

Co-Organizers

Center for Large Scale Data Systems (CLDS)
San Diego Supercomputing Center
University of California, San Diego, USA

Indian Statistical Institute (ISI)
Delhi Centre

Indian Institute of Public Health (IIPH)
Hyderabad

SEVENTH WORKSHOP ON

BIG DATA BENCHMARKING

(WBDB 2015.in)

Dates: **DECEMBER 14-15, 2015**

Venue: **India Habitat Centre, New Delhi, India**

Partners



NASSCOM

Sponsors



Program Committee

PC Chair

Prof. Saumyadipta Pyne, Indian Institute of Public Health (IIPH), Hyderabad

PC Members

Dr. Amitabha Bagchi, IIT Delhi
Dr. Chaitan Baru, UC San Diego
Dr. Milind Bhandarkar, BigData Top100 List
Prof. Pulak Ghosh, IIM Bangalore
Dr. Anil Goel, SAP
Bhaskar Gowda, Intel
Prof. Jayadeva, IIT Delhi
Dr. Rajendra Joshi, C-DAC, Pune, India
Prof. Arnab Laha, IIM Ahmedabad
Dr. Tridib Mukherjee, Xerox Research Center India
Raghunath Nambiar, Cisco

Dr. Ullas Nambiar, EMC India
Scott Pearson, Cray
Meikel Poess, Oracle
Dr. Tilmann Rabl, Technical University-Berlin
Dr. V.C.V. Rao, C-DAC, Pune, India
Prof. Yogesh Simmhan, IISc, India
Dr. Raghavendra Singh, IBM, New Delh
Prof. Dinkar Sitaram, PES University, Bangalore, India
Dr. Chiranjib Sur, Shell, India
Prof. Anil Vullikanti, Virginia Bioinformatics Institute, USA

Organizing Committee

- Prof. C.R. Rao, Eberly Professor Emeritus of Statistics, Penn State University & Research Professor, SUNY Buffalo
- Prof. Bimal K. Roy, Former Director & Professor, Indian Statistical Institute (ISI), Kolkata
- Dr. V. Rao Aiyagari, Senior Adviser, Public Health Foundation of India
- Prof. Sanghamitra Bandyopadhyay, Director, Indian Statistical Institute (ISI), Kolkata
- Prof. Bimal K. Roy, Former Director & Professor, Indian Statistical Institute (ISI), Kolkata
- Prof. Abhay G. Bhatt, Indian Statistical Institute (ISI), Delhi Centre
- Prof. Santanu Chaudhury, Dhananjay Chair Professor, Dept. of Electrical Engg., I.I.T., Delhi
- Dr. Manish Gupta, VP and Director, Xerox Research Center India
- Prof. Rajeeva L. Karandikar, Director, Chennai Mathematical Institute, Chennai
- Dr. Ajay Kumar, Joint Secretary, DeitY & DG, National Informatics Center (NIC), India
- Prof. Bipin Mehta, President, Computer Society of India
- Mr. H.R. Mohan, Former President, Computer Society of India & Chairman, IEEE Computer Society, Madras
- Prof. Rajat Moona, Director General, Centre for Development of Advanced Computing (C-DAC), Pune
- Prof. B.L.S. Prakasa Rao, Ramanujan Chair Professor, CRRao AIMSCS, Hyderabad
- Prof. S.B. Rao, Emeritus Professor, CRRao AIMSCS, Hyderabad
- Dr. Deepayan Sarkar, Indian Statistical Institute (ISI), Delhi Centre
- Prof. U.C. Sud, Director, ICAR - Indian Agricultural Statistics Research Institute, New Delhi

Health Systems Mini-Workshop Co-Chairs

Dr. Suresh Munuswamy, IIPH, Hyderabad
Dr. Yogesh Simmhan, IISc, Bangalore
Dr. Ullas Nambiar, EMC

Agriculture Mini-Workshop Co-Chairs

Prof. Sanjay Chaudhary, IET, Ahmedabad University
Prof. Rajender Parsad, ICAR-IASR, Delhi
Prof. Karmeshu, JNU, Delhi

Air Quality Management Mini-Workshop Co-Chairs

Dr. S.M. Shiva Nagendra, IIT Madras
Prof. Mukesh Khare, IIT Delhi
Prof. Prateek Sharma, TERI University, Delhi

Poster Co-Chairs

Prof. Dinkar Sitaram, PES University, Bangalore
Dr. Tridib Mukherjee, Xerox, Bangalore
Prof. Indranil Mukhopadhyay, ISI Kolkata
Dr. VCV Rao, C-DAC, Pune

UNIVERSITY OF CALIFORNIA, SAN DIEGO

UCSD

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO



SANTA BARBARA • SANTA CRUZ

MICHAEL L. NORMAN, DIRECTOR
SAN DIEGO SUPERCOMPUTER CENTER

DISTINGUISHED PROFESSOR,
DEPARTMENT OF PHYSICS

November 25, 2015

9500 GILMAN DRIVE, MC0505
LA JOLLA, CALIFORNIA 92093-0505
TELEPHONE: (858) 822-5461
FAX: (858) 534-5056



Dear Dr. Baru and Co-organizers of the 7th WBDB:

Congratulations on organizing the 7th Workshop on Big Data Benchmarking in New Delhi, India!

It gives me great pleasure to see that the WBDB workshop series, initiated in 2012 by the Center for Large-Scale Data Systems Research (CLDS) at the San Diego Supercomputer Center, University of California San Diego, is now in its seventh iteration and that WBDB 2015 is being jointly organized with the Indian Statistical Institute, New Delhi and the Indian Institute of Public Health, Hyderabad. I am encouraged to see that industry and academia have convened this ongoing series of workshops and meetings on the topic of Big Data benchmarking (<http://clds.ucsd.edu/bdbc/workshops>).

Research, development, and use of performance benchmarks in high-performance computing (HPC) and in database systems (TPC) has been active for over 20 years, as evidenced by the development of LINPACK, the emergence of the TOP500 list, and the creation of TPC-C, TPC-H and other TPC benchmarks. Recent benchmark efforts have sought to broaden the set of relevant benchmarks to cover performance under different application environments, such as Graph500 for graph processing. Energy efficiency has also emerged in recent years as a relevant and increasingly important area of measurement and profiling for HPC systems (e.g., Green500). With the explosion in Big Data applications, the community has also gained interest in developing reference approaches to measure and characterize system performance for large-scale data analysis hardware and software systems, such as the proposed Big Data Top 100. This remains an area of active research, development, and community discussion.

Given the emergence of inference-based computing, the growing role of data analysis, changes in scientific workflow due to dynamic availability of sensor and instrument data, the expanding use of large-scale computing in all scientific disciplines, the growing role of clouds, and a diversity of architectural approaches, this is a timely opportunity to engage the community in benchmark activities.

I offer my sincere, best wishes to you and your co-organizers for a successful workshop.

Sincerely yours,

A handwritten signature in blue ink that reads "Michael L. Norman".

Michael L. Norman
Director, San Diego Supercomputer Center
Distinguished Professor of Physics

भारतीय सांख्यिकीय संस्थान
INDIAN STATISTICAL INSTITUTE

प्रोफेसर संध्यामित्रा बंद्योपाध्याय
Professor Sanghamitra Bandyopadhyay
निदेशिका/Director



203, बैरकपूर ट्रंक रोड/Barrackpore Trunk Road
कोलकाता/Kolkata 700 108, भारत/India
दूरभाष/Telephones : 033-2577-3084
फैक्स/Fax : 033-2577-0925/0109
ईमेल/e-mail : director@isical.ac.in

No. D.O./2015/789
24 November 2015

Professor Saumyadipta Pyne
Indian Institute of Public Health (IIPH), Hyderabad.
Plot # 1, Road 44
Kavuri Hills, Madhapur
Hyderabad 500 033

Dear Professor Pyne,

On behalf of the Indian Statistical Institute, I would like to convey my sincere congratulations and greetings to the organizing committee of the Seventh International Workshop in Big Data Benchmarking (WBDB 2015) for organizing the Workshop in India in collaboration with Indian Statistical Institute, San Diego Supercomputer Center, UCSD, and Public Health Foundation of India.

As Big Data Analytics is a rapidly developing field, without such global efforts in benchmarking, it is difficult to assess the quality and the utility of the myriad technological platforms that are being designed for different types of data and applications.

Indian Statistical Institute is known for its historic leadership role in the areas of Statistical Quality Control and Operations Research that has long benefited Indian industry. Likewise, the design of Big Data solutions that meet the required standards of its domains of application will be beneficial to researchers, administrative officials as well as IT and other industries alike.

The WBDB 2015 programme is very rich in content and it will offer the participants a high quality event. I wish you and the Workshop participants an enriching and enjoyable experience.

Yours sincerely,

(Sanghamitra Bandyopadhyay)

प्रोफेसर संध्यामित्रा बंद्योपाध्याय
Prof. Sanghamitra Bandyopadhyay
निदेशिका/Director
भारतीय सांख्यिकीय संस्थान
INDIAN STATISTICAL INSTITUTE
203, बैरकपूर ट्रंक रोड, कोलकाता - 700 108
203, Barrackpore Trunk Road, Kol-700 108



PUBLIC
HEALTH
FOUNDATION
OF INDIA

December 1, 2015



Dear Prof. Pyne and the WBDB Organizers,

On behalf of Public Health Foundation of India, on the occasion of the Seventh International Workshop in Big Data Benchmarking 2015, I congratulate you for organizing this timely and important event in New Delhi. Congratulations to the co-Organizers from the Indian Statistical Institute in Delhi, and San Diego Supercomputing Center in University of California at San Diego.

The Big Data approach has the potential to help improve the health and economic conditions for large sections of our society. As data from different sources are getting larger and more varied, benchmarking efforts will assume increasing importance in the future. Efficient data-driven models may be used, for example, to predict potential disease outbreaks and their spread in the population, and thus provide key support to public health decision-making and policy.

Big Data projects may also help the society with risk estimation as needed to seek the right balance between its energy needs and environmental costs and its general dynamics. Therefore I am glad to note that many distinguished experts will speak on important areas such as Health Systems, Air Quality Management and Agriculture.

The excellent and extensive list of notable speakers and presenters will surely make the 7th WBDB a wonderful and memorable event. I wish all the participants and delegates a very successful workshop.

Sincerely yours,

Prof. K. Srinath Reddy
President
Public Health Foundation of India



Program

AGENDA

MONDAY, DECEMBER 14, 2015

SESSION 1

0930	0945	Welcome Address Inaugural Message	Prof. Saumyadipta Pyne Prof. C.R. Rao	Public Health Foundation of India Padma Vibhushan and Receptient of India Science Award
0945	1000	Introduction to WBDB	Dr. Chaitan Baru Dr. Tilmann Rabl	Associate Director, San Diego Supercomputer Center, UC San Diego, USA Research Director, Berlin Big Data Center, Technical University-Berlin, Germany
1000	1100	Keynote Talk: Emerging Trends in Big Data Software	Prof. Michael Franklin	Thomas M. Siebel Professor and Chair of Computer Science Division and Director of Algorithms, Machines, and People Laboratory (AMPLab), in University of California Berkeley, USA
1100	1130	BREAK		

SESSION 2 Chair: Dr. Milind Bhandarkar

1130	1200	Survey of Big Data Benchmarks	Dr. Chaitan Baru Dr. Tilmann Rabl	Associate Director, SDSC UC San Diego, USA Research Director, Berlin Big Data Center, Technical University Berlin, Germany
1200	1230	Accelerating BigBench on Hadoop	Tang Yang	Intel
1230	1300	Measuring Storage Access Performance in Array Databases	Prof. Peter Baumann Mr. George Merticariu	Computer Science, Jacobs University Bremen, Germany
1300	1315	Feature of Analytics with Ampool	Dr. Milind Bhandarkar	CEO and Founder, Ampool
1315	1415	LUNCH		

SESSION 3 Dr. Tilmann Rabl

1415	1515	Keynote Talk: Leveraging Web Technologies to Reimagine the Database Architecture	Dr. Mukund Deshpande	Head of Analytics, Persistent Systems, Pune, India
1515	1545	Benchmarking Fast Data Platforms for the Aadhaar Biometric Database	Dr. Yogesh Simmhan	Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore
1545	1630	Poster Session and Break		

SESSION 4 Dr. Chaitanya Baru

1630	1700	Featured Talk: Classifying Simulation and Data Intensive Applications and the HPC-Big Data Convergence	Prof. Geoffrey C. Fox	Distinguished Professor of Computer Science and Informatics and Physics, and Director of Digital Science Center, Indiana University, Bloomington, Indiana, USA
1700	1730	Invited Talk: Analytics and Learning on GPUs	Thomas Bradley	Senior Manager of Developer Technology; EMEA, Russia and India, NVIDIA

WORKSHOP DINNER

1900	1930	Special Address	Prof. Ashutosh Sharma	Secretary, Dept. of Science and Technology, Govt. of India
1930	2200	Dinner (Sponsored by NVIDIA)	India Habitat Center	Complimentary for registered participants, speakers and invited delegates

Program

AGENDA

TUESDAY DECEMBER 15, 2015

SESSION 5 Prof. GVS Murthy

0900	0920	Introduction to Day 2	Prof. Saumyadipta Pyne	Public Health Foundation of India
0920	0945	Featured Talk: Transformational Opportunities for Healthcare enabled by Sensing and Data Analytics	Dr. Manish Gupta	Director, Xerox Research Centre India, Bangalore
0945	1010	Featured Talk: The role of big data in public health and medicine	Dr. Anurag Agrawal	Principal Scientist, CSIR Institute of Genomics & Integrative Biology, Delhi
1010	1030	Invited Talk: Data Analytics Opportunities and Challenges in Medical Records	Gaur Sunder	Principal Technical Officer, VC&BA, Centre for Development of Advanced Computing, Pune
1030	1045	Contributed Paper: Open Standards in Developing Large Scale Healthcare Applications in India	Shelly Sachdeva, Shivani Batra, Subhash Bhalla	Jaypee Institute of Information Technology University, Noida
1045	1100	Contributed Paper: Assessment of weight gain in children admitted to Nutrition Rehabilitation Center (NRC) and afterwards for two years in Sangareddy, Medak District	Dr. Sarala Kumari, Stephen Raj, Dr. Suresh Munuswamy	Indian Institute of Public Health, Hyderabad
1100	1130	BREAK		

SESSION 6 Dr. V. Rao Aiyagari

1130	1145	Introductory Talk: Big data analytics in urban air quality management	Prof. Mukesh Khare	Indian Institute of Technology Delhi
1145	1210	Featured Talk: Fuzzy Logic Formalism for Big Data Analytics to Manage Air Quality in Smart Cities	Prof. Ashok Deshpande	Chair, Berkeley Initiative on Soft Computing, University of California, Berkeley
1210	1235	Featured Talk: Directional Statistics for Complex Data Analysis	Prof. Ashis SenGupta	Indian Statistical Institute, Kolkata
1235	1245	Contributed Paper: Assessment of ambient air quality compliance criteria using statistical distribution modelling approach	Prof. Prateek Sharma	TERI, New Delhi
1245	1300	Contributed Paper: Low cost sensors for air quality management in smart cities	Y. Pavan and Dr. SM Shiva Nagendra	Indian Institute of Technology Madras
1300	1315	Contributed Paper: Statistical distribution of NO2 concentrations at air quality control region in urban area – A case study of Delhi city	Dr. Sunil Gulia, IITD, Dr. S. M. Shiva Nagendra, IITM, Prof. Mukesh Khare, IITD	Indian Institute of Technology Delhi
1315	1415	LUNCH		

Program

AGENDA

SESSION 7 Prof. Rajender Parsad

1415	1425	Introductory Talk: Big Data Analytics in Agriculture	Dr. Sanjay Chaudhary	Professor and Head (Research), Institute of Engineering and Technology, Ahmedabad University
1425	1450	Featured Talk: Satellite Data use in Agriculture	Dr. Shibendu S. Ray	Director, Mahalanobis National Crop Forecast Centre, New Delhi
1450	1515	Featured Talk: Big Data in Agricultural Research and Digital Agriculture	Dr. Abhishek Rathore	Senior Scientist, ICRISAT Hyderabad
1515	1530	Invited Talk: ICAR Research Data Repository for Knowledge Management	Dr. A. Dhandapani	Principal Scientist, ICAR-National Academy of Agricultural Research Management, Hyderabad
1530	1545	Invited Talk: Big Data initiative in Agricultural Bioinformatics	Dr. K.K. Chaturvedi	Scientist, CABin, ICAR-Indian Agricultural Statistics Research Institute, New Delhi
1545	1600	Contributed Paper: Big Data Analytics for Crop Recommendation System	Deepak Hiremath	Institute of Engineering and Technology, Ahmedabad University
1600	1630	BREAK		

SESSION 8 Dr. Yogesh Simmhan

1630	1715	Panel discussion	Dr. Chaitan Baru	Senior Advisor for Data Science, CISE, National Science Foundation, USA
			Dr. Rajendra Joshi	Associate Director and Head, Bioinformatics Group, Centre for Development of Advanced Computing, Pune
			Mr. Mohana Krishnan	Director, Analytics Special Interest Group, NASSCOM
			Dr. K. Murali Mohan	Head, Big Data Initiative, Department of Science & Technology, Government of India
			Prof. G.V.S. Murthy	Director, Indian Institute of Public Health, Hyderabad
			Prof. Dinkar Sitaram	Head, Centre for Cloud Computing and Big Data, PES University, Bangalore
1715	1730	Wrap Up	Prof. Saumyadipta Pyne	Public Health Foundation of India
			Dr. Tilmann Rabl	Research Director, Berlin Big Data Center, Technical University-Berlin, Germany

Speaker Profiles



Prof. ASHUTOSH SHARMA is the Secretary, Department of Science and Technology, Government of India since January 09, 2015. He was a professor (1997-), an Institute Chair Professor (2007-) and the Head (2003-05) of Chemical Engineering, and the founding Coordinator of Nanosciences Center and Advanced Imaging Center at the Indian Institute of Technology at Kanpur. Ashutosh received his PhD from the State University of New York at Buffalo (SUNYAB; 1988) working with Prof. Eli Ruckenstein, his MS from the Pennsylvania State University (1984) and B.Tech. from IIT Kanpur (1982). Prof. Sharma is a recipient of numerous honors and awards including the inaugural Infosys Prize in Engineering and Computer Science, TWAS Science Prize of the World Academy of Sciences, Bessel Research Award of the Humboldt Foundation, J. C. Bose Fellowship, Bhatnagar Prize, Homi J. Bhabha Award of UGC, The Syed Husain Zaheer Medal of INSA, Distinguished Alumnus Award of IIT Kanpur and the Life-time Achievement Award of the Indian Science Congress. He is an elected Fellow of The Indian National Science Academy, The Indian Academy of Sciences, The National Academy of Sciences, India and Indian National Academy of Engineering, The World Academy of Sciences (TWAS) and the Asia-Pacific Academy of Materials.



Prof. MICHAEL J FRANKLIN is the Thomas M. Siebel Professor of Computer Science and Chair of the Computer Science Division in the EECS Department at the University of California, Berkeley. He has over 30 years of experience in the database, data analytics, and data management fields as an academic and industrial researcher, teacher, lab director, faculty member, entrepreneur, and software developer. Prof. Franklin is also the Director of the Algorithms, Machines, and People Laboratory (AMPLab) at UC Berkeley, a leading academic Big Data analytics research center. The AMPLab has 27 industrial partners including founding sponsors Amazon Web Services, Google, IBM, and SAP, and received a National Science Foundation CISE "Expeditions in Computing" award, which was announced as part of the White House Big Data Research initiative in March 2012. AMPLab has produced industry-changing open source software including Apache Spark and BDAS, the Berkeley Data Analytics Stack. Prof. Franklin is a co-PI and Executive Committee member for the Berkeley Institute of Data Science, a campus-wide initiative to advance Data Science Environments. He was founder and CTO of Truviso, a data analytics company that was subsequently purchased by Cisco Systems. He currently serves on the Technical Advisory Boards of a number of data-driven technology companies, including Databricks, an AMPLab spinout. He is a Fellow of the ACM and a two-time winner of the ACM SIGMOD "Test of Time" award, and received the outstanding Advisor Award from the Computer Science Graduate Student Association at Berkeley. He received the Ph.D. in Computer Science from the University of Wisconsin in 1993, a Master of Software Engineering from the Wang Institute of Graduate Studies in 1986, and the B.S. in Computer and Information Science from the University of Massachusetts in 1983.



Dr. MUKUND DESHPANDE has over 18 years of experience in the field of analytics, working in variety of different roles. Currently, he leads the Analytics Practice at Persistent and is responsible for driving the overall strategy, vision and delivery of the unit. Along with Social, Mobile & Cloud, Analytics is one of four focus areas for Persistent in what has been one of the largest & fastest growing units in the organization. At Persistent Systems, Mukund has been involved in several cutting edge projects and internationally recognized award-winning initiatives. One of the marquee engagements was creating the analytics around social media for the TV show Satyameva Jayate, building innovative apps based on cutting edge text analytics and big data technologies. Prior to Persistent, Mukund was with Oracle Corporation, USA, where he was involved in developing heuristic driven algorithms for problems in the supply chain domain, focusing on performance and flexibility. Mukund was awarded second place in the KDD Cup 2000 competition, which is an Annual competition organized by ACM SIG on Knowledge and Data Discovery. Mukund holds a PhD in Computer Science from the University of Minnesota, Minneapolis and has a bachelor's degree in Mechanical Engineering from the Government College of Engineering Pune and a Master's degree in Engineering in System Science & Automation from the Indian Institute of Science, Bangalore.

Speaker Profiles



Prof. GEOFFREY CHARLES FOX received a Ph.D. in Theoretical Physics from Cambridge University and is now distinguished professor of Informatics and Computing, and Physics at Indiana University where he is director of the Digital Science Center, Chair of Department of Intelligent Systems Engineering and Director of the Data Science program at the School of Informatics and Computing. He previously held positions at Caltech, Syracuse University and Florida State University after being a postdoc at the Institute of Advanced Study at Princeton, Lawrence Berkeley Laboratory and Peterhouse College Cambridge. He has supervised the PhD of 67 students and published around 1200 papers in physics and computer science with an hindex of 70 and over 26000 citations. He currently works in applying computer science from infrastructure to analytics in Biology, Pathology, Sensor Clouds, Earthquake and Ice-sheet Science, Image processing, Deep Learning, Manufacturing, Network Science and Particle Physics. The infrastructure work is built around Software Defined Systems on Clouds and Clusters. The analytics focuses on scalable parallelism. He is involved in several projects to enhance the capabilities of Minority Serving Institutions. He has experience in online education and its use in MOOCs for areas like Data and Computational Science. He is a Fellow of APS (Physics) and ACM (Computing).



Dr. CHAĪTANYA BARU is Senior Advisor for Data Science in the Computer and Information Science and Engineering Directorate at the National Science Foundation, Arlington, Virginia. He is on assignment from the San Diego Supercomputer Center, University of California San Diego, where he is a Distinguished Scientist and Associate Director of Data Initiatives. His research interests are in database systems, scientific data management and big data. He is a member of the Steering Group of the Workshops on Big Data Benchmarking. Baru has a B.Tech from IIT Madras and an ME and PhD in Electrical Engineering from the University of Florida.



Dr. TILMANN RABL is co-founder and CEO of bankmark and a senior researcher at TU Berlin. He is professional affiliate of the Transaction Processing Performance Council, member of the steering group of the Workshops on Big Data Benchmarking, and chair of the SPEC RG Working Group on Big Data. He has received his PhD in computer science from the University of Passau. His research interests include big data, database systems, and benchmarking.



Prof. SAUMYADĪPTA PYNE is Professor in Public Health Foundation of India and is based at the Indian Institute of Public Health, Hyderabad. His research interests include Big Data in Life Sciences and Health Informatics, Computational Statistics and High-dimensional Data Modeling, and has published extensively in top international journals. He taught one of the earliest MSc courses in the country on Big Data Analytics in the University of Hyderabad, and is currently editing a Springer title on the subject. Pyne is the founder convener of the Computer Society of India's Special Interest Group on Big Data Analytics and promotes Big Data research in India and abroad. He co-Chaired the 2nd and 3rd IEEE Big Data workshops in 2014 & 2015 in USA, and the 7th Workshop on Big Data Benchmarking 2015 and the IEEE International Workshop on Foundations of Big Data Computing 2015 in India. Pyne received his PhD at the State University of New York at Stony Brook and did postdoctoral research in Broad Institute of MIT and Harvard University, USA. Formerly, he worked as PC Mahalanobis Chair Professor at CRRao AIMSCS, and at Harvard Medical School.



THOMAS BRADLEY holds a first-class MEng from the University of Bristol, UK, and l'École Nationale Supérieure de Télécommunications in Brest, France. Starting as a digital hardware designer, he worked as a processor architect for video encoding processors at STMicroelectronics before moving to ClearSpeed Technology plc to lead their architecture development for general purpose parallel processors. Since then he has specialised in High Performance Computing software development at ClearSpeed and at NVIDIA, where he is Senior Manager of Developer Technology, heading the GPU computing group in EMEA, Russia and India.

Speaker Profiles



Dr. MANISH GUPTA is Vice President at Xerox Corporation and Director of Xerox Research Centre in India. Previously, Manish has served as Managing Director, Technology Division at Goldman Sachs India, and has held various leadership positions with IBM, including that of Director, IBM Research - India and Chief Technologist, IBM India/South Asia. From 2001 to 2006, he served as a Senior Manager at the IBM T.J. Watson Research Center in Yorktown Heights, New York, where he led the team developing system software for the Blue Gene/L supercomputer. IBM was awarded a National Medal of Technology and Innovation for Blue Gene by US President Barack Obama in 2009. Manish holds a Ph.D. in Computer Science from the University of Illinois at Urbana Champaign. He has co-authored about 75 papers, with over 6,000 citations in Google Scholar in the areas of high-performance computing, compilers, and virtual machine optimizations, and has been granted 19 US patents. Manish has received an Outstanding Innovation Award, two Outstanding Technical Achievement Awards and the Lou Gerstner Team Award for Client Excellence at IBM. He is an ACM Fellow, a recipient of a Distinguished Alumnus Award from IIT Delhi, and has been elected a Fellow of the Indian National Academy of Engineering.



Dr. ANURAG AGRAWAL is a principal scientist at the CSIR Institute of Genomics & Integrative Biology (IGIB), where he leads the Center of Excellence for Translational Research in Asthma & Lung disease (TRIAL) and the CSIR program for Enabling Affordable Community Health through Information Technology (EACH-IT). He graduated in medicine from the All India Institute of Medical Sciences in 1994, followed by specialization in Internal Medicine, Pulmonary Disease, and Critical Care from Baylor College of Medicine, Houston, (2003) and a PhD in Physiology. After serving as a faculty member at Baylor, he joined IGIB in 2007. His research in respiratory health has covered a full spectrum from molecular pathobiology to informatics and epidemiology, leading to his being awarded the Shanti Swaroop Bhatnagar award for Medical Sciences in 2014 and the Wellcome Trust India Alliance senior fellowship in 2015. He has previously received the Lady Tata Young Researcher Award and the Swarnjayanti Fellowship of the Department of Science and Technology. Dr. Agrawal believes health data analytics to be the new frontier of medicine and is actively engaged in promoting the confluence of medicine and informatics through research and as an academic council member of the Public Health Foundation of India



GAUR SUNDER is Principal Technical Officer in Visual Computing & Big-data Analytics Group (VC&BA) at Centre for Development of Advanced Computing (C-DAC), Pune. At C-DAC he has led several national programs for development of tools and technologies targeting Healthcare IT, some of which are; the Mercury EHR/EMR & Telemedicine Suite, Distributed Health Store, and Software Development Kits for DICOM, HL7, and SNOMED CT. He represents India in IHTSDO Member Forum, and also leads the activities of National Release Center for SNOMED CT and EHR Standards created by MoH&FW, Government of India. He is currently coordinating the efforts for building core technologies and infrastructures for nationwide standardized healthcare records system in India. He continues to lead and participate in national and international technology development projects funded by various agencies. Mr. Gaur Sunder has served on national committees dealing with e-Healthcare initiatives, which included; Committee on EHR Standard for India, National Telemedicine Taskforce, e-Governance Domain Committee for Health, and National Rural Telemedicine Network. He is also member (additional) of Health Informatics Sectional Committee of Bureau of India Standards (BIS). Mr. Gaur Sunder is co-author of book Electronic Health Record: Standards, Coding Systems, Frameworks, and Infrastructures published by IEEE-Wiley Press, along with several papers and invited lectures published in journals and conferences.



Dr. YOGESH SIMMHAN is an Assistant Professor at the SERC Department at IISc. Previously, he was a Research Assistant Professor in the Electrical Engineering Department (Computer Engineering) at the University of Southern California, Los Angeles and Associate Director of the USC Centre for Energy Informatics. His research explores abstractions, algorithms and applications on distributed data and computing systems. These span Cloud Computing, Scalable and Distributed Computing, distributed data and metadata management, and software architectures for large scale applications in eScience and eEngineering. His research advances fundamental knowledge, and offers a practitioner's insight, on building scalable, resilient systems to empower dynamic, distributed and Big Data applications. Yogesh has a Ph.D. in Computer Science from Indiana University and was earlier a Postdoc at Microsoft Research, San Francisco. He is a Senior Member of IEEE and ACM.

Speaker Profiles



Prof. MUKESH KHARE is Professor of Environmental Engineering and Management, Founder Coordinator of Indo French Unit for Water and Waste Technologies (IFUWWT) in Department of Civil Engineering and also holds charge as Dean, Alumni Affairs and International Program. He has worked on urban air quality management (local and episodic)/indoor air quality modelling and monitoring in urban built environment/exposure assessment/wind tunnel simulation of air dispersion since he joined IIT Delhi in 1991. He is recently led an expert team in the IITD which developed Comprehensive Environmental Pollution Index (CEPI) for Industrial Clusters for CPCB. He is also leading a task force on Indoor Air Quality in Central Pollution Control Board directed towards formulating IAQ protocol. In addition, Prof. Khare has authored more than 150 research publications primarily for peer reviewed journals and conference proceedings. He is specially known of his ongoing association with the Planet Earth Institute, London as Patron and Founder Member of the Global Scientific Committee (GSC) working for scientific independence of Africa. At IITD, he is currently leading research group on Sustainable Cities.



Prof. ASHOK DESHPANDE, founding Chair of Berkeley Initiative in Soft Computing (BISC)-Special interest Group (SIG)-Environment Management Systems (EMS) holds a PhD degree in Engineering and Technology. Dr. Ashok was Deputy Director, National Environmental Engineering Research Institute (NEERI) - CSIR. He is Visiting Professor in University of New South Wales, Canberra Australia and Indian Institute of Technology Mumbai India. Also Adjunct Professor in College of Engineering Pune India and National Institute of Technology Silchar, Assam. He has 4 decades of R&D experience and has over 120 publications in the Journals of international repute/conferences. In the past, Dr. Ashok was WHO Adviser, Common Wealth Science Council Resource Scientist, World Bank Project Director for the studies on Probabilistic Risk Assessment for Chemical Process industry. He is also guest faculty at the University of California Berkeley and visiting scientist at Lawrence Berkeley National Laboratory, California.



Prof. ASHIS SENGUPTA is currently Professor (HAG) of Applied Statistics Unit, Indian Statistical Institute, Kolkata, and Adjunct Professor of Dept. of Biostatistics & Epidemiology, Georgia Regents University, USA. He was born in Kolkata, India, and received Ph.D. from Ohio State University, USA, in Statistics. He has about one hundred research in a variety of areas such as Multivariate Statistical analysis, Directional Data analysis, Bioinformatics, Bayesian inference, Environmental sciences, Reliability, etc. He is co-author of the book, Topics in Circular Statistics, World Scientific, USA. He has been Vice-President of International Forum for Interdisciplinary Mathematics, President of International Indian Statistical Association (India chapter) successively for several terms, and President of Indian Science Congress Association. He was elected in 2010 a Fellow of the National Academy of Sciences, India and in 2009, a Fellow of the American Statistical Association. He has received "Distinguished Statistician" award by Indian Society for Probability and Statistics and the "Lifetime Achievements Award" by Indian Association for Reliability and Statistics in 2015.



Dr. SHIVA NAGENDRA is presently working as Associate Professor in Department of Civil Engineering, Indian Institute of Technology Madras, Chennai India. Dr. Nagendra worked as Senior Research Engineer at Fluidyn Software and Consultancy, Bangalore, involved in development of air quality software. He has published more than 35 research publications in international and national refereed journals, one reference book, and 50 papers in conferences. He is coauthor of the book titled 'Artificial Neural Networks in Vehicular Pollution Modelling' published by Springer-Verlag, Germany. He has been part of numerous grants involving multiple national and international partners. Much of his research interests focus on air quality management which includes monitoring, source apportionment, modelling, design and development of emission control system, development of air quality management system, exposure monitoring system, environmental impact assessment and indoor air quality management. He is a professional member of several technical institutions and organizations of India.

Speaker Profiles



Prof. SANJAY CHAUDHARY is working as a Professor and Head (Research) at Institute of Engineering and Technology (IET), Ahmedabad University. Earlier, he worked as a Professor and Dean (Academic Programs) at Dhirubhai Ambani Institute of Information and Communication Technology. His research areas are Data Analytics, Cloud Computing, and ICT Applications in Agriculture and Rural Development. He has authored four books and six book chapters. He has published more than eighty research papers and ten literary articles in international conferences, workshops and journals. He has served on many program committees of International conferences and workshops and he is also a member of review committees of leading journals. Four PhD candidates have completed PhD under his supervision and two PhD students are doing PhD research work with him. Thirty two students have completed their MTech research work under his guidance. He holds a doctorate degree in computer science from Gujarat Vidyapeeth. He has worked on various large-scale software development projects for corporate sector, co-operative sector, and government organizations. He is actively involved in various consultancy and enterprise application development projects. He has received research grants from IBM, Microsoft and Dept of Science and Technology, Govt. of India. Gujarat Sahitya Academy has awarded second prize to his book on 'Girnar' under 'Essays and Travelogue' category for the year 2009. He loves literature, music, travel and wild life.



Dr. SHIBENDU S. RAY is the Director of the Mahalanobis National Crop Forecast Centre (MNCFC) under Ministry of Agriculture & Farmers' Welfare, New Delhi, since its inception in March 2012. In MNCFC, he is leading a team of scientists and researchers on various aspects of space technology based agricultural assessment. Prior to this, he was the Head, Agro-Ecosystems Division, Space Applications Centre, ISRO, Ahmedabad, India. Dr. Ray has obtained his Doctoral and Masters Degree in Agricultural Physics from Indian Agricultural Research Institute, New Delhi, India. He has specialized in various fields of Remote Sensing Applications in Agriculture, through his more than twenty years of research activity in SAC (ISRO). He has been responsible for successful implementation of many national level projects, such as 'Crop Acreage and Production Estimation', 'Cropping System Analysis', 'Irrigation Management', 'Precision Farming', 'Climate Change Impact on Agriculture', 'Hyperspectral Remote Sensing Applications', etc. He was the Chair of the Working Group VIII/6 on 'Agriculture, Ecosystems and Biodiversity' of International Society of Photogrammetry and Remote Sensing during 2008-2012 and currently is the vice-President of Indian Society of Remote Sensing and Indian Society of Agrophysics. He is the recipient of P.R. Pisharoty Memorial Award (2005), Hari Om Ashram Prerit Vikram Sarabhai Award (2005) and ISRO Team Excellence Award (2007), ISRO Merit award (2015), National Geomatics Award (2015). He has more than 150 publications to his credit.



Dr. ABHISHEK RATHORE is a Senior Scientist (Biometrics) at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), headquartered at Patancheru, Telangana, India with extensive experience in Efficient Experimental Designs, Analysis of Experimental Data, Multi-Environment Analysis (GxE), Mixed Model Analysis, Nonlinear Statistical Modeling, Genome Wide Association Analysis (GWAS), QTL Mapping, Genetic Linkage Map Construction, Genomic Selection (GS), Marker Assisted Recurrent Selection (MARS), Programming in C, C++, Database Development, Management and Sharing. Excellent knowledge of statistical software including SAS, GenStat and R. Current responsibilities at ICRISAT includes Heading the Breeding Informatics Unit, leading a team of statisticians and programmers and providing research and analytical support to scientists, scholars and various research units based at the Headquarters and ICRISAT regional centers in ESA (Eastern and Sub-Saharan Africa) and WCA (Western and Central Africa). Has published more than 90 research papers in several national and international peer reviewed journals.

Posters Selected for Presentation in WBDB 2015.in

1. Application of Soft Computing Techniques in Modelling Real Time Vehicular Exhaust Emissions at Urban Roads – J. Rohit, S.M. Shiva Nagendra and R. Sivanandan
 2. Development of Fuzzy Inference System for Assessment of Air Quality in an Urban Area – C.G. Hrishikesh and S.M. Shiva Nagendra
 3. Big Data Analytics for Urban Heat Island Mitigation in Smart Cities – Y. Nithiyanandam
 4. An Efficient EAV Storage for Large Scale Standardized Electronic Healthcare Records – Shivani Batra and Shelly Sachdeva
 5. Using Autonomous Drones and Deep Learning in a Dengue Surveillance System – Raghvendra Jain, Helmut Prendinger, Suresh Munuswamy, Kai Yan, Sra Sontisirikit, Matthew Dailey and Serge Stinckwich
 6. Sunflower Commodity Forecasting and Strategy to Make Policy in Achieving Profitability for Farmers – Satya Syam Puppala
 7. Data-Driven Strategies to Address the Challenge of Air Pollution in India – Khaiwal Ravindra, Maninder Kaur, Suman Mor, Siby John and Saumyadipta Pyne
 8. Experiences in Setting up of eHealth Centers in India: Outcomes and Challenges – Viren Sardana, Aditya Nagori and Anurag Agrawal
 9. Making Universal Health Coverage a Reality in India – Ashwini Reddy and Dr. Suresh Munuswamy
 10. Statistical Data Stitching to Construct a High-dimensional Object from its Low - dimensional Representations – Daniel Ahfock, Sharon X. Lee, Geoffrey J. McLachlan and Saumyadipta Pyne
- 
- A dark silhouette of a person running is visible on the right side of the page, set against a background of a sunset or sunrise over water. The person is in a dynamic pose, with one leg forward and arms pumping, suggesting movement and energy.

Inaugural Message from Professor C.R. Rao, Padma Vibhushan and Recipient of the India Science Award.

I am delighted to learn that Prof. S. Pyne, well-known expert in Bioinformatics and processing of Big Data, and his Co-Organizers from Indian Statistical Institute and University of California San Diego, are organizing an international Workshop on Big Data Benchmarking (Seventh WBDB 2015.in) in New Delhi.

Big data systems address the challenges of capturing, storing, managing, analyzing and visualizing big data. Within this context, developing benchmarks to evaluate and compare big data systems has become an active topic for both research and industry communities. To date, most of the state-of-the-art big data benchmarks are designed for specific types of systems. However, considering the complexity, diversity and rapid evolution of big data systems, big data benchmarks must include diversity of data and workloads.

With rapid development of information technology, big data systems have emerged to manage and process data with high volume, velocity and variety. These new systems have given rise to various new requirements about how to develop a new generation of big data benchmarks. There are numerous problems for doing research and proposing methods of developing big data benchmarks. I hope the workshop will identify specific problems of current interest for investigation.

I wish the workshop a great success.



C.R. Rao

December 7, 2015.

ampool

Democratizing Fast Analytics

Accelerate Big Data analytical pipelines with memory-centric technologies

Extract data value with ease, deliver performance and maintain flexibility

Turbocharge data-driven applications with real-time analytics

SIMPLICITY

PERFORMANCE

FLEXIBILITY

www.ampool.io [/company/ampool-inc-](https://www.linkedin.com/company/ampool-inc-) [/ampoolio](https://www.facebook.com/ampoolio) [@AmpoolIO](https://twitter.com/AmpoolIO)



PERSISTENT

Don't Refresh. Reimagine!



Focus on the "How" of Digital Transformation -

- Think beyond IT
- Transform, starting from the middle
- Build a software driven business

www.persistent.com