

2011 IEEE 7th International Conference on E-Science

(e-Science 2011)

**Stockholm, Sweden
5 – 8 December 2011**



IEEE Catalog Number: CFP1106A-PRT
ISBN: 978-1-4577-2163-2

2011 Seventh IEEE International Conference on eScience

eScience 2011

Table of Contents

| | |
|--------------------------------------|-----|
| Message from the General Chairs..... | x |
| Message from the Program Chair..... | xi |
| Committee Lists..... | xii |
| Reviewers..... | xiv |

Session 1A (Bioinformatics and Health/Tools)

| | |
|--|----|
| A Virtual Research Environment for Cancer Imaging Research | 1 |
| <i>M.S. Avila-Garcia, X. Xiong, A.E. Trefethen, C. Crichton, A. Tsui, and P. Hu</i> | |
| An Open-source Collaboration Environment for Metagenomics Research | 7 |
| <i>Xiaoquan Su, Yongzheng Ma, Hongwei Yang, Xingzhi Chang, Kai Nan, Jian Xu, and Kang Ning</i> | |
| rCAD: A Novel Database Schema for the Comparative Analysis of RNA | 15 |
| <i>Stuart Ozer, Kishore J. Doshi, Weijia Xu, and Robin R. Gutell</i> | |

Session 1B (Tools)

| | |
|--|----|
| A National Grid Submission Gateway for eScience | 23 |
| <i>Shunde Zhang, Paul Coddington, and Andrew Wendelborn</i> | |
| Investigating the Use of Gadgets, Widgets, and OpenSocial to Build Science Gateways | 31 |
| <i>Zhenhua Guo, Raminderjeet Singh, Marlon Pierce, and Yan Liu</i> | |
| A Chemistry-Inspired Workflow Management System for Scientific Applications in Clouds | 39 |
| <i>Héctor Fernández, Cédric Tedeschi, and Thierry Priol</i> | |

Session 2A (Bioinformatics and Health/Tools)

| | |
|--|----|
| A Cloud-based Dynamic Workflow for Mass Spectrometry Data Analysis | 47 |
| <i>Ashish Nagavaram, Gagan Agrawal, Michael A. Freitas, Kelly H. Telu, Gaurang Mehta, Rajiv G. Mayani, and Ewa Deelman</i> | |
| Creating a Cloud-based Life Science Gateway | 55 |
| <i>Wenjun Wu, Hui Zhang, ZhenAn Li, and Yaokuan Mao</i> | |
| Optimizing Phylogenetic Analysis Using SciHm Cloud-based Scientific Workflow | 62 |
| <i>Kary A.C.S. Ocaña, Daniel de Oliveira, Jonas Dias, Eduardo Ogasawara, and Marta Mattoso</i> | |

Session 2B (Climate and Earth Sciences)

| | |
|--|----|
| Wireless Sensing Networks for Environmental Monitoring: Two Case Studies from Tropical Forests | 70 |
| <i>G. Arturo Sánchez-Azofeifa, Cassidy Rankine, Mario Marcos do Espirito Santo, Rob Fatland, and Milton Garcia</i> | |
| On the Powerful Use of Simulations in the Quake-Catcher Network to Efficiently Position Low-cost Earthquake Sensors | 77 |
| <i>K. Benson, T. Estrada, M. Tauffer, J. Lawrence, and E. Cochran</i> | |
| Communicating Coastal Risk Analysis in an Age of Climate Change | 85 |
| <i>Brian Blanton, John McGee, and Oleg Kapeljushnik</i> | |

Session 3A (Bioinformatics and Health/Tools)

| | |
|---|-----|
| Application of Data Mining in Research of Avian Influenza Virus Cross-Species Infection | 89 |
| <i>Shasha Li, Yuanchun Zhou, Jianhui Li, Ze Luo, Zheng Kou, Tianxian Li, and Baoping Yan</i> | |
| Fostering Scientific Workflow Preservation through Discovery of Substitute Services | 97 |
| <i>Khalid Belhajjame, Carole Goble, Stian Soiland-Reyes, and Davide De Roure</i> | |
| A Novel Framework for Monitoring and Analyzing Quality of Data in Simulation Workflows | 105 |
| <i>Michael Reiter, Uwe Breitenbücher, Shahram Dustdar, Dimka Karastoyanova, Frank Leymann, and Hong-Linh Truong</i> | |

Session 3B (Climate and Earth Sciences)

| | |
|--|-----|
| Transient Modeling of Permafrost Dynamics in Changing Climate Scenarios | 113 |
| <i>Andrew Frampton, Georgia Destouni, Ylva Sjöberg, and Scott Painter</i> | |
| Evolving Inversion Methods in Geophysics with Cloud Computing - A Case Study of an eScience Collaboration | 119 |
| <i>J. Craig Mudge, Pinaki Chandrasekhar, Graham S. Heinson, and Stephan Thiel</i> | |
| Assessing the Value of Cloudbursting: A Case Study of Satellite Image Processing on Windows Azure | 126 |
| <i>Marty Humphrey, Zach Hill, Catharine van Ingen, Keith Jackson, and Youngryel Ryu</i> | |

Session 4A (Data Management and Digital Repositories)

| | |
|---|-----|
| A Provenance Approach to Trace Scientific Experiments on a Grid Infrastructure | 134 |
| <i>Ammar Benabdelkader, Mark Santcroos, Souley Madougou, Antoine H.C. van Kampen, and Silvia D. Olabarriaga</i> | |
| The SinBiota 2.0 Biodiversity Information System | 142 |
| <i>Cleber Mira, Pedro Feijao, Tiago Duque-Estrada, Joao Meidanis, and Carlos A. Joly</i> | |

| | |
|--|-----|
| Large Scale Participatory Acoustic Sensor Data Analysis: Tools and Reputation Models to Enhance Effectiveness | 150 |
| <i>Anthony Truskinger, Haofan Yang, Jason Wimmer, Jinglan Zhang, Ian Williamson, and Paul Roe</i> | |
| Session 4B (Arts, Humanities and eSocial Science) | |
| Preparing DARIAH | 158 |
| <i>Tobias Blanke, Michael Bryant, Mark Hedges, Andreas Aschenbrenner, and Michael Priddy</i> | |
| e-Science and the Archaeological Frontier | 166 |
| <i>Aaron Gidding, Yuma Matsui, Thomas E. Levy, Tom DeFanti, and Falko Kuester</i> | |
| Building a Semantic Knowledge-base for Painting Conservators | 173 |
| <i>Jane Hunter and Suleiman Odat</i> | |
| Session 4C (Physical Sciences and Engineering) | |
| Building and Operating a Distributed Regional Centre for LHC Computing and Data Storage | 181 |
| <i>Oxana Smirnova, Gerd Behrmann, Michael Grønager, Daniel Johansson, Josva Kleist, Aleksandr Konstantinov, Balázs Kónya, and Henrik Thostrup Jensen</i> | |
| iGrid: Interactive Grid | 189 |
| <i>Marco Meoni</i> | |
| Chinese e-VLBI Network: A Multi-purpose e-science Platform | 197 |
| <i>Zheng Weimin, Chen Zhong, Wang Guangli, Shu Fengchun, Wang Weihua, Chen Xiao, and An Tao</i> | |
| Session 5A (Data Management and Digital Repositories) | |
| Adaptive Inference of Fine-grained Data Provenance to Achieve High Accuracy at Lower Storage Costs | 202 |
| <i>Mohammad Rezwanul Huq, Andreas Wombacher, and Peter M.G. Apers</i> | |
| A Scalable Architecture for e-Science Data Management | 210 |
| <i>Salman Toor, Manivasakan Sabesan, Sverker Holmgren, and Tore Risch</i> | |
| The Adaptive Collection and Analysis of Distributed Multimedia Sensor Data | 218 |
| <i>Mark Cottman-Fields, Anthony Truskinger, Jason Wimmer, and Paul Roe</i> | |
| Session 5B (Tools) | |
| Collaborative eResearch in a Social Cloud | 224 |
| <i>Ashfaq M. Thaufeeg, Kris Bubendorfer, and Kyle Chard</i> | |
| Dynamic Handling for Cooperating Scientific Web Services | 232 |
| <i>Reginald Cushing, Spiros Koulouzis, Adam Belloum, and Marian Bubak</i> | |
| <i>In Situ</i> Data Provenance Capture in Spreadsheets | 240 |
| <i>Hazeline U. Asuncion</i> | |

Session 5C (Physical Sciences and Engineering)

| | |
|---|-----|
| HELIO: Discovery and Analysis of Data in Heliophysics | 248 |
| <i>Robert Bentley, John Brooke, André Csillaghy, Donal Fellows, Anja Le Blanc, Mauro Messerotti, David Pérez-Suárez, Gabriele Pierantoni, and Marco Soldati</i> | |
| Scaling Dalton, A Molecular Electronic Structure Program | 256 |
| <i>Xavier Aguilar, Michael Schliephake, Olav Vahtras, Judit Gimenez, and Erwin Laure</i> | |
| Dynamic Pipeline Changes in Scientific Data Processing | 263 |
| <i>Johnson Mwebaze, Danny Boxhoorn, and Edwin Valentijn</i> | |

Session 6A (Data Management and Digital Repositories)

| | |
|--|-----|
| Enforcing Scientific Data Sharing Agreements | 271 |
| <i>Michael Wilson, Shirley Crompton, Brian Matthews, and Alexey Orlov</i> | |
| A Mosaic of Software | 279 |
| <i>Kenton McHenry, Rob Kooper, Michael Ondrejcek, Luigi Marini, and Peter Bajcsy</i> | |
| Creating Infrastructure for Tool-Independent Querying and Exploration of Scientific Workflows | 287 |
| <i>Valentina Ivanova and Lena Strömbäck</i> | |

Session 6B (Data Management/Climate and Earth Sciences)

| | |
|---|-----|
| Towards an e-Infrastructure for Urban Research across Australia | 295 |
| <i>Richard O. Sinnott, Gerson Galang, Martin Tomko, and Robert Stimson</i> | |
| A Linked Data Approach to Publishing Complex Scientific Workflows | 303 |
| <i>Arif Shaon, Sarah Callaghan, Bryan Lawrence, Brian Matthews, Andrew Woolf, Timothy Osborn, and Colin Harpham</i> | |
| Experiences Using GlideinWMS and the Corral Frontend across Cyberinfrastructures | 311 |
| <i>Mats Rynge, Gideon Juve, Gaurang Mehta, Ewa Deelman, Krista Larson, Burt Holzman, Igor Sfiligoi, Frank Würthwein, G. Bruce Berriman, and Scott Callaghan</i> | |

Session 6C (Physical Sciences and Engineering)

| | |
|--|-----|
| Large-scale Simulations of Turbulence: HPC and Numerical Experiments | 319 |
| <i>Philipp Schlatter, Johan Malm, Geert Brethouwer, Arne V. Johansson, and Dan S. Henningson</i> | |
| The Saaz Framework for Turbulent Flow Queries | 325 |
| <i>Alden King, Eric Arobone, Scott B. Baden, and Sutanu Sarkar</i> | |
| Parallel Finite Element Operator Application: Graph Partitioning and Coloring | 332 |
| <i>Katharina Kormann and Martin Kronbichler</i> | |

Session 7A (Education and eScience Practice)

| | |
|--|-----|
| Cyberinfrastructure Internship and its Application to e-Science | 340 |
| <i>David Abramson, Peter Arzberger, Gabriele Wienhausen, Jim Galvin, Susumu Date, Fang-Pang Lin, Kai Nan, and Shinji Shimojo</i> | |
| Cyberinfrastructure and e-Science Application Practices in Chinese Academy of Sciences | 348 |
| <i>Huang Xiangyang, Luo Ze, and Yan Baoping</i> | |
| e-Science Practice for Wild Birds' Monitoring and Protection in Qinghai Lake Region | 355 |
| <i>Luo Ze, Yan Baoping, Zhou Yuanchun, Li Jian, Tang Mingjie, Zhang Haiming, Shao Jing, Li Shasha, Lei Fumin, Cui Peng, Li Tianxian, Kou Zheng, Hou Yuansheng, He Yubang, and Xing Zhi</i> | |

Session 7B (Tools)

| | |
|--|-----|
| A Social Cloud for Public eResearch | 363 |
| <i>Koshy John, Kris Bubendorfer, and Kyle Chard</i> | |
| A User-orientated Electronic Laboratory Notebook for Retrieval and Extraction of Provenance Information for EUROCHAMP-2 | 371 |
| <i>Zulkifly Mohd Zaki, Peter M. Dew, Mohammed H. Haji, Lydia M.S. Lau, Andrew Rickard, and Jennifer Young</i> | |
| A Framework for Efficient Data Analytics through Automatic Configuration and Customization of Scientific Workflows | 379 |
| <i>Matheus Hauder, Yolanda Gil, and Yan Liu</i> | |

Session 7C (Physical Sciences and Engineering)

| | |
|---|-----|
| Design of a Parallel Hybrid Direct/Iterative Solver for CFD Problems | 387 |
| <i>Jonas Thies and Fred Wubs</i> | |
| CarbBuilder: An Adjustable Tool for Building 3D Molecular Structures of Carbohydrates for Molecular Simulation | 395 |
| <i>Michelle Kuttel, Yue Mao, Göran Widmalm, and Magnus Lundborg</i> | |
| An Approach to Optimise the Execution of RTM Algorithm in Multicore Machines | 403 |
| <i>Alexandre C. Sena, Aline P. Nascimento, Cristina Boeres, Vinod Rebello, and André Bulcão</i> | |

| | |
|---------------------------|-----|
| Author Index | 411 |
|---------------------------|-----|