



## Data Science for Social Good

Uncovering the hidden universe of rental units in Surrey

UBC fellows Jocelyn Lee, Andy Fink, Zhe Jiang and Tom Park present their project on hidden rental housing units in Surrey at a Microsoft networking event.

# Cascadia Urban Analytics Cooperative (CUAC) Program Report

## University of British Columbia:

**Jonathan Fink**, Visiting Professor of Urban Analytics; Co-Director, CUAC

**Raymond Ng**, Scientific Director, Data Science Institute; Professor, Computer Science

## University of Washington:

**Bill Howe**, Associate Professor, Information School; Director, Urbanalytics;  
Co-Director, CUAC

**Emily Keller**, Program Manager, Urbanalytics; Program Coordinator, DSSG

Published January 2020

**CASCADIA URBAN**  
ANALYTICS COOPERATIVE

**W**  
UNIVERSITY of  
WASHINGTON



THE UNIVERSITY  
OF BRITISH COLUMBIA

# INTRODUCTION

The **Cascadia Urban Analytics Cooperative** (CUAC) was originally funded to advance four goals, all of which were met. UBC and UW then invested in a broad portfolio of seed grants, initiating new foundational and interdisciplinary research collaborations between their faculty in critical areas across transportation, air quality, housing, responsible data use, and more. These projects are detailed in the remainder of this report.

## GOALS

1. Have UW help UBC **launch a Data Science for Social Good (DSSG) program.**
2. **Provide a venue for students and faculty to advertise their work** in this sector.
3. **Expand UW-UBC research ties** in urban analytics and related fields.
4. Develop systems and services to **facilitate responsible urban data science.**

---

In its first two years, CUAC attracted over \$5M in funding from government, corporate, philanthropic, and internal university sources.

---

## RESULTS

1. Since bootstrapping the DSSG program at UBC two years ago, CUAC has supported **15 DSSG projects** across UW and UBC, **engaging 63 students** and **22 mentors and data scientists.**
2. CUAC researchers routinely showcase results in the **Cascadia Innovation Corridor Conferences, at CUAC-sponsored events** and via the CUAC website.
3. UBC and UW **invested in a broad portfolio of seed grants** (see the remainder of this report.)
4. CUAC has developed and deployed an **Azure-hosted platform to support public-private university data sharing ecosystems** with strong data governance.

In its first two years, CUAC attracted over \$5M in funding from government, corporate, philanthropic, and internal university sources. Its work spans three data-intensive urban research areas (Housing Affordability, Mobility Access and Public Health) and two cross-cutting themes (Responsible Data Science and Legal Frameworks).

In this report, we provide an update of progress in these five research areas, show examples of DSSG projects and point to additional topics of collaborative interest to UBC, UW, and Microsoft, some of which connect them to cutting-edge projects in the City of Portland and at Portland State University (PSU). These latter topics, supported by both external grants and university investment, include large-scale urban data management, campus-based smart city testbeds, and homelessness research.

# TABLE OF CONTENTS

Executive summary	4
Urban research topics:	10
A CUAC housing study influencing statewide policy	10
Transportation access	12
Air quality and public health	14
Responsible data sharing: bias, equity, privacy	15
Cascadia Law Initiative (CLI)	16
Launching Data Science for Social Good (DSSG) as an independent program at UBC	16
Examples of urban-related DSSG projects	17
Interdisciplinary and regional convenings	20
Table 1 - 2017 and 2018 Data Science for Social Good programs at UW and UBC	21

---

In its urban research, CUAC focuses on the study of granular data to inform a holistic understanding of the Vancouver and Seattle metropolitan areas. Projects go beyond data averages and aggregations to identify nuanced localized distinctions that shed light on the complexity of the region as a whole.

---

# EXECUTIVE SUMMARY

In describing research and training activities associated with the Cascadia Urban Analytics Cooperative (CUAC), this report begins by covering three urban research topics: Housing Affordability, Transportation Access and Public Health/Air Quality. It then turns to two cross-cutting themes: Responsible Data Science and a legal framework for cross-border collaboration called the Cascadia Law Initiative. It next summarizes a few of the results from the UW and UBC Data Science for Social Good student internship programs. It wraps up by discussing ways that the generous Microsoft investment has been leveraged, including the launch of a new regional program that uses campuses as testbeds for “smart” urban innovation.

In its urban research, CUAC focuses on the study of granular data to inform a holistic understanding of the Vancouver and Seattle metropolitan areas. Projects go beyond data averages and aggregations to identify nuanced localized distinctions that shed light on the complexity of the region as a whole. These methods uncover inequities and disparities as well as model systems and success stories as they inform a big-picture view of the region’s urban areas. In the Housing Affordability discussion below, we start by showing how this work can lead directly to significant changes in public policy.

## COLLABORATIVE RESEARCH + CROSS-CUTTING THEMES:



1) Housing Affordability



2) Mobility Access



3) Air Quality and Public Health



4) Techniques and Technologies for Responsible Data Science



5) Cascadia Law Initiative



## How can knowledge-based economies have affordable housing?

### NEIGHBORHOOD DYNAMICS

UBC social science faculty Elvin Wyly, Penny Gurstein and Nathan Lauster collaborated with UW postdoc Tim Thomas and PhD student Mike Babb on comparative demographic studies of neighborhood change in Vancouver and Seattle. The UBC team are also continuing UBC's 2018 DSSG student project estimating informal rentals in Surrey based on social media, parking tickets and utility bills. Moreover, the 2018 UBC DSSG housing project also seeded another project in the School of Community and Regional Planning with Riley Iwamoto as the lead. The Data Science Institute provided seed funding as well as the scripts developed by the 2018 DSSG project fellows. The project seeks to better understand the housing market by comparing rental prices in Metro Vancouver with data from the Canada Mortgage and Housing Corporation, as well as the Seattle area.

### USING DATA SCIENCE TO REDUCE DISCRIMINATORY EVICTIONS POLICY

UW research led by Tim Thomas, who directs the CUAC Neighborhood Change Project and the Washington Evictions Research Project, directly contributed to a change in housing policy in Washington State. Thomas testified in front of the State Senate Housing Stability & Affordability Committee in early 2019 as part of a legislative effort that culminated in a series of amendments to landlord-tenant law. This included an extension to the required notice to vacate from 3 to 14 days for evictions based on overdue rent.

### INTEGRATING MULTI-AGENCY DATA IN HOUSING AND EDUCATION

Through a \$750k grant from the Bill and Melinda Gates Foundation, a UW team is working on integrating homelessness and education data from multiple jurisdictions to study equity and efficiency issues in housing service delivery.



## How equitable is transportation access?

### FORECASTING PUBLIC TRANSPORTATION DEMAND TO IMPROVE EQUITY

UW and UBC faculty have been using large data sets from public transit agencies in Metro Vancouver, Seattle and Portland to map out current and projected transit needs in these cities in order to evaluate issues related to access, equity and economic development. For instance, modeling by UBC's Martino Tran's group showed that the projected transit needs of Surrey's rapidly-growing population of seniors will be outstripped within ten years, compromising their ability to access health care facilities.

## INNOVATIONS IN TRANSPORTATION ACCESSIBILITY

Accessibility is a key focus for transportation researchers and staff throughout the region including those at Microsoft. For example, working with CUAC, UW's Anat Caspi developed "AccessMap," an app to help people in wheelchairs negotiate steep slopes, construction sites, and other obstacles, now available in Seattle, Bellingham and Mt. Vernon, Washington. UW and UBC researchers also attended the second PSU-hosted "Mobility Matters" conference on accessibility, transportation and city planning, March 11-12, 2019.

## DATA-DRIVEN POLICY FOR NEW MOBILITY SERVICES

Faculty at UW, UBC and their collaborators have been working with cities, startups and larger companies to evaluate new mobility options like bike-share, car-share, and e-scooters, along with associated phone apps and services. For instance, they have proposed policy interventions that can help apartment residents gain access to EV charging stations, and have used surveys to assess public attitudes about e-scooters.

---

Accessibility is a key focus for transportation researchers and staff throughout the region, including those at Microsoft.

---



## Who is most vulnerable to poor air quality?

### HIGH-RESOLUTION AIR QUALITY IN LIGHT OF INCREASED WILDFIRE SMOKE

Amanda Giang and Naomi Zimmerman at UBC, working with Julian Marshall at UW, funded by a UBC seed grant, have been correlating and modeling high spatial resolution urban air quality data with public health outcomes. This topic has become increasingly relevant due to the increased frequency and intensity of forest fires in Cascadia, which are beginning to have serious consequences for community health.

### NEW INSTRUMENT PACKAGES FOR MULTI-MODAL ENVIRONMENTAL SENSING

Seattle, Vancouver and Portland and partners at UW, UBC and PSU are deploying University of Chicago's "Array of Things" multi-sensor packages, built with NSF and DOE support, to correlate air quality with health outcomes.

### HEALTH CARE ACCESSIBILITY FOR TRANSIT USERS AND PEDESTRIANS

UBC researchers Martino Tran, Jerome R. Mayaud and Rohan Nuttall published the paper, "[An urban data framework for assessing equity in cities: Comparing accessibility to healthcare facilities in Cascadia](#)" in *Computers, Environment and Urban Systems*. Their research identified populations that are socially excluded from health care facilities based on commute times using public transit or foot travel.



## How can we assure more responsible use of data?

### PROTECTING PRIVACY THROUGH DATA SCIENCE

UW faculty and UBC collaborators deployed an Azure-based data sharing platform to help avoid discrimination and protect privacy through bias-corrected synthetic datasets.

### DATA GOVERNANCE

CUAC researchers presented the paper, “[Beyond Open vs. Closed: Balancing Individual Privacy and Public Accountability in Data Sharing](#),” based on the UW Transportation Data Collaborative, at the Association for Computing Machinery Conference on Fairness, Accountability, and Transparency in Atlanta, Georgia in January 2019.

### FAIRNESS

“[Capuchin: Causal Database Repair for Algorithmic Fairness](#),” which discusses synthetic data methods, won SIGMOD Conference’s best paper award for Bill Howe.

### INTERDISCIPLINARY LEADERSHIP

CUAC organized an invitation-only workshop at Data & Society Research Institute in New York City in October 2018 with participation from leading computer scientists, statisticians, and social scientists to articulate a new interdisciplinary field of responsible data science integrating technical methods, critical thinking in design processes, training of Ph.D. students and postdocs, and the development of partnerships with the public sector to account for policy implications of algorithmic systems.



*Data Science for Social Good fellows from UW and UBC meet to share joint mid-term presentations in Vancouver in July 2017.*



## How do cross-border law differences affect urban data use?

### **BINATIONAL LAW STUDY ASSESSES CROSS-BORDER DATA-SHARING ISSUES**

Law professors Hugh Spitzer (UW) and Margot Young (UBC) each recruited two law students to carry out a comparative analysis of the legal differences in data sharing regulations between the US and Canada. The resulting paper compares the basic legal powers available to cities in British Columbia and Washington State, and how these affect the communities' abilities to effectively address technology-related problems common across the border. Topics considered included the creation of a cross-border autonomous vehicle lane, and the exchange of policy lessons associated with transit systems, housing affordability and the opioid crisis.

### **INTER-UNIVERSITY COLLABORATIVE EDUCATION**

Besides providing a new compilation of information critical to cross-border policy collaboration, the Initiative demonstrated the benefits of inter-university educational programs between professional schools. Meetings in Vancouver and Seattle let the students and faculty sharpen their understanding of the contrasting legal systems. For one UBC student, the project sparked interest that led to a position with a municipal law firm and plans to write an academic article on the project for an American publication.

### **LEVERAGING THE MICROSOFT INVESTMENT**

CUAC work has helped attract over \$5M in new funding for urban analytics from various sources, including the Bill & Melinda Gates Foundation, NSF, Mitacs, UW, UBC and PSU.

To accelerate collaboration, UBC provided \$100K worth of seed grants for UBC faculty working with UW colleagues on urban analytic topics including new mobility services; neighborhood-scale variations in life expectancy; links between health outcomes and transportation access; high-resolution correlations between air quality and pulmonary health; and a program for integrating greenspace and health data.

CUAC has helped catalyze the growth of regional efforts. In 2018, the Bullitt Foundation funded the Emerald Corridor Collaboratory to support development of a regional network of city-university partnerships involving Vancouver and UBC; Seattle and UW; Portland and Portland State University; and Bellingham and Western Washington University.



UW DSSG fellows Sean Andrew Chen, An Yan and Haowen Zheng participate in a Human-Centered Design workshop at the eScience Institute.

## EMERGING COLLABORATIONS IN PORTLAND

### HOMELESSNESS RESEARCH AND ACTION

With \$1.5M of internal funding, PSU recently started a “Homelessness Research and Action Collaborative”; plans are underway to collaborate with UW and UBC.

### DATA PLATFORM DEPLOYMENT

The City of Portland, along with PSU, Microsoft, ESRI and two analytics companies, recently launched the Portland Urban Data Lake (PUDL), which is making it easier for cities to analyze large data sets about transportation and other urban infrastructure. We are currently exploring whether PUDL can be expanded to data sets from the cities of Seattle, Vancouver and Surrey, as well as campus data from PSU, UW and UBC.

### SMART CITY TESTBEDS

In order to evaluate data integration across multiple applications and sites, PSU, in partnership with UBC, UW and the City of Portland, launched the Digital City Testbed Center in 2018, which is leveraging \$1.5M of university funding to create a network of campus-based smart city testbeds where technology and policies can be assessed.

# URBAN RESEARCH TOPICS

## A CUAC HOUSING STUDY INFLUENCING STATEWIDE POLICY

A key strategy to achieve CUAC's mission of using data analytics to help solve complex urban problems is creating credible, policy-relevant results. The [Neighborhood Change Project](#) (NCP) stands out as CUAC's first success story, bridging academia and policy to contribute to legislative change. Led by UW Sociology and eScience Institute researcher Timothy A. Thomas and assisted by Geography doctoral candidate Michael Babb, the NCP investigates the effects of neighborhood change and segregation on racial and socioeconomic differences in health, housing, migration, and poverty. The project combines demographic trends and sociological theory with data science techniques to quantify displacement of Seattle area residents as home rental and real estate prices have recently skyrocketed.

To conduct their work, Thomas and Babb combine public data from the Census Bureau's American Community Survey and Decennial Census with real estate and rental listing data from Zillow. The data sets span 40 years, allowing the project to highlight distinctions before and after several economic shocks: the 2007-2009 recession, the ongoing local housing crisis and the growth of tech companies in the Seattle metropolitan area such as Microsoft, Amazon and Google. Aggregating these data sources allows the researchers to map the inflow and outflow of residents within the Seattle metropolitan area, beyond King County, and to other states and countries. Through these methods, Thomas and Babb have developed some of the first models to quantify residential displacement in the Puget Sound area.

In February 2019, Thomas and several co-authors released the report, "[The State of Evictions: Results from the University of Washington Evictions Project](#)," which highlights several demographic trends in home evictions and their impact on homelessness in Washington state:

- Between 2013 and 2017, 130,203 (1 in 55) adults in Washington State were evicted.
- In Pierce County, 1 in 6 black adults were evicted from 2013-2017, compared with 1 in 50 white adults. In King County, 1 in 11 black adults were evicted during that period, compared with 1 in 100 white adults.
- 80-90% of evictions are due to falling behind on rent, with more than 1/3 of defendants contributing over 80% of their income to rent.
- Only 8% of defendants had legal representation during the eviction process.

---

Thomas and Babb developed some of the first models to quantify residential displacement in the Puget Sound area

---

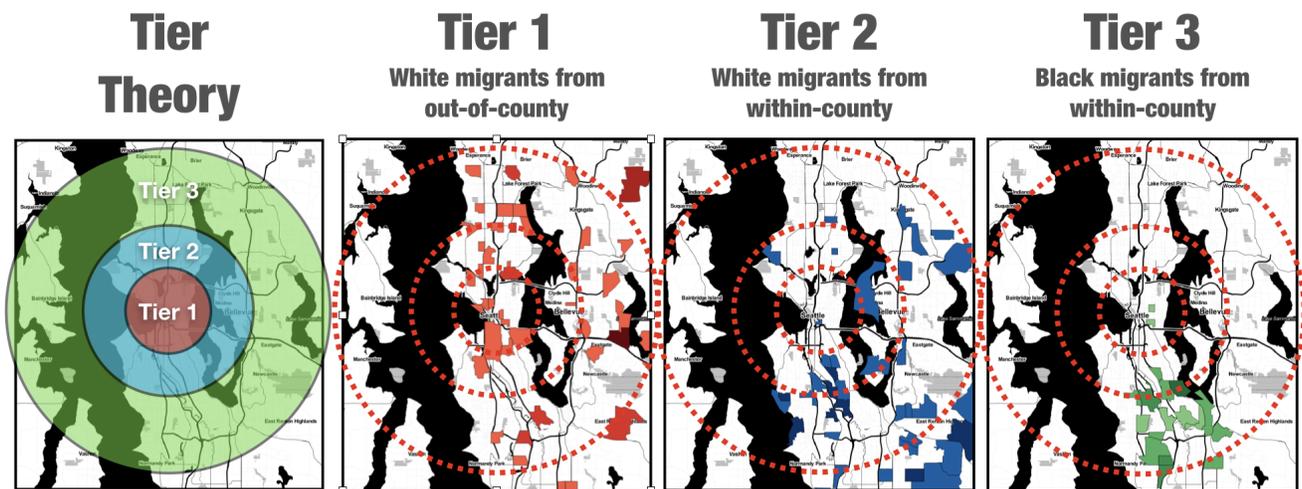


Figure 1. Maps of Greater Seattle illustrating migration patterns for different ethnic groups

Throughout his involvement with CUAC, Thomas, who also leads the [Washington Evictions Research Project](#), has given invited talks about equity, housing displacement and evictions to government, nonprofit, academic and corporate representatives. He has spoken to the Seattle Renters' Commission, the City of Seattle's Office of Arts & Culture, the American Sociological Association and the Population Association of America. His research and testimony have been cited in the [Washington Post](#), [The News Tribune](#), [Komo News](#), [KUOW](#), and [Seattle Met](#).

In January 2019, Thomas testified to the Senate Housing Stability & Affordability Committee at a work session about housing affordability and homelessness trends in Washington State. The following month, he joined [Rep. Nicole Macri](#) and others at a [press conference](#) to present a package of 34 bills introduced in the House of Representatives targeting affordable housing, homelessness and eviction reform. This culminated in the Washington State legislature's passage of Senate Bill 5600 in May 2019, which brings a [cluster of changes](#) to landlord-tenant law, renters' rights and affordable housing development.

According to the new law, which took effect in July 2019, landlords are required to provide 14 days' notice before evicting tenants based on overdue rent, as opposed to the prior 3-day requirement, through an amendment to the Residential Landlord-Tenant Act (RLTA). Landlords will also be required, through a standardized notice to vacate, to provide clear information on tenants' rights, obligations and options (such as low-cost legal support and how to acquire materials in multiple languages); to give tenants 60 days' notice, rather than the prior 30 days, before increasing their rent (with some exceptions); and to provide 120 days' notice to renters before ending their tenancy due to building changes such as demolition or rehabilitation.

Thomas's work is also emblematic of CUAC's mission to support collaborative research with UBC to create projects applicable to the Cascadia region. He has worked closely with UBC researchers studying similar topics in light of complex and even more extreme housing affordability challenges in the Vancouver metropolitan area. UBC faculty Elvin Wyly, Nathan Lauster, Penny Gurstein, Leanne Currie and Martino Tran are exploring relationships between social determinants of health (such as transportation access, neighborhood demography, and health behavior) and neighborhood health outcomes, to create a predictive model that will inform local and regional policies. The NCP's methods are expandable to cities throughout the Cascadia corridor, and have already been applied to Washington State cities surrounding the Seattle area.

# URBAN RESEARCH TOPICS

## TRANSPORTATION ACCESS

Three CUAC-affiliated transportation research projects have focused on neighborhood- and campus-scale variations in access: (1) the [Transportation Data Collaborative](#) (TDC), which utilizes institutional governance processes to facilitate privacy-protected sharing of public and private transportation data for research and transit planning purposes (Figure 2); (2) [Open Sidewalks](#), which seeks inclusion of sidewalks and other pedestrian routes as independent entries on street maps in order to help people with limited mobility negotiate their way through cities (Figure 3); and (3) a comprehensive [study of future accessibility for diverse populations in the City of Surrey](#), BC by UBC's [Urban Predictive Analytics Lab](#) (UPAL)(Figure 4). **These three programs are using advanced analytic tools to translate large, granular transportation data sets into policy, in Cascadia and beyond.** For instance, the TDC ingested and managed usage and geolocation data from three companies in the pilot phase of dockless bike share in Seattle, an approach that is now being applied to other cities in the Pacific Northwest and elsewhere in the U.S. Open Sidewalks is expanding beyond its initial deployment of the [AccessMap](#) sidewalk mapping tool in Seattle (Figure 3) to ten other cities in the U.S. UPAL's modeling results for Surrey show how the time needed to access the City's hospitals by public transport or walking will go up significantly in the next several years (Figure 4), an approach that is now being repackaged for application to other cities in the Cascadia region.

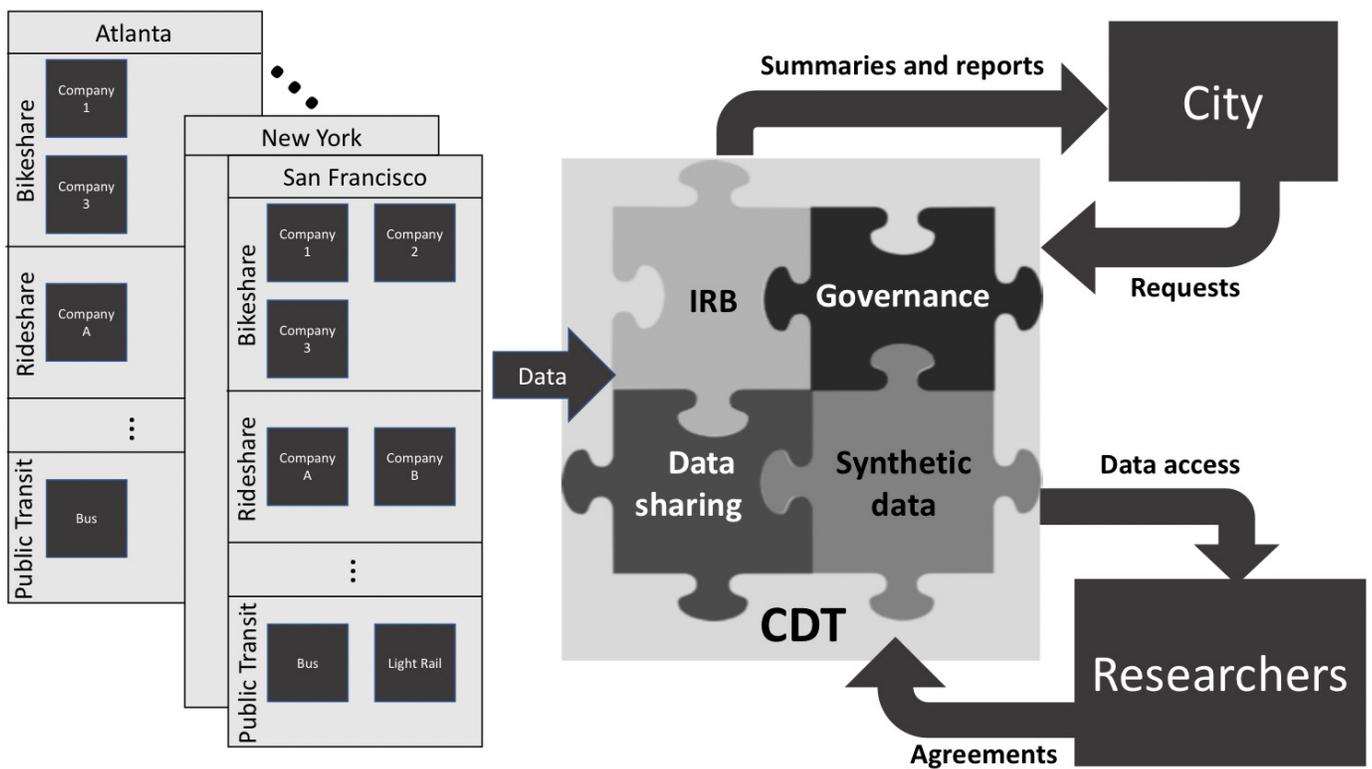


Figure 2. Diagram showing data flows and decision making associated with TDC's approach

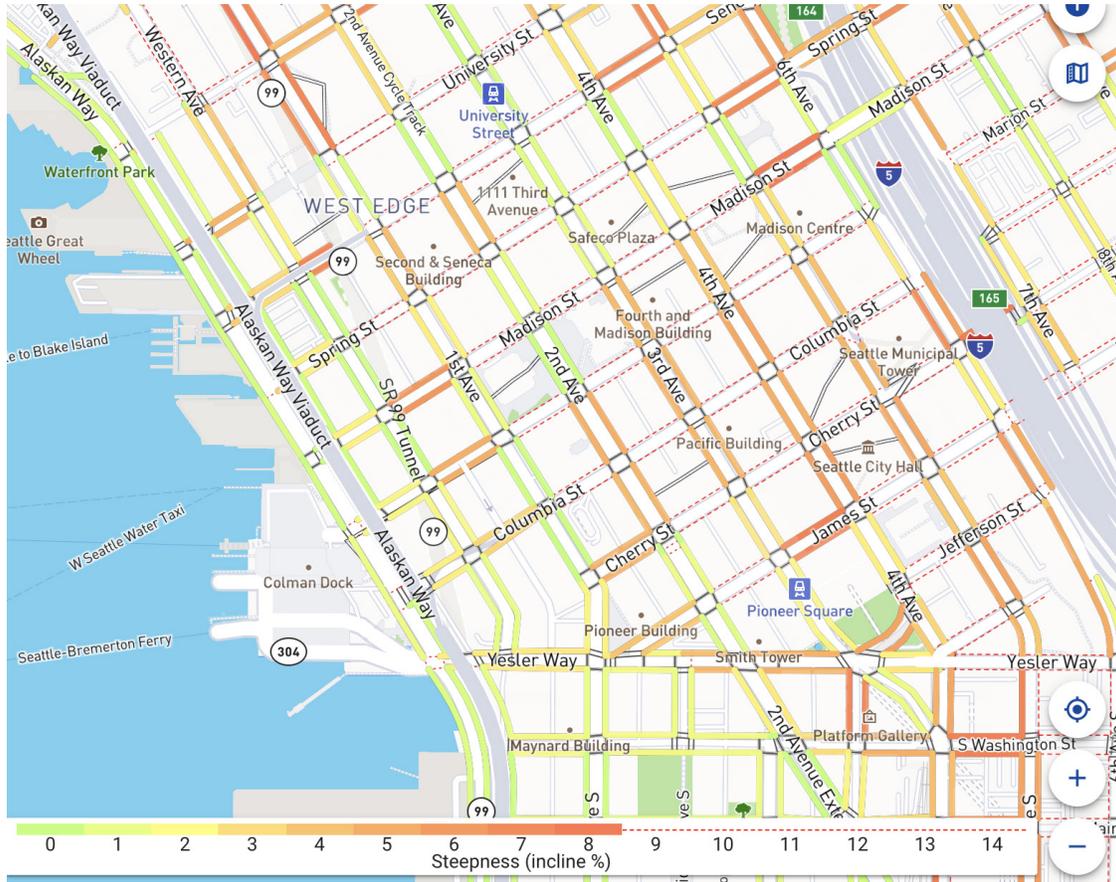


Figure 3. AccessMap helps mobility-challenged people negotiate steep slopes in Seattle

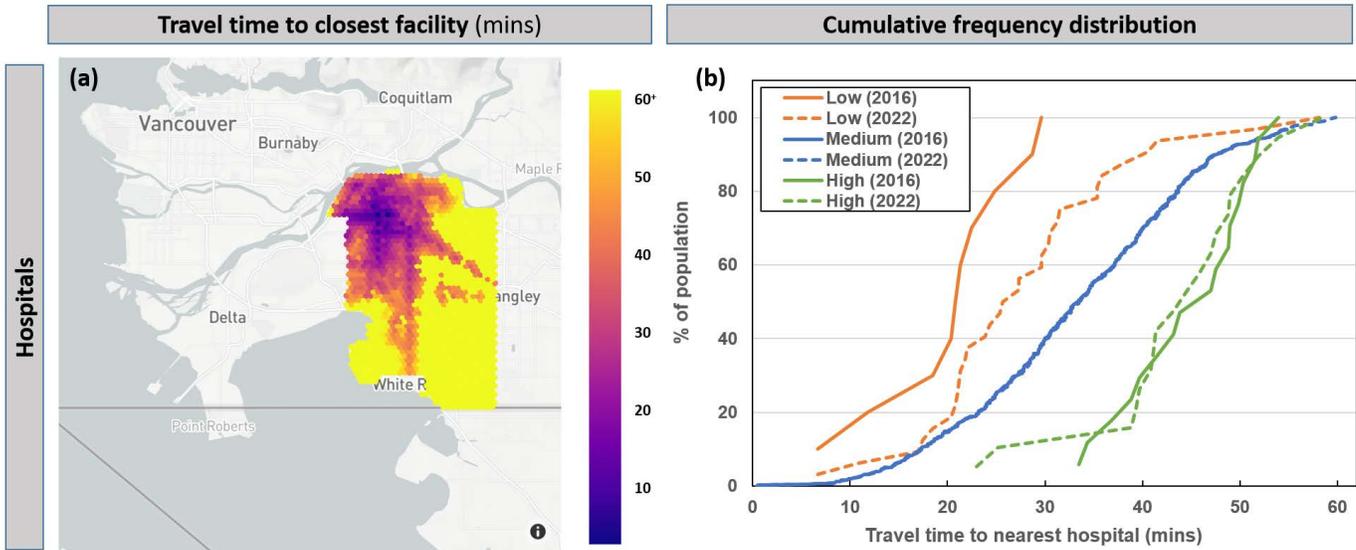


Figure 4. Map and graph showing current and projected future travel times to hospitals in Surrey

# URBAN RESEARCH TOPICS

## AIR QUALITY AND PUBLIC HEALTH

Two other UBC-based CUAC-funded projects are zeroing in on local variations in life expectancy in relation to environmental factors. Taking advantage of new neighborhood-level health, demographic and air quality data from Vancouver, the two groups, led by UBC Population and Public Health Professor Michael Brauer (in collaboration with the BC Centre for Disease Control) and Assistant Professor of Mechanical Engineering Amanda Giang, respectively, along with UW leads Laura Dwyer-Lindgren and Julian Marshall, are discovering correlations that had previously been concealed when looked for in county-scale data. This novel approach allows more targeted and equitable interventions to be proposed at the neighborhood level. Figure 5 shows the higher-resolution pollution and demographic data. These results can help city residents plan for increased urban smoke from wildfires.

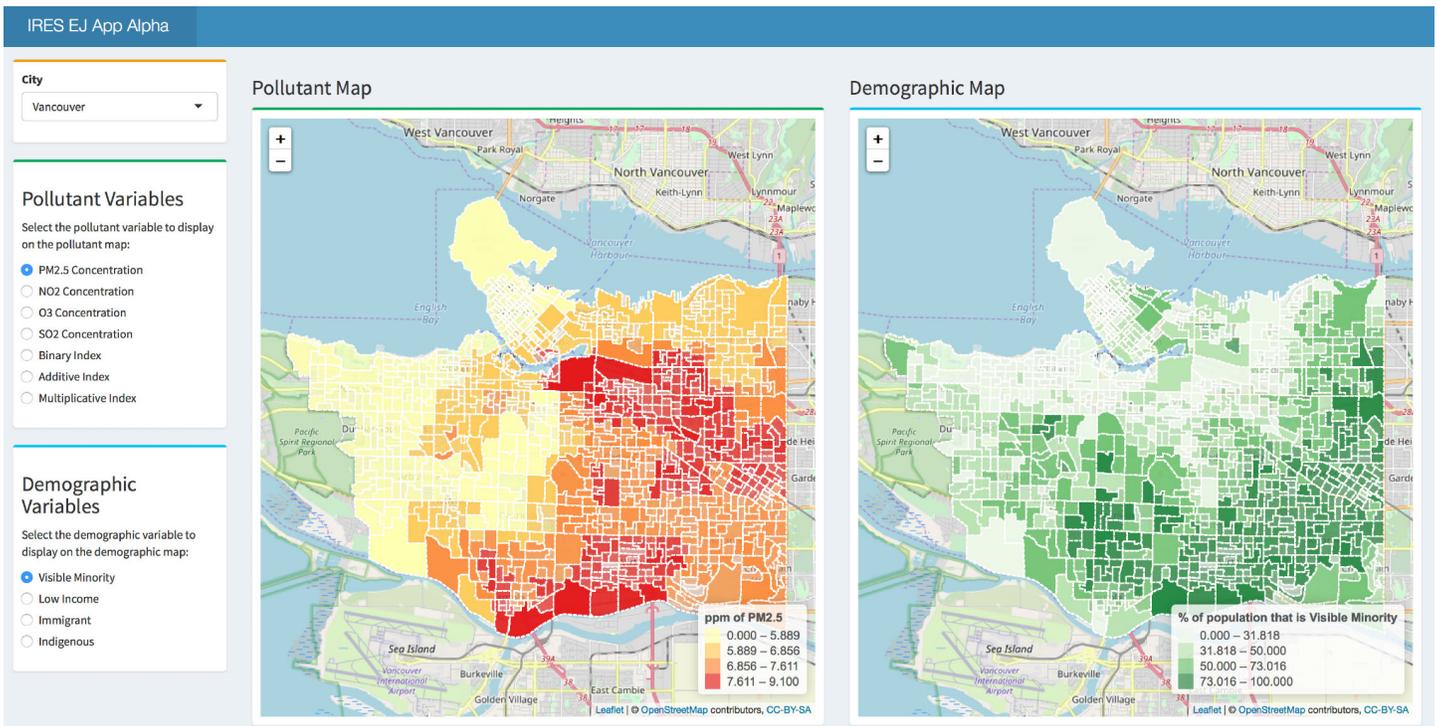


Figure 5. Neighborhood-scale data for pollutants and demography in Greater Vancouver, BC

“ Making air quality, demographic, and health data more accessible at a neighborhood scale can prioritize interventions and support community-level action. ”  
-Amanda Giang

## RESPONSIBLE DATA SHARING: BIAS, EQUITY, PRIVACY

Data science collaborations encompassing the Cascadia region face both legal and technical interoperability challenges around data sharing, storage and privacy. CUAC has engaged in multiple concurrent efforts to facilitate data comparisons and analysis across sectors, regions and the U.S.-Canadian border. Initially focusing on the housing, transportation and health sectors, the strategies and methods developed are applicable to a wide range of topics. These investments inform a broader CUAC effort to create a platform for data integration, sharing and governance to enable multi-sector data science collaborations that benefit the region.

With support from NSF, Microsoft, and the Bill and Melinda Gates Foundation, CUAC is developing Fides, a new data management platform coupled with an institutional governance framework for working with civic data in public-private partnerships to balance competing objectives: improving fairness, preserving privacy, protecting firms' proprietary information, and promoting transparency. The key idea is to repair the raw data to preserve individual privacy and selectively remove sources of bias. For example, unprocessed mobility data may have a geographic prejudice towards wealthy neighborhoods; transportation systems designed based on these data sets will propagate this tendency. Using techniques developed through CUAC, we can suppress the unwanted bias while retaining other patterns in the data. The resulting synthetic datasets retain the structural and statistical properties of the source data, but can be shared broadly without violating privacy, reinforcing discrimination, or destroying competitive advantage. CUAC researchers have applied these approaches to data on mobility, education and housing services, demonstrating their generality. These techniques are foundational to sharing data regionally and internationally to enable pilot collaborations *before* legal relationships are established---a key enabling technology for accelerating innovation in the Cascadia region. **CUAC researchers have deployed, and now sustainably operate cross-sector data-sharing ecosystems with strong data governance, a platform that enables data collaboratives. This technology combines new algorithms to control bias and protect privacy with advanced cloud services hosted on Azure (Figure 2).**



Mark Hallenbeck, Director of the Washington State Transportation Center, presents at UBC.

## CASCADIA LAW INITIATIVE (CLI)

A central component of the CUAC collaboration has been the [Cascadia Law Initiative](#), launched by law professors Margot Young (UBC) and Hugh Spitzer (UW), who worked with a team of four UBC and UW law student researchers to create a comparative study of the legal powers of municipal governments in Vancouver, Seattle, and their major suburbs like Surrey and Bellevue. To support cross-border collaboration on multi-city urban analytics projects, CLI presents a legal framework relevant to issues such as housing affordability, homelessness, transportation access and the opioid crisis, incorporating considerations of data privacy and data sharing. The project identifies potential legal barriers and tools available for city initiatives arising from urban analytics research collaboration. As a case study, the team examined how governmental and transportation agencies and their policies might influence a future autonomous vehicle trip from Vancouver to Bellevue. **The Cascadia Law Initiative explains legal differences associated with urban issues in B.C. and Washington and outlines how these variations affect government policies. This initiative can support CUAC projects across many domains.**

## LAUNCHING DATA SCIENCE FOR SOCIAL GOOD (DSSG) AS AN INDEPENDENT PROGRAM AT UBC

UBC's successful three-year experiment with the Data Science for Social Good (DSSG) program has generated considerable interest beyond the university. DSSG Director Raymond Ng and Vice President for Research and Innovation (VPRI) Gail Murphy leveraged the first two years of CUAC support to help establish the UBC DSSG program to operate independently with industry and government resources. CUAC co-director Jonathan Fink gave a presentation about UBC's DSSG at the "Data for Good" Conference in Ottawa in January 2019; representatives of Statistics Canada and other Canadian federal agencies were highly receptive to the idea of replicating DSSG across the country, something that looks promising for 2020 and beyond. DSSG was the first UBC-UW collaborative program in CUAC. Student participants from the two universities have gained considerable value from getting together with their counterparts.

Moving forward, the growth, scalability and possible independence of DSSG can become a model for how seed-funded collaborative programs like CUAC can generate their own support. The DSSG student presentations at the 2017 and 2018 Cascadia Innovation Corridor conferences were well-received as a positive demonstration of this potential. The faculty-based research collaborations that have developed in the first two years can similarly form building blocks for this emerging regional identity.

## EXAMPLES OF URBAN-RELATED DSSG PROJECTS

The majority of UBC and UW DSSG projects in 2017 and 2018 involved urban analytics. Here we describe a few of these. Table 1 lists all 15 projects at both universities.

### IMPROVING TRANSIT SERVICES USING ORCA DATA

[“Improving Transit Services Using ORCA Data”](#) (Figure 6), a 2017 DSSG project led by the Washington State Transportation Center, examined more than 40 million transit boarding records collected through electronic fare card usage to understand changes in transportation patterns following the opening of two new light rail stations. The students utilized machine learning to more accurately identify direct transfers versus “financial transfers,” which may include round-trip errands on a single fare rather than an extended trip; and improve estimations of the number of cash-paying transit customers who are missed by electronic fare card data collection. **The project generated an improved determination of travel patterns and a more accurate estimation of total transit ridership based on electronic fare card data and informed the development and research of the Transportation Data Collaborative.**

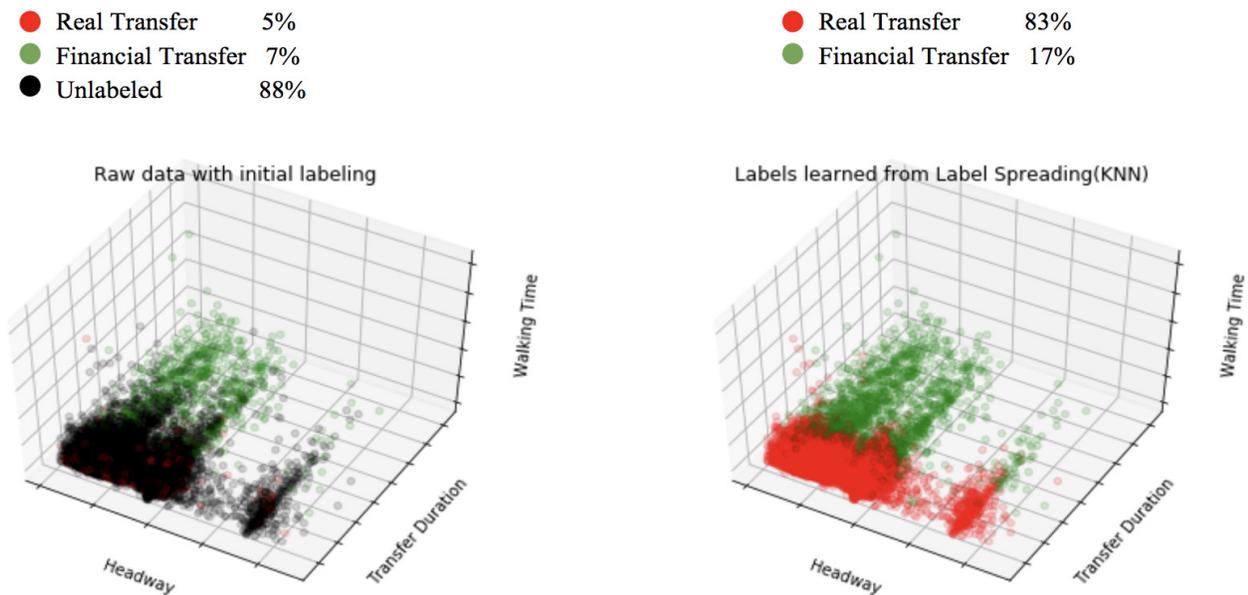


Figure 6. DSSG predictions of travel patterns based on fare card data.

## CAN TRAFFIC SENSOR DATA DETECT VEHICLE CRUISING?

A 2017 project led by the Seattle Department of Transportation (SDOT) called “[Can traffic sensor data detect vehicle cruising?](#)” examined the traffic congestion impacts of for-hire and transportation network services waiting for customers, and drivers who are searching for parking, in downtown Seattle. The team developed algorithms to quantify aggregated levels of vehicle traffic “cruising” using anonymous travel sensor data, paid parking transaction information, and parking occupancy surveys conducted by the City of Seattle and generated heat maps showing their results. **The team determined that more than one-third of all drivers in downtown Seattle are rideshare drivers waiting for passengers or drivers circling for parking.**

## SEATTLE MOBILITY INDEX PROJECT

The 2018 “[Seattle Mobility Index Project](#)” examined the affordability and reliability of travel to everyday destinations across Seattle through a partnership with the Seattle Department of Transportation. The team created a baseline measure for comparative analysis of transportation from the city’s 481 census block groups by travel mode (driving, public transit, biking and walking) to local destinations such as schools, hospitals, parks and grocery stores, and citywide destinations like employment centers, public colleges and landmarks. As part of the UW 2018 DSSG program, the team gathered trip data from the Google Distance Matrix API, which they calibrated using the Puget Sound Regional Council’s Household Travel survey containing travel and demographic information for 30,000 trips. They identified travel personas to understand mode choices associated with characteristics such as income and car ownership. **The project shed light on transportation disparities by block group across Seattle; and the team developed tools to predict with 77% accuracy whether a specific trip will be completed via driving versus other modes of travel.**

## UNCOVERING THE HIDDEN UNIVERSE OF RENTAL UNITS IN SURREY

The 2018 UBC-based project “Uncovering the hidden universe of rental units in Surrey” addressed the challenges associated with the growing population of informal renters, who couch-surf in basements and attics, rent single rooms or occupy lofts. These arrangements, which keep sizable populations off city governments’ radar, make it difficult to accurately provide adequate municipal services. **Using text from Craigslist and Kijiji advertisements, utility bills and parking tickets, the students derived an AI-based predictor of rental unit occupancy and type that the city checked using targeted surveys.**

## TRANSPORTATION ENERGY & EMISSIONS BASELINE AND FORECASTING FOR ONGOING MODELLING AND POLICY ANALYSIS

A second DSSG project, “Transportation Energy & Emissions Baseline and Forecasting for Ongoing Modelling and Policy Analysis,” established a greenhouse gas emissions baseline and forecast models for Surrey to 2050. Their model combined vehicle registration, demographic, transit and property use data. The resulting forecasting tool is helping Surrey design zero- and low-emission transportation infrastructure and strategies. This project returned in Year 3 of UBC’s DSSG program as well. **The student team developed an app that is being deployed for use internally by Surrey as a data visualization and exploration tool that integrates census data with city-collected data (e.g., transit, business licenses, buildings).**

## USING CAMPUSES IN CASCADIA TO TEST SMART CITY INNOVATION

Based on UBC’s successful use of its campus facilities and operations to test a variety of sustainability and smart city applications, PSU launched the Digital City Testbed Center (DCTC), which is creating a network of academic, corporate and nonprofit campuses in Portland, Seattle and Vancouver where these innovations can be more carefully evaluated before being deployed in cities at large. This network, supported by a 3-year, \$1.5M investment from PSU, provides UW, UBC, and PSU a platform through which they and their municipal and corporate partners can collaboratively test new data protocols, application interoperability, data sovereignty, and public education. Initial testbeds include the campuses of PSU, UBC and the Oregon Museum of Science and Industry (OMSI); future sites will likely include UW and Microsoft campuses and Portland International Airport. DCTC’s Advisory Board includes representatives from UBC, Microsoft, Bullitt Foundation, MetroLab Network, NIST, Intel, City of Portland, Technology Association of Oregon, and OMSI.



*Faculty and DSSG fellows participate in a project brainstorming session at UBC in July 2017 (left-right: Nathanael Lauster, Mina Park and Leanne Currie).*

## INTERDISCIPLINARY AND REGIONAL CONVENINGS

CUAC has sponsored a series of research presentations and meetings with participants from academic, industry, government and nonprofit sectors in Metro Seattle and Vancouver. At a two-day convening In July 2017, DSSG fellows, faculty and researchers from UW and UBC gathered at UBC to share [mid-term project presentations, provide lightning talks and brainstorm future collaborative projects](#). DSSG projects centered around transport, access to community resources, and economic development. **DSSG fellows from UW and UBC shared methods and domain knowledge, compared urban issues in Seattle and Vancouver, and learned from faculty about the design of urban data science projects.**

A two-day [CUAC Fall Symposium](#) was held at UW in September 2017 to identify collaborative research projects to pursue in the coming year in the areas of housing affordability, population health, responsible data management, neighborhood change and multi-modal transportation planning. The event drew 35 faculty, students and researchers from UW and UBC in the fields of landscape architecture, data science, urban analytics, neuroscience, community and regional planning, sociology, nursing and health care, urban design and planning, social work, statistics, computer science and engineering, real estate, population and public health, and civil engineering. Representatives from the cities of Surrey, Vancouver, Portland and Seattle, along with the World Council on City Data at the University of Toronto, joined the conversation. A subsequent CUAC Community Meeting was held at UW in winter 2017 to present ongoing projects in the Cascadia region. Presentation topics included responsible data, housing and homelessness, accessible routing, transportation data collection, and metrics for understanding equity and change across neighborhoods. **Participants included UW and UBC faculty and researchers, along with local stakeholders and city government representatives. Projects on neighborhood housing trends, responsible data management and transportation data sharing were advanced following these collaborative, interdisciplinary discussions.**

In July 2018, UW faculty and researchers, City of Seattle department representatives, UBC participants and local organizations joined a meeting at UW to discuss opportunities for collaboration between UW and the City of Seattle. Contributors discussed ideas for matching under-utilized City data sets with UW data and research scientists to create socially-relevant projects, including initiatives to improve city services and operations, using models such as the UBC-Surrey partnership. **Future discussions will be aimed at developing a streamlined process for scoping and coordination between project needs and data science expertise to support data-driven policy initiatives with stakeholder involvement.**

---

CUAC co-director Bill Howe led a meeting at Data & Society Research Institute to articulate a new foundational research agenda in responsible data science --- systems and methods that explicitly take into account the ethical and epistemic issues associated with the larger social and legal contexts in which they are deployed.

---

With support from the National Science Foundation (NSF), CUAC co-director Bill Howe led a meeting at [Data & Society Research Institute](#) in October 2018 to articulate a new foundational research agenda in responsible data science --- systems and methods that explicitly take into account the ethical and epistemic issues associated with the larger social and legal contexts in which they are deployed. Topics include fairness, accountability, transparency, privacy and reproducibility; the goal was to set a research agenda in sociotechnical systems that reflects shared societal values. **This meeting, led by CUAC researchers, connected emerging research in the Cascadia Region with the national community in responsible data science.**

Multiple presentations were made in 2019. In January, Jonathan Fink made a presentation about the DSSG program at Statistics Canada’s “Data for Good” conference in Ottawa. He also led two Cascadia sessions and a panel discussion at the National Institute of Standards and Technology’s Global City Teams Challenge in Portland in September, a session about accessibility tech at the MetroLab Network’s Annual Meeting in Boulder, and a wildfire and urban smoke session at the Fourth Cascadia Innovation Corridor conference in Seattle in October. In May, the first Northwest Data Science Summit was hosted by the UW eScience Institute and the UBC DSI as a direct result of collaborations through CUAC. The event, held at UW, had more than 200 participants. The second annual event is being organized for May 2020.

University of Washington			University of British Columbia		
Year	Project	Partner	Year	Project	Partner
2018	Automatic Damage Annotation on Post-Hurricane Satellite Imagery	UW Industrial & Systems Engineering	2018	Improving Early and Middle Childhood Outcomes	City of Surrey
2018	Seattle Mobility Index Project	Seattle Department of Transportation	2018	Transportation Energy and Emissions Baseline and Forecast for Ongoing Modelling and Policy Analysis	City of Surrey
2018	Access to Out-of-School Opportunities and Student Outcomes	Center on Reinventing Public Education	2018	Uncovering the Hidden Universe of Rental Units in Surrey	City of Surrey
2017	Improving Transit Services Using ORCA Data	Washington State Transportation Center	2018	Use of Machine Learning Techniques to Classify Laboratory Test Results	BC Centre for Disease Control
2017	Strengthening Capacities, Knowledge and Data Sharing Platforms for Sustainable Development	Vital Signs	2017	BC Tourism Resources Project	BC Tourism Authority
2017	Can Traffic Sensor Data Detect Vehicle Cruising?	Seattle Department of Transportation	2017	Avenues of Change - Early Child Education Project	Guildford West
2017	The "Equity Modeler": Examining Just Development in Seattle	Department of Urban Design and Planning	2017	Investment and Intergovernmental Project	City of Surrey
			2017	Surrey Transportation Project	City of Surrey

Table 1. 2017 and 2018 Data Science for Social Good programs at UW and UBC