

CONTACT
INFORMATION3416 Plateau Dr
Belmont, CA 94002(608) 338-4844
cornelia.ilin@berkeley.edu
<http://corneliailin.github.io>**Citizenship: U.S. permanent resident**

PROFILE

Faculty in AI/ML at UC Berkeley with 12 years of experience in academia, industry, and consulting. Doing research at the intersection of health and environment, employing a wide variety of **geospatial**, **statistical**, and **machine learning** methods. Teaching applied AI/ML classes to MS students in the Data Science program at UC Berkeley.

- Experience with supervised and unsupervised machine learning algorithms (incl., deep neural networks such as FNN, CNN, RNN/LSTM, Transformers), causal inference, geospatial analysis, temporal record linkage of vital records and hospital data
- Computer languages include **Python**, GIS, SQL, Bash shell; version control with Git(Hub)
- Experience with **big data analytics**: analyzed data using Hadoop/HDFS and Dask frameworks on CloudLab.us clusters and Google Cloud Platform
- Experience with packages such as TensorFlow2, SHAP for explainable AI, Scikit-Learn, Statsmodels, Geopandas, OSMnx, Rasterio
- Experience with IRB applications and working with **HIPAA protected** health data, including claims and electronic medical records (EMR, e.g., diagnosis, lab values, clinician notes, images).
- 2 years of **industry and consulting experience** with contributions to high-profile litigation cases in the healthcare industry
- Presentation and writing skills (I attend conferences and write papers for journal publications). Collaboration skills (direct interaction with faculty, postdocs, and graduate students). Highly rated Data Science lecturer.

PROFESSIONAL AND
RESEARCH
EXPERIENCE**UC-Berkeley**, School of Information, **Faculty in AI/ML**

2021 - present

- Research on maternal and child health: The effects of pollution exposure in-utero and first years of life on the likelihood of a couple of conditions in one's hospital/ER visit later in life. Includes personalized interpretations of predictions using diagnosis embeddings and SHAP.
- Data: EMR, vital statistics, atmospheric and wildfire data
- Methods: Linear regression model for building a pollution exposure instrument (proximity to wildfires) and causal inference, BERT model for predictive analysis
- Teaching: Applied Machine Learning, Capstone (final research project class)
- Topics: linear regression, logistic regression, random forests, unsupervised learning (clustering and dimensionality reduction), deep networks (FNN, CNN, RNN/LSTM, Transformers), embeddings, transfer learning, model interpretability and fairness

Stanford University, RegLab, **Research Scientist**

2021 - 2022

- Project (in partnership with the EPA): explore how machine learning can be used to protect human health with a focus on environmental justice and equity in California.

- Data: Satellite imagery to detect Concentrated Animal Feeding Operations (CAFOs), atmospheric data (wind patterns from NASA's MERRA-2 product), water pollutant discharge monitoring data (ICIS-NPDES), hospital/ER data (CDPH), census tract data (Geolytics Neighborhood Change Database)
- Methods: satellite imagery time series classification to detect building construction and/or expansion, causal inference with instrumental variables

UC-Berkeley, School of Information, Postdoctoral Fellow

2020 - 2021

- Advisor: Joshua Blumenstock, Ph.D.
- Project 1: Provide real-time feedback for managing the spread of COVID-19 at local, national, and global scale. Focus on the impact of non-pharmaceutical policies on human mobility, and the usefulness of cellphone data in predicting the spread of the pandemic.
- Project 2 (Cost-Benefit analysis in partnership with the CDC): Empirically estimate the impact of changes in non-pharmaceutical policy interventions, mobility, and other avoidance behaviors on growth rate of COVID-19 cases, COVID-19 deaths, and economic output for all countries in the world where GDP data is available.
- Data: COVID-19 cases and deaths, aggregate-level cellphone data (Google, Facebook, SafeGraph); individual-level human mobility data (InfoGroup), non-pharmaceutical policy interventions (CDC), quarterly GDP data.
- Methods: Causal inference, predictive analysis.

UW-Madison, Department of Applied Economics, Faculty Associate

2018 - 2020

- Teaching: Object Oriented Programming and Data Analytics with Python; Practicum for Applied Economists; Quantitative and Applied Economics Seminar.
- Topics: Data types, functions, classes, exceptions, IO files, data visualization, descriptive statistics, causal inference, cloud computing (incl. Bash), GIS with Python.
- Assistant Program Director: M.S. in Quantitative and Applied Economics program.

Analysis Group, Inc., Menlo Park, CA, Associate Economist

2017 - 2018

- Litigation consulting: Contributed to several high-profile litigation cases in the healthcare industry (e.g., Des Roches, et al. v. Blue Shield and Magellan).
- Research: Contributed to manuscripts and posters documenting the effectiveness of Leptin Replacement Therapy in treating Generalized and Partial Lipodystrophy.
- Data: Claims (mental health, substance abuse, orthopedics), quasi-experimental and survey.
- Methods: Surveys, discrete choice analysis (multinomial logistic and hierarchical Bayesian regressions), matching algorithms, Cox hazard models.

UW-Madison, Department of Applied Economics, Research Assistant

2012 - 2017

- Advisor: Guanming Shi, Ph.D.
- Projects: Modeled and analyzed the effect of competition and structural market changes on price dispersion in markets with firm capacity constraints and consumer brand loyalty. Explored the evolution of learning and uncertainty in the adoption of new technologies.
- Data: Individual-level data for conventional and genetically modified seed purchases (GFK).
- Methods: Two-stage game theory models, causal inference, quantile regression, instrumental variable, bootstrapping, structural estimation, myopic v. dynamic Bayesian learning models.

University of Zürich, Department of Economics, Research Assistant

2011

EPFL, Department of Computer Science, **Research Assistant** 2010

DHL European Headquarters, Belgium, **Intern** 2009

EDUCATION UW-Madison, Ph.D. in Applied Economics 2012 - 2017
 University of Lausanne, Switzerland, M.S. in Economics 2009 - 2011
 Academy of Economic Studies of Bucharest, Romania, B.S. in Economics 2004 - 2008

JOURNAL PUBLICATIONS Public Mobility Data Enables COVID-19 Forecasting and Management at Local and Global Scales (*Nature - Scientific Reports*, volume 11, article number: 13531, with S. Annan-Phan, X.H. Tai, S. Mehra, S. Hsiang, J. Blumenstock, 2021).

Global Health and Economic Impacts of Behavior Change During the COVID-19 Pandemic (*under review*, *Nature*, with J. Tseng, K.C. Coy, A.C. Ewing, T. Chong, S.M. Marks, I. Bolliger, N.M. Gonzalez, K. Bell, A.J. Hakim, S. Hsiang, 2021)

Evolution of Learning in Technology Adoption: The case of the U.S. Soybean Seed Industry (*under review*, *PLOS One*, with G. Shi, 2021).

Competition, Price Dispersion and Capacity Constraints: The Case of the U.S. Corn Seed Industry, *European Review of Agricultural Economics*, 2021 (with G. Shi).

in preparation:

The Role of Birth and Contemporaneous Pollution Exposure on Respiratory and Cardiovascular Conditions. Evidence from California (with D. Phaneuf, 2019).

MANUSCRIPTS AND POSTERS Longitudinal Matching. A Method for Generating Comparable Samples of Treatment and Treatment-Naive Patients with Progressive Conditions (Analysis Group Inc., 2018).

Effect of Leptin Replacement Therapy on Survival and Disease Progression in Generalized and Partial Lipodystrophy (study funded by Aegerion Pharmaceuticals Inc., 2018).

Patient Quality of Life and Benefits of Leptin Replacement Therapy in Generalized and Partial Lipodystrophy (study funded by Aegerion Pharmaceuticals Inc., 2018).

TEACHING EXPERIENCE

UC-Berkeley:

Lecturer, Capstone: 2022

Lecturer, Applied Machine Learning: 2020, 2021, 2022

UW-Madison:

Lecturer, Practicum for Applied Economists: Fall 2019

Lecturer, Object Oriented Programming and Data Analytics with Python: Summer 2019

Lecturer, Quantitative and Applied Economics Seminar: Spring 2019

TA, World Hunger and Malnutrition: Spring 2017

TA, Applied Econometric Analysis I: Fall 2016

TA, Applied Microeconomic Theory: Fall 2014

Lecturer, Math Camp for Incoming M.S. and Ph.D. Students: Summer 2014

FELLOWSHIPS, SCHOLARSHIPS AND GRANTS

Research Grant, American Bar Association, Section of Antitrust Law, 2016

Ph.D. Summer Program (competitive), Edgeworth Economics, Washington, DC, 2016

Kenneth and Pauline Parsons Graduate Fellowship Fund, UW-Madison, 2016

Best Paper Presentation Award, UW-Madison, 2016
SASC Graduate Funds, University of Lausanne, 2010 - 2011
Hessen Summer School (competitive), Goethe University of Frankfurt am Main, Germany, 2008
WU Summer School (competitive), Vienna University of Economics and Business, Austria, 2007
Excellency in Research Award, Academy of Economic Studies of Bucharest, Romania, 2007

SEMINAR AND
CONFERENCE
PRESENTATIONS

Stanford Maternal and Child Health Research Institute Symposium, 2021
Association of Environmental and Resource Economics, 2020
UW-Madison, Healthcare Economics Group seminar, 2019
University of Connecticut, Department of Economics, 2017
European Association for Research in IO (Rising Stars section), Lisbon, Portugal, 2016
Agricultural and Applied Economic Association Meeting, Boston, Massachusetts, 2016
UW-Madison, Department of Economics, IO seminar, 2016

PROFESSIONAL
ACTIVITIES

Reviewer for Nature Scientific Reports, 2022 - present
Reviewer for the American Public Health Association (APHA), 2019 - present
Social Chair, THC Club of AAE Department, UW-Madison, 2015 - 2016
Seminar Organizer, THC Club of AAE Department, UW-Madison, 2014 - 2015

LANGUAGE SKILLS

Romanian (native), English (fluent), French (basic)