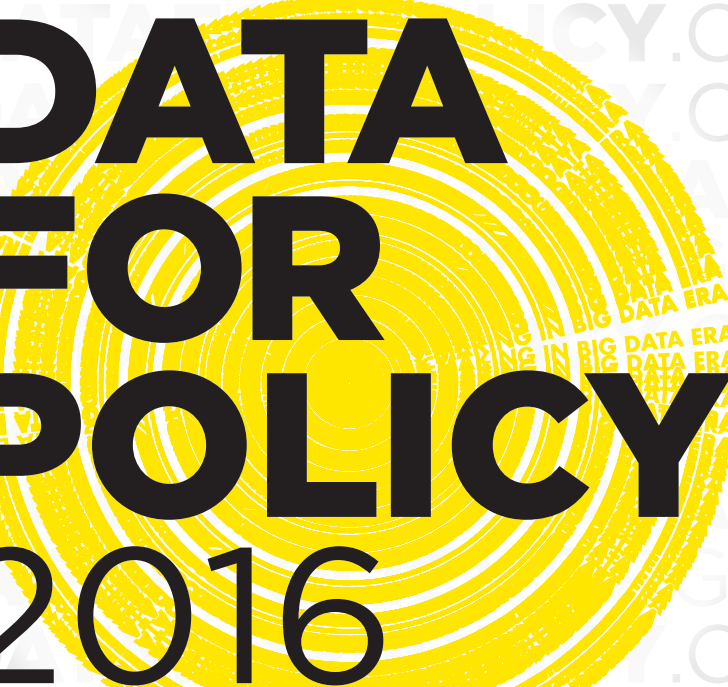


# DATAFORPOLICY.ORG



# DATA FOR POLICY 2016

**Frontiers of Data Science for Government:**  
Ideas, Practices, and Projections

15-16 September 2016  
University of Cambridge

[dataforpolicy.org](http://dataforpolicy.org) | [@dataforpolicy](https://twitter.com/dataforpolicy) | [#DataFrontiers](https://twitter.com/DataFrontiers)



# About Data for Policy Initiative

---

Launched at the inaugural conference that was hosted by the University of Cambridge in 2015, Data for Policy has been established as an independent initiative to promote interdisciplinary and cross-sectorial discussion on the use of Data Science research to enhance government operations and policy-making. The inaugural conference *"Policy-making in the Big Data Era: Opportunities and Challenges"* hosted top-level presentations from UK and EU government bodies, prestigious national and international universities, non-profit institutions, and major industrial stakeholders.

We are grateful to the following prestigious institutions for supporting this initiative:

- **University of Cambridge** – Computer Laboratory, Centre for Science and Policy, Cambridge Big Data Strategic Research Initiative, Digital Humanities Network, Cambridge Public Policy Initiative
- **European Commission**
- **Alan Turing Institute**
- **Imperial College London** – Data Science Institute
- **London School of Economics & Political Sciences** – Department of Methodology
- **University College London** – Department of Computer Science, UCL Public Policy, The Bartlett – UCL Faculty of the Built Environment
- **University of Oxford** – Oxford Internet Institute
- **Office for National Statistics**
- **Royal Statistical Society**
- **New York University** – The GovLab, Open Governance Research Exchange
- **Leiden University** – Centre for Innovation
- **Technopolis Group**

# Data for Policy 2016 Conference

---

Data Science is emerging as a key interdisciplinary research field to address major contemporary challenges across sectors. Particular focus on the government sector offers huge potentials to advance citizen services and collective decision-making processes. To reflect the diversity of skills and knowledge required to tackle challenges in this domain, the conference offers an open discussion forum for all stakeholders. Data for Policy 2016 invited individual and/or group submissions from all relevant disciplines and application domains. Topics covered included but were not limited to the following:

**Government & Policy:** Digital era governance and citizen services, public demand vs. government response, using data in the policy process, open source and open data movements, policy laboratories, citizen expertise for government, public opinion and participation in democratic processes, distributed data bases and data streams, information and evidence in policy context, case studies and best practices.

**Policy for Data & Management:** Data collection, storage, and access; psychology/behaviour of decision; privacy, trust, public rights, free speech, ethics and law; data security/ownership/linkage; provenance, curation, expiration; private/public sector/non-profit collaboration and partnership, etc.

**Data Analysis:** Computational procedures for data collection, storage, and access; large-scale data processing, dealing with biased/imperfect/uncertain data, human interaction with data, statistical/computational models, technical challenges, communicating results, visualisation, etc.

**Methodologies:** Qualitative/quantitative/mixed methods, gaps in theory and practice, secondary data analysis, web scraping, randomised controlled trials, sentiment analysis, Bayesian approaches and graphical models, biologically inspired models, real-time and historical data processing, simulation and modeling, small area estimation, correlation & causality based models, and other relevant methods.

**Data Sources:** Government administrative data, official statistics, commercial and non-profit data, user-generated web content (blogs, wikis, discussion forums, posts, chats, tweets, podcasting, pins, digital images, video, audio files, advertisements, etc.), search engine data, data gathered by connected people and devices (e.g. wearable technology, mobile devices, Internet of Things), tracking data (including GPS/geolocation data, traffic and other transport sensor data, CCTV images etc.), satellite and aerial imagery, and other relevant data sources.

**Policy/Application Domains:** Security, health, cities, public administration, economy, science and innovation, finance, energy, environment, social policy areas (education, migration, etc.) and other relevant domains.

# Organisers

---

## Advisory Committee:

Prof. Jean **Bacon**, Computer Laboratory, University of Cambridge  
Prof. Kenneth **Benoit**, Department of Methodology, London School of Economics and Political Science  
Dr. David **Cleavelly**, Centre for Science and Policy, University of Cambridge  
Prof. Jon **Crowcroft**, Computer Laboratory, University of Cambridge  
Prof. Anthony **Finkelstein**, Chief Scientific Advisor, Government Office for Science, UK  
Prof. Enrico **Giovannini**, Co-Chair, UN Data Revolution Group & University of Rome “Tor Vergata”, Italy  
Prof. David **Hand**, Winton Capital Management and Imperial College  
Prof. Jonathan **Haskel**, Data Science Institute, Imperial College London  
Prof. David **Howarth**, Department of Politics and International Studies, University of Cambridge  
Prof. Helen **Margetts**, Oxford Internet Institute, University of Oxford  
Prof. Beth **Noveck**, The GovLab, New York University & MIT Media Lab, US  
Mr. Bill **Oates**, Chief Data Scientist, Office for National Statistics, UK  
Prof. Rob **Procter**, Department of Computer Science, University of Warwick  
Prof. Graeme **Reid**, University College London  
Prof. John **Shawe-Taylor**, Department of Computer Science, University College London  
Prof. Peter **Smith**, ESRC Administrative Data Research Centre for England and University of Southampton  
Prof. John **Taysom**, University College London, University of Cambridge, Harvard University  
Prof. Philip **Treleaven**, Department of Computer Science, University College London  
Prof. Patrick **Wolfe**, Deputy Director, The Alan Turing Institute, UK  
Prof. Sir David **Wallace**, University of Cambridge  
Mr. Glen **Watson**, Director-General, Office for National Statistics, UK  
Prof. Robin **Williams**, School of Social and Political Science, University of Edinburgh  
Prof. Derek **Wyatt**, Chair, Royal Trinity Hospice; NHS Innovations South East, UK

## Executive Committee:

Dr. Zeynep **Engin**, Chief Organiser, Data for Policy 2016; University College London; University of Cambridge  
Dr. Anil **Bharath**, Data Science Institute & Dept. of Bioengineering, Imperial College London  
Dr. Bilal **Gokpinar**, School of Management, University College London  
Dr. Prabhat **Agarwal**, European Commission

## Programme Committee:

Dr. Anne **Alexander**, Digital Humanities Network, University of Cambridge, UK  
Mr. Daniel **Castro**, Director, Centre for Data Innovation, US  
Prof. Gabrielle **Demange**, Paris School of Economics, France  
Dr. Suleyman **Demirsoy**, Director, London Innovation Society, UK  
Dr. Yves-Alexandre **de Montjoye**, Harvard University, US

Dr. Clare **Dyer-Smith**, Cambridge Big Data SRI, University of Cambridge  
 Dr. Rayid **Ghani**, Centre for Data Science and Public Policy, University of Chicago, US  
 Dr. Ella **McPherson**, Department of Sociology, University of Cambridge  
 Prof. Eric **Meyer**, Oxford Internet Institute, University of Oxford  
 Dr. Slava **Michaylov**, Department of Political Science, University College London  
 Assoc. Prof. Suzy **Moat**, Warwick Business School, University of Warwick  
 Dr. Richard **Mortier**, Computer Laboratory, University of Cambridge  
 Mr. Will **Moy**, Director, Full Fact, UK  
 Dr. Mirco **Musolesi**, Department of Geography, University College London  
 Assoc. Prof. Tobias **Preis**, Warwick Business School, University of Warwick  
 Dr. Martijn **Poel**, Technopolis Group  
 Dr. Alison **Powell**, Media and Communications, London School of Economics and Political Science  
 Dr. Charlotte **Sausman**, Cambridge Public Policy SRI, University of Cambridge  
 Prof. Ralph **Schroder**, Oxford Internet Institute, University of Oxford  
 Drs. Gideon **Shimshon**, Director, Centre for Innovation, Leiden University, Netherlands  
 Dr. Jatinder **Singh**, Computer Laboratory, University of Cambridge  
 Dr. Olivia **Stevenson**, UCL Public Policy  
 Dr. Mesut **Tastan**, Department of Finance, London School of Economics and Political Science  
 Ms. Barbara **Ubaldi**, The Organisation for Economic Co-operation and Development (OECD)  
 Ms. Diana **Vlad-Calci**, European Commission  
 Dr. Indrė **Žliobaitė**, Uni. of Helsinki, Aalto Uni., Helsinki Inst. for Information Technology (HIIT), Finland

# Conference Programme

---

## Thursday, 15 September 2016

---

**09:30 – 10:00**                      **Arrivals & Registration**

**10:00 – 10:20**                      **Welcome & Introductions (Room – LT1)**

Andy **Hopper**, Professor of Computer Technology and Head of Department,  
Computer Laboratory, University of Cambridge

Zeynep **Engin**, Chief Organiser, Data for Policy 2016; Senior Research Associate,  
Urban Dynamics Lab, UCL; Policy Fellow, Centre for Science and Policy,  
University of Cambridge

**10:20 – 11:20**                      **Opening Plenary Session (Room – LT1)**

---

*“Big Data Analytics and EU Policy-making – huge potential, significant challenges”*  
Maive **Rute**, Deputy Director-General of Joint Research Centre, European  
Commission

*“Platform Government and a Data Science for Policy-making?”*  
Helen **Margetts**, Director of the Oxford Internet Institute, University of Oxford

*“Algorithmic Regulation: Using Blockchain smart contract technology to  
revolutionise Financial Regulation”*  
Philip **Treleaven**, Professor of Computing and Director of the Financial  
Computing Centre, University College London

**Chair:** Jon **Crowcroft**, Marconi Professor of Communications Systems, Computer  
Laboratory, University of Cambridge; The Alan Turing Institute

**11:20 – 11:40**                      **Break**

**11:40 – 12:40**                      **Parallel Session 1**

---

**Session 1a: Policy Process I (Room – LT1)**

**Chair:** Sir David **Wallace**, University of Cambridge

[67] *“Haze Gazer: A Crisis Analysis and Visualization Tool to Better Inform  
Peatland Fire and Haze Management”* – Derval **Usher**, George **Hodge**, Imaduddin  
**Amin** and Jong Gun **Lee**; UN Pulse Lab Jakarta, Indonesia

[95] *“Science for policy in the Big Data era”* – Panayotis **Christidis**, European  
Commission, Joint Research Centre, Spain

[103] *"Policy Compass - Prosperity Indicator-based Accountable Policy Analysis and Evaluation via Open Data Exploitation"* – Sotirios **Koussouris**, National Technical University of Athens, Greece; Panagiotis **Kokkinakos**, Greece; Ourania **Markaki**, National Technical University of Athens, Greece; Fabian **Kirstein**, Fraunhofer FOKUS, Germany; Yury **Glikman**, Fraunhofer FOKUS, Germany; Habin **Lee**, Brunel University London, UK.

#### **Session 1b: Methods I (Room – LT2)**

**Chair:** Kenneth **Benoit**, Head of the Department of Methodology, London School of Economics and Political Sciences

[18] *"The model is simple, until proven otherwise – how to cope in an ever changing world"* – Anita **Faul**, University of Cambridge, UK

[34] *"Early detection of properties at risk of blight using spatiotemporal data"* – Eduardo **Blancas Reyes**, Jennifer **Helsby**, Katharina **Rasch**, Talia **Kaufmann**, Paul **van der Boor**, Lauren **Haynes** and **Rayid Ghani**, University of Chicago, USA

[56] *"A systematic approach to Systemic Risk"* – Antoaneta **Sergueiva**, University College London and Bank of England; David **Bholat**, Bank of England, UK

[101] *"The future of human-machine decision-making"* – Claire **Craig**, The Royal Society; Susannah **Odell**, The Royal Society; Richard **Sandford**, GO-Science, UK

**Session 1c: Data-Driven Policy: Experimentation versus Regulation? (Room – FW26)**  
Organised by **The GovLab** at New York University and the **Centre for Innovation** at Leiden University

**Moderator:** Gideon **Shimshon**, Director, Centre for Innovation, Leiden University

#### **Debaters:**

Sarah **Giest**, Assistant Professor, The Institute of Public Administration, Leiden University

Andrew **Young**, Associate Director of Research, The GovLab, New York University

Yves-Alexandre **de Montjoye**, Research Scientist at the MIT Media Lab

TBC

**12:40 – 13:40      Lunch & Poster Session**



## Poster Presentations

---

- [8] *"Determinants of e-governance adoption practice in Public Administration of Sub Saharan Africa: Evidence from Cameroon"* – Soazic Elise **Wang Sonne**, United Nations University-MERIT, Netherlands
- [10] *"Exact and Approximate Methods for Public Payments: Who gets what anyway?"* – Charles **Rahal**, University of Birmingham, UK
- [11] *"Governance, the Commons and IoT: An Agriculture use case of Data for Policy"* – Harris **Moysiadis** and Nikos **Zotos**, Future Intelligence, Greece
- [12] *"Resurgence of Marx in Modern Day, Electronic Journalism"* – Aniruddha **Saha**, School of Communication, Manipal University, India
- [17] *"Informatics in Out of Hours Secondary Care Service Delivery: Methods and Applications to Inform Health Care Policy and Management"* – Iker **Perez**, Michael **Brown**, James **Pinchin**, Sarah **Martindale**, Sarah **Sharples**, Dominick **Shaw**, University of Nottingham; John **Blakey**, Liverpool School of Tropical Medicine, UK
- [19] *"Working hard or hardly working: Evaluating New Labour's active labour market policy, 1997 – 2010"* – Peter **Hill**, University of Warwick, UK
- [25] *"Developing a sociotechnical imaginary: The role of publics in the future of new forms of data"* – Emily **Rempel**, Julie **Barnett** and Hannah **Durrant**, University of Bath, UK
- [32] *"Linking foresight with simulation as part of an algorithmic impact assessment toolkit for decision-making, machine learning algorithms in the public sector"* – Michael **Veale**, University College London, UK
- [33] *"You can't use Data for Policy w/o Analytics Literate Leaders"* – Richard **Lee**, IMECS, LLC, USA
- [36] *"Evaluating the Impact of Arts-led Policy on the Gentrification in London Using Spatio-temporal Big Data"* – Xiao **Zhou**, Desislava **Hristova**, and Cecilia **Mascolo**, University of Cambridge; Anastasios **Noulas**, University of Lancaster, UK
- [58] *"When the chips come out: is our public interest research infrastructure fit for the future?"* – Jen **Persson**, defenddigitalme, UK
- [60] *"Generating Policy Insights from Micro-Narratives"*, Chiara **Amato**, Andreas **Karmhus**, Elena Cora **Magrini**, Charlie **Mealings**, Benjamin **Vigreux**, University College London, UK
- [69] *"Forecasting housing markets with Search keywords: A methodological 'horse race'"* – Donald **Stevens** and Franz **Fuerst**, University of Cambridge, UK
- [72] *"Discovery of complex anomalous patterns of sexual violence in El Salvador"* – Maria **De-Arteaga** and Artur **Dubrawski**, Carnegie Mellon University, USA

[77] *"Data for Science Policy: A SWOT Analysis"* – Arnaud **Vaganay**, London School of Economics and Political Sciences, UK

[78] *"The data science "pet" solution: Assessing solution acceptability and changing problem definitions"* – Kimberly **Gardner** and Eric **Lindquist**, Boise State University, USA

[80] *"Information Diffusion and Economic Development"* – Christopher **Smith-Clarke** and Licia **Capra**, University College London, UK

[88] *"Research directions for harvesting cross-sectorial correlations towards improved policy making"* – Nikolaos **Dimitriou** and Anastasios **Drosou**, CERTH/ITI, Greece; Nikolaos **Sarris**, ATC, Greece; Athanasios **Konstantinidis**, ICL, United Kingdom; Dimitrios **Tzovaras**, CERTH/ITI, Greece

[96] *"Data-Driven Discussions: A Social Platform for Open Data to discuss and visualize from heterogeneous data sources"* – Gennaro **Cordasco**, Seconda Università di Napoli; Renato **De Donato**, Delfina **Malandrino**, Pina **Palmieri**, Andrea **Petta**, Donato **Pirozzi**, Vittorio **Scarano**, Luigi **Serra**, Carmine **Spagnuolo** and Luca **Vicidomini**, Dipartimento di Informatica, Università degli Studi di Salerno, Italy

[104] *"Mass Replication of Policy Evaluation Through Citizen-Led Randomized Trials"* - J. Nathan **Matias**, MIT Media Lab Center for Civic Media, US

---

#### **13:40 – 14:40      Keynote Lecture 1 (Room – LT1)**

---

*"Data for Policy: a Myth or a Must?"*

Enrico **Giovannini**, Co-Chair, UN Data Revolution Group, University of Rome 'Tor Vergata'

**Chair:** Andrew **Blake**, Executive Director, The Alan Turing Institute

---

#### **14:45 – 15:45      Parallel Session 2**

---

##### **Session 2a: Official Statistics (Room – LT1)**

**Chair:** David **Johnson**, Interim Head of the Data Science Campus, Office for National Statistics

[15] *"Data Science and Public Policy: A Changing Landscape for the US Federal Government"* – Dan **Gaylin**, NORC at the University of Chicago, USA

[16] *"On the Evolution of the United Kingdom Price Distributions"* – Kim **Huynh**, Bank of Canada, Canada; Ba **Chu**, Carleton University, Canada; David **Jacho-Chavez**, Emory University, USA; Oleksiy **Kryvtsov**, Bank of Canada, Canada

[51] *"Role of Data science in official statistics (ONS)"* – Iva **Spakulova**, Office for National Statistics, UK

[90] *"Big Data as a Source for Official Statistics: Assessment of Using Mobile Phone Data for Population"* – Fernando **Reis**, Eurostat, Luxembourg; Freddy **de Meersman**, Proximus, Belgium; Marc **Debusschere**, Statistics Belgium, Belgium; Albrecht **Wirthmann**, European Commission – Eurostat, Luxembourg

**Session 2b: Health & Wellbeing (Room – LT2)**

**Chair:** Derek **Wyatt**, Chair, Royal Trinity Hospice; NHS Innovations South East, UK

[4] *"Fragile states and citizen data analytics: gathering voices and mapping beliefs for UNICEF"* – Claudia **Abreu Lopes**, Africa's Voices Foundation; Sharath **Srinivasan**, Centre of Governance and Human Rights, University of Cambridge, UK

[26] *"Measuring Organizational Performance from Customer Reviews: The case of National Health Service in England"* – Radoslaw **Kowalski**, Slava **Mikhaylov** and Marc **Esteve**; University College London, UK

[45] *"Historical Analysis of Policy and Economic Influence on National Subjective Wellbeing Using"* – Thomas **Hills**, Eugenio **Proto** and Daniel **Sgroi**; University of Warwick, UK

**Session 2c: Data for Land Governance Policy Making (Room – FW26)**

**Chair:** Denys **Nizalova**, Kyiv School of Economics, Ukraine

[41a] *"Public Policies for Creation of Land Records Big Data"* – Sachin **Garg**, School of Policy, Government & International Affairs, George Mason University, USA

[41b] *"Monitoring the Land Governance: case of Ukraine"* – Denys **Nizalova**, Kateryna **Ivinskaa**, Klaus **Deiningerb**, Sergiy **Kubakha**, Oleg **Niviyevskiy**; Kyiv School of Economics, Ukraine; World Bank, USA

**15:45 – 16:05                      Break**

**16:05 – 17:05                      Parallel Session 3**

---

**Session 3a: Open Data (Room – LT1)**

**Chair:** Peter **Smith**, Professor of Social Statistics, ESRC Administrative Data Research Centre for England and University of Southampton

[24] *"Towards open metadata and governance"* – Mandy **Chessell**, IBM, UK

[31] *"Crowdsourcing the collection and analysis of data on companies' Environmental, Social and Governance performance – WikiRate.org"* – Richard **Mills**, University of Cambridge, UK

[86] *"A Policy Oriented Methodology for the Publishing and Maintenance of Open Government Data"* – Andrea **Maurino** and Carlo **Batini**, Università degli Studi di Milano, Bicocca, Italy; Gianluigi **Viscusi**, École Polytechnique Fédérale de Lausanne, Switzerland

[106] *"Sharing more widely: data at the heart of evidence, policy and transformation at Defra"* – Vanessa **Pilley**, Stefan **Janusz**, Farhana **Amin**, Francesca **Gauntlett**, Steve **Wilkinson**, Kathleen **Cameron**, Gary **Kass**, Mike **Rose**, Jonathan **Hicks**, Emily **Miles** and Ian **Boyd**; Department for Environment Food & Rural Affairs (Defra), UK

### **Session 3b: Session organised by the Bank of England (Room – LT2)**

**Chair:** Gabrielle **Demange**, Directeur d'Etudes at EHESS; Associate Chair, Paris School of Economics, France

[100a] *"Gauging market dynamics using EMIR trade repository data: The case of the Swiss franc de-pegging"* – Olga **Cielinska**, Andreas **Joseph**, Ujwal **Shreyas**, John **Tanner** and Michalis **Vasios**, Bank of England, UK

[100b] *"News and narratives in financial systems: Exploiting big data for systemic risk assessment"* – Rickard **Nyman**, University College London; David **Gregory**, Bank of England; Sujit **Kapadia**, Bank of England; Paul **Ormerod**, University College London; David **Tuckett**, University College London; Robert **Smith**, University College London, UK

[100c] *"History Dependence in the Housing Market"* – Philippe **Bracke**, Bank of England; Silvana **Tenreyro**, London School of Economics and Political Sciences, UK

### **Session 3c: Urban Data I (Room – FW26)**

**Chair:** David **Howarth**, Professor of Law and Public Policy, Department of Politics and International Studies, University of Cambridge

[27] *"Online spaces for urban citizen-sourcing: A comparison of US civic engagement apps"* – Sarah **Giest**, Leiden University, Netherlands; Ansgar **Koene**, Nottingham University, UK; Elvira Perez **Vallejos**, Nottingham University, UK; Olli **Pitkänen**, Helsinki Institute for Information Technology, Finland; and Mattia **Fosci**, Gada, UK

[68] *"Big data visions: towards the development of low-carbon road transport policies"* – Elena **Paffumi**, European Commission JRC Institute for Energy and Transport, Italy; Michele **De Gennaro**, Austrian Institute of Technology G.m.b.H., Austria; **Giorgio Martini**, European Commission JRC Institute for Energy and Transport, Italy

[74] *"Boosting Tech Innovation Ecosystems in Cities, A framework for growth and sustainability of urban tech innovation ecosystems"* – Nga **Nguyen**, World Bank, Netherlands; Victor **Mulas** and Xin **Qian**, World Bank, USA

[105] *“Use of Emerging Data to Improve Passenger Vehicle Safety, Emissions, and Mobility Policy”* – H Scott **Matthews**, Paul **Fischbeck**, and Chris **Hendrickson**, Carnegie Mellon University; Dana **Peck**, US Department of Transportation, USA

**17:05 – 19:00                      Conference Reception**

*\*\* Please continue to the next page for the second day's programme \*\**

## Friday, 16 September 2016

---

09:15 – 09:30      Tea & Coffee

09:30 – 10:30      Parallel Session 4

---

### Session 4a: *Methods II* (Room – LT1)

**Chair:** Rayid **Ghani**, Director of the Center for Data Science and Public Policy, University of Chicago

[28] *“Rankings on data: the two-sided setting”* – Gabrielle **Demange**, Paris School of Economics - PSE-EHESS, France

[37] *“The role of mixed methods data in causal inference: Examples from three Australian trials”* – Nicholas **Biddle**, Australian National University, Australia

[54] *“Assessing the Effectiveness of EU Conditionality in Driving Change in Strategy Processes and Strategies: Application of text analytics for policy analysis”* – Rioh **Burke-Derby** and Alfio **Puglisi**, University College London, UK; Alexander **Kleibrink**, European Commission Joint Research Centre, Spain; Slava **Mikhaylov**, University College London, UK

[75] *“Drones and Data: UAVs as instrument of gathering data and its impact on intelligence”* – Clemens **Binder** and Isabella **Rebasso**, Austrian Institute for International Affairs, Austria

### Session 4b: *Big data for science and innovation policy* (Room – LT2)

Organised by **Technopolis Group**

**Chair:** Cristina **Rosemberg**, Technopolis Group

[42] *“Building an innovation data analytics platform for Welsh Government: Early findings from an ongoing project”* – Juan **Mateos-Garcia**, James **Gardiner** and Jen **Rae**, Nesta, UK

[108] *“The impact of public support for innovation on firm level outcomes: using data linking for policy evaluation”* – Mike **King**, National Physical Laboratory; Dan **Hodges**, Innovate UK

Discussant: Ian **Viney**, UK Medical Research Council

### Session 4c: *Workshop – Network of Innovators* (Room – FW26)

Organised by **The GovLab**, New York University

[93]

Beth **Noveck**, The Jerry Hultin Global Network Professor, Tandon School of Engineering; Co-Founder and Director of The GovLab, New York University

Andrew **Young**, Associate Director of Research, The GovLab, New York University

**10:30 – 10:50                      Break**

**10:50 – 11:50                      Keynote Lecture 2   (Room – LT1)**

---

*“Big Data and the Social Sciences: Can Accuracy and Privacy Co-Exist?”*

Jim **Waldo**, Gordon McKay Professor of the Practice of Computer Science and the Chief Technology Officer, School of Engineering and Applied Sciences, Harvard University

**Chair:** Anthony **Finklestein**, Chief Scientific Adviser for National Security to HM Government; Professor of Software Systems Engineering, The Alan Turing Institute

**11:50 – 12:10                      Student Poster Finalists – 3min presentations   (Room – LT1)**

**Chair:** Anil **Baharath**, Data Science Institute & Department of Bioengineering, Imperial College London

**12:10 – 13:10                      Lunch & Poster Session**

**13:10 – 14:10                      Parallel Session 5**

---

**Session 5a: Urban Data II   (Room – LT1)**

**Chair:** Michail **Skaliotis**, Head of Task Force - Big Data at EUROSTAT, European Commission

[44] *“Policy recommendations to promote the use of digital social innovation and open data for better addressing the mobility requirements of people with disabilities”* – Louise **Francis**, University College London; Mireia **Ferri**, Estrella **Dura**, and Maite **Ferrando** Polibienestar Research Institute, University of Valencia, Spain

[64] *“Urban Planning for Active and Healthy Public Spaces with User-Generated Big Data”* – Loes **van Renswouw**, Fontys University of Applied Sciences, School of Sport Studies; Sander **Bogers**, Eindhoven University of Technology; Steven **Vos**, Fontys University of Applied Sciences, School of Sport Studies, Netherlands

[87] *“Finding optimal sites for free schools using public sector data”* – Dan **Heron**, Cabinet Office, UK

**Session 5b: Politics (Room – LT2)**

**Chair:** Slava **Mikhaylov**, Department of Political Science, University College London

[7] *“Reading Between the Lines: Prediction of Political Violence Using Newspaper Text”* – Mueller **Hannes**, IAE (CSIC), Barcelona GSE, Spain; Christopher **Rauh**, University of Cambridge, INET Institute, UK

[40] *“The horizontalization of EU affairs in national parliaments: A dynamic automated content analysis of plenary debates in The Netherlands and the UK”* – Rik **de Ruiter** and Jelmer **Schalk**, Leiden university, Netherlands

[48] *“New and improved indicators for ‘fragile states’?”* – Isabel **Rocha de Siqueira**, IRI/PUC-Rio, Brazil

[52] *“Data-Driven Innovation for NGO’s: How to define and mobilise the Data Revolution for Sustainable Development?”* – Thomas **Baar**, Aikaterini **Deligianni** and Christoph Johann **Stettina**, Leiden University, Netherlands

**Session 5c: Big data in the online economy (Room – FW26)**

Organised by **European Commission**

**Chair:** TBC

[29] *“Taxation on Data and privacy protection”* – Francis **Bloch**, PSE-Paris 1; Gabrielle **Demange**, PSE-EHESS, France

[61] *“Financial data for the Good Society: an innovative data-science approach for Big Data regulatory reporting”* – Harald **Stieber**, European Commission, DG FISMA, Belgium; Prabhat **Agarwal**, European Commission, DG CONNECT, Belgium; Petros **Kavassalis**, University of the Aegean, Greece; Keith **Saxton**, KS Strategic; Francis **Gross**, European Central Bank, DG Statistics, Germany

[102] *“Living Lab Research as Intermediation”* – James **Stewart**, Ewan **Klein** and Catherine **McGill**, University of Edinburgh, UK

Discussant: Daniel **Castro**, Director, Centre for Data Innovation, US

---

**14:15 – 15:15      Parallel Session 6**

**Session 6a: Economy (Room – LT1)**

**Chair:** Barbara **Ubaldi**, The Organisation for Economic Co-operation and Development (OECD)

[35] *“Diverse uses of government contracting data to improve spending of public funds”* – Mihály **Fazekas** and Ágnes **Czibik**, University of Cambridge, UK

[49] *“Using new data to answer old questions on the extractive industries”* – David **Mihalyi** and David **Manley**, Natural Resource Governance Institute, UK



[76] *"Competitive Intensity and Corruption Risks in the Hungarian Public Procurement 2009-2015"* – Istvan Janos **Toth** and Miklos **Hajdu**, Corruption Research Center Budapest, Hungary

#### **Session 6b: Methods III (Room – LT2)**

**Chair:** David **Mair**, Head of Unit, Science Advice to Policy, European Commission Joint Research Centre

[22] *"Reviewing the worldwide trends of National Geoportals"* – Joep **Crompvoets**, Bruno **Broucker** and Geert **Bouckaert**, KU Leuven - Public Management Institute, Belgium

[98] *"Proxy Errors with Policy Consequences: Agricultural Productivity"* – Caroline Leigh **Anderson**, University of Washington Evans School; Travis **Reynolds**, Colby College; Pierre **Biscaye**, Evans School of Public Policy & Governance; Joshua **Merfeld**, University of Washington, USA

[107] *"Developing a semantic legal research interface for the OECD Services Trade Restrictiveness Index"* – Frédéric **Gonzales**, Sébastien **Miroudot** and Thierry **Vebr**, Organisation for Economic Cooperation and Development (OECD), France

#### **Session 6c: Policy Process II (Room – FW26)**

**Chair:** Daniel **Castro**, Director, Centre for Data Innovation, US

[6] *"Privacy - conscientious use of mobile phone data"* – Yves-Alexandre **de Montjoye**, Massachusetts Institute of Technology, USA

[47] *"The Next Step in Data-Driven Policy: Tensions in the Air and Progress on the Ground"* – Ralph **Schroeder**, Oxford Internet Institute, UK; Martijn **Poel**, Technopolis, Netherlands; Diana **Vlad Calcic**, European Commission, Belgium; Jerome **Treperman**, Technopolis, Germany; Arthur Thomas, Oxford Internet Institute, UK; Joost van Barneveld, Technopolis, Netherlands

[81] *"Humanitarian Data Responsibility: Distributed Risk - Holistic Governance"* – Johan Bernard **Berens**, Leiden University, Netherlands; Lucy **Bernholz**, PACS, Stanford University, USA; Nathaniel **Raymond**, Harvard Humanitarian Initiative, USA; Gideon **Shimshon**, Centre for Innovation, Leiden University, Netherlands; Stefaan **Verhulst**, The Governance Laboratory, New York University, USA

[83] *"Development of an engagement framework for data science and its use in policy"* – Michelle **Brook** and Anthony **Zacharzewski**, The Democratic Society, UK

**15:15 – 15:35      Break**

**15:35 – 16:35                      Closing Plenary Session**

---

*"Translating Rigorous Evidence into Policies that Benefit the Poor"*

Quentin **Palfrey**, Executive Director, J-PAL North America, Massachusetts  
Institute of Technology (MIT)

*"A Data-driven Public Sector for Sustainable and Inclusive Governance"*

Barbara **Ubaldi**, The Organisation for Economic Co-operation and Development  
(OECD)

*"Doing practical data science for social good and public policy"*

Rayid **Ghani**, Director of the Center for Data Science and Public Policy,  
University of Chicago

**Chair:** Jean **Bacon**, Professor Emerita of Distributed Systems, Computer Laboratory,  
University of Cambridge

**16:35 – 16:50                      Awards & Final Remarks**

**16:50                                      End**

# Terms & Conditions

---

Organisers cannot accept any liability for personal injuries or for loss or damage to property belonging to the delegates, either during, or as a result of the conference. Please check the validity of your own personal insurance before travelling.

Photography and video-recording will take place at conference venues to be used for post-conference publications and other related online/printed material to be produced by Data for Policy. Any reservations about this condition should be sent to [team@dataforpolicy.org](mailto:team@dataforpolicy.org) prior to the conference to avoid any disappointment in the future.

I agree to Data for Policy processing personal data contained within the registration process, or other data that may be obtained from me or other people whilst I am applying for the conference. I agree to the processing of such data for any purpose connected with my attendance at the conference, or my health and safety whilst on event premises.

The organisers reserve the right to change conference programme and venue details, and to cancel the conference in case of any unpredictable event.

## Notes

---

## Notes

---

## Notes

---



