, Vol. 8628, pp 56-67
109		Elag, M., Praveen Kumar, Margaret Hedstrom, James Myers, Beth Plale, Luigi Marini,and Robert McDonald (2014). Characterization of Emergent Data Networks Among Long-Tail Data, *European Geosciences Union General Assembly 2014*
108	2013	Jensen, S., B. Plale, M. Aktas, Y. Luo, P. Chen, and H. Conover (2013). Provenance Capture and Use in a Satellite Data Processing Pipeline, *IEEE Trans. on Geoscience and Remote Sensing*, special issue on Data Provenance, (51)11, pp. 5090-5097. DOI 10.1109/TGRS.2013.2266929
107		Aktas, M., Beth Plale, David Leake, Nirmal K. Mukhi (2013). Unmanaged Workflows: Their Provenance and Use, Chapter 3, *Data Provenance and Data Management in eScience, Studies in Computational Intelligence series*, Q. Bai, Q. Liu eds., Springer, Vol 426, pp. 59-81.
106		Chen, P., Beth Plale, Mehmet Aktas (2013). Temporal Representation for Mining Scientific Data Provenance, *Future Generation of Computer Systems*, Elsevier, Vol 36, pp. 363-378
105		Plale, B., Robert H. McDonald, Kavitha Chandrasekar, Inna Kouper, Stacy Konkiel, Margaret Hedstrom, James Myers, Praveen Kumar (2013). SEAD Virtual Archive: Building a Federation of Institutional Repositories for Long-Term Data Preservation in Sustainability Science, *Int'l Journal of Digital Curation*, 8(2), pp 172-180. DOI 10.2218/ijdc.v8i2.281
104		Chen, P., Beth Plale, and Tom Evans (2013). Dependency Provenance in Agent Based Modeling, *9th IEEE Int'l Conf on e-Science*, IEEE Computer Society, Beijing, China, Oct 2013.
103		Zeng, J. and Beth Plale (2013). Data pipeline in MapReduce, *Proc. 9th IEEE Int'l Conf on e-Science*, IEEE Computer Society, Beijing, China, Oct 2013.
102		Rahnemoonfar M., and Beth Plale (2013). Automatic Performance Evaluation of Dewarping Methods in Large Scale Digitization of Historical Documents, *Proc. Joint Conf on Digital Libraries (JCDL)* ACM, New York, NY. pp 331-334, DOI: 10.1145/2467696.2467744
101		Chen, M., Uma Pavalanathan, Scott Jensen, and Beth Plale (2013). Modeling Heterogeneous Data Resources for Social-Ecological Research: A Data-Centric Perspective, *Joint Conf on Digital Libraries (JCDL)*, ACM, New York, NY
100		Chakraborty, A., Milinda Pathirage, Isuru Suriarachchi, Kavitha Chandrasekar, Craig Mattocks and Beth Plale (2013). Storm Surge Simulation and Load Balancing in Azure Cloud, *Proc. High Performance Computing Symp,* Society for Computer Simulation International
99		Ruan, G., Hui Zhang, and Beth Plale (2013). Exploiting MapReduce and Data Compression for Data-intensive Applications, *Proc. XSEDE 2013*, ACM.
98		Cheah, Y-W., Richard Canon, Beth Plale, and Lavanya Ramakrishnan (2013). Milieu: Lightweight and Configurable Big Data Provenance for Science, *Proc. IEEE Second Int'l Congress on Big Data (BigData 2013)*

97		Ghoshal, D., and Beth Plale (2013). Provenance from Log Files: a BigData Problem, 1st Int'l Workshop on Managing and Querying Provenance Data at Scale, *Proc. 16th Conf on Extending Database Technology (EDBT),* ACM
96		Plale, B., P. Kumar, J. Myers, M. Hedstrom, R. McDonald, S. Konkiel, and K. Chandrasekar (2013). SEAD Virtual Archive: Building a Federation of Institutional Repositories for Long Term Data Preservation, *Proc. 8th Int'l Digital Curation Conf*, http://hdl.handle.net/2022/15247
95	2012	Withana, E.C. and Beth Plale (2012). Sigiri: Uniform Resource Abstraction for Grids and Clouds, *Concurrency and Computation, Practice and Experience*, Vol 24, Issue 18, pp. 2362-2380
94		Chen, P., Beth Plale, You-Wei Cheah, Devarshi Ghoshal, Scott Jensen, and Yuan Luo (2012). Visualization of Network Provenance Data, Workshop on Massive Data Analytics on Scalable Systems (DataMASS), *Proc. High Performance Computing Conf,* IEEE Computer Society
93		Cheah, You-Wei, and Beth Plale (2012). Provenance Analysis: Towards Quality Provenance, 8th IEEE Int'l Conf on e-Science, IEEE Computer Society 2012, Chicago, IL.
92		Plale, B. (2012). Managing the long tail of science: data and communities. *Proc. 1st Conf of Extreme Science and Engineering Discovery Environment (XSEDE'12),* ACM, 10.1145/2335755.2335866
91		Chen, P., and Beth Plale (2012). Poster Abstract: Visualizing Large-scale Data Provenance, *IEEE/ACM Supercomputing*, Salt Lake City, pp. 1385-1386 10.1109/SC.Companion.2012.205
90		Jensen, S., Beth Plale, Xiaozhong Liu, Miao Chen, David Leake, and Julie England (2012). Generalized Representation and Mapping for Social-Ecological Data: Freeing Data from the Database, *Proc. 8th IEEE Int'l Conf on e-Science,* IEEE Computer Society
89		Chen, P., Beth Plale, and Mehmet S. Aktas (2012). Temporal Representation for Scientific Data Provenance, *Proc. 8th IEEE Int'l Conf on e-Science,* IEEE Computer Society, 10.1109/eScience.2012.6404477
88		Beth Plale, Eran Chinthaka Withana, Chathura Herath, Kavitha Chandrasekar, and Yuan Luo (2012). Effectiveness of Hybrid Workflow Systems for Computational Science, *Proc. Int'l Conf on Computational Science (ICCS),* Procedia Computer Science, Elsevier, Vol 9, pp. 508-517
87		Chandrasekar, K., Milinda Pathirage, Samindra Wijeratne, Craig Mattocks, and Beth Plale (2012). Middleware Alternatives for Storm Surge Predictions in Windows Azure, *3rd Workshop on Scientific Cloud Computing*, ACM NY, NY pp 3-12, 10.1145/2287036.2287040
86		Jensen, S., Miao Chen, Xiaozhong Lin, Beth Plale, David Leake (2012). Short Paper: Mining Classifications from Social-Ecological Databases, *75th Annual Meeting of American Society for Information Science and Technology (ASIS&T)*
85		Chen, M. and Beth Plale (2012). Short Paper: From Metadata to Ontology Representation: A Case of Converting Severe Weather Forecast Metadata to an Ontology, *75th Annual Meeting of American Society for Information Science and Technology (ASIS&T)*
84	2011	Plale, B., Bin Cao, Chathura Herath, and Yiming Sun (2011). Data Provenance for Preservation of Digital Geoscience Data, Societal Challenges and Geoinformatics, A. Krishna Sinha, David Arctur, Ian Jackson, and Linda Gundersen, Eds., in *Geological Society of America (GSA), Special Paper 482*, 01 Nov 2011, ISBN 9780813724829
83		Simmhan, Y., and Beth Plale (2011). Using Provenance for Personalized Quality Ranking of Scientific Datasets, Artem Chebotko, Yogesh Simmhan and Paolo Missier, eds. *Int'l Journal of Computers and Their Applications: Special Issue on Scientific Workflows, Provenance and Their Applications*, 18(3), pp. 180-196
82		Luc Moreau, Ben Clifford, Juliana Freire, Joe Futrelle, Yolanda Gil, Paul Groth, Natalia Kwasnikowska, Simon Miles, Paolo Missier, James Myers, Beth Plale, Yogesh Simmhan, Eric Stephan, Jan Van den Bussche, The Open Provenance Model Core Specification

		(V1.1), *Future Gener. Comput. Syst.*, Elsevier Science Publishers B. V., Amsterdam Netherlands, Vol. 27, No. 6, June, 2011, pp. 743-756, DOI:10.1016/j.future.2010.07.005
81		Crippa, P., G. El Afandi, B.A. Plale, and S.C. Pryor (2011). Short Paper: Understanding the effects of boundary layer and synoptic meteorology on new particle formation based on WRF simulations and measurements in Southern Indiana, *American Geophysical Union Fall 2011 meeting*
80		Herath, C., and Beth Plale (2011). Programming abstraction for resource aware stream processing for scientific workflows, D3Science Workshop, *Proc. 9th IEEE Int'l Conf on e-Science*, Dec 201
79		Jensen, S., M. Cox, D. Bender, M. Chen, J. England, B. Plale, and D. Leake (2011). Spatial Data in an Ontology for Research on Forest Resources, *COSIT11 Workshop Ontology of Spatial Thinking and Reasoning: Multidisciplinary Reconciliation*
78		Cheah, Y-W., Beth Plale, Joey Kendall-Morwick, David Leake, and Lavanya Ramakrishnan (2011). A Noisy 10GB Provenance Database, 2nd Int'l Workshop on Traceability and Compliance of Semi-Structured Processes (TC4SP2011), *Proc. Business Process Management (BPM 2011)*
77		Luo, Y., Zenhua Guo, Yiming Sun, Beth Plale, Judy Qiu, and Wilfred W. Li (2011). Hierarchical Framework for Cross-domain MapReduce Execution, Workshop on Emerging Computational Methods for Life Sciences, *Proc. ACM High Performance Distributed Computing*
76		Plale, B. (2011). Challenges and Opportunities of Workflow Systems in Environmental Research, invited, Water Information Research and Development Alliance (WIRADA) Science Symposium, Melbourne, AU, Aug 2011
75		Guo, D., Beth Plale, Leon Welicki, and Eran Chinthaka (2011). Scientific Workflow Challenges, *Water Information Research and Development Alliance (WIRADA) Science Symposium*, Melbourne, AU, Aug 2011
74		Withana, E.C., Beth Plale, and Craig Mattocks (2011). Towards Enabling Mid-Scale Geoscience Experiments Through Microsoft Trident and Windows Azure, *Microsoft Cloud Futures Workshop*, Jun 2011
73	2010	Katz, D.S., Scott Callaghan, Robert Harkness, Shantenu Jha, Krzysztof Kurowski, Steven Manos, Sudhakar Pamidighantam, Marlon Pierce, Beth Plale, Carol Song, and John Towns (2010). Science on the Teragrid, *Computational Methods in Science and Technology*, Special Issue, Polish Academy of Sciences, pp. 81-97.
72		Jensen, S. and Beth Plale (2010). Trading Consistency for Scalability in Scientific Metadata, *Proc. 6th Int'l IEEE Conf on e-Science,* IEEE Computer Society Press
71		Withana, E.C. and Beth Plale (2010). Usage Patterns to Provision for Time Critical Scientific Experimentation in Clouds, *Proc. 2nd IEEE Int'l Conf on Cloud Computing Technology and Science (CloudCom 2010)*
70		Herath, C. and Beth Plale (2010). Streamflow - Programming Model for Data Streaming in Scientific Workflows, *Proc. 10th IEEE/ACM Int'l Symp on Cluster, Cloud, and Grid Computing (CCGrid 2010),* Melbourne Australia
69		Ramakrishnan, L., Dennis Gannon, and Beth Plale (2010). WORKEM: Representing and Emulating Distributed Scientific Workflow Execution State, *Proc. 10th IEEE/ACM Int'l Symp on Cluster, Cloud and Grid Computing (CCGrid 2010)*, Melbourne Australia 10.1109/CCGRID.2010.89
68		Ramakrishnan, L. and Beth Plale (2010). Multidimensional Classification Model for Scientific Workflow Characteristics, 1st Int'l Workshop on Workflow Approaches for New Data-Centric Science, *Proc. ACM SIGMOD International Conf on Management of Data*
67		Withana, E.C., Beth Plale, Roger Barga, and Nelson Araujo (2010). Versioning for Workflow Evolution, Data Intensive Distributed Computing Workshop (DIDC), *Proc. High Performance Distributed Computing*

66	2009	Cao, B., Beth Plale, Girish Subramanian, Paolo Missier, Carole Goble, and Yogesh Simmhan (2009). Semantically Annotated Provenance in the Life Science Grid, 1st Int'l Workshop on the role of Semantic Web in Provenance Management, *Proc. 8th Int'l Semantic Web Conference*, Washington D.C.
65		Cao, B., Beth Plale, Girish Subramanian, Ed Robertson, and Yogesh Simmhan (2009). Provenance Information Model of Karma Version 3, IEEE 2009 3rd Int'l Workshop on Scientific Workflows (SWF'09), *Proc. 2009 Congress on Services*, IEEE Computer Society, pp. 348-351 http://doi.ieeecomputersociety.org/10.1109/SERVICES-I.2009.54
64		Withana, E.C., Jalyia Ekanayke, David Leake, and Beth Plale (2009). CBR Based Workflow Composition Assistant, IEEE 2009 3rd Int'l Workshop on Scientific Workflows (SWF'09), *Proc. 2009 Congress on Services*, IEEE Computer Society, pp. 348-351 http://doi.ieeecomputersociety.org/10.1109/SERVICES-I.2009.51
63		Perera, S., Suresh Marru, Thilina Gunarathne, Dennis Gannon, and Beth Plale (2009). *Application of Management Frameworks: A Case Study on Managing Workflow-related Systems*, *Proc. IEEE Int'l Conf on Web Services (ICWS),* DOI 10.1109/ICWS.2009.52
62		Baker, K.M., Beth Plale, Ilya Zaslavsky, and Suresh Marru (2009). Towards Cyberinfrastructure for Multi-scale Crop Disease Early Warning Systems, *Proc. World Congress on Computers in Agriculture*
61		Plale, B., Dennis Groth, Bin Cao, Girish Subramanian, Carole Goble, and Paolo Missier (2009). Short Paper: Knowledge Discovery through Provenance Collection, Representation, and Use in the Life Science Grid, *Proc. 6th Int'l Conf. on Data Integration in the Life Sciences (DILS)*, Manchester, UK, 2009
60	2008	Vijayakumar, N. and Beth Plale (2008). Missing Event Prediction in Sensor Data Streams Using Kalman Filters. Book chapter in *Knowledge Discovery from Sensor Data,* Eds. A. R. Ganguly, J. Gama, O. A. Omitaomu, M. Gaber and R. R. Vatsavai, Taylor and Francis/CRC Press pp. 149-170
59		Simmhan, Y., Beth Plale, and Dennis Gannon (2008). Query Capabilities of the Karma Provenance Framework, *Concurrency and Computation: Practice and Experience*, John Wiley and Sons, Vol 20(5), pp. 441-451
58		Li, X., Beth Plale, Nithya Vijayakumar, Rahul Ramachandran, Sara Graves, and Helen Conover (2008). Real-time Storm Detection and Weather Forecast Activation through Data Mining and Events Processing, *Earth Science Informatics*, H.A. Babaie, Ed., Springer Berlin/Heidelberg, Vol 1(2), pp. 49-57. DOI 10.1007/s12145-008-0010-7
57		Simmhan, Y., B. Plale, and D. Gannon (2008). Karma2: Provenance Management for Data Driven Workflows, Extended and invited from ICWS 2006. *Int'l Journal of Web Services Research*, IGI Publishing, Vol 5 (2), pp. 1-22
56		Jensen, S. and B. Plale (2008). Schema-Independent and Schema-Friendly Scientific Metadata Management, *Proc. 4th Int'l IEEE Conf on e-Science,* IEEE Computer Society, pp. 428-429, http://doi.ieeecomputersociety.org/10.1109/eScience.2008.114
55		The Open Provenance Model (v1.01). Moreau, L. (Editor), B. Plale, S. Miles, C. Goble, P. Missier, R. Barga, Y. Simmhan, J. Futrelle, R. McGrath, J. Myers, P. Paulson, S. Bowers, B. Ludaescher, N. Kwasnikowska, J. Van den Bussche, T. Ellkvist, J. Frieire, P. Growth (2008). *Technical Report, Electronics and Computer Science, University of Southampton.* http://eprints.ecs.soton.ac.uk/16148
54		Sun, Y., S. Marru, and B. Plale (2008). Experience with Bursty Workflow-driven Workloads in LEAD Science Gateway, *Proc. 3rd Annual Teragrid Conference*
53		Jensen, S. and B. Plale (2008). Using Characteristics of Computational Science Schemas for Workflow Metadata Management, *Proc. 2008 IEEE Congress on Services*, Computer Society Press, Washington DC, http://doi.ieeecomputersociety.org/10.1109/SERVICES-1.2008.42
52	2007	Simmhan, Y., S.L. Pallickara, N. Vijayakumar, and B. Plale (2007). Data Management in Dynamic Environment-driven Computational Science. *Book chapter in Grid-Based Problem*

Solving Environments: Implications for Development and Deployment of Numerical Software, IFIP Int'l Federation for Information Processing Vol. 239, P.W. Gaffney and J.C.T. Pool, Eds, Springer Boston, pp. 317-333, DOI: 10.1007/978-0-387-73659-4_17

51		Gannon, D., B. Plale, M. Christie, Y. Huang, S. Jensen, N. Liu, S. Marru, S.L. Pallickara, S. Perera, S. Shirasuna, Y. Simmhan, A. Slominski, B. Plale, and D. Reed (2007). Service Architectures for e-Science Grid Gateways: Opportunities and Challenges, in *Lecture Notes in Computer Science Vol 4804*, pp. 1179-1185, DOI 10.1007/978-3-540-76843-2
50		Gannon, D., B. Plale, M. Christie, Y. Huang, S. Jensen, N. Liu, S. Marru, S.L. Pallickara, S. Perera, S. Shirisuna, Y. Simmhan, A. Slominski, Y. Sun, N. Vijayakumar (2007). Building Grid Portals for e-Science: A Service Oriented Architecture, Vol 16: *High Performance Computing and Grids in Action*, IOS Press - Amsterdam, Lucio Grandinetti editor
49		Ramakrishnan, L., Y. Simmhan, and B. Plale (2007). Realization of Dynamically Adaptive Weather Analysis and Forecasting in LEAD: Four Years Down the Road, *Dynamic Data-Driven Application Systems Workshop*, in *Proc International Conference on Computational Science (ICCS)*
48		Gannon, D., B. Plale, S. Marru, G. Kandaswamy, Y. Simmhan, and S. Shirasuna (2007). Dynamic, Adaptive Workflows for Mesoscale Meteorology. *Book chapter in Workflows for e-Science: Scientific Workflows for Grids*, Taylor, I.J.; Deelman, E.; Gannon, D.B.; Shields, M. (Eds.) Springer, pp. 129-145
47		Vijayakumar N., and B. Plale (2007). Tracking Stream Provenance in Complex Event Processing Systems for Workflow-driven Computing. *Proc. 2nd Int'l Workshop on Event-driven Architecture, Processing, and Systems (EDA-PS'07)*, in conjunction with VLDB'07, VLDB.
46		Vijayakumar, N. and B. Plale (2007). Prediction of Missing Events in Sensor Data Streams Using Kalman Filters, *Proc. 1st Int'l Workshop on Knowledge Discovery from Sensor Data*, in conjunction with ACM 13th Int'l Conf on Knowledge Discovery and Data Mining (KDD), ACM
45	2006	Plale, B., D. Gannon, J. Brotzge, K. Droegemeier, J. Kurose, D. McLaughlin, R. Wilhelmson, S. Graves, M. Ramamurthy, R. Clark, S. Yalda, D. Reed, E. Joseph, and V. Chandrasekar (2006). CASA and LEAD: Adaptive Cyberinfrastructure for Real-Time Multiscale Weather Forecasting, *Computer*, IEEE Computer Science Press, Vol. 39(11), pp. 56-63, http://doi.ieeecomputersociety.org/10.1109/MC.2006.375
44		Sun, Y., S. Jensen, S.L. Pallickara, and B. Plale (2006). Personal Workspace for Large-scale Data-driven Computational Experimentation, *Proc. 7th IEEE/ACM Int'l Conf on Grid Computing (Grid'06)*, IEEE Computer Society, pp. 112-119. DOI 10.1.1.142.4780
43		Liu, Y., N. N. Vijayakumar, and B. Plale (2006). Stream Processing in Data-driven Computational Science *Proc. 7th IEEE/ACM Int'l Conf on Grid Computing (Grid'06)* DOI 10.1109/ICGRID.2006.311011
42		Vijayakumar, N.N., B. Plale, R. Ramachandran, and X. Li (2006). Dynamic Filtering and Mining Triggers in Mesoscale Meteorology Forecasting, *Proc. IEEE Int'l Geoscience and Remote Sensing Symp (IGARSS'06)*, Denver, CO, Aug 2006.
41		Simmhan, Y., B. Plale and D. Gannon (2006). Towards a Quality Model for Effective Data Selection in Collaboratories, *Proc. IEEE Workshop on Workflow and Data Flow for Scientific Applications (SciFlow06)*, held in conjunction with ICDE, Atlanta, GA, April 2006.
40		Simmhan, Y., B. Plale, and D. Gannon (2006). A Framework for Collecting Provenance in Data-Centric Scientific Workflows, *Proc. IEEE Int'l Conf on Web Services (ICWS'06),* IEEE Computer Society Press, pp. 427-436, DOI 10.1109/ICWS.2006.5
39		Simmhan, Y., B. Plale, and D. Gannon (2006). Performance Evaluation of the Karma Provenance Framework for Scientific Workflows, *Provenance and Annotation of Data, Lecture Notes in Computer Science*, Springer Berlin/Heidelberg, Vol 4145, 2006. DOI 10.1007/11890850

38		Liu, Y. and B. Plale (2006). Multi-model Based Optimization for Stream Query Processing, *Proc. KSI 18th Int'l Conf on Software Engineering and Knowledge Engineering (SEKE'06)*, DOI 10.1.1.85.4855
37		Liu, Y and B. Plale (2006). Query Optimization for Distributed Data Streams, *Proc. ISCA 15th Int'l Conf on Software Engineering and Data Engineering (SEDE'06)* DOI 10.1.1.133.7271
36		Pallickara, S.L., B. Plale, L. Fang, and D. Gannon (2006). Trust Cell: Towards the End-to-End Trust in Data-Oriented Scientific Computing, *Proc. IEEE Cluster Computing and Grid (CCGrid),* IEEE Computer Society
35		Pallickara S.L. and B. Plale (2006). Enabling End-to-End Trustworthiness in Data-Oriented Scientific Computing, *Proc. Int'l Conf Workshops on Parallel Processing (ICPPW),* IEEE Computer Society, http://doi.ieeecomputersociety.org/10.1109/ICPPW.2006.76
34		Jensen, S., B. Plale, S.L. Pallickara, and Y. Sun (2006). A Hybrid XML-Relational Grid Metadata Catalog, *Proc. Int'l Conf Workshops on Parallel Processing (ICPPW),* IEEE Computer Society, pp. 15-24 http://doi.ieeecomputersociety.org/10.1109/ICPPW.2006.10
33		Vijayakumar N. and B. Plale (2006). Towards Low Overhead Provenance Tracking in Near Real-Time Stream Filtering, *Provenance and Annotation of Data, Lecture Notes in Computer Science,* Springer Berlin/Heidelberg, Vol 4145, 2006. DOI 10.1007/11890850
32		Siek, K.A., K. Connelly, A. Stephano, S. Menzel, J. Bauer, B. Plale (2006). Breaking the Geek Myth: Addressing Young Women's Misperceptions about Technology Careers, *Learning & Leading with Technology*, v. 33 no. 7 p19—22
31	2005	K. Droegemeier, K. Brewster, M. Xue, D. Weber, D. Gannon, B. Plale, D. Reed, L. Ramakrishnan, J. Alameda, R. Wilhelmson, T. Baltzer, B. Domenico, D. Murray, A. Wilson, R. Clark, S. Yalda, S. Graves, R. Ramachandran, J. Rushing, E. Joseph (2005). Service-Oriented Environments for Dynamically Interacting with Mesoscale Weather, *Computing in Science and Engineering*, IEEE Computer Society Press and American Institute of Physics, Vol. 7(6), pp. 12-29
30		Plale, B., D. Gannon, Y. Huang, G. Kandaswamy, S. Pallickara, and A. Slominski (2005). Cooperating Services for Managing Data Driven Computational Experimentation, *Computing in Science and Engineering,* IEEE Computer Society Press and American Institute of Physics, Vol. 7(5), pp. 34 – 43. http://doi.ieeecomputersociety.org/10.1109/MCSE.2005.91
29		Simmhan, Y., B. Plale, and D. Gannon (2005). A Survey of Data Provenance in e-Science, *ACM SIGMOD Record*, ACM Press, Vol. 34(3), pp. 31-36
28		Gannon, D., B. Plale, M. Christie, L. Fang,Y. Huang, S. Jensen, G. Kandaswamy, S. Marru, S. Lee Pallickara, S. Shirasuna, Y. Simmhan, A. Slominski, and Y. Sun (2005). Service Oriented Architectures for Science Gateways on Grid Systems, *Int'l Conf. on Service Oriented Computing 2005*, B. Benatallah, F. Casati, P. Traverso (Eds.), Lecture Notes in Computer Science 3826, Springer-Verlag Berlin Heidelberg pp. 21-32
27		Gannon, D., J. Alameda, O. Chipara, M. Christie, V. Dukle, L. Fang, M. Farellee, G. Fox, S. Hampton, G. Kandaswamy, D. Kodeboyina, C. Moad, M. Pierce, B. Plale, A. Rossi, Y. Simmhan, A. Sarangi, A. Slominski, S. Shirasauna, T. Thomas (2005). Building Grid Portal Applications from a Web-Service Component Architecture, *Proceedings of the IEEE*, IEEE Press, Vol. 93, No. 3, pp. 551-563
26		Plale, B., D. Gannon, J. Alameda, B. Wilhelmson, S. Hampton, A. Rossi, and K. Droegemeier (2005). Active Management of Scientific Data, *IEEE Internet Computing special issue on Internet Access to Scientific Data*, IEEE Computer Science Press, Vol. 9(1), pp. 27-34
25		Pallickara, S. L., B. Plale, S. Jensen, and Y. Sun (2005). Structure, Sharing, and Preservation of Scientific Experiment Data, *Proc. IEEE 3rd Int'l Workshop on Challenges of Large Applications in Distributed Environments (CLADE'05),* IEEE Computer Society Press
24		Plale, B., Gannon, D., Reed, D., Graves, S., Droegemeier, K., Wilhelmson, B., and Ramamurthy, M. (2005). Towards Dynamically Adaptive Weather Analysis and Forecasting in LEAD, *Proc. Computational Science - ICCS Workshop on Dynamic Data Driven Applications,*

Lecture Notes in Computer Science (LNCS), No. 3515, Part II, Springer-Verlag GmBH, pp. 624 – 631

23		Pallickara, S. L., B. Plale, S. Jensen, Y. Sun (2005). Monitoring Access to Stateful Resources in Grid Environments, *Proc. IEEE Int'l Conference on Services Computing (SCC'05)*, IEEE Computer Society Press
22		Vijayakumar, N. and Beth Plale (2005). Evaluation of Rate-based Adaptivity in Asynchronous Data Streams, *Proc. ACM/IEEE 19th Int'l Parallel and Distributed Processing Symposium (IPDPS)*, IEEE Computer Society Press, p. 69b, http://dx.doi.org/10.1109/IPDPS.2005.205
21	2004	Beth Plale (2004). Framework for Bringing Data Streams to the Grid, *Scientific Programming*, IOS Press, Amsterdam, Vol. 12(4), pp. 213-223
20		Plale, B., C. Jacobs, S. Jensen, Y. Liu, C. Moad, R. Parab, and P. Vaidya (2004). Understanding Grid Resource Information Management through a Synthetic Database Benchmark/Workload, *Proc. 4th IEEE/ACM Int'l Symp on Cluster Computing and the Grid (CCGrid2004)*, IEEE Computer Society Press
19		Plale, B. (2004). Using Global Snapshots to Access Data Streams on the Grid (2004). *Proc. 2nd European Across Grids Conf (AxGrids)* published as Lecture Notes in Computer Science, Springer-Verlag GmBH, Vol. 3165, pp. 191-201.
18	2003	Beth Plale and Karsten Schwan (2003). Dynamic Querying of Streaming Data with the dQUOB System, *IEEE Transactions of Parallel and Distributed Systems*, IEEE Computer Science Press, Vol. 14(3), pp. 422-432
17		Kodeboyina, D. and B. Plale (2003). Experiences with OGSA-DAI: Portlet Access and Benchmark, *Global Grid Forum Workshop on Designing and Building Grid Services*, Chicago, Illinois, http://www-unix.mcs.anl.gov/~keahey/DBGS/DBGS_files/dbgs_papers/kodeboyina.pdf September 2003.
16	2002	Plale, B., G. Turner, and A. Sharma (2002). Real Time Response to Streaming Data on Linux Clusters, *Proc. 3rd Int'l Conf on Linux Clusters: the HPC Revolution*, Linux Clusters Institute, http://www.linuxclustersinstitute.org/Linux-HPC-Revolution/Archive/2002techpapers.html
15		Plale, B. (2002). Leveraging Runtime Knowledge about Event Rates to Improve Memory Utilization in Wide Area Data Stream Filtering. *Proc. 11th IEEE Int'l Symp on High Performance Distributed Computing (HPDC)*, IEEE Computer Society, Washington, DC, p. 171, http://dx.doi.org/10.1109/IPDPS.2001.925038
14		Plale, B., P. Dinda, and G. von Laszewski (2002). Key Concepts and Services of a Grid Information Service, *Proc. 15th ISCA Int'l Parallel and Distributed Computing Systems (PDCS'02)*, Int'l Soc. for Computers and their Applications, pp. 437-442
13	2001	Plale, B., P. Widener, and K. Schwan (2001). Taking the Step from Meta Information to Communication Middleware in Computational Data Streams, *Proc. 10th Heterogeneous Computing Workshop*, IEEE Computer Society Press, Washington, DC, p. 20085b http://dx.doi.org/10.1109/IPDPS.2001.925038, 2001.
12		Plale, B. and Schwan, K. (2001). Optimizations Enabled by Relational Data Model View to Querying Data Streams, *Proc. 15th Int'l Parallel and Distributed Processing Symp (IPDPS)*, IEEE Computer Society Press, Washington, DC, p. 10022a http://dx.doi.org/10.1109/IPDPS.2001.924953
11	2000	Plale, B. and K. Schwan (2000). dQUOB: Managing Large Data Flows by Dynamic Embedded Queries, *Proc. IEEE High Performance Distributed Computing (HPDC'00)*, IEEE Computer Society Press, Washington DC, p. 263 http://dx.doi.org/10.1109/HPDC.2000.868658. Extended version available as Georgia Institute of Technology Technical Report GIT-TR-00-07.

10		Plale, B., G. Eisenhauer, L. K. Daley, P. Widener, and K. Schwan (2000). Fast Heterogeneous Binary Data Interchange for Event-based Monitoring, *Proc. ISCA Int'l Conf on Parallel and Distributed Computing Systems (PDCS)*
9		Oleson, V., K. Schwan, G. Eisenhauer, B. Plale, C. Pu and D. Amin (2000). Operational Information Systems - An Example from the Airline Industry, *Proc. 1st Workshop on Industrial Experiences with Systems Software (WIESS),* USENIX Advanced Computing Systems Association
8	1999	Plale, B., V. Elling, G. Eisenhauer, K. Schwan, D. King, and V. Martin (1999). Realizing Distributed Computational Laboratories, *Int'l Journal of Parallel and Distributed Systems and Networks*, Int'l Assn of Science and Technology for Development (IASTED) Press, Vol. 2(3)
7		Plale, B. and K. Schwan (1999). Run-time Detection in Parallel and Distributed Systems: Application to Safety-Critical Systems, *Proc. 19th IEEE Int'l Conference on Distributed Computing Systems (ICDCS)*, IEEE Computer Society Press, p. 0163, DOI:10.1109/ICDCS.1999.776517
6	1998	Plale, B., G. Eisenhauer, K. Schwan, J. Heiner, V. Martin, and J. Vetter (1998). From Interactive Applications to Distributed Laboratories, *IEEE Concurrency*, IEEE Computer Society Press, Vol. 6(2), pp. 78-90
5		Eisenhauer, G., B. Plale, K. Schwan (1998). DataExchange: High Performance Communications in Distributed Laboratories, *Journal of Parallel Computing*, Elsevier, Vol. 24(12-13), pp. 1713-1733
4	1997	Schroeder (Plale), B., S. Aggarwal, and K. Schwan, K. (1997). Software Approach to Hazard Detection Using On-line Analysis of Safety Constraints, *Proc. 16th IEEE Symp on Reliable Distributed Systems (SRDS)*, IEEE Computer Society Press, October, pp. 80, http://dx.doi.org/10.1109/RELDIS.1997.632801, 1997.
3		Eisenhauer, G., B. Plale-Schroeder, K. Schwan, V. Martin, and J. Vetter (1997). DataExchange: High Performance Communications in Distributed Laboratories, *Proc. IASTED Int'l Conf on Parallel and Distributed Computing and Systems (PDCS),* October 1997.
2	1996	Eisenhauer, G., B. Plale-Schroeder and K. Schwan (1996). From Interactive High Performance Programs to Distributed Laboratories: A Research Agenda, *Proc. IEEE SPDP'96 Workshop on Program Visualization and Instrumentation*, IEEE Computer Society Press.
1	1995	Schroeder (Plale) B. (1995). On-line Monitoring: A Tutorial, *Computer,* IEEE Computer Science Press, Vol. 28(6), pp. 72-78

INVITED TALKS (111) _____

	YEAR	TITLE (* is KEYNOTE)	HOST OR LOCATION
111	2024	* AI Research and Infrastructure in eScience: Proceed with Deliberation (and Enthusiasm), IEEE 20th Int'l Conf on eScience (eScience), Sep 2024	Osaka, Japan
110		* Navigating the Future of Data Driven, Computationally-based Research, Northwestern University Computation and Data Exchange 2024	Northwestern University
109		Opportunities of Shared AI Infrastructure in Chemistry Research, Beilstein Symposium	Rudesheim, Germany
108		NSF Open Research Resources for AI	NSF headquarters
107		Provenance and Trust in Scientific Workflows, Society for Industrial and Applied Math Conf. on Parallel Processing for Scientific Computing (SIAM PP24), Mar 2024	Baltimore, MD

106	2023	AI Research Infrastructure: its Equitable Creation and Use, Purdue Cyberinfrastructure Symposium, West Lafayette, IN, Sep 2023	Purdue University
105		Operationalizing AI Ethics through Model Cards in a Cyberinfrastructure Ecosystem, India Institute of Science Bangalore (IIS Bangalore), May 2023	Bangalore, India
104		Trustworthiness in the Edge-Cloud-HPC Continuum, 1st Int'l Workshop on Conversational AI Interfaces for HPC, co-located with PEARC'23, Portland, OR, Jul 2023	Portland, OR
103		AI Ethics and Model Bias, AI in Health Conference, Regenstrief Institute, Apr 2023	Indianapolis, IN
102		PIDs and Open Science, NISO Plus, National Information Standards Organization, Feb 2023	NISO
101	2022	ICICLE AI Institute: how infrastructure can drive democratized AI, Southeast Asia International Joint-Research and Training Program (SEAIP), Hshinchu City, Taiwan Dec 2022	Hshinchu City, Taiwan
100		ICICLE AI Institute AI Accountability, Summit for AI Institutes Leadership, Asilomar, Dec 2022	Asilomar, CA
99		Detect and Mitigate: role of AI Cyberinfrastructure in FAIRness, Computational and Autonomous Workflows (CAW) 2022, Oak Ridge National Labs, Sep 2022	Oak Ridge Nat'l Laboratories
98		* AI Trustworthiness: Detection and Minimization in AI Cyberinfrastructure, Helmholtz AI, Jun 2022, Dresden, Germany	Dresden Germany
97		Democratizing for Data Driven Innovation, American Meteorological Society Washington Forum, April 2022	Washington DC
96	2021	Pervasive Technology Institute: Shaping the Future of IT innovation for Indiana University, CENTRA 5, Sept 2021	Porto, Portugal
95		Trustworthy Artificial Intelligence and Open Science, Beilstein Open Science Symposium, Beilstein Institute, Oct 2021	Frankfurt Germany
94		Open Science: Trusted Products, Center for Informatics Research in Science and Scholarship, University of Illinois Urbana Champaign, May 2021	UIUC
93	2020	Communicating through Research and Researcher, SC20 Early Career Panel on Communications, International Conference for High Performance Computing, Networking, Storage, and Analysis, Nov 2020	Atlanta, Georgia
92		*Open Science: The Challenge and the Promise, Michigan Institute for Data Science, University of Michigan, Sep 2020	Univ of Michigan
91		Artificial Intelligence Research and NSF: a Conversation, Luddy AI Symposium, Jun 2020	Indiana Univ
90		*Open Science in Biodiversity, the Challenge and the Promise, 4th Digital Data Biodiversity Conference, Jun 2020	Bloomington, IN
89		Knowledge Graphs in AI: Trust and FAIR Reuse, AI Technology, Application, and Innovation for Digital Cities, 2020 Smart City Summit and Expo, Taipei, Taiwan, Jun 2020	Taipei, Taiwan
88		Generalist repositories: NSF policy and perspective, Establishing a FAIR Biomedical Data Ecosystem: The Role of Generalist and Institutional Repositories to Enhance Data Discoverability and Reuse, National Institutes of Health, Bethesda, MD, Feb 2020	National Institutes of Health (NIH)
87		*Open Science Research Ecosystem: a Multi-Legged Stool, PIDapalooza: Open Festival of Persistent Identifiers, Lisbon, Portugal, Jan 2020	Lisbon, Portugal
86	2019	Towards Fully Realized Open Science and Technology Trustworthiness, Halicioglu Data Science Institute, University of California San Diego, Dec 2019	UC San Diego
85		Transparency by Design in eScience Research (invited, visionary session	San Diego

		talk), 15th Int'l Conference on eScience, Sept 2019	California
84		*Open Science is (Good) Data Science, Indiana University Peebles Memorial Lecture, May 2019	Indiana Univ
83		Open Science, FAIR data, and Cyberinfrastructure, 36th Pacific Rim Applications and Grid Middleware (PRAGMA), Jeju South Korea, Mar 2019	Jeju, South Korea
82	2018	Open Science and Data Sharing: Policy and Infrastructure, Institute for Systems and Computer Engineering, Technology and Science (INESTEC), Mar 2018	INESTEC Porto, Portugal
81		Making Open Science work for science and society, IEEE Big Data Governance and Metadata Management workshop, Mar 2018	Berlin, Germany
80		Capsule Model for Open Science with Restricted Data, Data Preservation Alliance for the Social Sciences (Data-PASS), Sep 2018.	Boston, Massachusetts
79		*Open Science and Data Sharing: Policy and Infrastructure, Future Visions Symposium, Apr 2018	Colorado State Univ
78		*Open Science and Data Sharing: Policy and Infrastructure, 2nd Annual Texas A&M Research Computing Symposium, Jun 2018	Texas A&M
77		Capsule Framework for Open Science with Restricted Data, Technical Solutions to Advance Evaluation and Replication in the Social Sciences: What's New, What's Next workshop. Co-located with American Political Science Association (APSA) annual meeting, Aug 2018	Boston, Massachusetts
76		Capsule Computing: Safe Open Science, Computer Science Seminar, Binghamton University, Dec 2018	SUNY Binghamton
75		*Open Science as Roadmap to Better Data Science Research, Data Science Initiative Seminar, Binghamton University, Dec 2018	SUNY Binghamton
74	2017	*FAIR Open Science with PID Kernel Information: the RPID Testbed, Basarim 2017: 5th High Performance Computing Conference, Sep 2017	Istanbul, Turkey
73		Analyzing the HathiTrust Digital Library: 5.5 billion pages of knowledge, Earth Science Information Partners (ESIP) Summer Meeting, Jul 2017	Bloomington, Indiana
72		Foundations of a Data Ecosystem: Global Persistent IDs and Data Provenance, Int'l Symposium on Open Data and Innovation, Jul 2017	Beijing, China
71		PID Kernel Information: data handles and provenance, Collaborations to Enable Transnational Cyberinfrastructure Applications CENTRA2 All Hands Meeting, Apr 2017	Gainsville, Florida
70		Foundations of a Data Ecosystem: Global Persistent IDs and Data Provenance, Computer Science and Computer Engineering Colloquium, University of Arkansas, Mar 2017	Univ of Arkansas
69		*Data Mining Meets the Research Library, Computing and Humanity Speaker Series, Valparaiso University, Feb 2017	Valparaiso Univ
68	2016	Power of PID Kernel Information, Southeast Asia International Joint-Research and Training Program, Tainan City, Taiwan, Dec 2016	Tainan City, Taiwan
67		Opening a Massive Cultural Record, Dean's Advisory Council, School of Informatics and Computing, Indiana University Oct 2016	Indiana Univ
66		Data Science of Big Data, School of Public and Environmental Affairs, Indiana University, Oct 2016	Indiana Univ
65	2015	HathiTrust Research Center: Secure Commons, University of Toronto, Jun 2015	Univ of Toronto
64		*Trust Threads: Minimal Provenance for Data Publication and Reuse, National Conference on Data Integrity, Department of Information and Library Science, Colorado State University, May 2015	Colorado State Univ
63		Threads of Trust: Provenance of Data Reuse in Long Tail Science, Department of Information and Library Science, Indiana University, Apr	Indiana Univ

		2015	
62		*HathiTrust Research Center: Unlocking the Secrets of 4.6 Billion Pages, University of Missouri Cyberinfrastructure Days, Mar 2015	Univ of Missouri
61	2014	Big data, publishing, and data sharing, University of Tennessee i-School seminar, Dec 2014	Univ of Tennessee
60		Research Data Alliance and Big Data, Oak Ridge National Labs, Dec 2014	Knoxville, Tennessee
59		HathiTrust Data Capsules, CLIR/CNI Workshop on Expanding Access to Research Collections, Dec 2014	Washington, DC
58		Integrating Domain Repositories into the National Data Infrastructure, Workshop on National Data Infrastructures, ICPSR, Nov 2014	University of Michigan
57		Semantics, Data Provenance, Agent based Models, Ostrom Workshop on Political Theory and Policy Analysis, Oct 2014	Bloomington, Indiana
56		HathiTrust Research Center Data Capsule v1.0: An overview of functionality, with Robert McDonald, Miao Chen, IU Scholarly Data Commons, Sep 2014 http://hdl.handle.net/2022/18936	Bloomington, Indiana
55		Case Study in Big Data: the Socio-Technical Issues of HathiTrust Digital Texts, Women's Institute for Summer Enrichment (TRUST WISE), Jun 2014	Ithaka, New York
54		*Bridging Digital Humanities Research and Large Repositories of Digital Text, 2nd Encuentro de Humanistas Digitales, Biblioteca Vasconceles, May 2014 http://www.slideshare.net/BethPlale/keynote-2nd-encuentro-de-humanistas-digitales	Mexico City, Mexico
53		HathiTrust and HathiTrust Research Center: the Changing Digital Library, El Colegio de Mexico, May 2014	Mexico City, Mexico
52		HathiTrust Research Center: Challenges and Opportunities in Big Text Data, Miao Chen and Beth Plale, IU Digital Libraries Brown Bag, Mar 2014 http://hdl.handle.net/2022/17276	Bloomington, Indiana
51		Archiving a social-ecological database: challenges, solutions, and lessons learned, Beth Plale and Inna Kouper, IU Digital Libraries Brown Bag, Feb 2014 http://hdl.handle.net/2022/17301	Bloomington, Indiana
50		Research Data Alliance, Meeting of NSF EarthCube Community, Jan 2014	Washington, DC
49	2013	Big Data Opportunities and Challenges for IR, Text Mining, and NLP, The British Library, Dec 2013	London, UK
48		Big Data Opportunities and Challenges for IR, Text Mining, and NLP, Knowledge Media Institute (KMi), The Open University, Dec 2013	Milton Keynes, UK
47		Data Sets, Ensemble Cloud Computing, and the University Library, American Geophysical Union (AGU) Meeting, Dec 2013 (presented by co-author, Jim Myers)	San Francisco, California
46		Opportunities and Challenges of Text Mining HathiTrust Digital Library, The National Library of the Netherlands (Koninklijke Bibliotheek), Nov 2013	Den Haag, The Netherlands
45		*Big Data Opportunities and Challenges for IR, Text Mining and NLP, Int'l Workshop on Mining Unstructured Big Data Using Natural Language Processing (MNLP 2013), co-located with ACM Int'l Conf. on Information and Knowledge Management (CIKM), Oct 2013	San Francisco, California
44		*Big Data and Open Access: On Track for Collision of Cosmic Proportions?, 2nd Int'l LSDMA-Symposium - The Challenge of Big Data in Science - with a focus on Big Data Analytics, Sep 2013	Karlsruhe, Germany
43		HathiTrust Research Center (HTRC): Exploration of the World's First Massive Digital Library, with Miao Chen and Robert McDonald, Catapult	Bloomington, Indiana

		Center for Digital Humanities and Computational Analysis of Texts Digital History and Philosophy of Science (HPS) Workshop, Sep 2013	
42		Research Data Alliance: Researchers Sharing and Using Research Data Without Barriers, NIEHS-European Union Workshop on Identifying Opportunities for Global Integration of Toxicogenomics Databases, Jun 2013.	Research Triangle Park, North Carolina
41		Digital Humanities Text Mining at Scale: HathiTrust Research Center, Notre Dame Digital Humanities talk series, May 2013	South Bend, Indiana
40		Studies in Social-Ecological Systems Data Management, Interuniversity Consortium for Political and Social Research (ICPSR), Apr 2013	Ann Arbor, Michigan
39		International Data Sharing, Open Access and the Research Data Alliance, Advanced Regional Networks Envision Workshop on Big Data, Apr 2013	Boston, Massachusetts
38	2012	Research Data Alliance (RDA), EarthCube, and SEAD DataNet, Bridging Big Data Infrastructures Workshop, Dec 2012	Taichung, Taiwan
37		Explicit and Hidden Workflows in Environmental Science: Opportunities Enabled By, Microsoft Open Data for Open Science 2012 workshop, Apr 2012	Redmond, Washington
36		Metadata and Provenance: Fins in the Data Sea, Purdue University, Mar 2012	West Lafayette, Indiana
35		Digital Humanities at Scale: HathiTrust Research Center, University of Maryland College Park, Feb 2012	College Park, Maryland
34	2011	Creating Functionality Around Non-consumptive Research, American University School of Law, Jul 2011	Washington, DC
33		Metadata and Provenance Capture: Antecedent to Scientific Data Preservation, IU School of Library and Information Science course Data Curation Feb 2011.	Bloomington, Indiana
32		Provenance Collection of Unmanaged Workflows, Data To Insight Center Seminar, Indiana University, Jan 2011	Bloomington, Indiana
31		Metadata and Provenance Collection and Representation: Antecedent to Scientific Data Preservation, EECS Seminar, Oregon State University, Jan 2011	Corvallis, Oregon
30		Metadata and Provenance Collection and Representation: Antecedent to Scientific Data Preservation, EECS Seminar, Moratuwa University, May 2011	Columbo, Sri Lanka
29	2010	Keynote talk: LEAD II Hybrid Workflows for Timely Weather Products, 19th Pacific Rim Applications and Grid Middleware Assembly (PRAGMA), Sep 2010	Changchun, China
28		Metadata and Provenance Collection and Representation: Antecedent to Scientific Data Preservation, Open Data Seminar, University of Michigan, Nov 2010.	Ann Arbor, Michigan
27		Provenance and Workflows, Computer Network Information Center of Chinese Academy of Sciences, Oct 2010.	Beijing, China
26		LEAD II/Trident workflows for timely weather products: challenge of Vortex2, American-Chinese Cyberinfrastructure and E-science Workshop (ACCESS) on Data Intensive Sciences and Computing (DISC), Aug 2010.	Urbana, Illinois
25		LEADII: hybrid workflows in atmospheric science, DemoFest, Microsoft Faculty Research Summit, Jul 2010.	Seattle, Washington
24		LEAD II / Trident Workflows for Timely Weather Products: the Challenge of Vortex2, Microsoft External Research Symposium, Apr 2010	Seattle, Washington
23		Earth Systems Data in Real Time Applications: Low Latency, Metadata, and Preservation, Data-Intensive Research: how should we improve our ability	Edinburgh, Scotland

		to use data, e-Science Institute, University of Edinburgh, Mar 2010	
22	2009	Metadata and Preservation in Geosciences: Issues at Scale, IU Digital Libraries Program Brown Bag, Sep 2009	Bloomington, Indiana
21		Keynote talk: Discovery, Process and Preservation: the ABC's of Data in a Collaborative World, Indiana Universities Pervasive Technologies Institute Research Technologies Data Services Day, Sep 2009	Indianapolis, Indiana
20		It's All in the Data, Workflow Systems and the Weather, American Geophysical Union (AGU) Meeting, Geoinformatics Session, May 2009	Toronto, Canada
19		Integrating Weather/Climate Research Models and Data into End-Use Knowledge: Data to Insight, IBM Lecture Series, Dept. of Computer Science and Engineering, University of Notre Dame, Feb 2009	South Bend, Indiana
18		Integrating Weather/Climate Research Models and Data into End-Use Knowledge: Data to Insight, Computer Science Dept. Honors Lecture Series, Indiana University, Feb 2009	Bloomington, Indiana
17	2008	Beyond LEAD: Impact, Education, and Future Plans, Supercomputing '08, Nov 2008	Austin, Texas
16		*Research Instrumentation for Cyberinfrastructure, Data-intensive Computing and Weather Forecasting, NSF Workshop on Instrumentation Needs of Computer and Information Science Engineering, Snowbird, Jul 2008	Salt Lake City, Utah
15		Provenance and Metadata in Data-Intensive Computing, Seminar Series, Engineering, Architecture, and Computer Sciences, Howard University, Oct 2008	Washington, DC
14		Provenance of Digital Scientific Data, IEEE/ACM Supercomputing '08, Nov 2008	Austin, Texas
13	2007	Improving Data Capture in Science Discovery Cyberinfrastructure to Enable Educational Outcomes, Learning Sciences Professional Seminar, School of Education, Indiana University, Nov 2007	Bloomington, Indiana
12		Provenance of workflow generated data: bridge to future and aid to workflow interoperability, Science and Scholarly Workflows Workshop, Oct 2007	Baltimore, Maryland
11		Metadata, Provenance, and Search in e-Science, Complex Networks Seminar, Indiana University, Sep 2007	Bloomington, Indiana
10		Data Management, Metadata, and Search in Workflow-driven Computational e-Science, invited talk Renmin University, May 2007	Beijing, China
9		Data Integration, Search, and Analysis in Workflow-driven Computational Science, Workshop on Grid Portals and Data Management techniques for Earth Science Applications, Jun 2007	Sardinia, Italy
8		SOA and Events Processing for Next Generation Weather Forecasting: the LEAD Project, Object Management Group, Mar 2007	San Diego, California
7		Data Integration, Search, and Analysis in Workflow-driven Computational Science, Purdue University, June 2007	West Lafayette, Indiana
6	2006	*Distinguished Talk: Opening the Gates to Data Driven Computational Science through Cyberinfrastructure, Indiana University Office of Women's Affairs Distinguished Lecture Series, May 2006	Bloomington, Indiana
5		Metadata Catalogs and Stream Processing: Key Cyberinfrastructure for Data Driven Computational Science, University of Houston Computer Science Department Colloquia Series, Apr 2006	Houston, Texas
4		Transforming the Sensing and Prediction of Intense Local Weather Through Dynamic Adaptation, NSF Dynamic Data Driven Application Systems (DDDAS) Workshop, Jan 2006.	Washington, DC

3	2005	Calder, OGSA-DAI Access to Data in Streams, Argonne National Labs, Aug 2005	Chicago, Illinois
2		LEAD Data Subsystem: Overview, Current Approach to Integration, and Challenges, DIALOGUE Workshop, Aug 2005	Columbus, Ohio
1		Temporally Changing Geospatial Data in LEAD and DopplerSource, Microsoft e-Science Workshop, Aug 2005, https://www.microsoft.com/en-us/research/video/dopplersource-net-framework-for-accesing-doppler-radar-data/	Seattle, Washington

NON-PEER REVIEW PUBLICATIONS: INSTRUCTIONAL MATERIAL, DATASETS, PRE-PRINTS, WORKSHOP REPORTS (54) _____

	CITATION	*TYPE*
54	Envisioning National Resources for Artificial Intelligence Research: NSF Workshop Report, Shantenu Jha and Yolanda Gil, eds. arXiv 24.12.10278, 2024. https://arxiv.org/abs/2412.10278	Workshop report
53	Beth Plale, Achieving low barriers to entry in the FAIR Digital Objects (FDO) data space: a Use Case in Biodiversity Extended Specimen Networks, IU Scholarworks, Jul 2022, https://hdl.handle.net/2022/27837	Preprint
52	C.A. Stewart, V. Welch, T.G. Doak, T. Miller, B. Plale, et al. The Pervasive Technology Institute at 20: Two decades of success and counting. IU Scholarworks, 2022. http://hdl.handle.net/2022/22607	Project Report
51	Yu Luo and Beth Plale, Pilot evaluation of Collection API with PID Kernel Information (2019, Jul 3). ArXiv. arXiv:1905.02896v2 [cs.DL]	Preprint
50	Tobias Weigel, Beth Plale, Mark Parsons, Gabriel Zhou, Yu Luo, Ulrich Schwardmann, and Robert Quick. (2018, August 9). RDA Recommendation on PID Kernel Information (draft). Zenodo. http://doi.org/10.5281/zenodo.1462543	
49	Venice Margarette J Juanillas, Alexis Dereeper, Nicolas Beaume, Gaetan Droc, Joshua Dizon, John Robert Mendoza, Jon Peter Perdon, Locedie Mansueto, Lindsay Triplett, Jillian Lang, Gabriel Zhou, Kunalan Ratharanjan, Beth Plale, Jason Haga, Jan E Leach, Manuel Ruiz, Michael Thomson, Nickolai Alexandrov, Pierre Larmande, Tobias Kretzschmar, Ramil P Mauleon (2019), Rice Galaxy: an open resource for plant science, bioRxiv 358754; doi: https://doi.org/10.1101/358754	Preprint
48	Yu Luo, Kunalan Ratharanjan, Quan Zhou and Beth Plale (2018). Poster Abstract: Persistent IDs: Application to Workflow and Sensor Applications, 34th Pacific Rim Applications and Grid Middleware Assembly, http://hdl.handle.net/2022/22313	Poster
47	Inna Kouper, Yu Luo, Isuru Suriarachchi, Beth Plale (2017). Poster Abstract: Provenance Enriched PID Kernel Information as OAI-ORE Map Replacement for SEAD Research Objects, 17th ACM-IEEE-CS Joint Conf. on Digital Libraries, Toronto, CA Jun 2017	Poster
46	Beth Plale and Inna Kouper (2017). Tutorial: SEADTrain Data Analysis, ESIP Summer 2017 meeting, Bloomington IN, http://hdl.handle.net/2022/22312	Tutorial
45	Beth Plale (2016). Tutorial: Persistent Identifiers: their role in Data Quality. Data Quality in Era of Big Data Workshop, Bloomington, IN Sep 2016	Tutorial
44	Beth Plale, Matt Jones, Douglas Thain (2015), Workshop report: Software in Science: a Report of Outcomes of the 2014 National Science Foundation (NSF) Software Infrastructure for Sustained Innovation (SI2) Meeting, Final workshop report, http://hdl.handle.net/2022/19760, Mar 2015	Workshop report
43	Peng Chen and Beth Plale (2015). Big Data Provenance Analysis and Visualization, the	Doctoral

	Doctoral Symposium of 15th IEEE/ACM Int'l Symp. on Cluster, Cloud and Grid Computing (CCGrid), 2015	symposium
42	Beth Plale, Atul Prakash, Robert McDonald (2015). Final Project Report: The Data Capsule for Non-Consumptive Research, URI: http://hdl.handle.net/2022/19277	Project report
41	Guangchen Ruan and Beth Plale (2014), Evaluation of Data Storage in HathiTrust Research Center Using Cassandra. http://hdl.handle.net/2022/18472	Preprint
40	K. Caylor, T. Evans, L. Estes, J. Sheffield, B. Plale, and S. Attari (2014). Impacts of agricultural decision making and adaptive management on food security in Africa. AGU Fall Meeting Abstracts, 1:06	Preprint
39	Quan Zhou, Beth Plale, Keith Danielson, Robert J. Ping, Janae Cummings, and Alan Mauro (2013). Demo: Daily Regional Weather Forecasts in Support of Vortex2, IEEE Cluster 2013, Indianapolis, IN, Sep 2013	Demo
38	Luo, Yuan, Plale, Beth, Jensen, Scott, Cheah, You-Wei, Conover, Helen (2012). Provenance of AMSR-E Data from the National Snow and Ice Data Center (NSIDC). OPM XML Ver. 1.1., September 2 - October 4, 2011. Bloomington, Indiana: Data To Insight Center. http://dx.doi.org/10.5967/M0F47M2D (Dataset 13MB)	Dataset
37	Quan Zhou, Beth Plale (2012). Vortex2 Visualization, DOI: 10.5976/MOST7MST	Dataset
36	Scott Jensen, Devarshi Ghoshal, and Beth Plale (2011), Evaluation of Two XML Storage Approaches for Scientific Metadata , Indiana University Computer Science Technical Report TR698.	Preprint
35	Scott Jensen, Beth Plale, John Cobb, Rebecca Koskella (2011). Big Data Means Your Metadata Must Work, IEEE/ACM Supercomputing 2011, half day tutorial, Seattle, WA Nov 2011. 61 attendees.	Tutorial
34	Chathura Herath, Kavitha Chandrasekar, Beth Plale (2011). Integration of CEP into Scientific Workflows, 2 hour tutorial at ACM Distributed Event Based Systems (DEBS), Hawthorne, New York, 2011.	Tutorial
33	Jeff Cox and Beth Plale (2011), Improving Automatic Weather Observations with the Public Twitter Stream, Indiana University Computer Science Technical Report TR691.	Preprint
32	Scott Jensen and Beth Plale (2011), Schema-Independent and Schema-Friendly Scientific Metadata Management, Indiana University Computer Science Technical Report TR689. A (shorter) extended abstract of this paper was published in Proceedings of 4th IEEE Int'l Conf. on eScience.	Preprint
31	Eran Chinthaka, Suresh Marru, and Beth Plale (2009) Sigiri: Towards A Light-Weight Job Management System for Large Scale Systems, Indiana University Computer Science Technical Report TR681.	Preprint
30	Yiming Sun, Beth Plale, Chathura Herath, and Scott Jensen (2009). Tutorial: Event Processing in Weather Responsive Scientific Workflow Cyberinfrastructure: tutorial and abstract, 3rd Int'l Conference on Distributed Event Based Systems (DEBS), 2009.	Tutorial
29	Droegemeier, Kelvin, Beth Plale, Mohan Ramamurthy, and Craig Mattocks (2009). A New Approach for Using Webs, Grids, and Virtual Organizations in Mesoscale Meteorological Research, American Meteorological Society (AMS) 25th Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology, Jan 2009	Preprint
28	Bin Cao, Girish Subramanian, Sribabbu Doddapaneni, and Beth Plale (2008). Poster: Provenance Collection in an Industry Biochemical Discovery Cyberinfrastructure, 4th Int'l IEEE Conf. on e-Science (e-Science), Indianapolis, IN pp. 424-425, IEEE Computer Society, DOI 10.1109/eScience.2008.104 Dec 2008	Poster
27	Beth Plale, You-Wei Cheah, and Yiming Sun (2008). Towards Quantifying Limits in Automated Curation of e-Science Data, Indiana University Computer Science Technical Report TR672, Nov 2008.	Preprint

26	You-Wei Cheah and Beth Plale (2008). Representing LEAD Experiments in a FEDORA digital repository, Indiana University Computer Science Technical Report TR666, June 2008.	Preprint
25	Aparna Venkatraman, Vinay Pandey, Beth Plale, Shing-Shong Shei (2007). Benchmarking Effort of Virtual Machines on Multicore Machines, Indiana University Computer Science Technical Report TR654, Master's Research Project, Dec 2007.	Preprint (class project)
24	Xiang, X. and B. Plale (2007). Performance Evaluation of MySQL 5.0 and Berkeley DB XML as a Grid Resource Information Manager (GRIM) with a Benchmark/Workload, Indiana University Computer Science Technical Report TR-645.	Preprint
23	Plale, Beth, Rahul Ramachandran, and Steve Tanner (2006). Data Management Support for Adaptive Analysis and Prediction of the Atmosphere in LEAD, American Meteorological Society (AMS) 22nd Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology, 2006.	Preprint
22	Droegemeier, K. K., V. Chandrasekar, R. Clark, D. Gannon, S. Graves, E. Joseph, M. Ramamurthy, R. Wilhelmson, K. Brewster, B. Domenico, T. Leyton, V. Morris, D. Murray, B. Plale, R. Ramachandran, D. Reed, J. Rushing, D. Weber, A. Wilson, M. Xue, S. Yalda (2005). Linked Environments for Atmospheric Discovery (LEAD): Architecture, Technology Roadmap and Deployment Strategy, American Meteorological Society (AMS) 21st Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology, Jan 2005.	Preprint
21	Clark, Richard D., Sepideh Yalda, Dennis Gannon, Beth Plale, and Tom Baltzer (2005). Integrating LEAD Research in Undergraduate Education, American Meteorological Society (AMS) 22nd Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology, Jan 2006.	Preprint (undergrad education)
20	Yogesh L. Simmhan, Beth Plale and Dennis Gannon (2005). A Survey of Data Provenance Techniques, Indiana University Computer Science Technical Report TR618. Full length version of ACM SIGMOD Record, Sep 2005.	Preprint
19	Ying Liu, Beth Plale, and Nithya Vijayakumar (2005). Poster Abstract: Distributed Query Planner in the Calder System, 14th IEEE Int'l Symposium on High Performance Distributed Computing (HPDC), Research Triangle, NC, Jul 2005.	Poster
18	Nithya Vijayakumar, Ying Liu, and Beth Plale (2005). Poster Abstract: Calder: Enabling Grid Access to Data Streams, 14th IEEE Int'l Symposium on High Performance Distributed Computing (HPDC), Research Triangle, NC, Jul 2005.	Poster
17	Nithya Vijayakumar and Beth Plale (2005). dQUOBEC Event Channel Communication System, Indiana University Computer Science Technical Report TR-614.	Preprint
16	Liu, Y. B. Plale, and N. Vijayakumar (2005). Realization of GGF DAIS Data Service Interface for Grid Access to Data Streams, Indiana University Computer Science Technical Report TR-613.	Preprint
15	Droegemeier, K., V. Chandrasekar, R. Clark, D. Gannon, S. Graves, E. Joseph, M. Ramamurthy, R. Wilhelmson, K. Brewster, B. Domenico, T. Leyton, V. Morris, D. Murray, B. Plale, R. Ramachandran, D. Reed, J. Rushing, D. Weber, A. Wilson, M. Xue, S. Yalda (2004). Linked Environments for Atmospheric Discovery (LEAD): A Cyberinfrastructure for Mesoscale Meteorology Research and Education, American Meteorological Society (AMS) 20th Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology (IIPS), Jan 2004.	Preprint
14	Vijayakumar, N. and B. Plale (2004). RS-Algo: an Algorithm for Improved Memory Utilization in Continuous Query System under Asynchronous Data Streams, Indiana University Computer Science Technical Report TR-601.	Preprint
13	Moad, C. and B. Plale (2004). Portal Access to Parallel Visualization of Scientific Data on the Grid, Indiana University Computer Science Technical Report TR-593.	Preprint
12	Ying Liu and Beth Plale (2003). Survey of Publish-Subscribe Event Systems, Indiana University Computer Science Technical Report TR-574.	Preprint

11	Gannon, D., G. Fox, M. Pierce, B. Plale, G. von Laszewski, C. Severance, J. Hardin, J. Alameda, M. Thomas, J. Boisseau 2003. Grid Portals: A Scientist's Access Point for Grid Services, GGF Community Practice document, working draft 1.	Preprint
10	Plale, B., C. Jacobs, Y. Liu, C. Moad, R. Parab, and P. Vaidya 2003. Benchmark Details of Synthetic Database Benchmark/Workload for Grid Resource Information, Indiana University Computer Science Technical Report TR-583.	Preprint
9	Peter Dinda, and Beth Plale (2003). Poster Abstract: A Unified Relational Approach to Grid Information Services, 23rd IEEE Int'l Conference on Distributed Computing Systems (ICDCS), May 2003	Poster
8	Plale, B., C. Jacobs, Y. Liu, C. Moad, R. Parab, P. Vaidya, and N. Vijaykumar (2003). Poster Abstract: Understanding Grid Resource Information Management through a Synthetic Database Benchmark/Workload, Int'l Conference on High Performance Computing (HiPC), Hyderabad, India, Dec 2003.	Poster
7	Vaidya, P. and B. Plale (2003). Benchmark Evaluation of Xindice as a Grid Information Server, Indiana University Computer Science Technical Report TR-585.	Preprint
6	Plale, B. (2001). Performance Impact of Streaming Doppler Radar Data on Geospatial Visualization System, Georgia Institute of Technology Technical Report GIT-CC-01-07.	Preprint
5	Plale, B. and K. Schwan 2000. dQUOB: Managing Large Data Flows Using Dynamic Embedded Queries, Georgia Institute of Technology Technical Report GIT-CC-00-07.	Preprint
4	Isert, C., King, Schwan, K., Plale, B., and Eisenhauer, G. (1999). Poster Abstract: Steering Data Streams in Distributed Computational Laboratories, 19th IEEE Int'l Symposium on High Performance Distributed Computing (HPDC), IEEE Computer Science Press, Aug 1999.	Poster
3	Plale, B. and Schwan, K. (1998). Poster Abstract: Multi-level Steering in Distributed Laboratories, Proceedings of SIGMETRICS Symposium on Parallel and Distributed Tools, ACM Press, p. 162, Aug 1998.	Poster
2	Schroeder (Plale), B. and K. Schwan (1997). Language Issues in Hazard Detection Using Queries, Georgia Institute of Technology Technical Report GIT-CC-97-36.	Preprint
1	Schroeder (Plale), B., S. Aggarwal, and K. Schwan (1997). Software Approach to Hazard Detection Using On-line Analysis of Safety Constraints, Georgia Institute of Technology Technical Report GIT-CC-97-01.	Preprint

TEACHING and RESEARCH SUPERVISION

GRADUATE ADVISER (5):

Steve Holl – AI in engineering product design

Neelish Karthikeyan – AI model cards for edge cloud continuum

Jiannan Tian – data compression in heterogeneous hardware

GRADUATED PHD STUDENTS (PRIMARY ADVISOR*) (19):

GRADUATE	TITLE	YEAR
Sadia Khan	Institutional Analysis of Consumer Neurotechnology Regulation: Mediating Third Party Autonomy	2024

Punait Sujata	ANSTracker: A Comprehensive Tool for Monitoring Autonomic Nervous System Dynamics in Real-Time	2024
Sachith Withana*	Cyberinfrastructure Knowledge Network: Optimizing Edge-Cloud Continuum for Efficient Data Capture, Modeling, and Deployment	2024
Yu Luo*	Knowledge Enhanced Digital Objects	2022
Milinda Pathirage*	Efficient Hybrid Data Management System on Distributed Log Store	2018
Zong Peng*	Cloud-based Service for Access Optimization to Textual Big Data	2018
Isuru Suriarachchi*	Big Provenance Stream Processing for Data-Intensive Computing	2018
Guangchen Ruan*	Policy-based Analysis and Frameworks for Non-Consumptive Research	2016
Peng Chen*	Big Data Analytics in Static and Stream Processing	2016
Jiaan Zeng*	Electronic Arts	2015
Yuan Luo*	Facebook	2015
Devarshi Ghoshal*	Lawrence Berkeley National Labs	2014
You-Wei Cheah*	Lawrence Berkeley National Labs	2013
Chathura Herath*	Knight Capital Group	2011
Eran Chinthaka Withana*	Comprehend Systems	2011
Scott Jensen*	Professor, San Jose State Univ	2010
Ying Liu*	2007. Cisco	2007
Yogesh Simmhan*	Professor, India Institute of Science Bangalore	2007
Nithya Vijayakumar*	Apple	2007

POSTDOCTORAL ADVISOR (7):

Inna Kouper	Research Faculty, Indiana University Bloomington	2012 – 2016
Miao Chen	Alibaba, Hangzhou, China	2011 – 2015
Abhirup Chakraborty	Google	2012 – 2014
Stacy Kowalczyk	Professor, Dominican University, Illinois	2010 – 2012
Mehmet Aktas	Professor, Yildiz Teknik University, Budapest, Turkey	2010 – 2011
Bin Cao	Teradata Corp	2008 – 2009
Sangmi Lee Pallickara	Professor, Colorado State University	2005 – 2007

ADVISOR AND MENTOR (2):

Umashanthi Pavalanathan (2012-2013)	Junior visiting research scholar	Went on to do PhD from Georgia Institute of Technology
Kavitha Chandrasekar (2011-2014)	Data To Insight Center developer	Went on for PhD at Univ Illinois

MASTER'S THESIS ADVISOR (24):

Quan (Gabriel) Zhou (2018) eBay, San Francisco, CA	Aparna Venkatraman (2007) Cummins, Columbus IN
Liang Ran (2017), Amazon, Seattle, WA	Vinay Pandey (2007) Dell Technologies, Austin, TX
Charitha Madurangi Dandeniya Arachchi (2017), HathiTrust Research Center, Bloomington, IN	Srilatha Marru (2007) RightRez, Bloomington, IN
Pradeep Ravilla (2017), Global Network Operations Center, Bloomington, IN	Humin (Lily) Fang (2007)
Aravindh Varadharaju,	Ning Liu (2006) Audible, NYC
Yiming Sun (2013), Amazon, Seattle, WA	Ryan Baula (2005)
Bina Bhaskar (2012), Amazon, Orange County, CA	Craig Jacobs (2005)
Bimalee Salpitikorala, Indiana University, Knoxville, TN	Charlie Moad (2004) Costello, Indianapolis IN
Prajakta Purhoit (2011), Chef Software, Seattle, WA	Poornima Venkatakrishnan (2004) PayPal, San Francisco, CA
Shobana Krishnan (2011) Cummins, IN	Nithya Sivaraman (2004) Microsoft, Seattle, WA
Ashish Bhangale (2010) Microsoft, Seattle, WA	Deepti Kodeboyina (2004) ChargePoint, San Francisco, CA
Felix Terkhorn (2009) Cake Solutions Limited, NYC	
Ai Zhang (2008), Cisco, San Francisco, CA	

PHD EXTERNAL REVIEWER / EXTERNAL COMMITTEE MEMBER:

- Stian Soiland-Reyes, University of Amsterdam, 2024
- University of Melbourne, Melbourne, Australia
- Rensselaer Polytechnic University

TEACHING

INSTI-TUTION	COURSES TAUGHT (* NEW OR SUBSTANTIAL COURSE DEVELOPMENT)	DATES/ENROLLMENTS
IUB	*ENGR E222 Intelligent Systems Engineering II	Spr 2023 (35), Spr 2022 (32), Spr 2021 (35)
IUB	*ENGR E599, INFO I590, CSCI 659 Artificial Intelligence (AI) Trust and Safety	Fall 2020 (12)
IUB	* I535, B669, I435: Management Access and Use of Big and Complex Data. Hybrid course. Data lifecycle, data storage, cloud computing	Fall 2016 (290), Fall 2015 (60), Fall 2014 (40)
IUB	* CSCI P434: Distributed Systems. Foundations of distributed computing	Spr 2012 (20)
IUB	* CSCI B669: Scientific Data Management and Preservation. Data lifecycle, scientific workflows.	Spr 2013 (35), Spr 2011 (35)
IUB	* I590/B669: Topics in Data and Search Informatics. Data provenance, data indexing, case based reasoning, metadata models, and visualization. Indiana University, co-taught with Dennis Groth.	Fall 2008 (35)

IUB	* CSCI B534 Distributed Systems, Indiana University. Foundations of distributed computing including models, consistency, global time, architectures. Projects: 2010 - Parallel ray tracing application using Amazon Elastic Map Reduce and Web Services. 2007 - performance benchmark of virtual machines on identical Dell and HP hardware platforms. 2005: sponsored several teams to participate in a Kelley Business School and HP sponsored competition on course-wear for handheld devices.	2010, 2009, 2007, 2005, 2004, 2003 (as B649)
IUB	CSCI P436 Introduction to Operating Systems	2007, 2002, 2001
IUB	CSCI B649: Service Architecture and Science. Co-taught with Dennis Gannon	2006
IUB	CSCI B438 Fundamentals of Networking	2005, 2006
IUB	* CSCI B649 Topics in Systems: Systems Support for Wide Area Applications	2002
Georgia Tech	CS4210 Advanced Operating Systems	Sum 2001, Sum 2000
Georgia Perimeter College	Fortran for Scientists and Engineers	Fall 1995, Spr 1996
Georgia Perimeter College	Introduction to Pascal	Fall 1995, Spr 1996
North Springs High School	* Introduction to Visual Basic, Advanced Placement course taught at North Springs High School, Georgia Perimeter College	Fall 1995
Georgia Perimeter College	* Visual Basic for Professionals	Spr 1996
SUNY Binghamton	C Programming and the UNIX operating system. Off campus continuing education	Spr 1994
SUNY Binghamton	C Programming and the UNIX operating system, graduate course, SUNY Binghamton	Su 1994

SERVICE TO RESEARCH COMMUNITY _____

RESEARCH COMMUNITY LEADERSHIP

General Chair
or Program
Chair for peer

reviewed
conferences:

Program Chair, IEEE/ACM Int'l Symp on Cluster, Cloud, and Internet Computing (CCGRID 2024), Philadelphia PA 2024

Diversity and Belonging Track Chair, IEEE/ACM Int'l Symp on Cluster, Cloud, and Internet Computing (CCGRID 2023), Bangalore, India 2023

Chair, SC20 Transparency and Reproducibility Initiative, part of IEEE/ACM Int'l Conference on High Performance Computing, Networking, Storage, and Analysis (SC20), 2020

Program Chair, 2019 IEEE Int'l Congress on Big Data, Jul 2019

General Chair, ACM High Performance Distributed Computing (HPDC) 2014

General Chair, 5th Int'l Provenance and Annotation Workshop (IPAW) 2014

Program Chair, IEEE 20th Int'l Conference on Web Services (ICWS) 2013

Program Vice Chair, 2006 IEEE/ACM Int'l Conference on Grid Computing (Grid06)

Community
Leader:

IEEE Admission and Advancement Committee IEEE Senior Member application review, Jan 2024

Organized Workshop on Data Quality in Era of Big Data, Bloomington, IN Sep 2016

Organized 1st PRAGMA Int'l Clouds for Data Science, Indonesia, Oct 2015. NSF Funded.

Organized PRAGMA 27: Networking in a Big Data World, Bloomington, IN 2014

Organized 2014 NSF Software Infrastructure for Sustained Innovation (SI2) PI meeting. Report at https://hdl.handle.net/2022/19760

Organized Coming Together Around Data Workshop for NSF DataNet and INTEROP Projects, a Principal Investigator meeting, NSF funded, Indianapolis, IN, Jan 2012.

Organized 1st Int'l Workshop on Traceability and Compliance of Semi-Structured Processes (TC4SP2010), Sep 2010

Organized GeoCloud: Cloud Computing, Collaborative Technologies and the Geosciences, NSF funded. Sep 2009.

Planning Committee: National Forum for Geosciences Information Technology (FGIT), 2005

Organized Riding the Geoscience Cyberinfrastructure Wave of Data: Real Time Data Use in Education, Dec 2008

Organized Indiana Women in Computing (inWIC) 2006

Organized Central Indiana Celebration of Women in Computing (CICWIC) 2004

Co-lead, NCSA Alliance Scientific Portal Expedition, 2003-2005

Chair, Global Grid Forum NOMCOM, 2004

Co-Chair Global Grid Forum (GGF) Relational Grid Information Systems Group, 2001-2002

TECHNICAL PROGRAM COMMITTEE – FOR HIGHLY SELECTIVE PEER REVIEWED
CONFERENCES

VENUE	*YEAR SERVED*
ACM Conference on Reproducibility and Replicability (ACM REP '23)	2023
Int'l Conference on Distributed Computing Systems (ICDCS)	2022
Int'l Provenance and Annotation Workshop (IPAW)	2014, 2015, 2016, 2018
IEEE Int'l Conf. on Parallel and Distributed Systems (ICPADS)	2015
Int'l Conf. on Scientific and Statistical Database Management (SSDBM)	2014, 2015, 2019
Workshop on Scientific Cloud Computing (ScienceCloud)	2014, 2016, 2018
IEEE Int'l Conf. on eScience (e-Science)	2014, 2015, 2017, 2024
ACM Int'l Conf on Information and Knowledge Management (CIKM)	2016
IEEE/ACM Supercomputing (SC) tutorials committee,	2011, 2012
IEEE Int'l Conf. on High Performance Computing (HiPC)	2012
IEEE/ACM Int'l Parallel and Distributed Processing Symposium (IPDPS)	2011
Int'l ACM Symp. on High Performance Distributed Computing (HPDC)	2011, 2012
USENIX Conf. on File and Storage Technologies (FAST)	2010
Int'l Provenance and Annotation Workshop (IPAW)	2010, 2012
1st Int'l Workshop on Workflow Approaches for New Data-Centric Science (WANDS)	2010
3rd Int'l Workshop on Data Intensive Distributed Computing (DIDC)	2010
Challenges of Large Applications in Distributed Environments (CLADE)	2010
IEEE Int'l Conf. on e–Science (e-Science)	2010, 2011, 2012
Int'l Conf. on Distributed Event Based Systems (DEBS)	2008, 2009
IEEE Int'l Conf. on e-Science and Grid Computing	2007, 2009
IEEE Int'l Conf. on Information Reuse and Integration	2008
Grace Hopper Conference	2006, 2009
1st Int'l Workshop on the role of Semantic Web in Provenance Management	2009
Int'l Conf. for High Performance Computing and Communications (SC)	2006, 2007
Semantic Scientific Knowledge Integration, AAAI/SSS Workshop	2007
Second Int'l Workshop on Event-driven Architecture, Processing and Systems (EDA-PS'07)	2007
7th IEEE Int'l Symposium on Cluster Computing and the Grid (CCGrid)	2007
IEEE/ACM Int'l Parallel and Distributed Processing Symp. (IPDPS)	2006, 2005
IEEE High Performance Distributed Computing (HPDC)	2002
IEEE Heterogeneous Computing Workshop (HCW)	2002

SESSION CHAIR – FOR HIGHLY SELECTIVE PEER REVIEWED CONFERENCES

VENUE	*YEAR SERVED*
Int'l Conference on High Performance Computing, Networking, Storage, and Analysis (SC21),	2021
Int'l Conf. on Distributed Event Based Systems (DEBS)	2009
Int'l Conf. on Data Engineering (ICDE), Atlanta, GA	2006
IEEE/ACM Int'l Conf. on Grid Computing (Grid06)	2006
IEEE High Performance Distributed Computing (HPDC)	2002
ICSA Parallel and Distributed Computing Systems (PDCS)	2002

ICSA Parallel and Distributed Computing Systems (PDCS)	2000

FUNDER MERIT REVIEWS

ORGANIZATION	*YEAR SERVED*
US National Science Foundation, technical reviewer for OAC office and CISE, GEO, SBE, OISE, MPS directorates	2024, 2022, 2021, annually until 2017
Science Foundation Ireland	2021
Academy of Finland, Finland	2015
US Department of Energy	2024, 2014, 2013, 2010, 2009, 2007, 2006
US Department of Homeland Security	2013
Skoltech, Russia	2013
Engineering and Physical Sciences Research Council (EPSRC), United Kingdom	2006

SERVICE TO INDIANA UNIVERSITY

ROLE	*UNIVERSITY LEVEL ACTIVITY*	*YEAR SERVED*
Member	IU 2030 Strategic Plan Executive Committee	2023-2024
Member	IUB Promotion Committee	2020 – 2022
Member	IU Center of Excellence for Women and Technology (CEWIT) Advisory Board	2015 –
Chair	IU Associate Vice President for Research (AVPR) search committee	2014
Chair	IU Digitization Master Plan Task Force	2014
Member	IUB Provost Strategic Planning Faculty Team	2013
Chair	IUB Faculty Research Support Funding Program (FRSP) ($1M annual internal grant program)	2010
Panel member	IUB Faculty Research Support Funding Program (FRSP)	2005, 2006, 2007, 2008, 2009
Member	IUB Provost search committee	2012
Member	IU Vice President of Research search committee	2009, 2010
Member	IUB Dean of Libraries search committee	2008
Member	University Intellectual Property Council	2008 – 2011
Member	Bloomington Faculty Council (BFC) Faculty Mediation Committee	2009 – 2011
Member	University Conflict of Interest Committee	2008 – 2009
Member	University Information Technology Committee	2008
Member	Dean of School of Informatics, Computing, and Engineering search committee	2007
Faculty advisor	IUB Office of Women's Affairs Women in Science Program	2006 – 2007
Faculty advisor	Midwest Crossroads Alliance for Graduate Education and the Professoriate (AGEP)	2006 – 2009
Member	IU Restricted and Classified Task Force	2005 – 2006
Member	University IT Director of Systems search committee	2005
Member	University IT Director of Applications search committee	2005

ROLE	SCHOOL OR DEPT LEVEL ACTIVITY	YEAR SERVED
Chair	Luddy School Diversity and Inclusion Task Force	2020 – 2021
Chair	Data Science Program Curriculum Committee	2015 – 2017
Member	Data Science Program Admissions Committee	2016 – 2017
Member	SICE Promotion & Tenure Committee (School level)	2014 – 2015, 2015 – 2016
Member	Informatics Faculty hiring committee	2009 – 2010, 2010 – 2011, 2013 – 2014
ADR	Associate Dean of Research	2007 – 2009
Member	SICE School Visioning Task Force	2009
Member	SICE Structure Task Force	2008
Co-chair	Dept of Computer Science Diversity committee	2005 – 2006
Member	SICE Policy Committee	2005 – 2006
Member	SICE School Structure Committee	2005 – 2006
Organizer	Data and Search Informatics Seminar	2007 – 2009
Founder, Advisor	Women In Computing @ IU (WIC@IU)	2002 – 2007
Member	Computer Science Hiring Committee	2008 – 2011
Chair or member	Computer Science Ph.D. Qualifiers committee (chair or me	2002, 2003, 2004, 2005, 2006, 2007, 2008, 2010
Member	Computer Science Faculty Hiring committee	2004 – 2005, 2008 – 2009
Member	Dean's Faculty Advisory committee	2002 – 2003
Organizer	Systems Seminar	2001 – 2006
Member	Computer Science Dept Graduate Education committee (GEC)	2006 – 2008
Member	Undergraduate Education Committee (UEC)	2004 – 2005, 2012 – 2013