

2019 15th International Conference on eScience (eScience 2019)

**San Diego, California, USA
24 – 27 September 2019**



**IEEE Catalog Number: CFP1978F-POD
ISBN: 978-1-7281-2452-0**

**Copyright © 2019 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP1978F-POD
ISBN (Print-On-Demand):	978-1-7281-2452-0
ISBN (Online):	978-1-7281-2451-3

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2019 15th International Conference on eScience (eScience) eScience 2019

Table of Contents

Message from the eScience 2019 Chairs .xviii.....
Organizers .xx.....
Program Committee .xxii.....

Main Track

SOMOSPIE: A Modular SOil MOisture SPatial Inference Engine Based on Data-Driven Decisions .1.....
Danny Rorabaugh (University of Tennessee), Mario Guevara (University of Delaware), Ricardo Llamas (University of Delaware), Joy Kitson (University of Delaware), Rodrigo Vargas (University of Delaware), and Michela Taufer (University of Tennessee)

The International Forest Risk Model (INFORM): A Method for Assessing Supply Chain Deforestation Risk with Imperfect Data .11.....
Neil Caithness (University of Oxford), Cécile Lachaux (Man & Nature), and David C. H. Wallom (University of Oxford)

ForestEyes Project: Can Citizen Scientists Help Rainforests? .18.....
Fernanda Beatriz Jordan Rojas Dallaqua (ICT/UNIFESP), Álvaro Luiz Fazenda (ICT/UNIFESP), and Fabio Augusto Faria (ICT/UNIFESP)

Data Identification and Process Monitoring for Reproducible Earth Observation Research .28.....
Bernhard Gößwein (Vienna University of Technology), Tomasz Miksa (Vienna University of Technology & SBA Research), Andreas Rauber (Vienna University of Technology), and Wolfgang Wagner (Vienna University of Technology)

A Hybrid Algorithm for Mineral Dust Detection Using Satellite Data .39.....
Peichang Shi (University of Maryland), Qianqian Song (University of Maryland), Janita Patwardhan (University of Maryland), Zhibo Zhang (University of Maryland), Jianwu Wang (University of Maryland), and Aryya Gangopadhyay (University of Maryland)

Workflow Design Analysis for High Resolution Satellite Image Analysis .47.....
Ioannis Paraskevacos (Rutgers University), Matteo Turilli (Rutgers University), Bento Collares Gonçalves (Stony Brook, NY), Heather Lynch (Stony Brook, NY), and Shantenu Jha (Rutgers University and Brookhaven National Laboratory)

SATVAM: Toward an IoT Cyber-Infrastructure for Low-Cost Urban Air Quality Monitoring .57.....

Yogesh Simmhan (Indian Institute of Science), Srijith Nair (Indian Institute of Science), Sumit Monga (Indian Institute of Science), Ravi Sahu (Indian Institute of Technology), Kuldeep Dixit (Indian Institute of Technology), Ronak Sutaria (Indian Institute of Technology), Brijesh Mishra (Indian Institute of Technology), Anamika Sharma (Indian Institute of Technology), Anand SVR (Indian Institute of Science), Malati Hegde (Indian Institute of Science), Rajesh Zele (Indian Institute of Technology), and Sachchida N. Tripathi (Indian Institute of Technology)

Toward a Dynamic Network-Centric Distributed Cloud Platform for Scientific Workflows: A Case Study for Adaptive Weather Sensing .67.....

Eric Lyons (University of Massachusetts Amherst), George Papadimitriou (University of Southern California), Cong Wang (University of North Carolina at Chapel Hill), Komal Thareja (University of North Carolina at Chapel Hill), Paul Ruth (University of North Carolina at Chapel Hill), J. J. Villalobos (Rutgers Discovery Informatics Institute), Ivan Rodero (Rutgers Discovery Informatics Institute), Ewa Deelman (University of Southern California), Michael Zink (University of Massachusetts Amherst), and Anirban Mandal (University of North Carolina at Chapel Hill)

The Evolution of Bits and Bottlenecks in a Scientific Workflow Trying to Keep Up with Technology: Accelerating 4D Image Segmentation Applied to NASA Data .77.....

Scott Sellars (University of California San Diego), John Graham (University of California San Diego), Dmitry Mishin (University of California San Diego), Kyle Marcus (University of California San Diego), Ilkay Altintas (University of California San Diego), Thomas DeFanti (University of California San Diego), Larry Smarr (University of California San Diego), Camille Crittenden (University of California, Berkeley), Frank Wuerthwein (University of California San Diego), Joulien Tatar (University of California Irvine), Phu Nguyen (University of California Irvine), Eric Shearer (University of California Irvine), Soroosh Sorooshian (University of California Irvine), and F. Martin Ralph (Center for Western Weather and Water Extremes)

Out-of-the-Box Reproducibility: A Survey of Machine Learning Platforms .86.....

Richard Isdahl (Norwegian University of Science and Technology) and Odd Erik Gundersen (Norwegian University of Science and Technology)

dislib: Large Scale High Performance Machine Learning in Python .96.....

Javier Álvarez Cid-Fuentes (Barcelona Supercomputing Center), Salvi Solà (Barcelona Supercomputing Center), Pol Álvarez (Barcelona Supercomputing Center), Alfred Castro-Ginard (Dept. Física Quàntica i Astrofísica, Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB)), and Rosa M. Badia (Barcelona Supercomputing Center)

Recognition of Frog Chorus with Acoustic Indices and Machine Learning .106.....	
	<i>Hongxiao Gan (Queensland University of Technology), Jinglan Zhang (Queensland University of Technology), Michael Towsey (Queensland University of Technology), Anthony Truskinger (Queensland University of Technology), Debra Stark (The University of Queensland), Berndt van Rensburg (The University of Queensland), Yuefeng Li (Queensland University of Technology), and Paul Roe (Queensland University of Technology)</i>
Quality-Aware Human-Machine Text Extraction for Biocollections using Ensembles of OCRs .116.....	
	<i>Icaro Alzuru (University of Florida), Rhiannon Stephens (Australian Museum), Andréa Matsunaga (Advanced Computing and Information Systems Laboratory), Maurício Tsugawa (Advanced Computing and Information Systems Laboratory), Paul Flemons (Australian Museum), and José A.B. Fortes (University of Florida)</i>
Active Learning Yields Better Training Data for Scientific Named Entity Recognition .126.....	
	<i>Roselyne Tchoua (University of Chicago), Aswathy Ajith (University of Chicago), Zhi Hong (University of Chicago), Logan Ward (Argonne National Laboratory), Kyle Chard (University of Chicago), Debra Audus (National Institute of Standards and Technology), Shrayesh Patel (University of Chicago), Juan de Pablo (University of Chicago), and Ian Foster (Argonne National Laboratory)</i>
Reliability-Aware and Graph-Based Approach for Rank Aggregation of Biological Data .136.....	
	<i>Pierre Andrieu (Université Paris-Sud, CNRS, Université Paris-Saclay), Bryan Brancotte (Institut Pasteur), Laurent Bulteau (Université Paris-Est Marne-la-Vallée, CNRS), Sarah Cohen-Boulakia (Université Paris-Sud, CNRS, Université Paris-Saclay), Alain Denise (Université Paris-Sud, CNRS, Université Paris-Saclay), Adeline Pierrot (Université Paris-Sud, CNRS, Université Paris-Saclay), and Stéphane Vialette (Université Paris-Est Marne-la-Vallée, CNRS)</i>
Evaluation of Pilot Jobs for Apache Spark Applications on HPC Clusters .146.....	
	<i>Valerie Hayot-Sasson (Concordia University) and Tristan Glatard (Concordia University)</i>
Profit Optimization for Splitting and Sampling Based Resource Management in Big Data Analytics-as-a-Service Platforms in Cloud Computing Environments .156.....	
	<i>Yali Zhao (The University of Melbourne), Athanasios Vasilakos (Lulea University of Technology), James Bailey (The University of Melbourne), and Richard Sinnott (The University of Melbourne)</i>
On Distributed Information Composition in Big Data Systems .168.....	
	<i>Haifa AlQuwaiee (New Jersey Institute of Technology), Songlin He (New Jersey Institute of Technology), Chase Wu (New Jersey Institute of Technology), Qiang Tang (New Jersey Institute of Technology), and Xuewen Shen (New Jersey Institute of Technology)</i>
Dynamic Sizing of Continuously Divisible Jobs for Heterogeneous Resources .178.....	
	<i>Nicholas Hazekamp (University of Notre Dame), Benjamin Tovar (University of Notre Dame), and Douglas Thain (University of Notre Dame)</i>

Characterizing In Situ and In Transit Analytics of Molecular Dynamics Simulations for Next-Generation Supercomputers .188.....
Michela Taufer (The University of Tennessee), Stephen Thomas (The University of Tennessee), Michael Wyatt (The University of Tennessee), Tu Mai Anh Do (University of Southern California), Loïc Pottier (University of Southern California), Rafael Ferreira da Silva (University of Southern California), Harel Weinstein (Cornell University), Michel A. Cuendet (Cornell University; Lausanne University Hospital), Trilce Estrada (University of New Mexico), and Ewa Deelman (University of Southern California)

SPARCS: Stream-Processing Architecture Applied in Real-Time Cyber-Physical Security .199.....
Reinhard Gentz (Lawrence Berkeley National Laboratory), Sean Peisert (Lawrence Berkeley National Laboratory), Joshua Boverhof (Lawrence Berkeley National Laboratory), and Daniel Gunter (Lawrence Berkeley National Laboratory)

Timing is Everything: Identifying Diverse Interaction Dynamics in Scenario and Non-Scenario Meetings.203..
Chreston Miller (Virginia Tech) and Christa Miller (Virginia Tech)

Multi-model Investigative Exploration of Social Media Data with BOUTIQUE: A Case Study in Public Health .213.....
Junan Guo (University of California San Diego), Subhasis Dasgupta (University of California San Diego), and Amarnath Gupta (University of California San Diego)

Increasing Life Science Resources Re-Usability using Semantic Web Technologies .217.....
Marine Louarn (INSERM & Univ Rennes, Inria, CNRS, IRISA), Fabrice Chatonnet (INSERM, Univ Rennes, CHU Rennes, EFS), Xavier Garnier (Univ Rennes, Inria, CNRS, IRISA), Thierry Fest (INSERM, Univ Rennes, CHU Rennes, EFS), Anne Siegel (Univ Rennes, Inria, CNRS, IRISA), and Olivier Dameron (Univ Rennes, Inria, CNRS, IRISA)

Data Encoding in Lossless Prediction-Based Compression Algorithms .226.....
Ugur Cayoglu (Karlsruhe Institute of Technology (KIT)), Frank Tristram (Karlsruhe Institute of Technology (KIT)), Jörg Meyer (Karlsruhe Institute of Technology (KIT)), Jennifer Schröter (Karlsruhe Institute of Technology (KIT)), Tobias Kerzenmacher (Karlsruhe Institute of Technology (KIT)), Peter Braesicke (Karlsruhe Institute of Technology (KIT)), and Achim Streit (Karlsruhe Institute of Technology (KIT))

Social Sciences & Humanities Track

defoe: A Spark-Based Toolbox for Analysing Digital Historical Textual Data .235.....
Rosa Filgueira (University of Edinburgh), Michael Jackson (University of Edinburgh), Anna Roubickova (University of Edinburgh), Amrey Krause (University of Edinburgh), Ruth Ahnert (Queen Mary University of London), Tessa Hauswedell (University College London), Julianne Nyhan (University College London), David Beavan (The Alan Turing Institute), Timothy Hobson (The Alan Turing Institute), Mariona Coll Ardanuy (The Alan Turing Institute), Giovanni Colavizza (The Alan Turing Institute), James Hetherington (The Alan Turing Institute), and Melissa Terras (University of Edinburgh)

Understanding a Rapidly Expanding Refugee Camp Using Convolutional Neural Networks and Satellite Imagery .243.....
Susanne Benz (UC San Diego), Hogeun Park (UC San Diego), Jiaxin Li (UC San Diego), Daniel Crawl (UC San Diego), Jessica Block (UC San Diego), Mai Nguyen (UC San Diego), and Ilkay Altintas (UC San Diego)

Social Media Intelligence and Learning Environment: an Open Source Framework for Social Media Data Collection, Analysis and Curation .252.....
Chen Wang (University of Illinois at Urbana-Champaign), Luigi Marini (University of Illinois at Urbana-Champaign), Chieh-Li Chin (University of Illinois at Urbana-Champaign), Nickolas Vance (University of Illinois at Urbana-Champaign), Curtis Donelson (University of Illinois at Urbana-Champaign), Pascal Meunier (Purdue University), and Joseph T. Yun (University of Illinois at Urbana-Champaign)

Software Tools & Infrastructure Track

Toward an Elastic Data Transfer Infrastructure .262.....
Joaquin Chung (Argonne National Laboratory), Zhengchun Liu (Argonne National Laboratory), Rajkumar Kettimuthu (Argonne National Laboratory), and Ian Foster (Argonne National Laboratory)

Scalable Performance Awareness for In Situ Scientific Applications .266.....
Matthew Wolf (Oak Ridge National Laboratory), Jong Choi (Oak Ridge National Laboratory), Greg Eisenhauer (Georgia Institute of Technology), Stéphane Ethier (Princeton Plasma Physics Laboratory), Kevin Huck (University of Oregon), Scott Klasky (Oak Ridge National Laboratory), Jeremy Logan (Oak Ridge National Laboratory), Allen Malony (University of Oregon), Chad Wood (University of Oregon), Julien Dominski (Princeton Plasma Physics Laboratory), and Gabriele Merlo (University of Texas, Austin)

ENVRI-FAIR - Interoperable Environmental FAIR Data and Services for Society, Innovation and Research.277
Andreas Petzold (Forschungszentrum Jülich GmbH), Ari Asmi (University of Helsinki), Alex Vermeulen (Lund University), Gelsomina Pappalardo (CNR Institute of Methodologies for Environmental Analysis), Daniele Bailo (Istituto Nazionale di Geofisica e Vulcanologia), Dick Schaap (MARIS B.V.), Helen M. Graves (British Geological Survey), Ulrich Bundke (Forschungszentrum Jülich GmbH), and Zhiming Zhao (University of Amsterdam)

Custom Execution Environments with Containers in Pegasus-Enabled Scientific Workflows .281.....
Karan Vahi (University of Southern California), Mats Rynge (University of Southern California), George Papadimitriou (University of Southern California), Duncan Brown (Syracuse University), Rajiv Mayani (University of Southern California), Rafael Ferreira da Silva (University of Southern California), Ewa Deelman (University of Southern California), Anirban Mandal (University of North Carolina), Eric Lyons (University of Massachusetts at Amherst), and Michael Zink (University of Massachusetts at Amherst)

SciInc: A Container Runtime for Incremental Recomputation .291.....	
	<i>Andrew Youngdahl (DePaul University), Dai-Hai Ton-That (DePaul University), and Tanu Malik (DePaul University)</i>
Usage Patterns of Wideband Display Environments In e-Science Research, Development and Training .301.....	
	<i>Jason Leigh (University of Hawaii at Manoa), Dylan Kobayashi (University of Hawaii at Manoa), Nurit Kirshenbaum (University of Hawaii at Manoa), Troy Wooton (University of Hawaii at Manoa), Alberto Gonzalez (University of Hawaii at Manoa), Luc Renambot (University of Illinois at Chicago), Andrew Johnson (University of Illinois at Chicago), Maxine Brown (University of Illinois at Chicago), Andrew Burks (University of Illinois at Chicago), Krishna Bharadwaj (University of Illinois at Chicago), Arthur Nishimoto (University of Illinois at Chicago), Lance Long (University of Illinois at Chicago), Jason Haga (National Institute of Advanced Industrial Science and Technology), John Burns (University of Hawaii at Hilo), Francis Cristobal (University of Hawaii at Hilo), Jared McLean (University of Hawaii at Hilo), Roberto Pelayo (University of Hawaii at Hilo), and Mahdi Belcaid (University of Hawaii at Manoa)</i>
Comprehensible Control for Researchers and Developers Facing Data Challenges .311.....	
	<i>Malcolm Atkinson (University of Edinburgh), Rosa Filgueira (University of Edinburgh), Iraklis Klampanos (National Centre for Scientific Research Demokritos), Antonis Koukourikos (National Centre for Scientific Research Demokritos), Amrey Krause (University of Edinburgh), Federica Magnoni (Istituto Nazionale di Geofisica e Vulcanologia), Christian Pagé (Université de Toulouse, CNRS), Andreas Rietbrock (Karlsruhe Institute of Technology), and Alessandro Spinuso (Koninklijk Nederlands Meteorologisch Instituut)</i>
BBBlockchain: Blockchain-Based Participation in Urban Development .321.....	
	<i>Robert Muth (Technische Universität Berlin), Kerstin Eisenhut (Technische Universität Berlin), Jochen Rabe (Technische Universität Berlin), and Florian Tschorsch (Technische Universität Berlin)</i>
A Framework for Model Search Across Multiple Machine Learning Implementations .331.....	
	<i>Yoshiki Takahashi (Tokyo Institute of Technology), Masato Asahara (NEC System Platform Research Laboratories), and Kazuyuki Shudo (Tokyo Institute of Technology)</i>
Enhanced Interactive Parallel Coordinates using Machine Learning and Uncertainty Propagation for Engineering Design .339.....	
	<i>Wiktor Piotrowski (University of Cambridge), Timoleon Kipouros (University of Cambridge), and P John Clarkson (University of Cambridge)</i>
OKG-Soft: An Open Knowledge Graph with Machine Readable Scientific Software Metadata .349.....	
	<i>Daniel Garijo (University of Southern California), Maximiliano Osorio (University of Southern California), Deborah Khider (University of Southern California), Varun Ratnakar (University of Southern California), and Yolanda Gil (University of Southern California)</i>

Efficient Runtime Capture of Multiworkflow Data Using Provenance .359.....	<i>Renan Souza (COPPE/UFRJ & IBM Research), Leonardo Azevedo (IBM Research), Raphael Thiago (IBM Research), Elton Soares (IBM Research), Marcelo Nery (IBM Research), Marco A. S. Netto (IBM Research), Emilio Vital (IBM Research), Renato Cerqueira (IBM Research), Patrick Valduries (Inria & U. Montpellier), and Marta Mattoso (COPPE/UFRJ)</i>
AdaptLidarTools: A Full-Waveform Lidar Processing Suite .369.....	<i>Ravi Shankar (Boise State University), Nayani Ilangakoon (Boise State University), Aaron Orenstein (Treasure Valley Math and Science Center), Floriana Ciaglia (Boise State University), Nancy Glenn (Boise State University), and Catherine Olschanowsky (Boise State University)</i>
SDM: A Scientific Dataset Delivery Platform .378.....	<i>Illyoung Choi (University of Arizona), Jude Nelson (Blockstack PBC), Larry Peterson (Open Networking Foundation), and John Hartman (University of Arizona)</i>

Computing & Physics Track

Photon Propagation using GPUs by the IceCube Neutrino Observatory .388.....	<i>Dmitry Chirkin (University of Wisconsin-Madison), Juan Carlos Diaz-Vélez (University of Wisconsin-Madison), Claudio Kopper (Michigan State University), Alexander Olivas (University of Maryland), Benedikt Riedel (University of Wisconsin-Madison), Martin Rongen (RWTH Aachen University), David Schultz (University of Wisconsin-Madison), and Jakob van Santen (Deutsches Elektronen-Synchrotron-Zeuthen)</i>
Simulating Data Access Profiles of Computational Jobs in Data Grids .394.....	<i>Volodimir Begy (CERN, University of Vienna), Joeri Hermans (University of Liège), Martin Barisits (CERN), Mario Lassnig (CERN), and Erich Schikuta (University of Vienna)</i>
Towards Exascale: Measuring the Energy Footprint of Astrophysics HPC Simulations .403.....	<i>Giuliano Taffoni (INAF - OATs), Luca Tornatore (INAF - OATs), David Goz (INAF - OATs), Antonio Ragagnin (INAF - OATs), Sara Bertocco (INAF - OATs), Igor Coretti (INAF - OATs), Manolis Marazakis (FORTH - Foundation For Research & Technology), Fabien Chaix (FORTH - Foundation For Research & Technology), Manolis Plumidis (FORTH - Foundation For Research & Technology Hellas), Manolis Katevenis (FORTH - Foundation For Research & Technology Hellas), Renato Panchieri (EnginSoft S.p.A. (EnginSoft)), and Gino Perna (EnginSoft S.p.A. (EnginSoft))</i>

Visionary Track

The Future of Swedish e-Science: SeRC 2.0 .413.....	<i>Erwin Laure (SeRC & KTH), Olivia Eriksson (SeRC & KTH), Erik Lindahl (SeRC & KTH), and Dan Henningson (SeRC & KTH)</i>
---	---

Understanding ML Driven HPC: Applications and Infrastructure .421.....	<i>Shantenu Jha (Rutgers University and Brookhaven National Laboratory) and Geoffrey Fox (Indiana University)</i>
Transparency by Design in eScience Research .428.....	<i>Beth Plale (Indiana University)</i>
Serverless Science for Simple, Scalable, and Shareable Scholarship .432.....	<i>Kyle Chard (University of Chicago; Argonne National Laboratory) and Ian Foster (University of Chicago; Argonne National Laboratory)</i>
Learning Everywhere: A Taxonomy for the Integration of Machine Learning and Simulations .439.....	<i>Geoffrey Fox (Indiana University) and Shantenu Jha (Rutgers University)</i>
Cyberinfrastructure Center of Excellence Pilot: Connecting Large Facilities Cyberinfrastructure .449.....	<i>Ewa Deelman (University of Southern California), Anirban Mandal (University of North Carolina at Chapel Hill), Valerio Pascucci (University of Utah, Salt Lake City), Susan Sons (Indiana University), Jane Wyngaard (University of Notre Dame), Charles Vardeman (University of Notre Dame), Steve Petruzza (University of Utah), Ilya Baldin (University of North Carolina), Laura Christopherson (University of North Carolina), Ryan Mitchell (University of Southern California), Loic Pottier (University of Southern California), Mats Rynge (University of Southern California), Erik Scott (University of North Carolina), Karan Vahi (University of Southern California), Marina Kogan (University of Utah), Jasmine Mann (University of Southern California), Tom Gulbransen (Battelle Ecology, Inc.), Daniel Allen (Battelle Ecology, Inc.), David Barlow (Battelle Ecology, Inc.), Santiago Bonarrigo (Battelle Ecology, Inc.), Chris Clark (Battelle Ecology, Inc.), Leslie Goldman (Battelle Ecology, Inc.), Tristan Goulden (Battelle Ecology, Inc.), Phil Harvey (Battelle Ecology, Inc.), David Hulsander (Battelle Ecology, Inc.), Steve Jacobs (Battelle Ecology, Inc.), Christine Laney (Battelle Ecology, Inc.), Ivan Lobo-Padilla (Battelle Ecology, Inc.), Jeremy Sampson (Battelle Ecology, Inc.), John Staarmann (Battelle Ecology, Inc.), and Steve Stone (Battelle Ecology, Inc.)</i>
The Research Assistant and AI in eScience .458.....	<i>Dennis Gannon (Indiana University)</i>

Workshop on Platform-Driven e-Infrastructure Innovations (EINFRA)

Transkribus. A Platform for Automated Text Recognition and Searching of Historical Documents .463.....	<i>Sebastian Colutto (University of Innsbruck), Philip Kahle (University of Innsbruck), Hackl Guenter (University of Innsbruck), and Guenter Muehlberger (University of Innsbruck)</i>
Unlocking the LOFAR LTA .467.....	<i>Hanno Spreeuw (Netherlands eScience Center), Souley Madougou (Netherlands eScience Center), Ronald Van Haren (Netherlands eScience Center), Berend Weel (Netherlands eScience Center), Adam Belloum (University of Amsterdam), and Jason Maassen (Netherlands eScience Center)</i>

European HPC Landscape .471.....	
	<i>Florian Berberich (PRACE aisbl and Jülich Supercomputing Center, Forschungszentrum Juelich GmbH), Janina Liebmann (Jülich Supercomputing Center, Forschungszentrum Juelich GmbH), Jean-Philippe Nominé (ETP4HPC and Commissariat à l'énergie atomique et aux énergies alternatives), Oriol Pineda (PRACE aisbl and Barcelona Supercomputing Center), Philippe Segers (Grand équipement national de calcul intensif), and Veronica Teodor (Jülich Supercomputing Center, Forschungszentrum Juelich GmbH)</i>
Reference Exascale Architecture .479.....	
	<i>Martin Bobák (Slovak Academy of Sciences), Ladislav Hluchy (Slovak Academy of Sciences), Adam Belloum (University of Amsterdam), Reginald Cushing (University of Amsterdam), Jan Meizner (AGH University of Science and Technology), Piotr Nowakowski (AGH University of Science and Technology), Viet Tran (Slovak Academy of Sciences), Ondrej Habala (Slovak Academy of Sciences), Jason Maassen (Netherlands eScience Center), Balázs Somosköi (Lufthansa Systems), Mara Graziani (University of Applied Sciences, Western Switzerland (HES-SO)), Matti Heikkurinen (University of Applied Sciences, Western Switzerland (HES-SO)), Maximilian Hüb (Ludwig-Maximilians Universität), and Jan Schmidt (Ludwig-Maximilians Universität)</i>
The AllScale API .488.....	
	<i>Philipp Gschwandtner (University of Innsbruck), Herbert Jordan (University of Innsbruck), Peter Thoman (University of Innsbruck), and Thomas Fahringer (University of Innsbruck)</i>
ESiWACE: On European Infrastructure Efforts for Weather and Climate Modeling at Exascale .498.....	
	<i>Philipp Neumann (German Climate Computing Center) and Joachim Biercamp (German Climate Computing Center)</i>

Workshop on Research Objects 2019 (RO 2019)

Reproducibility by Other Means: Transparent Research Objects .502.....	
	<i>Timothy McPhillips (University of Illinois at Urbana-Champaign), Craig Willis (University of Illinois at Urbana-Champaign), Michael R. Gryk (University of Illinois at Urbana-Champaign), Santiago Nunez-Corrales (University of Illinois at Urbana-Champaign), and Bertram Ludascher (University of Illinois at Urbana-Champaign)</i>
Interactivity, Distributed Workflows, and Thick Provenance: A Review of Challenges Confronting Digital Humanities Research Objects .510.....	
	<i>Katrina Fenlon (University of Maryland, College of Information Studies)</i>

Application of BagIt-Serialized Research Object Bundles for Packaging and Re-Execution of Computational Analyses .514.....	
	<i>Kyle Chard (University of Chicago), Niall Gaffney (University of Texas at Austin), Matthew B. Jones (University of California at Santa Barbara), Kacper Kowalik (University of Illinois at Urbana-Champaign), Bertram Ludäscher (University of Illinois at Urbana-Champaign), Timothy McPhillips (University of Illinois at Urbana-Champaign), Jarek Nabrzyski (University of Notre Dame), Victoria Stodden (University of Illinois at Urbana-Champaign), Ian Taylor (University of Notre Dame), Thomas Thelen (University of California at Santa Barbara), Matthew J. Turk (University of Illinois at Urbana-Champaign), and Craig Willis (University of Illinois at Urbana-Champaign)</i>
Data Quality Issues in Current Nanopublications .522.....	
	<i>Imran Asif (Heriot-Watt University, Edinburgh), Jessica Chen-Burger (Heriot-Watt University, Edinburgh), and Alasdair J. G. Gray (Heriot-Watt University, Edinburgh)</i>

Advanced Knowledge Technologies for Science in a FAIR World (AKTS)

Describing Datasets in Wikidata .528.....	
	<i>Denny Vrandečić (Google)</i>
Making Data FAIR Requires More than Just Principles: We Need Knowledge Technologies .530.....	
	<i>Mark Musen (Stanford University)</i>
Iterative Document Retrieval via Deep Learning Approaches for Biomedical Question Answering .533.....	
	<i>Ibrahim Burak Ozyurt (UC San Diego) and Jeffrey Grethe (UC San Diego)</i>
Incorporating New Concepts Into the Scientific Variables Ontology .539.....	
	<i>Maria Stoica (University of Colorado, Boulder) and Scott Peckham (University of Colorado, Boulder)</i>

5th Workshop on Curricula and Teaching Methods in Cloud Computing, Big Data, and Data Science (DTW 2019)

Data Science Model Curriculum Implementation for Various Types of Big Data Infrastructure Courses .541.....	
	<i>Tomasz Wiktorski (University of Stavanger, Norway), Yuri Demchenko (University of Amsterdam, The Netherlands), and Oleg Chertov (National Technical University of Ukraine)</i>
Teaching DevOps and Cloud Based Software Engineering in University Curricula .548.....	
	<i>Yuri Demchenko (University of Amsterdam), Zhiming Zhao (University of Amsterdam), Jayachander Surbiryala (University of Stavanger), Spiros Koulouzis (University of Amsterdam), Zeshun Shi (University of Amsterdam), Xiaofeng Liao (University of Amsterdam), and Jelena Gordiyenko (Agile Telecom)</i>
EDISON Data Science Framework (EDSF) Extension to Address Transversal Skills Required by Emerging Industry 4.0 Transformation .553.....	
	<i>Yuri Demchenko (University of Amsterdam), Tomasz Wiktorski (University of Stavanger), Juan Cuadrado Gallego (University of Alcala), and Steve Brewer (University of Southampton)</i>

Bridging from Concepts to Data and Computation for eScience (BC2DC'19)

Active Provenance for Data-Intensive Workflows: Engaging Users and Developers .560.....	
<i>Alessandro Spinuso (Koninklijk Nederlands Meteorologisch Instituut), Malcolm Atkinson (University of Edinburgh), and Federica Magnoni (Istituto Nazionale Geofisica e Vulcanologia)</i>	
Modeling and Matching Digital Data Marketplace Policies .570.....	
<i>Sara Shakeri (University of Amsterdam), Valentina Maccatrozzo (Netherlands eScience Center), Lourens Veen (Netherlands eScience Center), Rena Bakhshi (Netherlands eScience Center), Leon Gommans (University of Amsterdam), Cees de Laat (University of Amsterdam), and Paola Grosso (University of Amsterdam)</i>	
DARE: A Reflective Platform Designed to Enable Agile Data-Driven Research on the Cloud .578.....	
<i>Iraklis Klampanos (NCSR "Demokritos"), Athanasios Davvetas (NCSR "Demokritos"), André Gemünd (Fraunhofer SCAI), Malcolm Atkinson (University of Edinburgh), Antonios Koukourikos (NCSR "Demokritos"), Rosa Filgueira (University of Edinburgh), Amrey Krause (University of Edinburgh), Alessandro Spinuso (KNMI), Angelos Charalambidis (NCSR "Demokritos"), Federica Magnoni (INGV), Emanuele Casarotti (INGV), Christian Pagé (CERFACS), Mike Lindner (KIT), Andreas Ikonomopoulos (NCSR "Demokritos"), and Vangelis Karkaletsis (NCSR "Demokritos")</i>	
Ease Access to Climate Simulations for Researchers: IS-ENES Climate4Impact .586.....	
<i>Christian Pagé (Université de Toulouse, CNRS), Wim Som de Cerff (Koninklijk Nederlands Meteorologisch Instituut), Maarten Plieger (Koninklijk Nederlands Meteorologisch Instituut), Alessandro Spinuso (Koninklijk Nederlands Meteorologisch Instituut), and Xavier Pivan (Université de Toulouse, CNRS)</i>	
Managing Scientific Literature with Software from the PORTAL-DOORS Project .588.....	
<i>Shiladitya Dutta (Brain Health Alliance), Pooja Kowshik (Brain Health Alliance), Adarsh Ambati (Brain Health Alliance), Sathvik Nori (Brain Health Alliance), S. Koby Taswell (Brain Health Alliance), and Carl Taswell (Brain Health Alliance)</i>	
Towards a Computer-Interpretable Actionable Formal Model to Encode Data Governance Rules .594.....	
<i>Rui Zhao (University of Edinburgh) and Malcolm Atkinson (University of Edinburgh)</i>	
Towards a New Paradigm for Programming Scientific Workflows .604.....	
<i>Reginald Cushing (University of Amsterdam), Onno Valkering (University of Amsterdam), Adam Belloum (University of Amsterdam), and Cees de Laat (University of Amsterdam)</i>	
Bridging Concepts and Practice in eScience via Simulation-Driven Engineering .609.....	
<i>Rafael Ferreira da Silva (University of Southern California), Henri Casanova (University of Hawaii), Ryan Tanaka (University of Hawaii), and Frédéric Suter (IN2P3 Computing Center, CNRS)</i>	

Poster Session

Accelerating Scientific Discovery with SCAIGATE Science Gateway	615
<i>Chao Jiang (University of Florida), David Ojika (University of Florida), Bhavesh Patel (Dell EMC), Ann Gordon-Ross (University of Florida), and Herman Lam (University of Florida)</i>	
The Engagement and Performance Operations Center: EPOC	617
<i>Edward Moynihan (Indiana University), Jennifer Schopf (Indiana University), and Jason Zurawski (Lawrence Berkeley National Lab)</i>	
Streaming Graph Ingestion with Resource-Aware Buffering and Graph Compression	619
<i>Subhasis Dasgupta (University of California San Diego), Aditya Bagchi (RKMV Educational and Research Institute), and Amarnath Gupta (University of California San Diego)</i>	
Streaming Workflows on Edge Devices to Process Sensor Data on a Smart Manufacturing Platform	621
<i>Prakashan Korambath (University of California, Los Angeles), Haresh Malkani (University of California, Los Angeles), and Jim Davis (University of California, Los Angeles)</i>	
Enabling Transparent Access to Heterogeneous Architectures for IS-ENES Climate4Impact using the DARE Platform	623
<i>Christian Pagé (Université de Toulouse, CNRS), Wim Som de Cerff (Koninklijk Nederlands Meteorologisch Instituut), Maarten Plieger (Koninklijk Nederlands Meteorologisch Instituut), Alessandro Spinuso (Koninklijk Nederlands Meteorologisch Instituut), and Xavier Pivan (Université de Toulouse, CNRS)</i>	
Enabling Server-based Computing and FAIR Data Sharing with the ENES Climate Analytics Service	627
<i>S. Bendoukha, T. Weigel, S. Fiore, D. Elia</i>	
Support for HTCondor high-Throughput Computing Workflows in the REANA Reusable Analysis Platform ...	630
<i>Rokas Maiulaitis (CERN), Paul Brenner (University of Notre Dame), Scott Hampton (University of Notre Dame), Michael D. Hildreth (University of Notre Dame), Kenyi Paolo Hurtado Anampa (University of Notre Dame), Irena Johnson (University of Notre Dame), Cody Kankel (University of Notre Dame), Jan Okraska (CERN), Diego Rodriguez (CERN), and Tibor Šimko (CERN)</i>	
Effective Digital Object Access and Sharing Over a Networked Environment using DOIP and NDN	632
<i>Cas Fahrenfort (University of Amsterdam) and Zhiming Zhao (University of Amsterdam)</i>	
Contextual Linking between Workflow Provenance and System Performance Logs	634
<i>Elias el Khaldi Ahanach (University of Amsterdam), Spiros Koulouzis (University of Amsterdam), and Zhiming Zhao (University of Amsterdam)</i>	

A Historical Big Data Analysis to Understand the Social Construction of Juvenile Delinquency in the United States .636.....	
	<i>Sandeep Puthanveetil Satheesan (University of Illinois at Urbana-Champaign), Alan B. Craig (Extended Collaborative Support Services, Extreme Science and Engineering Discovery Environment), and Yu Zhang (State University of New York at Brockport)</i>
Workflow Automation in Liquid Chromatography Mass Spectrometry .638.....	
	<i>Reinhard Gentz (Lawrence Berkeley National Laboratory), Hector Garcia Martin (Joint BioEnergy Institute), Edward Baidoo (Joint BioEnergy Institute), and Sean Peisert (Lawrence Berkeley National Laboratory)</i>
A Vision Towards Future eScience .640.....	
	<i>Shinji Shimojo (Osaka University) and Susumu Date (Osaka University)</i>
HUBzero© Goes OneSciencePlace: The Next Community-Driven Steps for Providing Software-as-a-Service .642	
	<i>David Benham (Purdue University) and Sandra Gesing (University of Notre Dame)</i>
Sharing and Archiving Data Science Course Projects to Support Pedagogy for Future Cohorts .644.....	
	<i>Stephanie Labou (Library, University of California San Diego), Ho Jung Yoo (Library, University of California San Diego), David Minor (Library, University of California San Diego), and Ilkay Altintas (San Diego Supercomputer Center, University of California San Diego)</i>
Expanding Library Resources for Data and Compute-Intensive Education and Research .646.....	
	<i>Stephanie Labou (University of California San Diego) and Reid Otsuji (University of California San Diego)</i>
Predicting Eating Events in Free Living Individuals .648.....	
	<i>Jiue-An Yang (University of California San Diego), Jiayi Wang (University of California San Diego), Supun Nakandala (University of California San Diego), Arun Kumar (University of California San Diego), and Marta M. Jankowska (University of California San Diego)</i>
Author Index 651	