Data Science for Social Good

Sarah Stone - co-Executive Director, eScience Institute
website - http://escience.washington.edu/dssg/
@uwescience #DSSG2017
Goal: This program brings together students and researchers with data science and domain expertise to work on focused, collaborative projects for societal benefit.

Modeled after similar programs at the University of Chicago and Georgia Tech with elements from our own Data Science Incubator.

Through the Cascadia Urban Analytics Cooperative (CUAC) we worked with the University of British Columbia to set up their pilot DSSG program in 2017.
Project Teams

4 projects supported each summer

Each team consists of:

Project Leads (1-2)
eScience Data Scientist Leads (1-2)
DSSG Student Fellows (4)
Data Science Ethnographers

• study the culture & practice of data science
• provide programmatic insight
• stakeholder identification/collaboration
• data science ethics

Brittany Fiore-Gartland
eScience Director of Data Science Ethnography

Anissa Tanweer
Graduate Student, Communications
Project Proposals – call is NOW OPEN!

We invite short proposals for 10-week data-intensive research projects that have an applied social good dimension. => take advantage of drop-in office hours to hone/discuss project ideas

Proposals may be submitted by academic researchers, public agencies, non-profit entities, and industry.
Project Proposals - what we were looking for:

• strong science, strong methods

• availability, commitment – **Project Proposer** => **Project Lead**

• clarity and shovel-readiness

• capacity for measurable outcomes

• sustained engagement
Project criteria

• strong science, strong methods
• new directions, new questions
• availability, engagement, commitment
• “do only what we can only do together”
  – with apologies to Djikstra
• alignment with sponsor and program goals
• clarity and shovel-readiness
• capacity for measurable outcomes and sustained engagement
• properly scoped, such that significant progress can be made in the summer time frame
Domain Areas of Interest

• Poverty, equality, income
• Crime, justice, legal
• Housing
• Public Education
• City planning
• Transportation
• Hazards / Resilience
• Utilities
• Economics
Technical Areas of Interest

• new platforms, new algorithms, new methods, new datasets
• working with large, heterogeneous, and noisy datasets
• scalable analytics and predictive models
• interactive visualization
• code review, publishing, and reproducibility
• online teaching materials, tutorials
Project Lead (PL) Expectations

• PL submits the proposal
• Together with the Data Scientist Leads, the PL manages the student team setting expectations, goals, timelines
• PL bears primary responsibility for project design and execution
• **Project Leads commit to being in Studio at least 16 hours/week**
  - Project leads and data scientists should coordinate schedules
• Arrange a fieldtrip for the project team to appropriate stakeholder
Project Teams

4 projects supported each summer

Each team consists of:

Project Leads (1-2)
eScience Data Scientist Leads (1 primary, 1 secondary)
DSSG Student Fellows (4)
Stakeholders
Student Fellow Selection Criteria

40 hours per week for 10 weeks, $6500 stipend

- 16 DSSG Fellows selected
- Current student, advanced undergrad or graduate
  - Spring graduates okay
- Educational experience, GPA
- Technical/Analytical experience
- Demonstrated interest in “social good” through research projects, paid and/or volunteer experience
- Experience with collaborative projects
Program Activities

Tutorials - front-loaded, then weekly
Weekly “stand up” meetings, happy hours
Stakeholder fieldtrips

Housed on UW campus in the WRF Data Science Studio
Project expectations at end of summer

• Online data product delivered
  – an online visualization, a public library on github
• Findings summarized in a brief online report
• All code hosted on github
• Acknowledgment on papers resulting from the project
• Findings presented to the eScience and broader community via Final Presentations
2018 DSSG Timeline

January 8\textsuperscript{th} – Student Fellow applications open NOW
February 12\textsuperscript{th} – Student Fellow applications close

\begin{itemize}
  \item January 15\textsuperscript{th} – Call for proposals open NOW
  \item January 26\textsuperscript{th} – Proposal info session
  \item February 25\textsuperscript{th} – Proposals due online by EOD
  \begin{itemize}
    \item => light weight \sim 2 page proposal
  \end{itemize}
  \item March 19\textsuperscript{th} – Project short-list notification
\end{itemize}

June 11\textsuperscript{th} – Program kickoff
- 1\textsuperscript{st} & 2\textsuperscript{nd} weeks the focus is on tutorials, skill building, and project introduction
August 17\textsuperscript{th} – Final Day