SAVING THE WORLD WITH DATA

THE CASE FOR CIVIC DATA SCIENCE

Andrew Therriault
Chief Data Officer, City of Boston

City of Boston
Mayor Martin J. Walsh

Innovation & Technology

A FEW PRELIMINARIES...

WHAT I DO NOW

Chief Data Officer for the City of Boston. I lead Boston's Citywide Analytics Team, which compiles and uses data for projects with nearly every department & agency in the city.

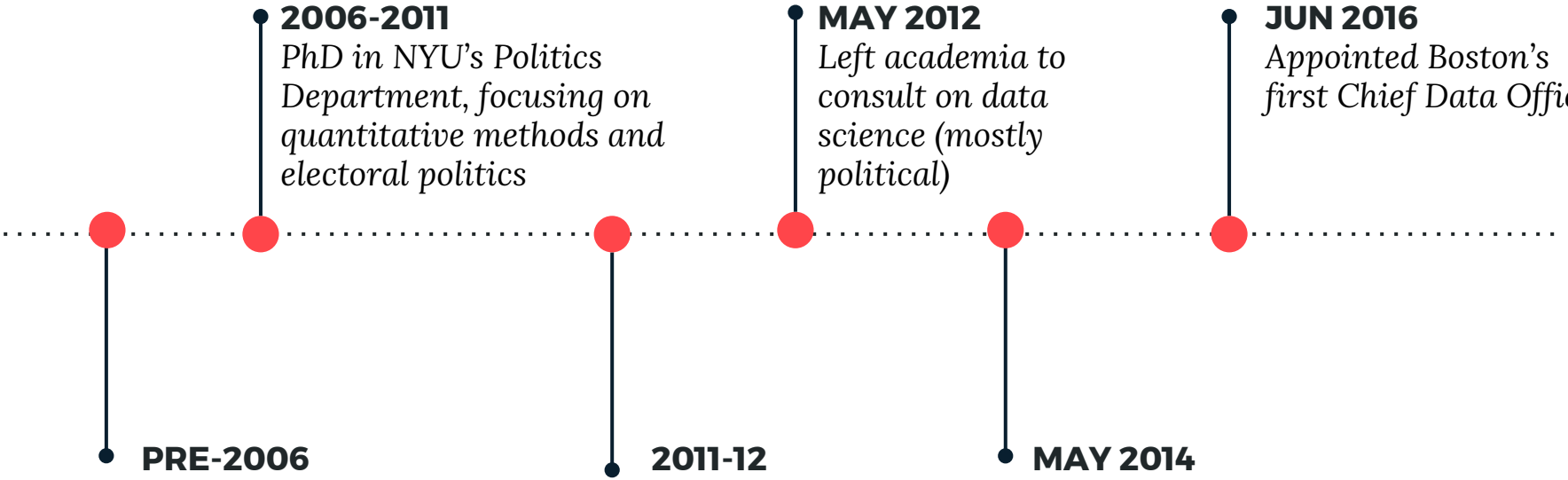
WHAT I DID BEFORE

Poli-sci PhD, then 5 years of applied research and consulting for political campaigns, issue advocacy groups, and nonprofits. In 2014, founded the Democratic National Committee's Data Science team and led until June 2016.

WHAT YOU CAN EXPECT

A high-level discussion of how data science can help civic organizations. We'll look at why data science is a natural fit for these applications, see some real-world examples, and talk about how to get involved. (Zero code or prerequisites.)

BEFORE THAT: MY JOURNEY INTO DATA SCIENCE



AN INTRO TO CIVIC DATA SCIENCE

What it is, what it's good for, and why we should care

Civic data science: the application of data science methods* to support the work of governments, nonprofits, advocacy groups, political campaigns, and other organizations which exist to deliver positive social impacts

** broadly defined—everything from fundamental data management, analysis, and reporting to predictive models, experimental tests, and complex visualizations*

THE CASE FOR CIVIC DATA SCIENCE

CIVIC ORGS MATTER

The tech industry talks about making the world better (and sometimes succeeds), but their ultimate motives are still financial. Civic orgs use money primarily as a tool to promote social goals, and their direct impacts on public well-being are orders of magnitude larger.

BIG POTENTIAL IMPACTS

Two core principles of applied data science: efficiency and effectiveness. With resources almost always scarce for civic organizations, this kind of optimization can be the difference between success and failure. And even if not, marginal improvements are still very meaningful.

EXCITING CHALLENGES

Beyond just being meaningful, civic data science projects are also some of the most exciting you can work on. Subjects are varied and unique, people are smart and highly motivated, and opportunities for creativity and innovation are unparalleled.

REAL-WORLD CASE STUDIES

Five examples of civic data science projects from my own work

PREVENTING FOODBORNE ILLNESS AND OTHER RISKS

Since July 2016, the City of Boston has used predictive models of critical health code violations to prioritize restaurant inspections.

Using inspection histories, citizen complaints, property records, and more, these models have helped to improve inspection efficiency by more than 20%, so problems are addressed earlier and more often.

A similar approach can also be applied to other risks such as students dropping out and individuals becoming homeless, where early intervention is key.

TARGETING FOR HEALTHCARE ENROLLMENT CAMPAIGNS

IDENTIFYING NEED AND ELIGIBILITY

Data-driven targeting is crucial to making outreach programs practical

COORDINATING OUTREACH

These campaigns are often massive, so skillful data management is essential

TRACKING PERFORMANCE

Monitoring in real time is key to spotting problems and developing best practices

LEARNING OVER TIME

The most effective campaigns use every new result to refine their strategy and tactics

IMPROVING SUBSTANCE ABUSE AND MENTAL HEALTH TREATMENT

We're working to expand Boston's "311 for Recovery Services" system to coordinate care and keep patients on track by integrating data and making it accessible to providers.

We've also piloted a system to help identify patients in our public health systems who are likely to overdose, in order to provide proactive interventions before that happens.

If successful, these program could be expanded to include a broader range of mental health issues, in order to better connect those in need with specialists who can help.

OPTIMIZING NON-PROFIT FUNDRAISING

IDENTIFYING NEW PROSPECTS

Integrating data from a variety of sources gives organizations a more complete picture of past donors and helps find new prospects

LEARNING WHO TO TARGET

Traditional fundraising often relies on "look-alike" models of existing donors, but a much more effective approach is to treat every ask as a new data point

TESTING THE BEST APPEALS

Whether it's via direct mail, phone, digital ads, or email, non-profit fundraising benefits from the same kind of experimental tests have become common in advertising

WHAT TO ASK FOR AND WHEN

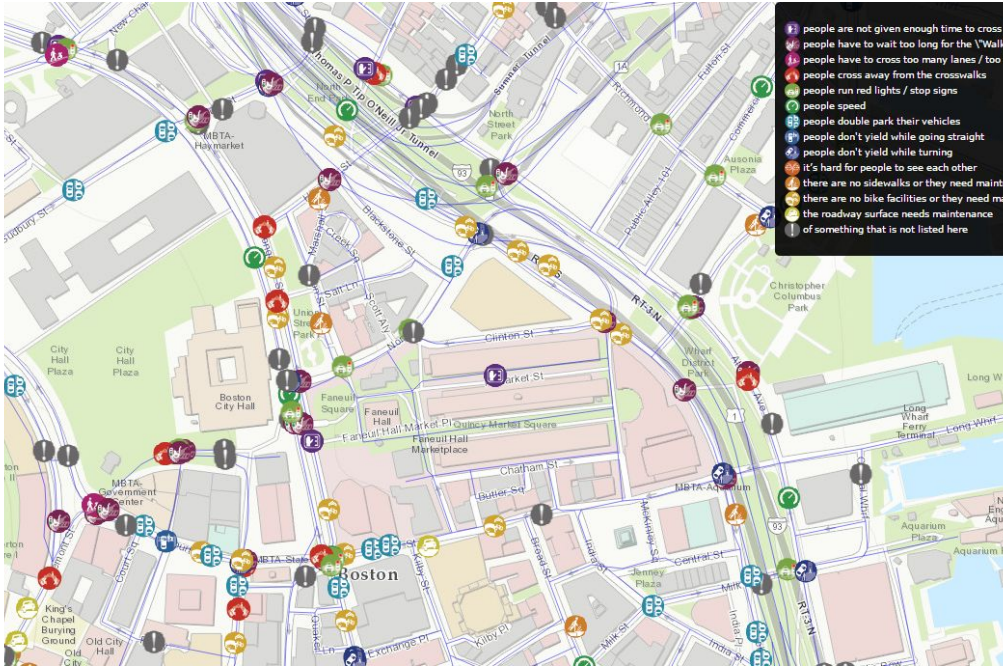
The biggest questions in fundraising are about who to ask for money and how to ask them, but the finer details—how much to ask for, how often, etc.—are key to long-term success

PREDICTING CRASHES ON OUR STREETS

We're gathering data to build crash prediction models for Boston's streets to address structural problems and prevent tragedies.

These models will combine data on traffic patterns and hazards from a wide variety of sources, including traffic sensors, GPS data, historical crash records, traffic incident reports, and residents' reports of safety concerns across the city.

We would like to make this project a collaborative effort with community volunteers, leading to an open source framework that could be adopted by other cities as well.



HOW YOU CAN GET INVOLVED

(Yes, you!)

SOME SUGGESTIONS...

LEARN

Start by learning about the areas you're interested in:

Explore the data landscape

Hear about what's being done with data already

Get to know the relevant experts in your community

VOLUNTEER

Find an org that needs your help or can connect you.

Join a group (Data for Democracy, DataKind, CfA)

Contribute to open source or build with open data

Offer your help to civic orgs directly (but don't be surprised if they say no!)

FIND A JOB

You'll make the biggest impact if you can commit to an organization full-time.

Find orgs that interest you and make connections

Do fellowships with civic orgs or groups like DSSG

Work for respected vendors or consultants for civic orgs

My final pitch: Whatever your long-term career plans, spending a few years in civic data science will make you a much better data scientist.

You'll work on more varied projects, learn more new skills, advance more quickly, have more autonomy, and get more opportunities to share your work with the world.

Even if your main reason for doing data science is to make money, the experience will almost surely be more valuable than the modestly higher salary you'd otherwise get now.

For more information about the City of Boston's Analytics Team, go to boston.gov/analytics.

(NOTE: Summer Fellowship applications now open!)

And if you want to talk more:
andrew.therriault@boston.gov (work)
andrew.therriault@gmail.com (personal)
@therriaultphd (twitter)

THANK YOU!