



PROFESSIONAL TRAINING WORKSHOP

on Big Data and development



DATA-POP
ALLIANCE

June 27–30, 2016
MIT Media Lab
Cambridge, MA

AGENDA

DAY 0 | MONDAY, JUNE 27

SESSION 0 | 05:30pm—07:30pm

Welcome

Welcome reception and opening remarks

[Alex 'Sandy' Pentland](#)

Prof. Alex 'Sandy' Pentland, Academic Director of Data-Pop Alliance, will welcome participants and introduce Data-Pop Alliance's Professional Training Program, followed by an informal reception with light refreshments.

DAY 1 | TUESDAY, JUNE 28

Breakfast and registration at 8:00 am

SESSION 1 | 08:45am—09:30am

Opening

Welcome and overview

[Natalie Shoup](#) | 08:45am—09:00am

This session will explain the rationale, background, modalities and objectives of the workshop.

Participant lightning introductions

[Juan Casasbuenas](#) | 09:00am—09:30am

SESSION 2 | 09:30am—10:30pm

Presentation: Framing Big Data and development

Policies, societies and literacy in the age of Big Data: building capacities to foster social change

[Emmanuel Letouzé](#) | 09:30am—10:30am

This session will provide participants with a foundation of the context and concepts of Big Data and development.

SESSION 3 | 10:45am—12:30pm

Presentation: Synthesis of research and practice

Leveraging cell-phone data and satellite imagery for sustainable development and in humanitarian crises: The experience of Flowminder

[Linus Bengtsson & Erik Wetter](#) | 10:45am—11:45am

Technical tutorial preview

[Gabriel Pestre](#) | 12:15pm—12:30pm

Through interactive presentations, this session will showcase case studies of using Big Data in humanitarian action, development research and practice, as well as provide a preview of the datasets that participants will use during the workshop.

Q&A and group discussion

[Juan Casasbuenas](#) | 11:45am—12:15pm

Lunch served from 12:30pm—01:30pm

SESSION 4 | 02:00pm—03:30pm

Keynote presentation

Social physics: how social networks can make us smarter

[Alex 'Sandy' Pentland](#) | 02:00pm—03:00pm

During this session, Prof. Alex 'Sandy' Pentland, Academic Director of Data-Pop Alliance, will discuss 'social physics' and its potential to build predictive, computational theory of human behavior towards better social systems.

Q&A and group discussion

[Emmanuel Letouzé](#) | 03:00pm—03:30pm

SESSION 5 | 03:45pm—06:00pm

Group work: Big Data and development in action

Interactive workshop

[David Dabscheck & David Sangokoya](#)

03:45pm—05:30pm

Through interactive, design-thinking group activities, this session will help participants synthesize and connect the day's key terms and concepts and start thinking about how to apply them to concrete problems in their work.

Group discussion

[Juan Casasbuenas](#) | 05:30pm—06:00pm

DAY 2 | WEDNESDAY, JUNE 29

Breakfast at 8:00 am

SESSION 6 | 08:30am—10:00am

Technical tutorial 1

Machine learning and language classification using Yelp data from Montreal

[Gabriel Pestre, Andrés Clavijo, & Rodrigo Lara Molina](#)

08:30am—09:30am

This facilitator-led tutorial session will introduce basic technical concepts (e.g. APIs, text mining, etc.) and give participants a hands-on training in coding in Python by manipulating data from Yelp and demonstrating a simple machine learning language classification model.

Group discussion on take-aways

Moderated by [Andrés Clavijo](#) | 09:30am—10:00am

SESSION 7 | 10:15am—11:45am

Technical tutorial 2

Sentiment analysis using geo-tagged Twitter data from Colombia

[Gabriel Pestre, Andrés Clavijo & Rodrigo Lara Molina](#)

10:15am—11:15am

Group discussion on take-aways

Moderated by [Andrés Clavijo](#) | 11:15am—11:45am

This facilitator-led tutorial session will give participants a hands-on training in building more complex sentiment analysis models using geo-tagged Twitter data from Colombia.

SESSION 8 | 11:45am—01:00pm

Technical tutorial 3

Analyzing mobile phone metadata with Bandicoot

[Yves-Alexandre de Montjoye & Luc Rocher](#)

11:45am—12:45pm

This session will re-introduce the topic of mobile phone data and development and give participants hands-on experience in metadata analysis through the Bandicoot toolkit.

Q&A and group discussion

12:45pm—01:00pm

Lunch served from 01:00pm—02:00pm

SESSION 9 | 02:00pm—07:30pm

Data Expedition: Crime, conflict and social media in Colombia

Ideation, exploration, visualization and presentation

[Sergio Araiza & Emmanuel Letouzé](#)

02:00pm—07:30pm

Through a series of guided lessons and independent group work, the Data Expedition will give participants the opportunity to build on the tools and lessons presented earlier in the day by exploring research questions using real world datasets. Casual dinner will be served.

DAY 3 | TUESDAY, JUNE 29

Breakfast at 8:00 am

SESSION 10 | 09:00am—10.30am

Presentation

The ethics and politics of Big Data

[Patrick Vinck](#) | 09:00am—10:00am

This session will provide an overview of existing frameworks and ongoing discussions on data ethics and implications for development projects and policies.

Q&A and group discussion

10:00am—10:30am

SESSION 11 | 10:45am—12:15pm

Technical tutorial 4

Big, small, and popular data—engaging communities with data

[Rahul Bhargava & Catherine D'Ignazio](#) | 10:45am—11:15pm

Using simple online and offline tools, this session will help participants engage wider populations with data analysis and presentation.

Group activities around WTFcsv and WordCounter

[Rahul Bhargava & Catherine D'Ignazio](#) | 11:15am—12:15pm

Lunch served during session 12

SESSION 12 | 12:30pm—02:00pm

Presentation: Lightning talks

Mixed methods in complex contexts + Q&A

[Phuong Pham](#) | 12:30pm—01:30pm

This session will explore methodological challenges and advances in using data-driven approaches in fragile settings.

Analyzing urban vitality and criminality through Big Data

[Bruno Lepri \(via video\)](#) | 01:30pm—02:00pm

SESSION 13 | 02:00pm—03:30pm

Data Expedition: Feedback

Presentation of Data Expedition projects

[Sergio Araiza](#) | 02:00pm—03:00pm

During this session, participants will present the findings from the Data Expedition projects and discuss implications and requirements for designing and developing Big Data and development projects in their work.

Facilitator led group discussion

[Natalie Shoup & Juan Casasbuenas](#)

03:00pm—03:30pm

SESSION 14 | 03:45pm—05:00pm

Wrap-up

Debrief

Moderated by [Juan Casasbuenas](#)

03:45pm—04:15pm

This session provides participants the opportunity to share feedback on the workshop and hear about Data-Pop Alliance's ongoing and planned training activities.

Roadmap of future training workshops and activities

Presented by [Emmanuel Letouzé](#)

04:15pm—04:30pm

Feedback on the pilot

Moderated by [Juan Casasbuenas](#)

04:30pm—05:00pm