SIGMOD 2018
New Researcher Symposium

Advice from Early PhD to Early Career

Juliana Freire
New York University



Reviewing Papers

How to be a constructive reviewer that helps advance the field, while still gatekeeping for quality?

- Our review process needs to be improved and you can help!
- DB community addresses a broad set of problems from different: theory, algorithms, systems, applications-- there is no one-size-fits-all paper template
- Embrace diversity
- Recognize and reward novelty-- new ideas are hard to publish
- Simplicity is good, if solution is effective
  - Combining existing ideas to solve a new problem is a valuable contribution
- Do not nitpick -- papers need not (and often cannot) be perfect

Pitfalls to in Writing Research Proposals

- You can learn how to write a good proposal, takes practice like writing papers
- Volunteer to participate in review panels
- Meet program managers
- Ask senior colleagues for feedback (Michael Benedict reviewed my NSF Career proposal)
- Easier to write proposals for projects you have already started -- or completed (Serge Abiteboul)
- But even good proposals get rejected
  - Submit faster than they can reject! (I.V. Ramakrishnan)
  - Target different funding sources

Time Management

- I need advice on this!

How to make your research impactful?

- Choose the *right* problems
- You need to be passionate about what you are doing
  - There will be many failures
  - There is no instant success -- https://www.npr.org/podcasts/510313/how-i-built-this
- Advertise you work (Ben Shneiderman)
- Success has different flavors
  - Many papers
  - Many users
- The VisTrails project
  - Papers rejected from DB conferences, including CIDR; many proposals declined initially
  - But then many papers accepted -- including a best paper award at IEEE Vis
  - Many grants, and users

How to make your research impactful?

- Choose the *right* problems
- You need to be passionate about what you are doing
  - There will be many failures
  - There is no instant success -- https://www.npr.org/podcasts/510313/how-i-built-this
- Success has different flavors
  - Many papers
  - Many users
- The VisTrails project
  - Papers rejected from DB conferences, including CIDR; many proposals declined initially
  - But then many papers accepted -- including a best paper award at IEEE Vis
  - Many grants, and users

How to make your research impactful?

- Choose the *right* problems
- You need to be passionate about what you are doing
  - There will be many failures
  - There is no instant success -- https://www.npr.org/podcasts/510313/how-i-built-this
- Success has different flavors
  - Many papers
  - Many users
- The VisTrails project
  - Papers rejected from DB conferences, including CIDR; many proposals declined initially
  - But then many papers accepted -- including a best paper award at IEEE Vis
  - Many grants, and users

Opportunity for “Data” People

- Data is central to everything
- We have the luxury to select from a broad range of problems that can have practical impact
- But real problems are complex – they require techniques from different areas and collaboration with domain experts
  - High risk and high payoff
  - You have the opportunity to learn new things
  - You make lots of cool friends!

Opportunity for “Data” People

DEFENSE ADVANCED

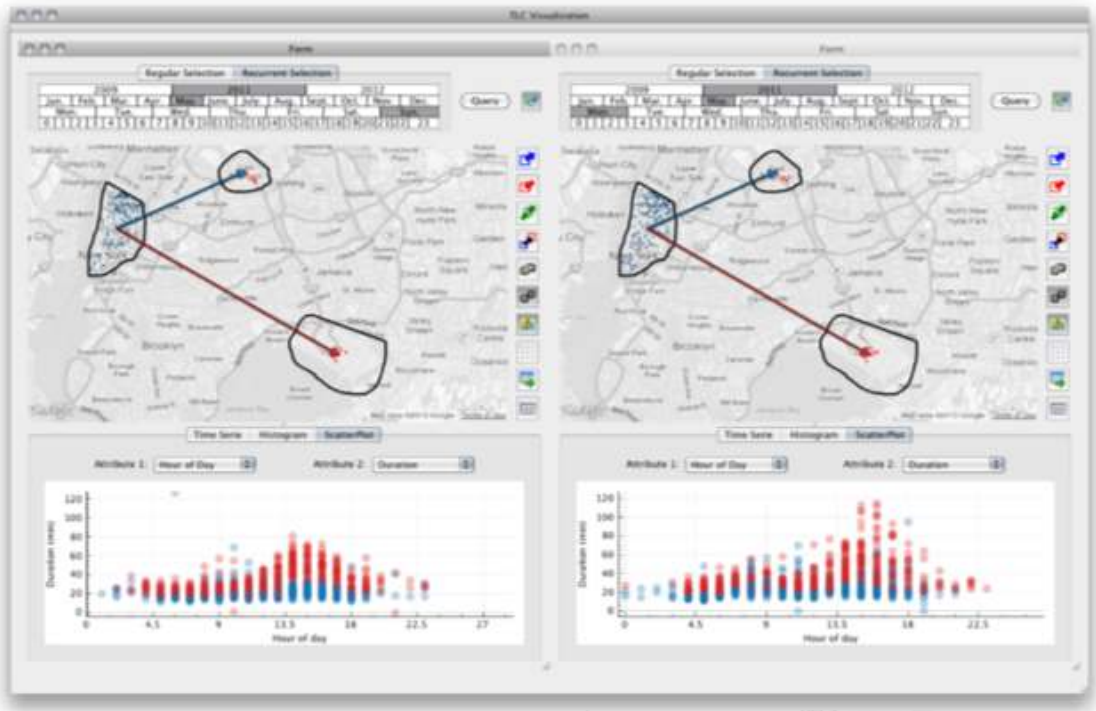
EXPLORE BY TAG

TaxiVis

Overview Video Paper Code Team Sponsors

Visual Exploration of Big Spatio-Temporal Urban Data: A Study of New York City Taxi Trips

As increasing volumes of urban data are captured and become available, new opportunities arise for data-driven analysis that can lead to improvements in the lives of citizens through evidence-based decision making and policies. In this project, we focus on a particularly important urban data set: taxi trips. Taxis are valuable sensors and information associated with taxi trips can provide unprecedented insight into many different aspects of city life, from economic activity and human behavior to mobility patterns. But analyzing these data presents many challenges. The data are complex, containing geographical and temporal components in addition to multiple variables associated with each trip. Consequently, it is hard to specify exploratory queries and to perform comparative analyses (e.g., compare different regions over time). This problem is compounded due to the size of the data---there are on average 500,000 taxi trips each day in NYC. We propose a new model



same operating system!

These are just my opinions,
I could be wrong.

Thanks!