

2024 IEEE 20th International Conference on e-Science (e-Science 2024)

**Osaka, Japan
16-20 September 2024**



**IEEE Catalog Number: CFP2406A-POD
ISBN: 979-8-3503-6562-7**

**Copyright © 2024 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:	CFP2406A-POD
ISBN (Print-On-Demand):	979-8-3503-6562-7
ISBN (Online):	979-8-3503-6561-0
ISSN:	2325-372X

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

Full Papers

Evaluating Active-learning Based Performance Prediction of Parallel Applications.....	1
<i>Shivam Aggarwal and Preeti Malakar</i>	
Autonomous Electrochemistry Platform with Real-Time Normality Testing of Voltammetry Measurements Using ML.....	11
<i>Anees Al-Najjar, Nageswara Rao, Craig Bridges, Sheng Dai and Alex Walters</i>	
Streamlined Edge Computing for Fire Science and Management using WIFIRE Edge....	21
<i>Ilkay Altintas, Shweta Purawat, Ismael Perez, Jenny Lee, Melissa Floca, Jessica Block, Josh Breslow and Daniel Crawl</i>	
An Empirical Investigation of Container Building Strategies and Warm Times to Reduce Cold Starts in Scientific Computing Serverless Functions.....	31
<i>André Bauer, Maxime Gonthier, Ryan Chard, Haochen Pan, Daniel Grzenda, Martin Straesser, J. Gregory Pauloski, Alok Kamatar, Matthew E. Baughman, Nathaniel Hudson, Ian Foster and Kyle Chard</i>	
Predicting Gas in 3D Dark Matter N-body Simulations with Convolutional Neural Networks.....	41
<i>Miguel Conceição, Alberto Krone Martins and Antonio da Silva</i>	
Reproducibility in Named Entity Recognition: A Case Study Analysis.....	50
<i>Carlos Cuevas Villarmin, Sarah Cohen-Boulakia and Nona Naderi</i>	
CuttleFlow: Infrastructure-Specific Workflow Adaption for Improved Reusability.....	60
<i>Ninon De Mecquenem, Simon Bosse, Vasilis Bountris, Somayeh Mohammadi, Knut Reinert and Ulf Leser</i>	
AI in Space for Scientific Missions: Strategies for Minimizing Neural-Network Model Upload.....	70
<i>Jonah Ekelund, Ricardo Vinuesa, Yuri Khotyaintsev, Pierre Henri, Gian Luca Delzanno and Stefano Markidis</i>	
Toward Reliable Biodiversity Information Extraction From Large Language Models....	80
<i>Michael Elliott and Jose Fortes</i>	

AstroMAE: Redshift Prediction Using a Masked Autoencoder with a Novel Fine-Tuning Architecture.....	90
<i>Geoffrey C. Fox and Amirreza Dolatpour Fathkouhi</i>	
CHARM-SYCL & IRIS: A Toolchain for Performance Portability on Extremely Heterogeneous Systems.....	100
<i>Norihisa Fujita, Beau Johnston, Narasinga Rao Miniskar, Ryohei Kobayashi, Mohammad Alaul Haque Monil, Keita Teranishi, Seyong Lee, Jeffrey S. Vetter and Taisuke Boku</i>	
A General Framework for Error-controlled Unstructured Scientific Data Compression.....	110
<i>Qian Gong, Zhe Wang, Viktor Reshniak, Xin Liang, Jieyang Chen, Qing Liu, Tushar Athawale, Yi Ju, Anand Rangarajan, Sanjay Ranka, Rick Archibald and Scott Klasky</i>	
A Personalized AI Assistant For Intuition-Driven Visual Explorations.....	120
<i>James Hammer, Tanner Hobson, David Pugmire, Scott Klasky, Ken Moreland and Jian Huang</i>	
Predicting Alzheimer’s Disease and Mild Cognitive Impairment with Off-line and On-line House Drawing Tests.....	130
<i>Nina Hosseini-Kivanani, Elena Salobrar-García, Lorena Elvira-Hurtado, Mario Salas, Christoph Schommer and Luis A. Leiva</i>	
MTP: A Cloud-based Real-time Mobile Tele-medicine Platform - Systems Paper.....	140
<i>Glenn Tesla Jayaputera, Andrew Bivard, Ivo Widjaya, James Jayaputera, Rio Susanto, Yunjie Jia, Yao Pan and Richard Sinnott</i>	
Enhancing Autonomous Intrusion Detection System with Generative Adversarial Networks.....	149
<i>Kevin Kostage, David West, Tim Meinert, Chengyi Qu, Prasad Calyam and Luca Mazzola</i>	
Ponder: Online Prediction of Task Memory Requirements for Scientific Workflows....	159
<i>Fabian Lehmann, Jonathan Bader, Ninon De Mecquenem, Xing Wang, Vasilis Bountris, Florian Friederici, Ulf Leser and Lauritz Thamsen</i>	
Deriva-ML: A Continuous FAIRness Approach to Reproducible Machine Learning Models.....	169
<i>Zhiwei Li, Carl Kesselman, Mike D'Arcy, Michael Pazzani and Benjamin Xu</i>	

Morbius: Motif Discovery Hardware Acceleration via Gibbs Sampling.....	179
<i>Se-Min Lim, Esmerald Aliaj and Sang-Woo Jun</i>	
Generative AI for Research Data Processing: Lessons Learnt From Three Use Cases...	189
<i>Modhurita Mitra, Martine de Vos, Nicola Cortinovia and Dawa Ometto</i>	
Griffin: Fast Transactional Database Index with Hash and B+-Tree.....	199
<i>Sho Nakazono, Yutaro Bessho, Hideyuki Kawashima and Tatsuhiko Nakamori</i>	
FaaSr: Cross-Platform Function-as-a-Service Serverless Scientific Workflows in R.....	209
<i>Sungjae Park, R. Quinn Thomas, Cayelan C. Carey, Austin D. Delany, Yun-Jung Ku, Mary E. Lofton and Renato J. Figueiredo</i>	
TaPS: A Performance Evaluation Suite for Task-based Execution Frameworks.....	219
<i>J. Gregory Pauloski, Valerie Hayot-Sasson, Maxime Gonthier, Nathaniel Hudson, Haochen Pan, Sicheng Zhou, Ian Foster and Kyle Chard</i>	
Dynamic Tracking, MLOps, and Workflow Integration: Enabling Transparent Reproducibility in Machine Learning.....	229
<i>Hamza Safri, George Papadimitriou and Ewa Deelman</i>	
There is Strength in Numbers: A Comprehensive Study of Machine Learning Algorithms for Sex Identification on Animal Bone Remains.....	239
<i>Nadine-Sarah Schöler, Ptolemaios Dimitrios Paxinos, Jing Yuan, Maximilian von Zastrow, Joris Peters and Peer Kröger</i>	
A Performance Comparison of Convolutional Neural Networks and Transformer-Based Models for Classification of the Spread of Bushfires.....	249
<i>Richard Sinnott, Chaojun Tang and Glenn Jayaputera</i>	
Mission Reproducibility: An Investigation on Reproducibility Issues in Machine Learning and Information Retrieval Research.....	258
<i>Moritz Staudinger, Bettina M. J. Kern, Tomasz Miksa, Lukas Arnhold, Peter Knees, Andreas Rauber and Allan Hanbury</i>	
Patra ModelCards: AI/ML Accountability in the Edge-Cloud Continuum.....	267
<i>Sachith Withana and Beth Plale</i>	

Advancing frances: New Heritage Textual Ontology, Enhanced Knowledge Graphs, and Refined Search Capabilities.....	277
---	-----

Lilin Yu, Ash Charlton, Melissa Terras and Rosa Filgueira

Multi-Level AI-Driven Analysis of Software Repository Similarities.....	287
---	-----

Honglin Zhang, Leyu Zhang, Lei Fang and Rosa Filgueira

Poster Extended Abstracts

FAIR Metrics for motivating excellence in peer review.....	297
--	-----

Adam Craig and Carl Taswell

CrowdAL: Towards a Blockchain-empowered Active Learning System in Crowd Data Labeling.....	299
--	-----

Shaojie Hou, Yuandou Wang and Zhiming Zhao

Estimating Metadata of Research Artifacts to Enhance their Findability.....	301
---	-----

Koichiro Ito and Shigeki Matsubara

Traffic-aware GTS Allocation Interval Determination for IEEE 802.15.4 DSME Networks.....	303
--	-----

Inhyeok Kang, Sol-Bee Lee, Jung-Hyok Kwon and Eui-Jik Kim

A Collaborative Framework for Facilitating Federated Learning among Jupyter Users.....	305
--	-----

Anandan Krishnasamy, Yuandou Wang and Zhiming Zhao

Enhancing Bayesian Optimization with NLML-Based Transfer Learning.....	307
--	-----

Yanchen Li and Taku Tsuzuki

X Marks the Spot? Applying Recent Keypoint Detection Methods to Paleozoological Landmarking.....	309
--	-----

Nadine-Sarah Schöler, Maximilian Archimedes Xaver von Zastrow, Nadja Pöllath, Claudius Zelenka and Joris Peters

Mitigating memory latency in FM Index search.....	311
---	-----

Igor Sfiligoi, Daniel McDonald and Rob Knight

Protocol-Level Kafka Controls in a Customizable Proxy.....	313
<i>Rein Spanjer and Adam Belloum</i>	
Project-Specific Research Data Management Beyond Repositories for FAIR Research Practice.....	315
<i>Steffen Strohm, Yannick Wölker and Matthias Renz</i>	
An Enhanced Credit-Based Shaper for Resilience to Time-Sync Misalignment.....	317
<i>Kohei Taniguchi, Arata Endo, Hirotake Abe, Lee Chonho, Kundjanasith Thonglek, Wassapon Watanakesuntorn, Junya Yamamoto and Susumu Date</i>	
SFETEC: Split-FEDerated Learning Scheme Optimized for Thing-Edge-Cloud Environment.....	319
<i>Le Van An, Jason Haga, Yusuke Tanimura and Truong Thao Nguyen</i>	
I/O of Scientific Workflows Monitored in Detail.....	321
<i>Joel Witzke, Ansgar Lößer, Vasilis Bountris, Florian Schintke and Björn Scheuermann</i>	
Automated Calibration of a Simulator of MPI Application Executions.....	323
<i>Yick Ching Wong, Frederic Suter, Kshitij Mehta, Henri Casanova and Jesse McDonald</i>	

Workshop on Generative AI and FAIR Principles in Science Communication (AI4SC)

AI4SC 2024 Workshop Foreword and Organisation.....	325
Discovering Research Areas in Dataset Applications through Knowledge Graphs and Large Language Models.....	327
<i>Irina Gerasimov, Armin Mehrabian, Binita Kc, Jerome Alfred and Michael McGuire</i>	
Data Drift for Automatic FAIR-compliant Dataset Versioning in Large Repositories....	337
<i>Alba Gonzalez-Cebrian, Iulian Ciolacu, Michael Bradford, Ciprian Dobre and Horacio Gonzalez-Velez</i>	
On Modeling User-Aware Proactive Communication in Conversational Recommendation Systems.....	347
<i>Amarnath Gupta</i>	

On the Design of a Searchable Encryption Protocol for Keyword Search using Proactive Secret Sharing.....	353
--	-----

Praveen K, Gabriel Anand K S, Indranil Ghosh Ray, Avishek Adhikari, Sabyasachi Datta and Arnab Kumar Biswas

AI for precision medicine must keep non-random complexity in mind to support equity in outcomes.....	361
--	-----

Benjamin Smarr

Workshop on Impact from eScience (IFE)

IFE 2024 Workshop Foreword and Organisation.....	368
--	-----

Data Skills for the Future: Leveraging University-Industry Collaborations.....	370
--	-----

Rhetta Chappell

How to Make eScience Work in the Real-world: from AVs to AI.....	373
--	-----

Miles Elsdon

Bridging the Data Skills Gap: The Role of the National Innovation Centre for Data and its Data Innovation Bootcamp.....	376
---	-----

Barry Hodgson

Creating Thriving Data-Centric Communities from Basic Research to Commercial Applications.....	379
--	-----

Robert Schuler and Carl Kesselman

Workshop on Citizen Science Engagement based on ICT Solutions (INSTIL)

INSTIL 2024 Workshop Foreword and Organisation.....	382
---	-----

Federated Learning and Crowdsourced Weather Data: Practice and Experience.....	384
--	-----

Ciro Giuseppe De Vita, Gennaro Mellone, Angelo Casolaro, Massimiliano Giordano Orsini, Jose Luis Gonzalez-Compean and Angelo Ciaramella

The Yeti in us – An app to facilitate behaviour change step by step.....	393
--	-----

Joy Ommer, Saša Vranić, Milan Kalas, Muhammad Adnan, Carlo Trozzi and Antonio Parodi

Citizens4Climate - A dashboard to support citizen science activities for climate action.....	400
<i>Saša Vranić, Antonella Galizia, Joy Ommer, Ulrike Falk, Paolo Mazzetti, Antonio Parodi and Milan Kalas</i>	

Workshop on Near Real-time Data Processing for Interconnected Scientific Instruments (NRDPISI)

NRDPISI 2024 Workshop Foreword and Organisation.....	407
An Analysis of Research Data Storage Systems.....	409
<i>David Abramson, Jake Carroll and Bronis de Supinski</i>	
Accelerating In-transit Isosurface Generation With Topology Preserving Compression.....	418
<i>Yanliang Li and Jieyang Chen</i>	
Diaspora: Resilience-Enabling Services for Real-Time Distributed Workflows.....	421
<i>Bogdan Nicolae, Justin Wozniak, Tekin Bicer, Hai Nguyen, Parth Patel, Haochen Pan, Amal Gueroudji, Maxime Gonthier, Valerie Hayot-Sasson, Eliu Huerta, Kyle Chard, Ryan Chard, Matthieu Dorier, Nageswara S. V. Rao, Anees Al-Najjar, Alessandra Corsi and Ian Foster</i>	
Top Research Challenges and Opportunities for Near Real-Time Extreme-Scale Visualization of Scientific Data.....	430
<i>David Pugmire, Kenneth Moreland, Tushar Athawale, James Hammer and Jian Huang</i>	
A Generic Performance Metric for Scientific Data Compression.....	436
<i>Zhenlu Qin, Jessie Gu, Maggie Liu, Jinzhen Wang and Hongjian Zhu</i>	
Towards Resilient Near Real-Time Analysis Workflows in Fusion Energy Science.....	439
<i>Frederic Suter, Norbert Podhorszki and Scott Klasky</i>	
Towards Development of University-wide Data Aggregation and Management Infrastructure for Research Data Utilization.....	445
<i>Hideyuki Tanushi, Hiroshi Furutani, Takeo Hosomi, Naoto Kai, Kaname Harumoto and Susumu Date</i>	

Workshop on Reproducible Workflows, Data Management, and Security (ReWorDS)

ReWorDS 2024 Workshop Foreword and Organisation.....	449
KS+: Predicting Workflow Task Memory Usage Over Time.....	451
<i>Jonathan Bader, Ansgar Lößer, Lauritz Thamsen, Björn Scheuermann and Odej Kao</i>	
Economy-based Greedy Resource Bidding for CAE Workflows in Hybrid Cloud Infrastructure.....	458
<i>Srishti Dasgupta, Tähvend Uustalu, Michael Gerndt and Babak Gholami</i>	
FAIR Assessment of Cloud-based Experiments.....	465
<i>Krishna Kamath, Nicole Brewer and Tanu Malik</i>	
Processing FHIR in modern Data Lakehouse.....	471
<i>Shreyas Patil and Adam Belloum</i>	
Workflow Provenance in the Computing Continuum for Responsible, Trustworthy, and Energy-efficient AI.....	478
<i>Renan Souza, Silvina Caino-Lores, Mark Coletti, Tyler Skluzacek, Alexandru Costan, Frederic Suter, Marta Mattoso and Rafael Ferreira da Silva</i>	

Workshop on Sustainable Project Pathways for HPC Software and Applications

2024 Workshop Foreword and Organisation.....	485
--	-----

Workshop on Global Research Platform

2024 Workshop Foreword and Organisation.....	487
--	-----