Course: **Statistics and Modeling with Novel Data Streams**
Instructors: **Alex Vespignani and Mauricio Santillana**
When: **July 18,19,20**
Where: **University of Washington, Seattle.**

Software Requirements: **Python, R and Matlab**

Total of hours in the course: **15 hours**

General plan

++++++++++++++++++

Monday

8:30am-10am **General introduction** (45 mins Alex and 45 Mauricio)

- Prediction
- What we need (input)
- Big data streams: Passive (Google, Twitter); Active (RFID TAGS: contact network patterns, Flu Near You)
- Models

10:30am - 12pm **Models introduction** (Alex)

- Generative models
- Compartmental disease representation
- Computational implementation of binomial Markov chain models
- Stochastic models

LUNCH

1:30pm - 3pm **Multivariable linear models and Google searches**. **A tutorial** (Mauricio)

- Coding a simple version of Google Dengue Trends
- Static vs dynamic
- Multivariable models (variable aggregation vs individual contribution)
- Adding auto-regressive information: ARGO
- Technicalities: CDC (Data acquisition, revisions), Google (Sampling issues)
- Google Correlate

3:30pm - 5pm **Homogeneous model —> Network, RFID experiments** (Alex)

- Individual based & Networks models
- Network data
- RFID experiments
- Large-scale network models

++++++++++++++++++

Tuesday

8:30am-10am **Data streams introduction: beyond Google searches, and beyond Flu** (Mauricio) Healthmap Flu Trends and Dengue Trends

- Real-time prediction. Development of a website to scrape/centralize information
- Beyond Google searches (Flu).
- UpToDate and Electronic Health Records as proxies for disease incidence
- Participatory surveillance as a way to track diseases.
- “Together we are stronger”: ensemble approaches lead to more robust systems.
- Beyond Flu. Dengue and Zika prediction using Google searches
- News reports (success stories with Ebola and Zika)

10:30am - 12pm **Human mobility Data and Modeling**

- Human Mobility (mobile call records, airline data etc.)
- Spatially structured models
- Metapopulation networks
- Data-driven large scale simulations

LUNCH

1:30pm - 3pm **Flu Near You and participatory surveillance and Epicore** (Mauricio)

- What is Flu Near You?
- Demographics of FNY.
- Consistent users vs sporadic users.
- How many reports do we need to see a signal?

3:30pm - 5pm **Global Epidemic and Mobility Platform**

++++++++++++++++++

Wednesday

8:30am-10am **Hands on** (Mauricio) Healthmap Flu Trends

- Coding ARGO a near-time and forecasting model

10:30am - 12pm **Data-Driven Simulations and Forecast** (Alex)

- Near-time and Real-time forecast
- Seasonal Flu
- Emerging infectious diseases
- More than forecast