

**Workshop on the Intersection of Data Science & Population Science:
Future Directions for Collaboration & Accelerating Scientific Discoveries**

February 28 – March 1, 2019
Duke University

This workshop brings together leaders from U.S. population research centers for a thorough discussion of the possibilities presented through greater collaborations and interactions between data scientists and population scientists. The workshop emphasis lies primarily with advancing population science through closer collaboration with data scientists, although the mutual benefit of the interaction will also be discussed. There are multiple goals for the workshop. The first is to deepen and ‘seed’ our workshop discussions by bringing leading researchers at the forefront of collaborations between data scientists and social scientists. These researchers will offer their own perspectives from their unique vantage points. They include: Dr. Emilio Zagheni (Max Planck Institute of Demographic Research), Dr. Elizabeth Bruch (University of Michigan), Dr. Ingmar Weber (Qatar Computing Research Institute) and Dr. Chris Bail (Duke University). The second is to chart a future course for building capacities for collaboration and accelerating scientific work at the intersection of data science and population science. An outcome of this second goal may be proposals for several small grants that could fund summer workshops or collaborations and to which we would invite co-investigators from amongst those participating in the workshop. The third is to elaborate and hone clear articulations for advancing population science presented by more collaborations and better training at the intersection with data science. It is our hope that one outcome would be a white paper about data science and demography, sponsored by the Association of Population Centers and prepared for submission to a leading scientific publishing outlet.

To maximize workshop success, participants are invited to read the following papers in preparation for the meetings and discussions:

Bruch, Elizabeth, Fred Feinberg, and Kee Yeun Lee. 2016. “Extracting multistage screening rules from online dating activity data.” *Proceedings of the National Academy of Sciences of the United States of America* 113: 10530-10535.

Cesare, Nina, Hedwig Lee, Tyler McCormick, Emma Spiro, Emilio Zagheni. 2018. Promises and Pitfalls of Using Digital Traces for Demographic Research. *Demography* 55(5):1979-1999.

Kreuger, Kurt and Elizabeth Bruch. 2018. System Science and Data Science. White Paper.

Harron, Katie, Chris Dibben, James Boyd et al. 2017. Challenges in administrative data linkage for research. *Big Data & Society*. 7:1-12.

Zagheni, Emilio and Ingmar Weber. 2015. Demographic research with non-representative internet data. *International Journal of Manpower* Vol. 36 No. 1, 2015 pp. 13-25.

**Workshop on the Intersection of Data Science & Population Science:
Future Directions for Collaboration & Accelerating Scientific Discoveries**

February 28 – March 1, 2019
Duke University

Agenda

February 28 Drinks and Dinner: 6:30pm in the Bolton Family Tower Room, Broadhead Center, Duke University

Participants' Introduction - including a description of the data science projects at their institution

March 1 Workshop: Gross Hall Room 230E, Duke University

8:30-8:45 -- Welcome and introductions

8:45-10:05 -- Presentations (20 min each)

8:45-9:05: Emilio Zagheni, Max Planck Institute for Demographic Research: "Reflections from the perspective of a demographer –How do you do better population science using novel data and data science methods?"

9:05-9:25: Elizabeth Bruch, University of Michigan: "Reflections from the perspective of a sociologist – How do you better model social behavior with data science tools and big data?"

9:25-9:45 Ingmar Weber, Qatar Computing Research Institute: "Reflections from the perspective of a computer scientist – How do social science collaborations advance data science and computer science? "

9:45-10:05: Chris Bail, Duke University – "Reflections on computational social science and demography."

10:05-11:00 -- Q&A with panelists

11:00-11:15 – Coffee break

11:15-12:30 -- Structured Discussions around the following questions:

What are the promises and pitfalls of big data relevant for population science research questions?

How can data science and sources of big data provide the evidence for revisiting the major frameworks of population science? - *e.g. Are there persistent puzzles faced by population scientists that might be addressed better with new data and data science approaches and systems and that may revolutionize our frameworks and generate new paradigms?*

How does one tackle the challenges of accurate descriptions and inference and lack of structure in big data to answer population science questions?

12:30-1:30 -- Lunch

1:30-3:30 – Continue Structured Discussions

How do we adapt training in demography for the next generation to catalyze cutting-edge population science that relies on an ever-growing variety of data sources? What are the incentives for young researchers to embark on a career at the intersection of demography and data science?

How can population scientists (and population centers and associations of population centers) favor collaborations with the industry to develop approaches for sharing data that are mutually beneficial, ethical and preserve the privacy of individuals?

3:30-4:00 – Wrap up

Additional Participants

Emily Agree (Hopkins)	Giovanna Merli (Duke University)
Jennie Brand (UCLA)	Jim Moody (Duke University)
Kyle Crowder (U of Washington)	Dana Pasquale (Duke University)
Sara Curran (U of Washington)	Jerry Reiter (Duke University)
Dennis Feehan (UC Berkeley)	Marcos Rangel (Duke University)
Lingxin Hao (Johns Hopkins U)	Jennie Romich (U of Washington)
Stephane Helleringer (Johns Hopkins U)	Mark Yacoub (Duke University)
Tyler McCormick (U of Washington)	

Acknowledgements: Thank you to the Duke Population Research Center for hosting the workshop and for covering Drs. Zagheni's and Weber's travel expenses and to the population centers of the Johns Hopkins University, the University of California at Berkeley, the University of California at Los Angeles, and the University of Washington for supporting participants' travel.