

## **Tentative Schedule**

### **Workshop on 'Big and Open Data**

#### ***Evolving Data Science Standards & Citation Attribution Practices***

**November 5-6, 2015**

**Venue: Indian National Science Academy**

**Bahadur Shah Zafar Marg, New Delhi - 110002**

<b>November 5, 2015 (Thursday)</b>		
9.00 – 9.50 hrs	<b>Registration</b>	
10.00 – 11.00 hrs	<b>Inaugural Session</b>	
	Welcome & Introduction	<b>Dr. Chandrima Saha</b> Vice-President Elect, International , INSA, & Director National Institute of Immunology
	Address by	<b>Prof. Krishan Lal</b> Workshop Chair & President, AASSA, Former Director, NPL, & Immediate Past President, INSA
	About the Current Initiative- Data Citation Principles & Practices	<b>Dr. Jan Brase</b> Co-Chair, CODATA-ICSTI Data Citation TG
	Address by	<b>Dr. Tishyarakshit Chatterjee</b> Former Secretary, MoEF & Director, IIPA
	Keynote Address <i>Data Collection and Management: Practices in Government</i>	<b>Dr. T.C.A. Anant</b> Secretary, MoSPI and Chief Statistician of India
	Presidential Address	<b>Shri T.N. Chaturvedi</b> Chairman , IIPA and Former Governor of Karnataka & Kerala
	Inaugural Address	<b>Dr. R. Chidambaram</b> Principal Scientific Advisor to GoI & Chairman, SAC-C
	Vote of Thanks	<b>Dr. Usha Munshi</b> Workshop Coordinator
11.00 – 11.30 hrs	<b>Tea / Coffee Break</b>	
11.30- 13.30 hrs	<b>Theme I: Data Science, Technology, Research and Applications</b>	
	<b>Chairman : Professor Ramakrishna Ramaswamy</b> Professor, School of Physical Sciences, Jawaharlal Nehru University, Former Vice Chancellor, University of Hyderabad	
	<i>DST Report: Data Science, Technology, Research and Applications</i>	<b>Dr. K.R. Murali Mohan</b> Advisor & Head, Big Data Initiative, DST, GoI
	<i>Data Science Research -Technological and other Issues : Current Scenario &amp; Future Directions</i>	<b>Prof. Yadati Narahari</b> Chairman, Division of Electrical Sciences, IISc, Bengaluru
	<i>e-infrastructure for research data management: Strategy and Implementation Plan</i>	<b>Dr. B.K. Murthy</b> Executive Director, C-DAC, Noida

	<i>Application of Data Techniques across domains: Some Insights</i>	<b>Prof. Ajit Kembhavi</b> Director, IUCAA, Pune
13.30 – 14.30 hrs	<b>Lunch Break</b>	
14.30 – 16.30 hrs	<b>Panel Discussion</b> <b>Data Science, Technology, Research and Applications : Developments in Indian Context</b>	
	<b>Chairman: Dr. Prahlada Rao B.B.</b> Executive Director, C-DAC, Bengaluru	
	<b>Prof. Alok Bhattacharya</b> , Professor, School of Life Sciences, JNU <b>Prof. Santanu Chaudhury</b> , Dhananjay Chair Professor, Department. of Electrical Engineering, Indian Institute of Technology , Delhi <b>Dr. Sudha Midha</b> , Additional Director General, National Sample Survey Organisation <b>Prof. Suresh Mishra</b> , Chair Professor & Chairman Centre for Consumer Studies, IIPA <b>Dr. K J Ramesh</b> , Scientist G & Advisor, Ministry of Earth Sciences, GoI <b>Prof. Vinod K. Sharma</b> , Executive Vice-Chairman, Sikkim State Disaster Management Authority <b>Dr. Durga Toshniwal</b> , Department of Computer Science .& Engineering, .Indian Institute of Technology, Roorkee	
16.30 – 16.45hrs	<b>Tea</b>	
18.00 -	<p style="text-align: center;"><b>Networking over Dinner Discussion</b></p> <p style="text-align: center;"><b>Invited Dignitaries</b></p> <p style="text-align: center;"> <b>Dr. Jitendra Singh*</b>  MoS (Independent Charge),  Science and Technology and Earth Sciences and MoS for Prime Minister Office  Personnel, Public Grievances &amp; Pensions Department of Atomic Energy and  Department of Space </p> <p style="text-align: center;"> <b>Dr. Girish Sahni*</b>  DG, CSIR &amp; Secretary, DSIR, GoI. </p> <p style="text-align: center;"> <b>Dr. Ajay Kumar*</b>  DG, NIC &amp; Additional Secretary, Deity, GoI </p> <p style="text-align: center;"> <b>Shri Pawan Agarwal</b>,  Jt. Secretary, Ministry of Skill Development and Entrepreneurship    &amp; many more dignitaries are expected to join </p>	
	<b>Dinner</b>	

**November 6, 2015 (Friday)**

10.00 - 13.00 hrs	<b>Theme II: Data Attribution &amp; Citation Practices</b>	
	<b>Chairman: Dr. Jan Brase</b> Co-Chair, CODATA-ICSTI Data Citation Task Group <b>Co-Chair : Dr. Nafees Meah</b> Director, Research Councils, UK.(RCUK) India, British High Commission	
	<i>Big and Open Data in Government</i>	<b>Dr. Neeta Verma</b> DDG, NIC
	<i>Current Status of Managing Research Data by R&amp;D community</i>	<b>Dr. S Ganesan</b> Raja Ramanna Fellow of the DAE, Reactor Physics Design Division & Professor, Homi Bhabha National Institute, Mumbai, BARC & Member, NDPCL.
	<i>Data Management &amp; rolling out strategy and implementation Plan: Country Experience from UK Research Council</i>	<b>Dr. Paul Meller</b> Strategic Lead for Data and Resources, Economic and Social Science Research Council (ESRC)
13.00 - 14.00 hrs	<b>Lunch Break</b>	
	<b>Panel Discussion</b> <b>Data - Management, Sharing and Services : Current &amp; Future Role of Data Stakeholders</b>	
	<b>Chairman: Dr. R. Gopichandran</b> Director, Vigyan Prasara <b>Co-Chair : Dr. Paul Meller</b> Strategic Lead for Data and Resources, Economic and Social Science Research Council (ESRC)	
14.00-16.00 hrs	<b>Dr. Jagdish Arora Director</b> , INFLIBNET <b>Dr. Gopal Bhushan, Director</b> , DESIDOC, DRDO <b>Dr. A.K. Garg, Director</b> , International Cooperation, DeitY, GoI <b>Prof. Hiranmay Ghosh</b> , Principal Scientist, TCS <b>Col. Haridas M.</b> , CLAWS <b>Dr. H.K. Kaul, Director</b> , DELNET <b>Dr. Debasisa Mohanty</b> , Scientist, NII <b>Prof. A.R.D. Prasad</b> , Head, DRTC, ISI, Bangalore <b>Prof. Roshan Raina</b> , IIM, Lucknow & VC <b>Dr. Harpreet Singh</b> , ICMR <b>Dr. Biplav Srivastava</b> , Senior Researcher, IBM Master Inventor IBM Research, India <b>Ms. Neeta Verma</b> , DDG, NIC, GoI <b>Prof. O.P. Wali</b> , Professor, IIFT	
16.00-17.30 hrs	<b>Concluding Session</b>	
	Welcome Address	<b>Prof. Alok Bhattacharya</b> Professor, JNU & Chair- CODATA, National Committee
	Opening Remarks	<b>Dr. T. Chatterjee</b> Former Secretary, MoEF and Director, IIPA
	Introduction of Guest Speakers	<b>Prof. Vinod K. Sharma</b> Executive Vice-Chairman, Sikkim-Disaster Management Authority & Senior Professor, IIPA
	Address by the Chief Guest	<b>Dr. P.K. Mishra</b> Addl. Principal Secretary to PM *

	Recommendations for framing Policy Guidelines	<b>Prof. Krishan Lal</b> Workshop Chair & President, AASSA, Former Director, NPL, & Immediate Past President, INSA
	Wrap up	<b>Dr. Usha Mujoo Munshi</b> Workshop Coordinator
	<b>High Tea</b>	

**\*To be confirmed**

### Heritage Visits

November 7, 2015 (Payment Basis)

- a) Delhi Monuments - Heritage Walk
- b) Agra -Taj Mahal
- c) Jaipur -Heritage Walk



### For Participation Please Contact

Dr. Usha Mujoo Munshi  
Indian Institute of Public Administration  
I.P. Estate, Ring Road,  
New Delhi- 110 002 (INDIA)

e-mail: bigdatacitprotocol@yahoo.com; umunshi@gmail.com  
Tel: 011-91-23702431; 23468320; (M) 9717967686

## Workshop

on

## 'Big and Open Data'

Evolving Data Science Standards & Citation Attribution Practices

**November 5-6, 2015**



*Organised by*



*Indian Institute of Public Administration  
International Council for Science-Committee on Data for Science & Technology  
Indian National Science Academy  
Department of Science & Technology  
Department of Electronics and Information Technology, MCIT  
Department of Biotechnology  
Ministry of Earth Sciences*

**Knowledge Partners**





## Background

Recent decades have seen an unprecedented explosion in the human capacity to acquire, store and manipulate data and information and to instantaneously communicate them globally, irrespective of location. Research is rapidly moving from an era of data scarcity, in which, with some exceptions, data have been small in volume and sparse in distribution, with statistical techniques optimised to extract information from such limited data, to one of abundance, in which an unprecedented deluge of data offers major opportunities and profound challenges.

Many of these opportunities and challenges arise from 'Big Data'. The challenges are to create infrastructure, methodologies, policies and practices that enable researchers to identify patterns and processes that have hitherto been beyond our capacity to resolve and analyse to predict the behaviour of complex systems.

Effective exploitation of Big Data depends fundamentally upon the international culture of 'Open Data' which involves sharing of data and their availability for re-use and re-purposing. It not only offers novel possibilities for commercial innovation, but also for greater involvement of a wider ranges of stakeholders and citizens in co-production of knowledge, and for deeper democratic engagement with the ways that knowledge is created and used.

New possibilities and new challenges of the data explosion of recent decades is spreading, in what some experts quote as a 'second revolution of discovery'. The contemporary trends of rapid development and change, a strategy must be a living document that is able to adapt to novel discoveries and approaches and the evolving needs of international research and science community.

## Workshop Title

***'Big and Open Data'***  
***Evolving Data Standards & Citation Attribution Practices***

## International Workshop – Need and Objectives

There is a crying need for: *a framework of international agreements, practices or standards* codified in the context of subject-specific conventions; *national policies and practices* for funding and incentivising research, and for physical and software infrastructure; *institutional policies and practices*

## National Organizing Committee

**Chairman:** Prof. Krishan Lal - President, AASSA, Former Director, NPL and Immediate Past President, INSA

### Members:

Prof. Alok Bhattacharya -JNU, New Delhi  
Prof. Sanghamitra Bandyopadhyay- Director, ISI, Kolkata  
Shri Gopal Bhushan - Director, DESIDOC, DRDO  
Dr. Jagdish Arora-Director- INFLIBNET, Ahmedabad  
Prof. Santanu Chaudhury – IIT, Delhi  
Dr. S. Ganesan- Bhabha Atomic Research Centre, Mumbai  
Prof. M.P. Gupta – IIT Delhi  
Prof. Ajit K. Kembhavi – Director, IUCAA, Pune  
Dr. Anil Kumar -NCL, Pune  
Prof. U. Dinesh Kumar - IIM, Bangalore  
Dr. K.R. Murali Mohan – Advisor & Head, Big Data Initiatives Division, DST, New Delhi  
Dr. Debasisa Mohanty - National Institute of Immunology, New Delhi  
Dr. Alok Moitra - Executive Director, INSA, New Delhi  
Dr. U. Kamachi Mudali - Indira Gandhi Centre for Atomic Research, Kalpakkam, Tamil Nadu  
Prof. Y. Narahari - Chairman, CSA, IISc, Bangalore  
Prof. R. Nagaraj- IGIDR, Mumbai  
Dr. G. Padmanabhan - Emergency Analyst, UNDP  
Smt. Alka Panda (IAS) - Director General, Bureau of Indian Standards  
Prof. K K Pandey - IIPA, New Delhi  
Prof. Sankar K. Pal – Distinguished Scientist, ISI, Kolkata  
Prof. Rajendra Prasad - JNU, New Delhi & Vice President (INSA)  
Dr. K.J. Ramesh- Ministry of Earth Sciences, New Delhi  
Ms. Mousmi Roy - Sr. Director, Urban Infrastructure & Smart Cities, FICCI  
Dr. Sameer Sharma - Additional Secretary, MoUD, Smart Cities  
Prof. V.K. Sharma – IIPA, New Delhi & Member, UNISDR, Asia Science, Technology & Academia Advisory Group  
Dr. Harpreet Singh – Indian Council of Medical Research, New Delhi  
Dr. V.M. Tiwari - CSIR-NGRI, Hyderabad  
Dr. Neeta Verma - DDG, NIC, New Delhi  
Prof. Sushma Yadav – Pro-Vice Chancellor, IGNOU, New Delhi

**Workshop Coordinator:** Dr. Usha Mujoo Munshi – IIPA, New Delhi

policy document on data citation protocols and standards as a collective endeavour of 11 workshops as the tangible output.

### **Registration**

No Registration Fee. Registration on First Come First Serve Basis

### **Workshop Chair**

Prof. Krishan Lal – President, AASSA, Former Director, NPL and Immediate Past President, INSA

**Patron :** Prof. Raghavendra Gadagkar – President, INSA

Dr. Tishyarakshit Chatterjee – Former Secretary, MoEF & Director, IIPA

### **International Advisory Committee**

Prof. Ashutosh Sharma - Secretary, DST, New Delhi

Prof. K. VijayRaghavan -Secretary, DBT, New Delhi

Dr. Shailesh Nayak - Secretary, Ministry of Earth Sciences, New Delhi

Prof. Geoffrey Boulton- FRS, President, CODATA, UK

Dr. Sarah Callaghan - Editor-in-Chief, Data Science Journal, Ireland

Dr. Jan Brase - Head of Research and Development, Georg-August-University, Germany

Prof. Christine Borgman - Distinguished Professor & Presidential Chair in Information Studies, University of California, USA

Dr. Theodora Bloom - Executive Editor of BMJ, UK

Dr. Todd Vision -Associate Director of Informatics, National Evolutionary Synthesis Center, Department of Biology, University of North Carolina at Chapel Hill, USA

Prof. Sukhadeo Thorat - Chairman, ICSSR, New Delhi

Prof. Rajat Moona - Director General, C-DAC, Pune

Mr. BVR. Mohan Reddy - Chairman, NASSCOM, New Delhi

Mr. Chandrajit Banerjee – Director General, CII, New Delhi

Dr. T. C. A. Anant - Secretary and Chief Statistician of India, MoSPI, New Delhi

Dr. Mustapha Mokrane – Executive Director, World Data System, Japan

Dr. Simon Hudson - Executive Director, CODATA, Paris

through which open data curation and support for researchers are managed.

The proposed important international event on Data Citation and attribution, is part of world initiative led by ICSU, CODATA - *Data Citation Task Group*. The issue at hand focuses on an important gap in the research data infrastructure – the absence of broad availability and use of data citation protocols. In the global research enterprises, there are many improvements that can and should be made to make research data more efficient and effective. Among those is to make research data activities better evaluated, recognized, and rewarded; research data easier to locate and share; and, generally, to improve the stewardship of data resources and the return on research support. The lack of established practices for referring to portions or subsets of data has resulted in the complicating the problem of citing online data.

As funding sources for research in general have begun to require data management plans as part of their selection and approval processes, it is important that the necessary standards, incentives, and convictions to support data citation, preservation, and accessibility be put into place. There are, in fact a number of initiatives in different organizations, countries, and disciplines already underway. One important group is DataCite, while others remain ad hoc and uncoordinated. An important milestone in this area is the joint declaration of the *Data Citation Principles* available at <http://www.force11.org/datacitation>.

The main objective in the next one year is to promote the implementation of the data citation principles in the research policy and funding communities throughout the world.

### **Issues Requiring Attention**

The issues of technical, scientific, institutional, financial, sustainability, persistent identifiers, legal issues, intellectual property rights, socio-cultural and community norms and other issues that need attention related to data citation, coordinate activities in this area, and promote common practices and standards in the research community, globally are being examined by the Data Citation TG of CODATA with representatives from several other organizations.

It is essential for our country to take stock of what is being done in the data citation area and to begin formulating and adopting a formal data citation protocol for use by different discipline areas or the research

community. Equally important are prioritizing addressing the scientific and technical research issues in this arena.

Thus the workshop would broadly take in its ambit the following key issues in the data arena:

### **Theme I: Prioritising, Addressing & Establishing Standards & Good Practices**

#### **Scientific & Technical Issues:**

- Interoperability & Facilitation of Re-use
- Development of Persistent Identifier Standards & Models to Address Diversity of (i) Databases, (ii) Needs
- Data Science Research: Intellectual Property Rights
- Development & Implementation of Discipline Specific Data Citation Standards & Good Practices
- Data Citation Index

### **Theme II: Data Attribution & Citation Practices**

#### **Key Issues:**

- Data Citation Practices Worldwide
- Promoting Adoption of Data Citation Practices
- Recognize Data Compilation, Analysis and Publishing Practices
- Incentivize Disclosure of Data in Interoperable & Quality Control Ways (Open Format)

### **Workshop Series**

It has been decided to hold discussions on this important issue of Data Citation standards and Protocols and 11 countries (including India) have been identified to hold workshops on the centric issue of big and open data citation protocols.

Around 11 national and regional workshops dedicated to this focused objectives for the policy makers and funding community in the respective countries, that includes *India, China, Taiwan, Japan, Australia, South Africa, Israel, Brazil, EU, France, USA, and the Group on Earth Observations* and to come out with a synthesis paper that integrates the findings. The aim is to tailor-fit the workshops to the needs and interests of the respective country or region. In countries/regions with an existing culture of data sharing and citation the targeted audience will mainly consist of funders, policy makers, journal editors and learned societies, whereas in countries/regions that have just begun to focus on data related

issues, the targeted audience will mainly be the science and social science communities. In those countries/regions it might also be a useful approach to concentrate on individual disciplines whereas the more high-level approach should be inter-disciplinary in nature.

### **Workshop Schedule**

The dates for the workshops in many countries have been finalized. The workshops in **China**: August 25, 2015; **South Africa**: October 15, 2015; **Japan**: October 29, 2015; **Taiwan**: Spring 2016, **Israel**: Spring 2016, **USA**: (Washington, DC) Fall 2016; **Europe** : (Brussels) Summer 2016; **Australia**: 2016 (date not confirmed)

**Workshop in India: November 5–6, 2015**

### **Venue**

Indian National Science Academy  
Bahadur Shah Zafar Marg, New Delhi 110 002

### **Workshop Participants**

Around 100 participants are expected and the participation is primarily by invitation. The participants would broadly include:

- Researchers
- University Administrators/Policy Makers
- Data Centre Managers
- Research Librarians
- Research Publishers and Professional Societies

### **Workshop Outcomes**

After every workshop a brief summary will be made available to the other organisers and the group. The expected outputs of the workshops are:

- To raise general awareness for data science and data citation in the individual regions
- To report back to the group and CODATA the state-of-the-art of data science and data citation from the individual regions/ countries
- To gain an overview on obstacles and challenges concerning data science and data citation in the specific region/ country

The organisers in each country would write a detailed summary on the topics mentioned above, to share it with the larger group within 3 months after the workshop. The prestigious event would eventually have a global