Paul Groth

Professor of Algorithmic Data Science
Informatics Institute
University of Amsterdam

✉: p.groth@uva.nl
: http://pgroth.com

Experience

2018 - present — Full Professor - Algorithmic Data Science
Informatics Institute, University of Amsterdam

- Lead: Intelligent Data Engineering Lab
  - The lab (indelab.org) focuses on intelligent systems that support people in their work with data and information from diverse sources.
- Founding scientific director: UvA's Data Science Centre (2021 - present)
  - The centre (dsc.uva.nl) accelerates the use of data science methods across the university.
- Co-scientific director: AIRLab (2019 - present) & the Discovery Lab (2020 - present)
  - These two Innovation Center for Artificial Intelligence labs work with industry to develop AI research to improve retail (icai.ai/airlab/) with Ahold-Delhaize and scientific discovery (icai.ai/discovery-lab/) with Elsevier.
- Director: Graduate School of Informatics (2023 - present)
  - Responsible for 8 computing masters programmes with over 1100 students.

2023 - present — Co-Founder
longform.ai

- AI startup that turns longform conversations into structured actionable data.

2015 - 2018 — Disruptive Technology Director
Elsevier Labs

- Performed research on knowledge graph construction, intelligent systems in science, and data provenance.
- Established new research relationships with academic collaborators.
- Led creation of a company-wide internal best practices for dataset metadata.

- Co-author of company-wide internal architectural principles for knowledge graphs.
- Delivered multiple in-depth technology reports for C-level executives.
- Led technology input during development of Elsevier-wide precision medicine strategy.

2011 - 2015 Assistant Professor
Department of Computer Science, Vrije Universiteit Amsterdam

- Lead architect for the Open PHACTS (`http://www.openphacts.org`) semantic data integration platform for pharmacology. The €16 million project involves working with major pharmaceutical companies and leading academics.
- Co-chair W3C Provenance Working Group that created a standard for provenance on the Web used in organizations ranging from Oracle to NASA.
- I acted as Associate Director of Amsterdam Data Science from Fall 2013 until the end of 2014.

2009 - 2011 Postdoctoral Researcher
Knowledge Representation and Reasoning Group, Department of Computer Science, Vrije Universiteit Amsterdam

- Led the interdisciplinary Semantically Mapping Science project where I developed novel network analysis based methods applied to knowledge about scientific activity extracted from the Web.
- I developed the technical underpinning for nanopublications - a way of representing fine-grained provenance.
- Initiated the altmetrics research area (`http://altmetrics.org`) with social scientists to promote the study of alternative measures of science. Altmetrics has been covered by Nature, Wired Science and The Guardian.

2007 - 2009 Postdoctoral Research Associate
Intelligent Systems Division, Information Sciences Institute, University of Southern California

- As part of the DARPA Bootstrap Learning project, I developed algorithms to learn procedures from human instruction. (`http://www.sri.com/work/projects/bootstrapped-learning`)
- Designed and helped develop the provenance framework for the Windward: Scalable Knowledge Discovery Through Grid Workflows project. (`http://www.wings-workflows.org`)
- Led the integration of the Wings Workflow System with MIT/Harvard Broad Institute's GenePattern workflow system

2004-2007 PhD Computer Science
University of Southampton

- During my PhD, I was a member of the UK e-Science Provenance Aware Service Oriented Architecture (PASOA) Project (http://www.pasoa.org) and the EU Provenance Project (http:// http://www.gridprovenance.org).
- Co-developed PReServ, an open source web service that allows for the storage and retrieval of provenance information, which used by other research groups including at The University of Leeds and the China Research and Development environment Over Wide-area Network project.
- Heavily participated in the design and specification of the architecture used by the EU Provenance Project.

2003 - 2004 Internship
Fraunhofer-Institut für Produktionstechnik und Automatisierung

- Co-developed a prototype multi-agent system to control semiconductor factories, which successfully integrated with a commercial simulator.

2001 - 2002 Internship
University of Ulm, Distributed Systems Department

- Designed and implemented a directory service for the Mobility And Service Adaptation in Heterogeneous Mobile Networks (MASA) project. Project partners were the University of Ulm, NEC Europe Ltd. and Siemens AG.

1998 - 2003 Student Research Assistant
Florida Institute for Human & Machine Cognition

- Developer on NOMADS (`http://www.ihmc.us/research/projects/Nomads`) a Java-based mobile agent distributed system based on a clean room implementation of the Java virtual machine.
- Developer on CMapTools (`http://cmap.ihmc.us/`). The software empowers users to easily construct, navigate, share, and criticize knowledge models through concept maps.

External Appointments

- Senior Research Fellow - Netherlands School for Information and Knowledge Systems (SIKS)
- Supervisory Board - Amsterdam Data Science (2019 - 2023)
- Board of Directors - FORCE11 (2014 - 2019)

Education

2004-2007 PhD Computer Science, University of Southampton

- Thesis title: "The Origin of Data: Enabling the Determination of Provenance in Multi-institutional Scientific Systems through the Documentation of Processes"
- My research focused on provenance in e-Science applications. It develops novel concepts for creating, recording, and organizing provenance information. These concepts were applied to a large scale bioinformatics application.

1999-2002 B.S. Computer Science, University of West Florida

- Magna Cum Laude and John C Pace Jr. Scholar
- Study Abroad - University of Ulm, Germany (Fall 2011 – Spring 2012)

  Baden-Württemberg-Stipendium recipient
- Study Abroad - Fukuoka Communication Arts, Japan (Summer 2001)

Service

Standardization Activities

- Co-chair W3C Provenance Interchange Working Group - April 1, 2011 - September 30, 2013

  `http://www.w3.org/2011/prov/wiki/Main_Page`
- Member W3C Incubator Group on Provenance - October 1, 2009 - November 30, 2010

Editorships

- Co-editor Spring Synthesis Lectures on Data, Semantics, and Knowledge
- International Provenance and Annotation Workshop Steering Committee (Chair)
- Transactions on Graph Data and Knowledge (TGDK) - Editorial Board
- Journal of Web Semantics - Editorial Board
- ACM Transactions on the Web- Distinguished Reviewers Board
- Journal of Web Semantics, Using Provenance in the Semantic Web - special issue editor
- Future Generation Computing Systems, The third provenance challenge on using the Open Provenance Model for interoperability special issue editor
- Organizer PLoS One, Altmetrics Collection

Selected Program Committees

- I have been on the PC of the following conferences. I highlight senior roles.
- General Chair - 2022 Extended Semantic Web Conference

  SPC and track chair in several editions
- Program Chair - 2016 International Semantic Web Conference

  SPC and track chair in several additions
- Program Chair - 2021 Semantics
- Program Chair - 2012 International Provenance & Annotation Workshop
- World World Wide Web Conference

  SPC 2018, 2022, 2023

  Posters & Demos Chair 2018

  Tutorials Chair 2021

• Very Large Data Bases • International Conference on Extending Database Technology • International Joint Conference on Artificial Intelligence • ACM International Conference on Intelligent User Interfaces • ACM Knowledge Capture Conference • Theory & Practice of Digital Libraries • NeurIPS Datasets and Benchmarks • ACM Conference on Information and Knowledge Management • ACM Web Science Conference • IEEE Internaitonal Conference on e-Science

Reviewing

Beyond my editorial board memberships, I have reviewed for a wide variety of journals including:
Communications of the ACM, IEEE Intelligent Systems, Data & Knowledge Engineering, Future Generation Computer Systems, IEEE Internet Computing, IEEE Transactions on Automation Science and Engineering, Journal of Web Semantics, Semantic Web Journal, Interacting with Computers, Concurrency Computation Practice and Experience, Computers & Geosciences, Journal of Systems and Software, Interacting with Computers, Semantic Web Journal, Journal of Data & Information Quality, Applied Ontology, Journal of Parallel & Distributed Systems

I have been on the PhD committee of: Yang Fang, Xu Wang, Shuo Chen, Nora Abdelmageed, Sirko Schindler, Russa Bissau, Lu Zhang, Imran Asif, Wei Zeng, Nikos Voskarides, Jie Zou, Armel Lefebvre, Marzieh Fadaee, Hamidreza Ghader , Chaun Wu, Gosia Migut, Carlos Sáenz Adán, Tom De Nies, Marcin Wylot, Matthew Gamble, Krystyna Milan, Mohammad Rezwanul Huq

I have reviewed for a wide variety of funders, institutes and book publishers.

Teaching & Supervision

- I have obtained University Teaching Qualification the recognised qualification by Dutch universities for teaching within higher eduction.
- I have served on admissions board, examination boards, as well as various education coordinator roles with the UvA.

Courses

- Information Organization - University of Amsterdam - Fall 2019 - 2022
- Modern Databases - University of Amsterdam - Fall 2019 - 2022
- Data, Sensors and Complex Service - Spring 2018, 2019
- Information Retrieval - Vrije Universiteit Amsterdam - Fall 2013, 2014
- Knowledge & Media - Vrije Universiteit Amsterdam - Fall 2013, 2014
- Web-based Knowledge Representation & Semantic Web - Vrije Universiteit Amsterdam - Fall 2009 - 2012
- I regularly give guest lectures.

PhD Supervision

- Madelon Hulsebos - University of Amsterdam - 2024

  "Table Representation Learning"

  https://hdl.handle.net/11245.1/74d21daa-16a3-488f-a618-3f49300137e9

  Madelon is now at University of California, Berkeley.
- Manolis Stamatogiannakis - Vrije Universiteit Amsterdam - 2021

  "High-Fidelity Provenance"

  https://hdl.handle.net/1871.1/c8f23f90-9d64-4446-966e-3739df32280a

  Manolis is now at Wundermart.
- Kathleen Gregory - University of Maastricht - 2021

  "Findable and reusable? Data discovery practices in research"

  https://doi.org/10.26481/dis.20210302kg

  Kathleen graduated cum laude. She is now at the University of Leiden.
- Sara Magliacane - Vrije Universiteit Amsterdam - 2017

  "Logics for causal inference under uncertainty"

  https://hdl.handle.net/1871/55267

  Sara is now at the University of Amsterdam & IBM Research.

Funding Profile

2023	PI. Longform.ai (40k €). NWO
2022-2025	Co-PI. ENEXA: Efficient Explainable Learning on Knowledge Graphs (521,162 € UvA). EU Horizon Europe.
2019-2025	Co-Scientific Director. Innovation Center for Artificial Intelligence - AI for Retail Lab. Ahold Delhaize.
2019-2024	Co-Scientific Director. Innovation Center for Artificial Intelligence - Discovery Lab. Elsevier.
2020-2024	Co-PI. Making the hidden visible: Co-designing for public values in standards-making and governance (750,000 € total) . NWO.
2022 - 2025	CausalFusion: Causal Knowledge Extraction and Fusion from Multiple Modalities. AFOSR.
2019 – 2020	Knowledge Graph Construction from Situated Multimodal Dialogue, Award Number. AFOSR.
2015	Industry Partner - Elsevier. Re-SEARCH: Contextual Search for Scientific Research Data (750,000 € total). NWO Funded.
2014	Lead acquisition of Big Data Europe project `http://www.big-data-europe.eu` for the Open PHACTS foundation. H2020 Funded.
2014 - 2018	Named staff. RISIS: Research infrastructure for research and innovation policy studies (700,000 € - VU University Amsterdam). FP7 Funded.
2011 - 2014	Work package Leader and led project acquisition for the VU University Amsterdam - Open Pharmacological Space (Open PHACTS). (615,880 € - VU University Amsterdam). EU IMI Funded.
2011 - 2015	Work package Leader – Data2Semantics – A COMMIT Dutch National Project (1 PhD student)
2012 - 2013	VU University Amsterdam Network Institute Student Assistantships (2 paid students for 10 months)

Outreach

Invited Talks

- Knowledge (Graphs) in the Language Model Era - Center for Informatics Research in Science and Scholarship UIUC speaker series - 2023
- Knowledge Engineering in the Language Model Era - 7th ISKO UK biennial conference 2023
- Minimal Viable Data Reuse - Vogin IP Conference 2022
- Keynote - Knowledge Graph Development - Knowledge Graph and Semantic Web Conference 2022
- Knowledge Graph Maintenance - NEC Labs Europe - 2021
- Keynote - Data Communities - ConTech 2021
- The Information Need that is Data - University of Zurich - 2020

- CSE Distinguished Seminar Series - The Challenge of Constructive Data Search - Technical University of Lisbon - 2020
- The Information Need that is Data - University of Fribourg - 2020
- Panel - European Semantic Web Conference 2020
- Keynote - Content and Signals - ConTech 2020
- The Making of Data - WU Wien - 2019
- Keynote - Standards and AI - Standards Technology Forum 2019
- Flexible and Transparent Data Reuse - University of Manchester - 2019
- The Making of Data - University of Bologna - 2019
- Diversity and Depth: Implementing AI across many long tail domains. Industry Day - International Joint Conference on AI (IJCAI) 2018
- Panel - Increasing Scientific Productivity through Artificial Intelligence. EuropeanScience Open Forum. July 13 2018
- Keynote - From Text to Data to the World: The Future of Knowledge Graphs. Paul Groth. Integrative Bioinformatics.June 15, 2018
- Panel - Do we need a Google for Data? International Workshop on Profiling and Searching Data on the Web. Web Conference April 24, 2018
- Keynote - Combining Explicit and Latent Web Semantics for Maintaining Knowledge Graphs. Keynote. Big Net Workshop at The Web Conference. April 24, 2018.
- Panel - How does scholarly publishing move towards digital/Web based publishing? Digital Publishing Summit Europe. May 17, 2018
- Keynote - Leveraging AI to construct large scale scientific knowledge graphs. Connected Data London. Nov. 16, 2017
- Keynote - The Roots: Linked Data and the foundations of successful agriculture data. G20 Meeting of Agricultural Chief Scientists (MACS) Workshop Linked Open Data in Agriculture. Sept. 27, 2017
- Keynote - Machines are People Too. 21st Theory and Practice of Digital Libraries. Sept. 18, 2017
- Robotic Labs and Science as a Service. Information Sciences Institute. May 30, 2017
- Data: The Driver of Precision Medicine. Internet Medicine: Medicine in the Year 2045 RWTH Aachen. January 18, 2017

- Panel - Data, metadata and partnerships. STM event Publishing in a larger world, or is it a smaller one?. Dec. 8, 2016
- Tuning Search: Making the most of Open Source Search and SOLR at Elsevier. BigDataLDN. November 3-4, 2016
- Panel - Deep Learning to Solve Business Problems. BigDataLDN. November 3-4, 2016
- Structured Data & the Future of Educational Material. Education & Linked Data - Platform Linked Data Netherlands. September 9, 2016
- Data For Science: How Elsevier Is Using Data Science To Empower Researchers. European Data Forum. June 29-30, 2016
- Knowledge Graph Construction and the Role of DBPedia. 6th DBpedia Community Meeting 2016
- What does that column mean? Semantics and the messy part of data science. Leiden Center for Data Science Symposium. Jan 14, 2016.
- The future of linked data in an open world. Linking Life Science Data: Design to Implementation, and Beyond. Feb 19, 2016
- Panel - NSF Northeast Big Data Hub Kickoff workshop. December 2015
- Data Integration and Transparency tackling the tension. University of Fribourg Informatics Colloquim. June 15, 2015
- Panel - Celebrating 5 years of the altmetrics manifesto. Almetrics Conference 2015
- Machine Reading - what it means for publishers. Society of Scholarly Publishing Conference 2015
- Panel - Knowledge Graphs in Industry. World Wide Web Conference 2015.
- Increasing the Productivity of Scholarship: The Case for Knowledge Graph - Semantics, Analytics, Visualisation: Enhancing Scholarly Data (SAVE-SD) Workshop at World Wide Web Conference 2015.
- Open PHACTS – A Platform For Drug Discovery – European Data Forum. March 18-19, 2014
- Altmetrics: Building a Borader Picture of Impact - Academic Publishing in Europe 9 - January 20 - 21, 2014
- "Open Phacts - Increasing the Effectives of Open Data" - Royal Society Workshop on Revaluing Science in the Digital Age - September 2, 2012
- "Altmetrics: New approaches to measuring and tracking research" - Elsevier Labs Online Lectures on Science Informatics - November 22, 2011

- "Authenticity, provenance, and trust - maintaining the scholarly value chain" - National Academies of Science, Developing Data Attribution and Citation Practices and Standards - August. 22, 2011
- "People as Provenance: The Key to Validating, Reproducing and Measuring New Science" - Second Annual VIVO Conference - August 24, 2011
- "Buy microscopes - don't build them: The scientific demand for HPC as an instrument and how the Semantic Web can help" - 1st Workshop on High-performance Computing for the Semantic Web (HPCSW2011) - May 29 2011
- "Total Impact: Aggregating alternative measures of impact " - Beyond the PDF Workshop, San Diego, CA - January 20, 2011
- "W3C Incubator Group summary and open questions" - UK e-Science Institute, Edinburgh, UK Workshop Understanding Provenance and Linked Open Data - March 20, 2011
- "Seminar on the Findings of the W3C Provenance Incubator Group" -UK e-Science Institute, Edinburgh, UK - February 18, 2011
- "Enabling Data DJs" - Florida Institute for Human and Machine Cognition - June 24, 2010
- "Background Knowledge for Automated Experimentation" - February 23-24, 2010; International Workshop on Automated Experimentation, e-Science Institute, Edinburgh, UK.
- "Nanopublications" at the Scientific Discourse Task Force, of the W3C Health Care and Life Science Interest Group– February 15, 2010
- "CMapTools: Designing a useable and research oriented learning technology". Learning Societies Lab, University of Southampton. June 6, 2006, Southampton, UK.
- "What Happened? Using Provenance for Compliance and Verification." Florida Institute for Human and Machine Cognition. Jan. 11, 2006, USA
- "The Origin of Data: Determining the Provenance of Data Produced by Multi-institutional Applications." At the AI Seminar Series. Intelligent Systems Division. Information Sciences Institute. Nov. 11, 2008.

Invited Participation

- Dagstuhl Seminar: Are Knowledge Graphs Ready for the Real World? Challenges and Perspective. Feb. 4-9, 2024
- Dagstuhl Seminar: Co-organiser. Knowledge Graphs and their Role in the Knowledge Engineering of the 21st Century. September 12–14, 2022.

- Lorenz Center: Automated Composition of Workflows in the Life Sciences . March 9 - 13, 2020
- Dagstuhl Seminar: Dagstuhl Seminar on Knowledge Graphs. Sept. 9 - 14, 2018
- Lorentz Center: Workshop on Online Information Quality. March 26 - 29, 2018
- The Royal Society & KNAW Bilateral Meeting on Responsible Data Science. Feb. 20 - 22, 2018
- Dagstuhl Seminar: Citizen Science: Design and Engagement. July 2 – 5 , 2017
- Observatory for Knowledge Organization Systems Workshop. Feb. 1 - 3, 2017
- Scholarly Commons Working Group Infrastructure and Policy Workshop. Sept. 18 - 20, 2016
- Evolution and Variation of Classification Systems Workshop. March 4-5, 2015
- NSF & Sloan Foundation. Supporting Scientific Discovery through Norms and Practices for Software and Data Citation and Attribution. Jan 29 - 30, 2015
- Dagstuhl Perspectives Workshop: Digital Scholarship and Open Science in Psychology and the Behavioral Sciences - July 19 -2 24, 2015
- Preserving Linked Data Workshops (PRELIDA). April 2 - 4 and subsequently October 17 – 18, 2014
- NSF Computational Challenges in Data Citation. Philadelphia. April 18, 2014
- Lorentz Center. Jointly designing a data FAIRPORT. Jan 13 - 16, 2014
- Dagstuhl Seminar: Semantic Data Management - April 22 - 27, 2012
- Dagstuhl Seminar: Principles of Provenance - February 26 - March 2, 2012
- Data Citation Principles Workshop - IQSS at Harvard University - May 16 - May 17, 2011
- Dagstuhl Perspectives Workshop: The Future of Research Communication - August 15 - 18, 2011
- Elsevier Life Sciences Strategy Meeting - September 21, 2011
- Beyond Impact Workshop - Open Society Foundation and the Wellcome Trust - May 9 - 11, 2011.
- NSF/JSMF Workshop on Mapping of Science and Semantic Web - Indiana University Bloomington - March 4-5, 2010
- Elsevier Workshop on the Journal Workflow of the Future. July 8, 2010

- Subject Matter Expert for AstraZenca Design and Interpretation - Structured Authoring Workshop, Moindal, Sweden - August 31- September 2, 2010
- The Role of Biomedical Informatics in Overcoming Barriers in Cancer Research Workshop. May 21-23, 2008. Hosted by the NIH-National Cancer Institute
- Panel discussion on UK-Sino e-Science Collaborations. UK e-Science All Hands Meeting. September 2005

Memberships

- Phi Kappa Phi
- Association of Computing Machinery (ACM)
- The Institute of Electrical and Electronics Engineers (IEEE)
- Association for the Advancement of Artificial Intelligence (AAAI)

Broader Impact

- 2600 followers on Twitter – `http://twitter.com/pgroth`
- My blog `http://thinklinks.wordpress.com` has had over 30,000 views
- My presentations have been viewed more than 14000 times on slideshare.com
- I have been interviewed by The Times Higher Education Supplement, Science Magazine-Careers, De Volkskrant and published articles in SURF Magazine – the national Dutch magazine for ICT in Higher Education.
- I have participated in a number of workshops organized by policy makers such as the US National Institutes of Health, the Welcome Trust, and JISC.
- Article "Reducing Friction" in SURF Magazine for ICT in Dutch Higher Education
- Article in Chemistry World ' Data Ex Machina'
- Multiple articles for Elsevier Connect.

Publications

According to Google Scholar[1](as of April 2024), my h-index is 53 and my publications have received over 26000 citations. My top ten publications have each received more than 270 citations. All publications listed below except reports are peer-reviewed.

Books & Theses

1. Luc Moreau and Paul Groth. *Provenance: An Introduction to PROV*. Morgan & Claypool, 2013

2. Grigoris Antoniou, Paul Groth, Frank van Harmelen, and Rinke Hoekstra. *A Semantic Web Primer*. MIT press, third edition, 2012

3. Paul T Groth. *The Origin of Data: Enabling the Determination of Provenance in Multi-institutional Scientific Systems through the Documentation of Processes*. Ph.d thesis, University of Southampton, sep 2007

Edited Volumes

4. Paul Groth, Maria-Esther Vidal, Fabian M. Suchanek, Pedro A. Szekely, Pavan Kapanipathi, Catia Pesquita, Hala Skaf-Molli, and Minna Tamper, editors. *The Semantic Web - 19th International Conference, ESWC 2022, Hersonissos, Crete, Greece, May 29 - June 2, 2022, Proceedings*, volume 13261 of *Lecture Notes in Computer Science*. Springer, 2022

5. Mehwish Alam, Paul Groth, Victor de Boer, Tassilo Pellegrini, Harshvardhan J. Pandit, Elena Montiel, Víctor Rodríguez Doncel, Barbara McGillivray, and Albert Meroño-Peñuela, editors. *Further with Knowledge Graphs: Proceedings of the 17th International Conference on Semantic Systems, 6–9 September 2021, Amsterdam, The Netherlands*, volume 53 of *Studies on the Semantic Web*. IOS Press, August 2021

6. Eva Blomqvist, Paul Groth, Victor de Boer, Tassilo Pellegrini, Mehwish Alam, Tobias Käfer, Peter Kieseberg, Sabrina Kirrane, Albert Meroño-Peñuela, and Harshvardhan J. Pandit. *Semantic Systems. In the Era of Knowledge Graphs - 16th International Conference on Semantic Systems, SEMANTiCS 2020, Amsterdam, The Netherlands, September 7-10, 2020, Proceedings*, volume 12378 of *Lecture Notes in Computer Science*. Springer, 2020

7. Paul Groth and Michel Dumontier. Introduction – fair data, systems and analysis. *Data Science*, 3(1):1–2, June 2020

[1]http://scholar.google.com/citations?user=0tHSHCIAAAAJ

8. Paul Groth, Elena Simperl, Alasdair Gray, Marta Sabou, Markus Krötzsch, Freddy Lecue, Fabian Flöck, and Yolanda Gil, editors. *The Semantic Web – ISWC 2016*, volume 9981 of *Lecture Notes in Computer Science*. Springer International Publishing, 2016

9. Marcelo Arenas, Oscar Corcho, Elena Simperl, Markus Strohmaier, Mathieu D'Aquin, Kavitha Srinivas, Paul Groth, Michel Dumontier, Jeff Heflin, Krishnaprasad Thirunarayan, Krishnaprasad Thirunarayan, and Steffen Staab, editors. *The Semantic Web - ISWC 2015*, volume 9366 of *Lecture Notes in Computer Science*. Springer International Publishing, Cham, 2015

10. Peter Mika, Tania Tudorache, Abraham Bernstein, Chris Welty, Craig Knoblock, Denny Vrandečić, Paul Groth, Natasha Noy, Krzysztof Janowicz, and Carole Goble, editors. *The Semantic Web – ISWC 2014*, volume 8796 of *Lecture Notes in Computer Science*. Springer International Publishing, Cham, 2014

11. Harith Alani, Lalana Kagal, Achille Fokoue, Paul Groth, Chris Biemann, Josiane Xavier Parreira, Lora Aroyo, Natasha Noy, Chris Welty, and Krzysztof Janowicz, editors. *The Semantic Web – ISWC 2013*, volume 8218 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013

12. Paul Groth and James Frew, editors. *Provenance and Annotation of Data and Processes: 4th International Provenance and Annotation Workshop, IPAW 2012*, volume 7525 of *Lecture Notes in Computer Science,*. Springer, 2012

Journal Publications

13. Erkan Karabulut, Salvatore F. Pileggi, Paul Groth, and Victoria Degeler. Ontologies in digital twins: A systematic literature review. *Future Generation Computer Systems*, 153:442–456, April 2024

14. Bradley P. Allen, Lise Stork, and Paul Groth. Knowledge Engineering Using Large Language Models. *Transactions on Graph Data and Knowledge*, 1(1):3:1–3:19, dec 2023

15. Daniel Daza, Dimitrios Alivanistos, Payal Mitra, Thom Pijnenburg, Michael Cochez, and Paul Groth. Bioblp: a modular framework for learning on multimodal biomedical knowledge graphs. *Journal of Biomedical Semantics*, 14(1), December 2023

16. Stefan Grafberger, Shubha Guha, Paul Groth, and Sebastian Schelter. Mlwhatif: What if you could stop re-implementing your machine learning pipeline analyses over and over? *Proc. VLDB Endow.*, 16(12):4002–4005, sep 2023

17. James Nevin, Paul Groth, and Michael Lees. An approach for analysing the impact of data integration on complex network diffusion models. *Journal of Complex Networks*, 11(4):cnad025, 07 2023

18. Enrico Daga and Paul Groth. Data journeys: Explaining AI workflows through abstraction. *Semantic Web*, pages 1–27, June 2023

19. Sami Jullien, Mozhdeh Ariannezhad, Paul Groth, and Maarten de Rijke. A simulation environment and reinforcement learning method for waste reduction. *Transactions on Machine Learning Research*, 2023

20. Lucas Prieto, Jeroen Den Boef, Paul Groth, and Joran Cornelisse. Parameter efficient node classification on homophilic graphs. *Transactions on Machine Learning Research*, 2023

21. Melika Ayoughi, Pascal Mettes, and Paul Groth. Self-contained entity discovery from captioned videos. *ACM Trans. Multimedia Comput. Commun. Appl.*, 19(5s), jun 2023

22. Thiviyan Thanapalasingam, Lucas van Berkel, Peter Bloem, and Paul Groth. Relational graph convolutional networks: a closer look. *PeerJ Computer Science*, 8:e1073, November 2022

23. Gytė Tamašauskaitė and Paul Groth. Defining a knowledge graph development process through a systematic review. *ACM Transactions on Software Engineering and Methodology*, 32(1), feb 2022

24. Stian Soiland-Reyes, Genís Bayarri, Pau Andrio, Robin Long, Douglas Lowe, Ania Niewielska, Adam Hospital, and Paul Groth. Making Canonical Workflow Building Blocks Interoperable across Workflow Languages. *Data Intelligence*, pages 1–16, 03 2022

25. Stefan Grafberger, Paul Groth, Julia Stoyanovich, and Sebastian Schelter. Data distribution debugging in machine learning pipelines. *The VLDB Journal*, January 2022

26. Max Schröder, Susanne Staehlke, Paul Groth, J. Barbara Nebe, Sascha Spors, and Frank Krüger. Structure-based knowledge acquisition from electronic lab notebooks for research data provenance documentation. *Journal of Biomedical Semantics*, 13(1), January 2022

27. Stian Soiland-Reyes, Peter Sefton, Mercè Crosas, Leyla Jael Castro, Frederik Coppens, José M. Fernández, Daniel Garijo, Björn Grüning, Marco La Rosa, Simone Leo, and et al. Packaging research artefacts with ro-crate. *Data Science*, page 1–42, January 2022

28. James Nevin, Michael Lees, and Paul Groth. The non-linear impact of data handling on network diffusion models. *Patterns*, page 100397, November 2021

29. Weixin Zeng, Xiang Zhao, Jiuyang Tang, Xuemin Lin, and Paul Groth. Reinforcement learning–based collective entity alignment with adaptive features. *ACM Trans. Inf. Syst.*, 39(3), may 2021

30. Laura Koesten, Pavlos Vougiouklis, Elena Simperl, and Paul Groth. Dataset reuse: Toward translating principles to practice. *Patterns*, page 100136, November 2020

31. Laura Koesten, Kathleen Gregory, Paul Groth, and Elena Simperl. Talking datasets – understanding data sensemaking behaviours. *International Journal of Human-Computer Studies*, 146:102562, 2021

32. Friso Selten, Cameron Neylon, Chun-Kai Huang, and Paul Groth. A longitudinal analysis of university rankings. *Quantitative Science Studies*, 1(3):1109–1135, 2020

33. Kathleen Gregory, Paul Groth, Andrea Scharnhorst, and Sally Wyatt. Lost or Found? Discovering Data Needed for Research. *Harvard Data Science Review*, 2(2), apr 30 2020

34. Paul Groth, Helena Cousijn, Tim Clark, and Carole Goble. FAIR data reuse – the path through data citation. *Data Intelligence*, 2(1-2):78–86, January 2020

35. Kathleen M Gregory, Helena Cousijn, Paul Groth, Andrea Scharnhorst, and Sally Wyatt. Understanding data search as a socio-technical practice. *Journal of Information Science*, 46(4):459–475, April 2019

36. Kathleen Gregory, Paul Groth, Helena Cousijn, Andrea Scharnhorst, and Sally Wyatt. Searching data: A review of observational data retrieval practices in selected disciplines. *Journal of the Association for Information Science and Technology*, 70(5):419–432, March 2019

37. Adriane Chapman, Elena Simperl, Laura Koesten, George Konstantinidis, Luis-Daniel Ibáñez, Emilia Kacprzak, and Paul Groth. Dataset search: a survey. *The VLDB Journal*, 29(1):251–272, August 2020

38. P. Groth and J. Cox. Indicators for the use of robotic labs in basic biomedical research: A literature analysis. *PeerJ*, 2017(11), 2017

39. Manolis Stamatogiannakis, Elias Athanasopoulos, Herbert Bos, and Paul Groth. PROV 2R: Practical Provenance Analysis of Unstructured Processes. *ACM Transactions on Internet Technology*, 17(4):1–24, aug 2017

40. M. Wylot, P. Cudre-Mauroux, M. Hauswirth, and P. Groth. Storing, Tracking, and Querying Provenance in Linked Data. *IEEE Transactions on Knowledge and Data Engineering*, 29(8), 2017

41. Michael Lauruhn and Paul Groth. Sources of Change for Modern Knowledge Organization Systems. *Knowledge Organization*, 43(8):622–629, 2016

42. Davide Ceolin, Paul Groth, Valentina Maccatrozzo, Wan Fokkink, Willem Robert Van Hage, and Archana Nottamkandath. Combining User Reputation and Provenance Analysis for Trust Assessment. *Journal of Data and Information Quality*, 7(1-2):1–28, jan 2016

43. Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim

Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J.G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A.C 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3:160018, mar 2016

44. Michel Dumontier, Alasdair J.G. Gray, M. Scott Marshall, Vladimir Alexiev, Peter Ansell, Gary Bader, Joachim Baran, Jerven T. Bolleman, Alison Callahan, José Cruz-Toledo, Pascale Gaudet, Erich A. Gombocz, Alejandra N. Gonzalez-Beltran, Paul Groth, Melissa Haendel, Maori Ito, Simon Jupp, Nick Juty, Toshiaki Katayama, Norio Kobayashi, Kalpana Krishnaswami, Camille Laibe, Nicolas Le Novère, Simon Lin, James Malone, Michael Miller, Christopher J. Mungall, Laurens Rietveld, Sarala M. Wimalaratne, and Atsuko Yamaguchi. The health care and life sciences community profile for dataset descriptions. *PeerJ*, 4:e2331, aug 2016

45. Luc Moreau, Paul Groth, James Cheney, Timothy Lebo, and Simon Miles. The rationale of PROV. *Web Semantics: Science, Services and Agents on the World Wide Web*, 35(P4):235–257, dec 2015

46. Sara Magliacane, Philip Stutz, Paul Groth, and Abraham Bernstein. foxPSL: A Fast, Optimized and eXtended PSL implementation. *International Journal of Approximate Reasoning*, jun 2015

47. Antonis Loizou, Renzo Angles, and Paul Groth. On the formulation of performant SPARQL queries. *Web Semantics: Science, Services and Agents on the World Wide Web*, nov 2015

48. Paul Groth, Antonis Loizou, Alasdair J.G. Gray, Carole Goble, Lee Harland, and Steve Pettifer. API-centric Linked Data Integration: The Open PHACTS Discovery Platform Case Study. *Web Semantics: Science, Services and Agents on the World Wide Web*, 29:12–18, apr 2014

49. Alyssa Goodman, Alberto Pepe, Alexander W Blocker, Christine L Borgman, Kyle Cranmer, Merce Crosas, Rosanne Di Stefano, Yolanda Gil, Paul Groth, Margaret Hedstrom, David W Hogg, Vinay Kashyap, Ashish Mahabal, Aneta Siemiginowska, and Aleksandra Slavkovic. Ten simple rules for the care and feeding of scientific data. *PLoS computational biology*, 10(4):e1003542, apr 2014

50. Christine Chichester, Pascale Gaudet, Oliver Karch, Paul Groth, Lydie Lane, Amos Bairoch, Barend Mons, and Antonis Loizou. Querying neXtProt nanopublications and their value for insights on sequence variants and tissue expression. *Web Semantics: Science, Services and Agents on the World Wide Web*, 29:3–11, 2014

51. Christine Chichester, Daniela Digles, Ronald Siebes, Antonis Loizou, Paul Groth, and Lee Harland. Drug discovery FAQs: workflows for answering multidomain drug discovery questions. *Drug discovery today*, nov 2014

52. Alasdair J. G. Gray, Paul Groth, Antonis Loizou, Sune Askjaer, Christian Brenninkmeijer, Kees Burger, Christine Chichester, Chris T. Evelo, Carole Goble, Lee Harland, Steve Pettifer, Mark Thompson, Andra Waagmeester, and Antony J. Williams. Applying linked data approaches to pharmacology: Architectural decisions and implementation. *Semantic Web*, 5(2):101–113, 2014

53. Paul Groth. Transparency and Reliability in the Data Supply Chain. *Internet Computing, IEEE*, 17(2):2–4, 2013

54. Paul Groth. The Knowledge-Remixing Bottleneck. *IEEE Intelligent Systems*, 28(5):44–48, sep 2013

55. Paul Groth, Yolanda Gil, James Cheney, and Simon Miles. Requirements for Provenance on the Web. *International Journal of Digital Curation*, 7(1), 2012

56. Yolanda Gil, Varun Ratnakar, Timothy Chklovski, Paul Groth, and Denny Vrandecic. Capturing Common Knowledge about Tasks. *ACM Transactions on Interactive Intelligent Systems*, 2(3):1–35, sep 2012

57. Jason Priem, Paul Groth, and Dario Taraborelli. The Altmetrics Collection. *PLoS ONE*, 7(11):e48753, nov 2012

58. F. Harmelen, G. Kampis, K. Börner, P. Besselaar, E. Schultes, C. Goble, P. Groth, B. Mons, S. Anderson, S. Decker, C. Hayes, T. Buecheler, and D. Helbing. Theoretical and technological building blocks for an innovation accelerator. *The European Physical Journal Special Topics*, 214(1):183–214, dec 2012

59. Antony J. Williams, Lee Harland, Paul Groth, Stephen Pettifer, Christine Chichester, Egon L. Willighagen, Chris T. Evelo, Niklas Blomberg, Gerhard Ecker, Carole Goble, and Barend Mons. Open PHACTS: Semantic interoperability for drug discovery. *Drug Discovery Today*, 17(21-22):1188–1198, jun 2012

60. Yolanda Gil, Varun Ratnakar, Jihie Kim, Pedro Antonio Gonzalez-Calero, Paul Groth, Joshua Moody, and Ewa Deelman. Wings: Intelligent Workflow-Based Design of Computational Experiments. *IEEE Intelligent Systems*, 26(1):62 – 72, 2011

61. Simon Miles, Paul Groth, Steve Munroe, and Luc Moreau. PrIMe: A Methodology for Developing Provenance-Aware Applications. *ACM Transactions on Software Engineering and Methodology*, 20(3):1–42, aug 2011

62. Luc Moreau, Ben Clifford, Juliana Freire, Joe Futrelle, Yolanda Gil, Paul Groth, Natalia Kwasnikowska, Simon Miles, Paolo Missier, Jim Myers, Beth Plale, Yogesh Simmhan, Eric Stephan, and Jan Van den Bussche. The Open Provenance Model core specification (v1.1). *Future Generation Computer Systems*, 27(6):743–756, jul 2011

63. Paul Groth and Luc Moreau. Representing distributed systems using the Open Provenance Model. *Future Generation Computer Systems*, 27(6):757–765, oct 2011

64. Christophe Guéret, Shenghui Wang, Paul Groth, and Stefan Schlobach. Multi-scale analysis of the Web of Data: a challenge to the Complex System's community. *Advances in Complex Systems*, 14(4):589–607, 2011

65. Barend Mons, Herman van Haagen, Christine Chichester, Peter-Bram 't Hoen, Johan T den Dunnen, Gertjan van Ommen, Erik van Mulligen, Bharat Singh, Rob Hooft, Marco Roos, Joel Hammond, Bruce Kiesel, Belinda Giardine, Jan Velterop, Paul Groth, and Erik Schultes. The value of data. *Nature Genetics*, 43(4):281–3, jan 2011

66. Yolanda Gil and Paul Groth. Using Provenance in the Semantic Web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(2):147–148, 2011

67. Paul Groth, Andrew Gibson, and Jan Velterop. The anatomy of a nanopublication. *Information Services & Use*, 30:51–56, 2010

68. Paul Groth, Simon Miles, and Luc Moreau. A Model of Process Documentation to Determine Provenance in Mash-ups. *Transactions on Internet Technology (TOIT)*, 9(1), 2009

69. Paul Groth and Luc Moreau. Recording Process Documentation for Provenance. *IEEE Transactions on Parallel and Distributed Systems*, 20(9):1246–1259, 2009

70. Simon Miles, Paul Groth, Steve Munroe, Sheng Jiang, Thibaut Assandri, and Luc Moreau. Extracting Causal Graphs from an Open Provenance Data Model. *Concurrency and Computation: Practice and Experience*, 20(5):577–586, 2008

71. Jie Xu, Paul Townend, Nik Looker, and Paul T Groth. FT-Grid: a system for achieving fault tolerance in grids. *Concurrency and Computation: Practice and Experience*, 20(3):297–309, 2008

72. Luc Moreau, Paul Groth, Simon Miles, Javier Vazquez-Salceda, John Ibbotson, Sheng Jiang, Steve Munroe, Omer Rana, Andreas Schreiber, Victor Tan, and Laszlo Varga. The provenance of electronic data. *Communications of the ACM*, 51(4):52–58, 2008

73. Simon Miles, Paul Groth, Ewa Deelman, Karan Vahi, Gaurang Mehta, and Luc Moreau. Provenance: The Bridge Between Experiments and Data. *Computing in Science and Engineering*, 10:38–46, 2008

74. David W. Eccles and Paul T. Groth. Wolves, bees, and football: Enhancing coordination in sociotechnological problem solving systems through the study of human and animal groups. *Computers in Human Behavior*, 23(6):2778–2790, nov 2007

75. Simon Miles, Sylvia C Wong, Weijian Fang, Paul Groth, Klaus-Peter Zauner, and Luc Moreau. Provenance-Based Validation of e-Science Experiments. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 5(1):28–38, 2007

76. Simon Miles, Paul Groth, Miguel Branco, and Luc Moreau. The Requirements of Using Provenance in e-Science Experiments. *Journal of Grid Computing*, 5(1):1–25, 2007

77. Luc Moreau, Bertram Ludäscher, Ilkay Altintas, Roger S Barga, Shawn Bowers, Steven Callahan, George Chin Jr., Ben Clifford, Shirley Cohen, Sarah Cohen-Boulakia, Susan Davidson, Ewa Deelman, Luciano Digiampietri, Ian Foster, Juliana Freire, James Frew, Joe Futrelle, Tara Gibson, Yolanda Gil, Carole Goble, Jennifer Golbeck, Paul Groth, David A Holland, Sheng Jiang, Jihie Kim, David Koop, Ales Krenek, Timothy McPhillips, Gaurang Mehta, Simon Miles, Dominic Metzger, Steve Munroe, Jim Myers, Beth Plale, Norbert Podhorszki, Varun Ratnakar, Emanuele Santos, Carlos Scheidegger, Karen Schuchardt, Margo Seltzer, Yogesh L Simmhan, Claudio Silva, Peter Slaughter, Eric Stephan, Robert Stevens, Daniele Turi, Huy Vo, Mike Wilde, Jun Zhao, and Yong Zhao. The First Provenance Challenge. *Concurrency and Computation: Practice and Experience*, 20:409–418, 2007

78. David W Eccles and Paul T Groth. Agent coordination and communication in sociotechnological systems: Design and measurement issues. *Interacting with Computers*, 18:1170–1185, 2006

79. David W. Eccles and Paul Groth. Problem Solving Systems Theory: Implications for the Design of Socio-technological Systems. *Technology Instruction Cognition and Learning*, 3:323–343, 2006

Book Chapters

80. Kathleen Gregory, Paul Groth, Andrea Scharnhorst, and Sally Wyatt. *The Mysterious User of Research Data: Knitting Together Science and Technology Studies with Information and Computer Science*, pages 191–211. Springer International Publishing, Cham, 2023

81. Manfred Hauswirth, Marcin Wylot, Martin Grund, Paul Groth, and Philippe Cudré-Mauroux. Linked Data Management. In *Handbook of Big Data Technologies*, pages 307–338. Springer International Publishing, Cham, 2017

82. Ewa Deelman, Bruce Berriman, Ann Chervenak, Oscar Corcho, Paul Groth, and Luc Moreau. Metadata and Provenance Management. In Arie Shoshani and Doron Rotem, editors, *Scientific Data Management: Challenges, Technology, and Deployment*, chapter 12, pages 433–466. Chapman & Hall, 1 edition, 2010

83. Paul Groth, Steve Munroe, Simon Miles, and Luc Moreau. Applying the Provenance Data Model to a Bioinformatics Case. In Lucio Grandinetti, editor, *HPC and Grids in Action*. IOS Press, 2008

Conference Publications

84. Fina Polat, Ilaria Tiddi, Paul Groth, and Piek Vossen. Improving graph-to-text generation using cycle training. In Sara Carvalho, Anas Fahad Khan, Ana Ostroški Anić, Blerina Spahiu, Jorge Gracia, John P. McCrae, Dagmar Gromann, Barbara Heinisch, and Ana Salgado, editors, *Proceedings of the 4th Conference on Language, Data and Knowledge*, pages 256–261, Vienna, Austria, September 2023. NOVA CLUNL, Portugal

85. James Nevin, Paul Groth, and Michael Lees. Data integration landscapes: The case for non-optimal solutions in network diffusion models. In Jiří Mikyška, Clélia de Mulatier, Maciej Paszynski, Valeria V. Krzhizhanovskaya, Jack J. Dongarra, and Peter M.A. Sloot, editors, *Computational Science – ICCS 2023*, pages 494–508, Cham, 2023. Springer Nature Switzerland

86. Stefan Grafberger, Paul Groth, and Sebastian Schelter. Automating and optimizing data-centric what-if analyses on native machine learning pipelines. *Proceedings of the ACM on Management of Data (SIGMOD)*, 1(2):1–26, June 2023

87. Madelon Hulsebos, undefinedagatay Demiralp, and Paul Groth. Gittables: A large-scale corpus of relational tables. *Proceedings of the ACM on Management of Data (SIGMOD)*, 1(1):1–17, May 2023

88. Corey A. Harper, Ron Daniel, and Paul Groth. Question Answering with Additive Restrictive Training (QuAART): Question Answering for the Rapid Development of New Knowledge Extraction Pipelines. In Oscar Corcho, Laura Hollink, Oliver Kutz, Nicolas Troquard, and Fajar J. Ekaputra, editors, *Knowledge Engineering and Knowledge Management (EKAW)*, volume 13514, pages 51–65, Cham, September 2022. Springer International Publishing

89. Xue Li, Sara Magliacane, and Paul Groth. The challenges of cross-document coreference resolution for email. In *Proceedings of the 11th on Knowledge Capture Conference*, K-CAP '21, page 273–276, New York, NY, USA, 2021. Association for Computing Machinery

90. Daniel Daza, Michael Cochez, and Paul Groth. Inductive entity representations from text via link prediction. In *Proceedings of The Web Conference*, April 2021

91. Corey Harper, Jessica Cox, Curt Kohler, Antony Scerri, Ron Daniel Jr., and Paul Groth. SemEval-2021 task 8: MeasEval – extracting counts and measurements and their related contexts. In *Proceedings of the 15th International Workshop on Semantic Evaluation (SemEval-2021)*, pages 306–316, Online, August 2021. Association for Computational Linguistics

92. Marieke van Erp and Paul Groth. Towards entity spaces. In *Proceedings of The 12th Language Resources and Evaluation Conference*, pages 2129–2137, Marseille, France, May 2020. European Language Resources Association

93. Paul T Groth, Michael Lauruhn, Antony Scerri, and Ron Daniel. Open Information Extraction on Scientific Text: An Evaluation. In *Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018*, pages 3414–3423, 2018

94. Marcin Wylot, Philippe Cudre-Mauroux, and Paul Groth. Executing Provenance-Enabled Queries over Web Data. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1275–1285, may 2015

95. Marcin Wylot, Philippe Cudré-Mauroux, and Paul Groth. TripleProv: Efficient Processing of Lineage Queries over a Native RDF Store. In *Proceedings of the 23rd International World Wide Web Conference*, 2014

96. Anca Dumitrache, Paul Groth, and Peter van den Besselaar. Identifying Research Talent Using Web-Centric Databases. In *Web Science 2013*, 2013

97. Fabian Eikelboom, Paul Groth, Victor De Boer, and Laura Hollink. A Comparison between Online and Offline Prayer. In *Web Science 2013*, number 2, 2013

98. Paul Groth and David A. Shamma. Spinning data: remixing live data like a music dj. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13*, pages 3063–3066, New York, New York, USA, apr 2013. ACM Press

99. Philippe Cudré-Mauroux, Iliya Enchev, Sever Fundatureanu, Paul Groth, Albert Haque, Andreas Harth, Felix Leif Keppmann, Daniel P Miranker, Juan Sequeda, and Marcin Wylot. NoSQL Databases for RDF: An Empirical Evaluation. In Harith Alani, Lalana Kagal, Achille Fokoue, Paul T Groth, Chris Biemann, Josiane Xavier Parreira, Lora Aroyo, Natasha F Noy, Chris Welty, and Krzysztof Janowicz, editors, *International Semantic Web Conference (2)*, volume 8219 of *Lecture Notes in Computer Science*, pages 310–325. Springer, 2013

100. Christophe Guéret, Paul Groth, Claus Stadler, and Jens Lehman. Assessing Linked Data Mappings using Network Measures. In Elena Simperl, Philipp Cimiano, Axel Polleres, Óscar Corcho, and Valentina Presutti, editors, *9th Extended Semantic Web Conference, ESWC 2012*, volume 7295 of *Lecture Notes in Computer Science*, pages 87–102, Heraklion, Crete, Greece, may 2012. Springer

101. Shenghui Wang and Paul Groth. A Framework for Longitudinal Influence Measurement between Communication Content and Social Networks. In Toby Walsh, editor, *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI'11)*, pages 2758–2763, Barcelona, Catalonia, Spain, jul 2011. AAAI Press

102. Paul Groth and Thomas Gurney. Studying Scientific Discourse on the Web using Bibliometrics: A Chemistry Blogging Case Study. In *WebSci10: Extending the Frontiers of Society On-Line*, page In Press, Raleigh, NC, 2010

103. Bruce Berriman, Ewa Deelman, Paul Groth, and Gideon Juve. The Application of Cloud Computing to the Creation of Image Mosaics and Management of Their

Provenance. In *SPIE Conference 7740: Software and Cyberinfrastructure for Astronomy*, 2010

104. Shenghui Wang and Paul Groth. Measuring the dynamic bi-directional influence between content and social networks. In *The 9th International Semantic Web Conference (ISWC 2010)*, pages 814–829, 2010

105. Christophe Guéret, Paul Groth, Frank Van Harmelen, and Stefan Schlobach. Finding the Achilles Heel of the Web of Data : using network analysis for link-recommendation. In *The 9th International Semantic Web Conference (ISWC 2010)*, 2010

106. Yolanda Gil, Paul Groth, Varun Ratnakar, and Christian Fritz. Expressive Reusable Workflow Templates. In *Proceedings of the Fifth IEEE International Conference on e-Science*, Oxford, UK, 2009

107. Paul T Groth. A Distributed Algorithm for Determining the Provenance of Data. In *Proceedings of the fourth IEEE International Conference on e-Science (e-Science'08)*, 2008

108. Simon Miles, Ewa Deelman, Paul Groth, Karan Vahi, Gaurang Mehta, and Luc Moreau. Connecting Scientific Data to Scientific Experiments with Provenance. In *Proceedings of the third IEEE International Conference on e-Science and Grid Computing (e-Science'07)*, Bangalore, India, 2007

109. Paul Townend, Paul Groth, and Jie Xu. A Provenance-Aware Weighted Fault Tolerance Scheme for Service-Based Applications. In *Proc. of the 8th IEEE International Symposium on Object-oriented Real-time distributed Computing (ISORC 2005)*, may 2005

110. Sylvia C Wong, Simon Miles, Weijian Fang, Paul Groth, and Luc Moreau. Provenance-based Validation of E-Science Experiments. In *Proceedings of 4th International Semantic Web Conference (ISWC'05)*, volume 3729 of *Lecture Notes in Computer Science*, pages 801–815, Galway, Ireland, nov 2005. Springer-Verlag

111. Paul Groth, Simon Miles, Weijian Fang, Sylvia C Wong, Klaus-Peter Zauner, and Luc Moreau. Recording and Using Provenance in a Protein Compressibility Experiment. In *Proceedings of the 14th IEEE International Symposium on High Performance Distributed Computing (HPDC'05)*, jul 2005

112. Paul Groth, Michael Luck, and Luc Moreau. A protocol for recording provenance in service-oriented Grids. In *Proceedings of the 8th International Conference on Principles of Distributed Systems (OPODIS'04)*, Grenoble, France, 2004

113. David W Eccles and Paul T Groth. Wolves, football, and ambient computing: facilitating collaboration in problem solving systems through the study of human and animal groups. In *Proceedings of the Third Nordic conference on Human-Computer interaction*, pages 269–275. ACM Press, oct 2004

114. Niranjan Suri, Jeffrey M Bradshaw, Marco M Carvalho, Thomas B Cowin, Maggie R Breedy, Paul T Groth, and Raul Saavedra. Agile Computing: Bridging the Gap between Grid Computing and Ad-hoc Peer-to-Peer Resource Sharing. In *Proceedings of the 3st International Symposium on Cluster Computing and the Grid*, page 618. IEEE Computer Society, 2003

115. Niranjan Suri, Jeffrey Bradshaw, Andrzej Uszok, Maggie Breedy, Marco Carvalho, Paul Groth, Renia Jeffers, Matt Johnson, Shri Kulkarni, James Lott, Mark Burstein, Brett Benyo, and David Diller. Towards DAML-based policy enforcement for semantic data transformation and filtering in multi-agent systems. In *Proceedings of the Second International joint conference on Autonomous agents and multiagent systems (AAMAS 2003)*, pages 1132–1133. ACM Press, 2003

116. Niranjan Suri, Paul T Groth, and Jeffrey M Bradshaw. While You're Away: A System for Load-Balancing and Resource Sharing Based on Mobile Agents. In *Proceedings of the 1st International Symposium on Cluster Computing and the Grid (CCGrid'01)*, page 470. IEEE Computer Society, 2001

117. Niranjan Suri, Jeffrey Bradshaw, Maggie R Breedy, Paul T Groth, Gregory A Hill, and Renia Jeffers. Strong Mobility and Fine-Grained Resource Control in NOMADS. In Friedemann Mattern David Kotz, editor, *Proceedings of the Second International Symposium on Agent Systems and Applications and Fourth International Symposium on Mobile Agents, ASA/MA 2000, Zurich, Switzerland*, volume 1882 / 200 of *Lecture Notes in Computer Science*, pages 2–15. Springer-Verlag, 2000

118. Niranjan Suri, Jeffrey M Bradshaw, Maggie R Breedy, Paul T Groth, Gregory A Hill, Renia Jeffers, Timothy S Mitrovich, Brian R Pouliot, and David S Smith. NOMADS: toward a strong and safe mobile agent system. In *Proceedings of the fourth international conference on Autonomous agents*, pages 163–164. ACM Press, 2000

Workshop Publications

119. Stefan Grafberger, Bojan Karlaš, Paul Groth, and Sebastian Schelter. Towards declarative systems for data-centric machine learning. In *Proceedings of the Data-Centric Machine Learning Research work- shop (DMLR) at ICML, 2023*, 2023

120. Xue Li, Fina Polat, and Paul Groth. Do instruction-tuned large language models help with relation extraction? In *KBC-LM'23: Knowledge Base Construction from Pre-trained Language Models workshop at ISWC 2023*, November 2023

121. Xue Li, Anthony Hughes, Majlinda Llugiqi, Fina Polat, Paul Groth, and Fajar J. Ekaputra. Knowledge-centric prompt composition for knowledge base construction from pre-trained language models. In *KBC-LM'23: Knowledge Base Construction from Pre-trained Language Models workshop at ISWC 2023*, November 2023

122. Erkan Karabulut, Victoria Degeler, and Paul Groth1. Semantic association rule learning from time series data and knowledge graphs. In *SemIIM'23: 2nd International Workshop on Semantic Industrial Information Modelling co-located with 22nd International Semantic Web Conference (ISWC 2023)*, November 2023

123. Qingzhi Hu, Daniel Daza, Laurens Swinkels, Kristina Usaite, Robbert-Jan 't Hoen, and Paul Groth. Harnessing the web and knowledge graphs for automated impact investing scoring. In *KDD Fragile Earth Workshop*, Aug 2023

124. Valentina Anita Carriero, Paul Groth, and Valentina Presutti. Towards improving wikidata reuse with emerging patterns. In Lucie-Aimée Kaffee, Simon Razniewski, Gabriel Amaral, and Kholoud Saad Alghamdi, editors, *Proceedings of the 3rd Wikidata Workshop 2022 co-located with the 21st International Semantic Web Conference (ISWC2022), Virtual Event, Hanghzou, China, October 2022*, volume 3262 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2022

125. Stefan Grafberger, Paul Groth, and Sebastian Schelter. Towards data-centric what-if analysis for native machine learning pipelines. In *Proceedings of the Sixth Workshop on Data Management for End-To-End Machine Learning*, DEEM '22, New York, NY, USA, june 2022. Association for Computing Machinery

126. Daniel Daza, Michael Cochez, and Paul Groth. SlotGAN: Detecting mentions in text via adversarial distant learning. In *Proceedings of the Sixth Workshop on Structured Prediction for NLP*, pages 32–39, Dublin, Ireland, May 2022. Association for Computational Linguistics

127. Kinga Szarkowska, Veronique Moore, Pierre-Yves Vandenbussche, and Paul Groth. Quality Assessment of Knowledge Graph Hierarchies using KG-BERT. In Mehwish Alam, Davide Buscaldi, Michael Cochez, Francesco Osborne, Diego Reforgiato Recupero, and Harald Sack, editors, *Proceedings of the Workshop on Deep Learning for Knowledge Graphs (DL4KG 2021)*, volume 3034 of *CEUR Workshop Proceedings*, Virtual Conference, online, October 2021. CEUR

128. Jeroen Den Boef, Joran Cornelisse, and Paul Groth. GraphPOPE: Retaining Structural Graph Information Using Position-aware Node Embeddings. In Mehwish Alam, Davide Buscaldi, Michael Cochez, Francesco Osborne, Diego Reforgiato Recupero, and Harald Sack, editors, *Proceedings of the Workshop on Deep Learning for Knowledge Graphs (DL4KG 2021)*, volume 3034 of *CEUR Workshop Proceedings*, Virtual Conference, online, October 2021. CEUR

129. Natasha Shroff, Pierre-Yves Vandenbussche, Véronique Moore, and Paul Groth. Supporting ontology maintenance with contextual word embeddings and maximum mean discrepancy. In Sarra Ben Abbès, Rim Hantach, Philippe Calvez, Davide Buscaldi, Danilo Dessì, Mauro Dragoni, Diego Reforgiato Recupero, and Harald Sack, editors, *Joint Proceedings of the 2nd International Workshop on Deep Learning meets Ontologies and Natural Language Processing (DeepOntoNLP 2021) & 6th International Workshop on Explainable Sentiment Mining and Emotion Detection (X-SENTIMENT 2021) co-located with co-located with 18th Extended Semantic Web Conference 2021, Hersonissos, Greece, June 6th - 7th, 2021 (moved online)*, volume 2918 of *CEUR Workshop Proceedings*, pages 11–19. CEUR-WS.org, 2021

130. Ryan Brate, Paul Groth, and Marieke van Erp. Towards olfactory information extraction from text: A case study on detecting smell experiences in novels. In

Proceedings of the The 4th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, pages 147–155, Online, December 2020. International Committee on Computational Linguistics

131. Mark Berger, Jakub Zavrel, and Paul Groth. Effective distributed representations for academic expert search. In *Proceedings of the First Workshop on Scholarly Document Processing at EMNLP*, pages 56–71, Online, November 2020. Association for Computational Linguistics

132. Manolis Stamatogiannakis, Herbert Bos, and Paul Groth. Pandacap: a framework for streamlining collection of full-system traces. In *Proceedings of the 13th European Workshop on Systems Security*, EuroSec '20, page 1–6, New York, NY, USA, 2020. Association for Computing Machinery

133. Anthi Symeonidou, Viachaslau Sazonau, and Paul Groth. Transfer learning for biomedical named entity recognition with biobert. In *Proceedings of the Posters and Demo Track of the 15th International Conference on Semantic Systems co-located with 15th International Conference on Semantic Systems (SEMANTiCS 2019), Karlsruhe, Germany, September 9th - to - 12th, 2019.*, 2019

134. Paul T. Groth, Antony Scerri, Ron Daniel, and Bradley P. Allen. End-to-end learning for answering structured queries directly over text. In *Proceedings of the Workshop on Deep Learning for Knowledge Graphs (DL4KG2019) Co-located with the 16th Extended Semantic Web Conference 2019 (ESWC 2019), Portoroz, Slovenia, June 2, 2019.*, pages 57–70, 2019

135. Michael Lauruhn, Paul T Groth, Corey A Harper, and Helena F Deus. Use of Internal Testing Data to Help Determine Compensation for Crowdsourcing Tasks. In Anna Lisa Gentile, Lora Aroyo, Gianluca Demartini, and Chris Welty, editors, *Proceedings of the 2nd International Workshop on Augmenting Intelligence with Humans-in-the-Loop*, volume 2169 of *{CEUR} Workshop Proceedings*, pages 27–37. CEUR-WS.org, 2018

136. Manolis Stamatogiannakis, Hasanat Kazmi, Hashim Sharif, Remco Vermeulen, Ashish Gehani, Herbert Bos, and Paul Groth. Trade-Offs in Automatic Provenance Capture. In *International Provenance and Annotation Workshop*, pages 29–41. Springer International Publishing, 2016

137. Paul Groth, Sujit Pal, Darin Mcbeath, Brad Allen, and Ron Daniel. Applying Universal Schemas for Domain Specific Ontology Expansion. In *5th Workshop on Automated Knowledge Base Construction (AKBC)*, 2016

138. Manolis Stamatogiannakis, Paul Groth, and Herbert Bos. Decoupling provenance capture and analysis from execution. In *Proceedings of the 7th USENIX Conference on Theory and Practice of Provenance*. USENIX Association, 2015

139. M Stamatogiannakis, P T Groth, and H J Bos. Looking Inside the Black-Box: Capturing Data Provenance using Dynamic Instrumentation. In *5th International Provenance and Annotation Workshop (IPAW'14)*, 2014

140. A Wibisono, Peter Bloem, G K D de Vries, P T Groth, A Belloum, and M Bubak. Generating Scientific Documentation for Computational Experiments Using Provenance. In *5th International Provenance and Annotation Workshop (IPAW'14)*, 2014

141. R Hoekstra and P T Groth. PROV-O-Viz - Understanding the Role of Activities in Provenance. In *5th International Provenance and Annotation Workshop (IPAW'14)*, 2014

142. Christian Y A Brenninkmeijer, Carole A Goble, Alasdair J G Gray, Paul T Groth, Antonis Loizou, and Steve Pettifer. Including Co-referent URIs in a SPARQL Query. In Olaf Hartig, Juan Sequeda, Aidan Hogan, and Takahide Matsutsuka, editors, *Proceedings of the Fourth International Workshop on Consuming Linked Data, COLD 2013*, volume 1034 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2013

143. Rinke Hoekstra and Paul Groth. Linkitup: Link Discovery for Research Data. In *Proceedings of the AAAI Fall Symposium on Discovery Informatics*, 2013

144. Sara Magliacane and Paul Groth. Repurposing Benchmark Corpora for Reconstructing Provenance. In *Workshop on Semantic Publications (Sepublica 2013) at Extended Semantic Web Conference 2013*, 2013

145. Paul Groth, Yolanda Gil, and Sara Magliacane. Automatic Metadata Annotation through Reconstructing Provenance. In Khalid Belhjjame, José-Manuel Gomez-Perez, Paolo Missier, Satya S. Sahoo, and Jun Zhao, editors, *Third International Workshop on the role of Semantic Web in Provenance Management*, volume 856 of *CEUR Workshop Proceedings*. CEUR Proceedings, may 2012

146. Davide Ceolin, Paul T. Groth, Willem Robert van Hage, Archana Nottamkandath, and Wan Fokkink. Trust Evaluation through User Reputation and Provenance Analysis. In *Proceedings of the 8th International Workshop on Uncertainty Reasoning for the Semantic Web, Boston, USA, November 11, 2012*, volume 900 of *CEUR Workshop Proceedings*, pages 15–26. CEUR-WS.org, nov 2012

147. Christian Brenninkmeijer, Chris Evelo, Carole Goble, Alasdair J G Gray, Paul Groth, Steve Pettifer, Robert Stevens, and Antony J Williams. Scientific Lenses over Linked Data: An approach to support task specific views of the data. A vision. In Tomi Kauppinen, Line C. Pouchard, and Carsten Keßler, editors, *Proceedings of the Second International Workshop on Linked Science 2012 - Tackling Big Data - In conjunction with the International Semantic Web Conference (ISWC2012)*, volume 951 of *CEUR Workshop Proceedings*, pages 2–5, Boston, MA, 2012. CEUR-WS

148. J Priem, C Parra, H Piwowar, P Groth, and A Waagmeester. Uncovering impacts: a case study in using altmetrics tools. In *Workshop on the Semantic Publishing (SePublica 2012) at the 9th Extended Semantic Web Conference*, 2012

149. Paul Groth. The Rise of the Verb. In *What will the Semantic Web look like 10 years from now? In conjunction with the 11th International Semantic Web Conference 2012*, 2012

150. Paul Groth and Yolanda Gil. Linked Data for Network Science. In Tomi Kauppinen, Line C. Pouchard, and Carsten Keßler, editors, *Proceedings of the First International Workshop on Linked Science 2011 in Co*, Bonn, Germany, oct 2011. CEUR Workshop Proceedings

151. Christophe Guéret, Spyros Kotoulas, and Paul Groth. TripleCloud : An infrastructure for exploratory querying over Web-scale RDF data. In *Proceedings of Workshop on Web-scale Knowledge Representation, Retrieval, and Reasoning (Web-KR3 2011)*, page 4, Lyon, 2011

152. Davide Ceolin, Paul Groth, and Willem Robert Van Hage. Calculating the Trust of Event Descriptions using Provenance. In *Proceedings Of The SWPM 2010, Workshop At The 9th International Semantic Web Conference, ISWC-2010*, nov 2010

153. Paul Groth, Simon Miles, Sanjay Modgil, Nir Oren, Michael Luck, and Yolanda Gil. Determining the Trustworthiness of New Electronic Contracts. In *Proceedings of the Tenth Annual Workshop on Engineering Societies in the Agents' World, (ESAW-09)*, Utrecht, The Netherlands, 2009

154. Paul Groth and Yolanda Gil. Scaffolding Instructions to Learn Procedures from Users. In *AAAI 2009 Spring Symposium Agents that Learn from Human Teachers*, 2009

155. Paul Groth and Yolanda Gil. Analyzing the Gap between Workflows and their Natural Language Descriptions. In *Proceedings of the IEEE Third International Workshop on Scientific Workflows (SWF'09)*, page 6, 2009

156. Paul Groth, Ewa Deelman, Gideon Juve, Gaurang Mehta, and Bruce Berriman. A Pipeline-Centric Provenance Model. In *The 4th Workshop on Workflows in Support of Large-Scale Science*, 2009

157. Paul Groth. Exposing Privacy Obligation Policies in Social Networking Sites. In *AAAI 2009 Spring Symposium on Social Semantic Web*, 2009

158. Yolanda Gil, Paul Groth, and Varun Ratnakar. Leveraging Social Networking Sites to Acquire Rich Task Structure. In *Proceedings of the IJCAI 2009 Workshop on User-Contributed Knowledge and Artificial Intelligence: An Evolving Synergy (WikiAI09), held in conjunction with the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI-09)*, Pasadena, CA, 2009

159. Simon Miles, Paul Groth, and Michael Luck. Handling Mitigating Circumstances for Electronic Contracts. In *Proceedings of the AISB 2008 Symposium on Behaviour Regulation in Multi-agent Systems*, pages 37–42, Aberdeen, UK, apr 2008. The Society for the Study of Artificial Intelligence and Simulation of Behaviour

160. Simon Miles, Paul Groth, Steve Munroe, Michael Luck, and Luc Moreau. Agent-PrIMe: Adapting MAS Designs to Build Confidence. In *Proceedings of the 8th International Workshop on Agent Oriented Software Engineering*, 2007

161. Paul Groth, Simon Miles, and Steven Munroe. Principles of High Quality Documentation for Provenance: A Philosophical Discussion. In Luc Moreau and Ian Foster, editors, *Proceedings of Third International Provenance and Annotation Workshop (IPAW'06)*, volume 4145 of *Lecture Notes in Computer Science*, Chicago, IL, 2006. Springer

162. Victor Tan, Paul Groth, Simon Miles, Sheng Jiang, Steve Munroe, Sofia Tsasakou, and Luc Moreau. Security Issues in a SOA-based Provenance System. In Luc Moreau and Ian Foster, editors, *Proceedings of Third International Provenance and Annotation Workshop (IPAW'06)*, volume 4145 of *Lecture Notes in Computer Science*, Chicago, IL, 2006. Springer

163. Paul Groth, Simon Miles, and Luc Moreau. PReServ: Provenance Recording for Services. In *Proceedings of the UK OST e-Science Fourth All Hands Meeting (AHM05)*, sep 2005

164. Sylvia C Wong, Simon Miles, Weijian Fang, Paul Groth, and Luc Moreau. Validation of E-Science Experiments using a Provenance-based Approach. In *Proceedings of Fourth All Hands Meeting (AHM'05)*, Nottingham, 2005

165. Liming Chen, Victor Tan, Fenglian Xu, Alexis Biller, Paul Groth, Simon Miles, John Ibbotson, Michael Luck, and Luc Moreau. A Proof of Concept: Provenance in a Service Oriented Architecture. In *Proceedings of the Fourth All Hands Meeting (AHM'05)*, sep 2005

166. Paul Townend, Paul Groth, Nik Looker, and Jie Xu. FT-Grid: A Fault-Tolerance System for e-Science. In *Proceedings of the UK OST e-Science Fourth All Hands Meeting (AHM05)*, sep 2005

167. Paul Groth, Michael Luck, and Luc Moreau. Formalising a protocol for recording provenance in Grids. In *Proceedings of the UK OST e-Science Second All Hands Meeting 2004 (AHM'04)*, Nottingham, UK, 2004

168. Niranjan Suri, Marco Carvalho, Jeffrey M Bradshaw, Maggie R Breedy, Thomas B Cowin, Paul T Groth, Raul Saavedra, and Andrzej Uszok. Enforcement of Communications Policies in Software Agent Systems through Mobile Code. In *Proceedings of the 4th IEEE International Workshop on Policies for Distributed Systems and Networks*, page 247. IEEE Computer Society, 2003

169. Niranjan Suri, Jeffrey M Bradshaw, Maggie R Breedy, Paul T Groth, Gregory A Hill, Renia Jeffers, and Timothy S Mitrovich. An Overview of the NOMADS Mobile Agent System. In *6th ECOOP WORKSHOP ON MOBILE OBJECT SYSTEMS*, 2000

170. Paul T Groth and Niranjan Suri. CPU Resource Control and Accounting in the NOMADS Mobile Agent System. In *Proceedings of the ACM OOPSLA Workshop on Experiences with Autonomous Mobile Objects and Agent Based Systems, Minneapolis, USA, Oct. 2000.*, 2000

Posters & Demos

171. Stefan Grafberger, Paul Groth, and Sebastian Schelter. Provenance tracking for end-to-end machine learning pipelines. In *Companion Proceedings of the ACM Web Conference 2023*, WWW '23 Companion, page 1512, New York, NY, USA, 2023. Association for Computing Machinery

172. Madelon Hulsebos, Sneha Gathani, James Gale, Isil Dillig, Paul Groth, and Çagatay Demiralp. Making table understanding work in practice. In *12th Conference on Innovative Data Systems Research, CIDR 2022, Chaminade, CA, USA, January 9-12, 2022*. www.cidrdb.org, 2022

173. Jessica Cox, Curt Kohler, Anthony Scerri, Corey Harper, Paul Groth, and Ron Daniel Jr. Measuring Scholarly Discourse Change with Respect to Citations-A Nobel Prize Case Study. In *23rd International Conference on Science and Technology Indicators (STI 2018), September 12-14, 2018, Leiden, The Netherlands*. Centre for Science and Technology Studies (CWTS), 2018

174. Peter Cotroneo, Wouter Haak, Gabriel Oscares, Eleonora Presani, Abhinav Rohatgi, and Paul T Groth. From Data Search to Data Showcasing: The Role of Semantic Technologies in a New Service. In Marieke van Erp, Medha Atre, Vanessa López, Kavitha Srinivas, and Carolina Fortuna, editors, *Proceedings of the ISWC 2018 Posters & Demonstrations, Industry and Blue Sky Ideas Tracks*, volume 2180 of *{CEUR} Workshop Proceedings*. CEUR-WS.org, 2018

175. Alex DeJong, Radmila Bord, Will Dowling, Rinke Hoekstra, Ryan Moquin, Charlie O, Mevan Samarasinghe, Paul Snyder, Craig Stanley, Anna Tordai, Michael Trefry, and Paul T Groth. Elsevier's Healthcare Knowledge Graph and the Case for Enterprise Level Linked Data Standards. In Marieke van Erp, Medha Atre, Vanessa López, Kavitha Srinivas, and Carolina Fortuna, editors, *Proceedings of the ISWC 2018 Posters & Demonstrations, Industry and Blue Sky Ideas Tracks*, volume 2180 of *{CEUR} Workshop Proceedings*. CEUR-WS.org, 2018

176. Marcin Wylot, Philippe Cudré-Mauroux, and Paul Groth. A demonstration of TripleProv. In *Proceedings of the VLDB Endowment*, volume 8, pages 1992–1995. VLDB Endowment, aug 2015

177. Wouter Beek, Paul Groth, Stefan Schlobach, and Rinke Hoekstra. A web observatory for the machine processability of structured data on the web. In *Proceedings of the 2014 ACM conference on Web science - WebSci '14*, pages 249–250, New York, New York, USA, jun 2014. ACM Press

178. C Brenninkmeijer, C Evelo, C Goble, A J G Gray, P Groth, S Pettifer, R Stevens, A Williams, and E Willighagen. Scientific Lenses: An Approach to Dynamically Vary the Relationships between Datasets. In *Intelligent Systems for Molecular Biology and European Conference on Computational Biology (ISMB/ECCB 2013)*, Berlin, Germany, jul 2013

179. A J G Gray, S Askjaer, C Brenninkmeijer, K Burger, C Chichester, J Eales, C T Evelo, C Goble, P Groth, L Harland, A Loizou, Stephen Pettifer, Rishi Ramgolam, Mark Thompson, Andra Waagmeester, and Antony J. Williams. The Pharmacology Workspace: A Platform for Drug Discovery. In *3rd International Conference on Biomedical Ontology*, volume 897 of *CEUR Workshop Proceedings*, jul 2012

180. Sara Magliacane and Paul Groth. Towards Reconstructing the Provenance of Clinical Guidelines. In *Proceedings of the 5th International Workshop on Semantic Web Applications and Tools for Life Sciences (SWAT4LS)*, volume 952 of *CEUR Workshop Proceedings*, Paris, France, 2012

181. Paul Groth and Yolanda Gil. LinkedDataLens: Linked Data as a Network of Networks. In *Proceedings of the ACM International Conference on Knowledge Capture (K-CAP)*, pages 191–192, Banff, Alberta, 2011

182. Carina Haupt, Paul Groth, and Marc Zimmermann. Representing Text Mining Results for Structured Pharmacological Queries. In *The 10th International Semantic Web Conference*, 2011

183. Christophe Guéret, Paul Groth, Claus Stadler, and Jens Lehmann. Linked Data Quality Assessment through Network Analysis. In *The 10th International Semantic Web Conference*, 2011

184. Paul Groth. ProvenanceJS: Revealing the Provenance of Web Pages. In *International Provenance and Annotation Workshop (IPAW'10)*, jun 2010

185. Paul Groth and Yolanda Gil. A Scientific Workflow Construction Command Line. In *International Conference on Intelligent User Interface 2009 (IUI2009)*, 2009

Reports

186. Paul Groth, Elena Simperl, Marieke van Erp, and Denny Vrandečić. Knowledge Graphs and their Role in the Knowledge Engineering of the 21st Century (Dagstuhl Seminar 22372). *Dagstuhl Reports*, 12(9):60–120, 2023

187. Mehwish Alam, Mehdi Ali, Paul Groth, Pascal Hitzler, Jens Lehmann, Heiko Paulheim, Achim Rettinger, Harald Sack, Afshin Sadeghi, and Volker Tresp. Proceedings of machine learning with symbolic methods and knowledge graphs colocated with european conference on machine learning and principles and practice of knowledge discovery in databases (ECML PKDD 2021), virtual, september 17, 2021. 2997, 2021

188. Robert West, Smriti Bhagat, Paul Groth, Marinka Zitnik, Francisco M. Couto, Pasquale Lisena, Albert Meroño Peñuela, Xiangyu Zhao, Wenqi Fan, Dawei Yin, Jiliang Tang, Linjun Shou, Ming Gong, Jian Pei, Xiubo Geng, Xingjie Zhou, Daxin Jiang, Benjamin Ricaud, Nicolas Aspert, Volodymyr Miz, Jennifer Dy, Stratis Ioannidis, undefinedlkay Yıldız, Rezvaneh Rezapour, Samin Aref, Ly Dinh, Jana Diesner, Alexey Drutsa, Dmitry Ustalov, Nikita Popov, Daria Baidakova, Shubhanshu

Mishra, Arjun Gopalan, Da-Cheng Juan, Cesar Ilharco Magalhaes, Chun-Sung Ferng, Allan Heydon, Chun-Ta Lu, Philip Pham, George Yu, Yicheng Fan, Yueqi Wang, Florian Laurent, Yanick Schraner, Christian Scheller, Sharada Mohanty, Jiawei Chen, Xiang Wang, Fuli Feng, Xiangnan He, Irene Teinemaa, Javier Albert, Dmitri Goldenberg, Flavian Vasile, David Rohde, Olivier Jeunen, Amine Benhalloum, Otmane Sakhi, Yu Rong, Wenbing Huang, Tingyang Xu, Yatao Bian, Hong Cheng, Fuchun Sun, Junzhou Huang, Shobeir Fakhraei, Christos Faloutsos, Onur Çelebi, Martin Müller, Manuel Schneider, Olesia Altunina, Wolfram Wingerath, Benjamin Wollmer, Felix Gessert, Stephan Succo, Norbert Ritter, Evann Courdier, Tudor Mihai Avram, Dragan Cvetinovic, Levan Tsinadze, Johny Jose, Rose Howell, Mario Koenig, Michaël Defferrard, Krishnaram Kenthapadi, Ben Packer, Mehrnoosh Sameki, and Nashlie Sephus. Summary of tutorials at the web conference 2021. In *Companion Proceedings of the Web Conference 2021*, WWW '21, page 727–733, New York, NY, USA, 2021. Association for Computing Machinery

189. Anna-Lena Lamprecht, Magnus Palmblad, Jon Ison, Veit Schwämmle, Mohammad Sadnan Al Manir, Ilkay Altintas, Christopher J. O. Baker, Ammar Ben Hadj Amor, Salvador Capella-Gutierrez, Paulos Charonyktakis, Michael R. Crusoe, Yolanda Gil, Carole Goble, Timothy J. Griffin, Paul Groth, Hans Ienasescu, Pratik Jagtap, Matúš Kalaš, Vedran Kasalica, Alireza Khanteymoori, Tobias Kuhn, Hailiang Mei, Hervé Ménager, Steffen Möller, Robin A. Richardson, Vincent Robert, Stian Soiland-Reyes, Robert Stevens, Szoke Szaniszlo, Suzan Verberne, Aswin Verhoeven, and Katherine Wolstencroft. Perspectives on automated composition of workflows in the life sciences. *F1000Research*, 10:897, September 2021

190. Altmetric Engineering, Stacy Konkiel, Jason Priem, Euan Adie, Gemma Derrick, Fereshteh Didegah, Paul Groth, Cameron Neylon, Shenmeng Xu, Zohreh Zahedi, Timothy Bowman, Vanash M Patel, Robin Haunschild, Lutz Bornmann, Mike Taylor, Liesa Ross, Yin-Leng Theng, Saeed-Ul Hassan, and Naif R. Aljohani. The state of altmetrics: a tenth anniversary celebration, 2020

191. Mehwish Alam, Paul Groth, Pascal Hitzler, Heiko Paulheim, Harald Sack, and Volker Tresp. Cssa'20: Workshop on combining symbolic and sub-symbolic methods and their applications. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, CIKM '20, page 3523–3524, New York, NY, USA, 2020. Association for Computing Machinery

192. Yolanda Gil, James Cheney, Paul Groth, Olaf Hartig, Simon Miles, Luc Moreau, and Paolo Pinheiro DaSilva. Final Report of the W3C Provenance Incubator Group. Technical report, 2011

193. Jason Priem, Dario Taraborelli, Paul Groth, and Cameron Neylon. Altmetrics: a manifesto, 2010

194. Paul Groth, Sheng Jiang, Simon Miles, Steve Munroe, Victor Tan, Sofia Tsasakou, and Luc Moreau. An Architecture for Provenance Systems. Technical report, University of Southampton, 2007

Last updated: April 21, 2024 • Typeset in XeTeX

http://pgroth.com/cv.pdf