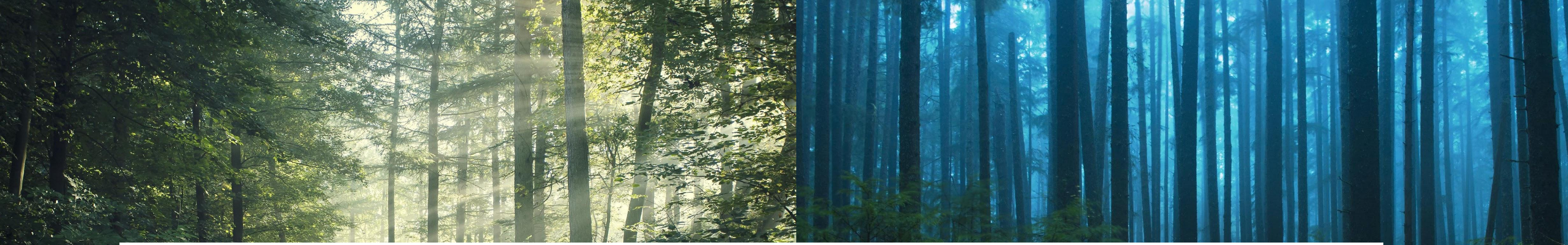


Auditing and Designing for Equity in Government Service Allocation

Nikhil Garg, Cornell Tech (ngarg@cornell.edu)

Special thanks to Urban Tech Hub at Cornell Tech & NYC Department of Parks and Recreation

Government service allocation

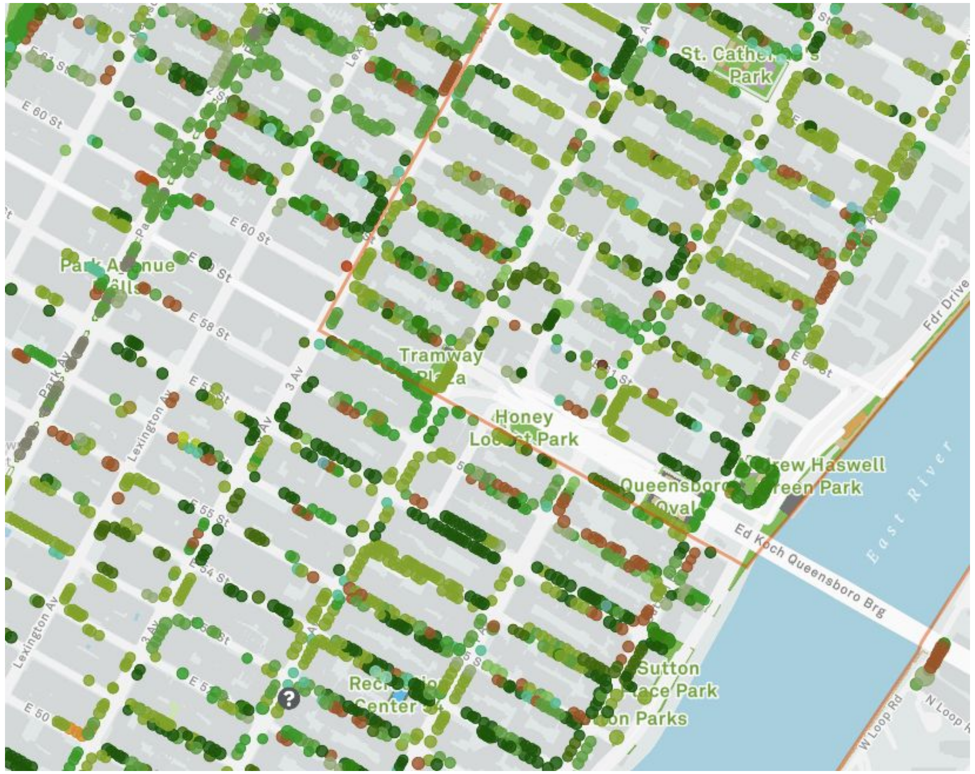
Local government manages many services

- ~8k miles of streets in NYC
- ~700k trees lining streets in NYC
- Housing, sanitation, transportation, etc.

Operational tasks

- [Learning] What problems are there?
- [Allocation] Which ones to address?
- [Auditing] Did we do a good job?

Desiderata: Efficiency & Equity



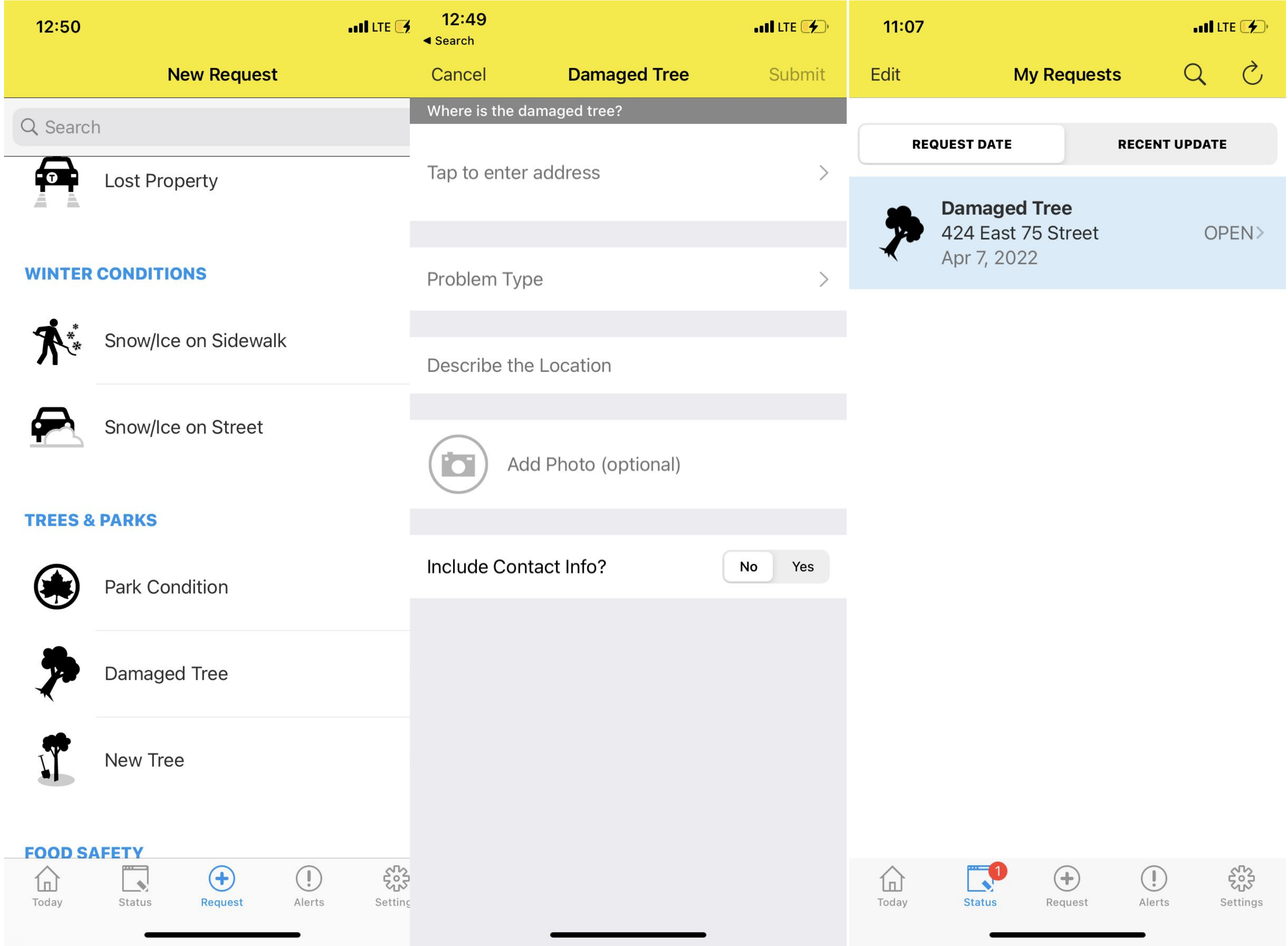
Street trees on Upper East Side in NYC

311 (crowdsourcing) systems

Cities have a phone number & app to complain to the local government

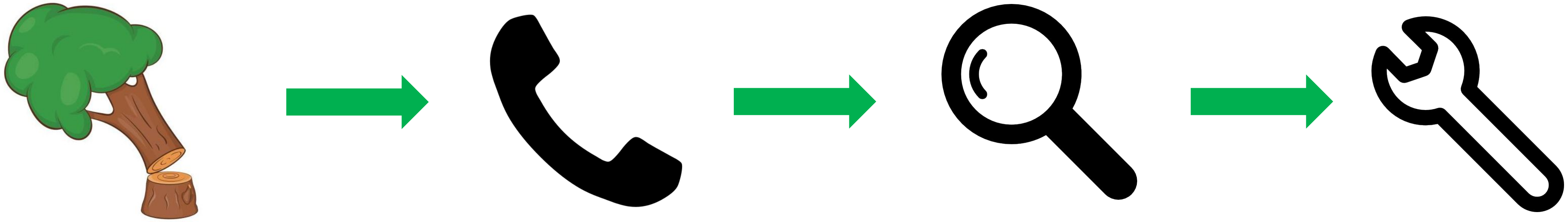
NYC's 311 system received about 2.7 million requests 2021

These are the primary way the government learns about problems



Pipeline: from incident to work orders

Incident | **311 report** | **Inspection** | **Work order**



70-100k/year to forestry unit of NYC DPR

~60% of reports

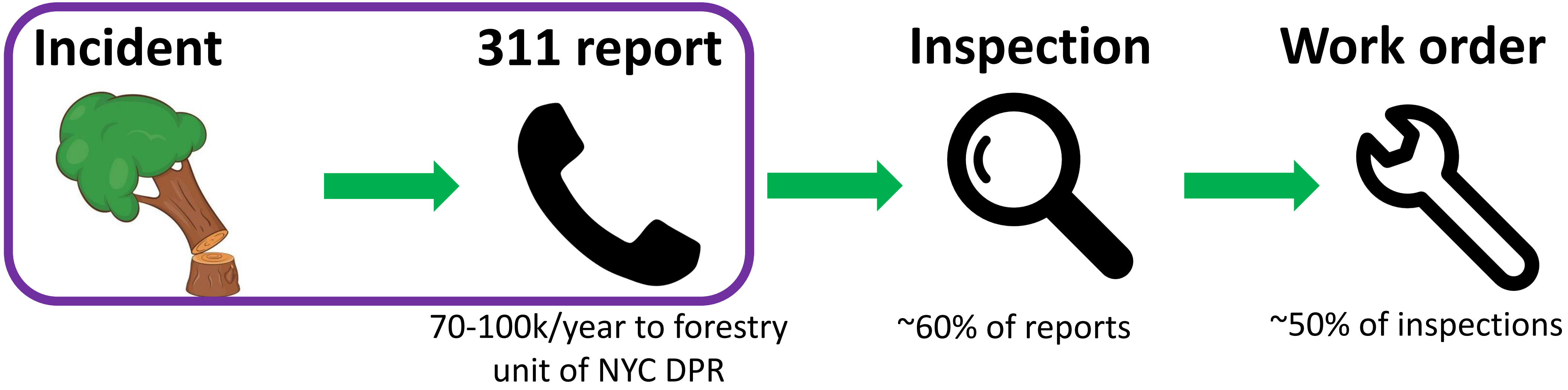
~50% of inspections

Why is this hard? Uncertainty, heterogeneous + strategic behavior, distribution shifts over time, capacity constraints, pipelined decisions

Research agenda: Audit and improve process's efficiency and equity

Existing collaboration: NYC Department of Parks and Recreation

Understanding reporting behavior



Why? If there are disparities in who reports problems, there will be disparities in what work gets done

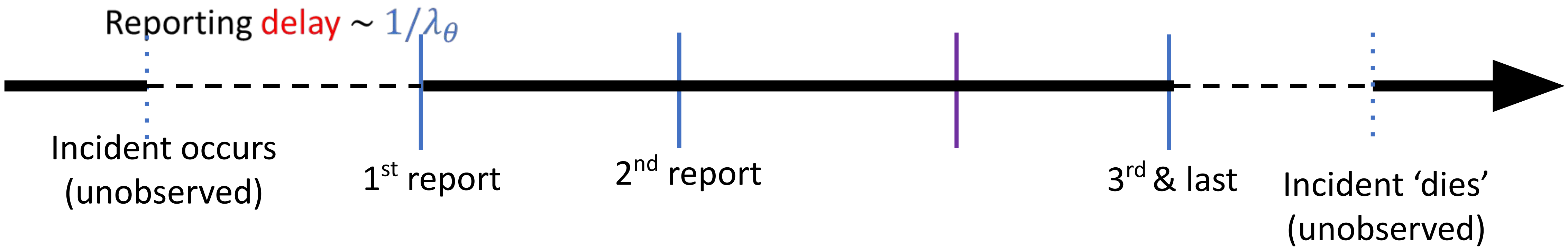
"Equity in Resident Crowdsourcing:
Measuring Under-reporting without Ground Truth Data"
w/ **Zhi Liu** (ACM EC 2022)

Model + Method summary

How long does it take for an incident of type θ to be reported?

(Hard because we never observe anything before the first report)

Step 1: Write down a system model where the estimand corresponds to some identifiable quantity



Model + Method summary

How long does it take for an incident of type θ to be reported?

Step 1: Write down a system model where the estimand corresponds to some identifiable quantity

Step 2: Computationally + statistically tractable estimation

$$\#\,\text{reports}(i) \sim \text{Poisson}\big(\lambda_\theta \times (b_i - a_i)\big)$$

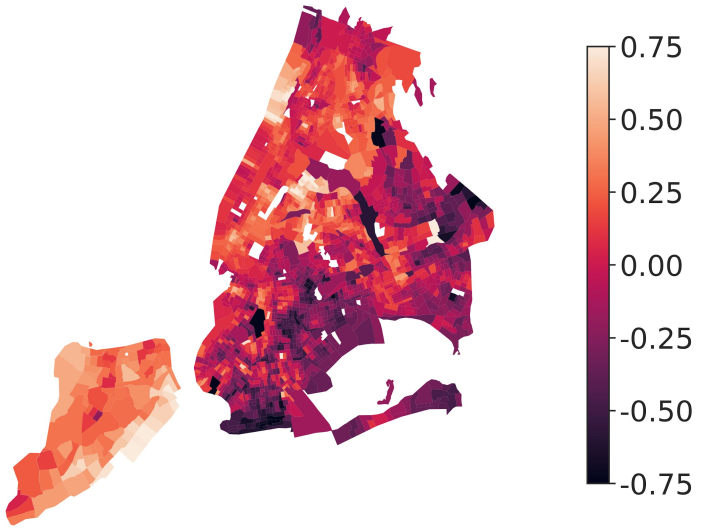
Spatial smoothing: ICAR Model [Morris et al. 2019]
Type θ contains an indicator for census tract (2000+ in NYC)
Then, α_k for each tract is drawn with mean of α_j of neighboring tracts

Results

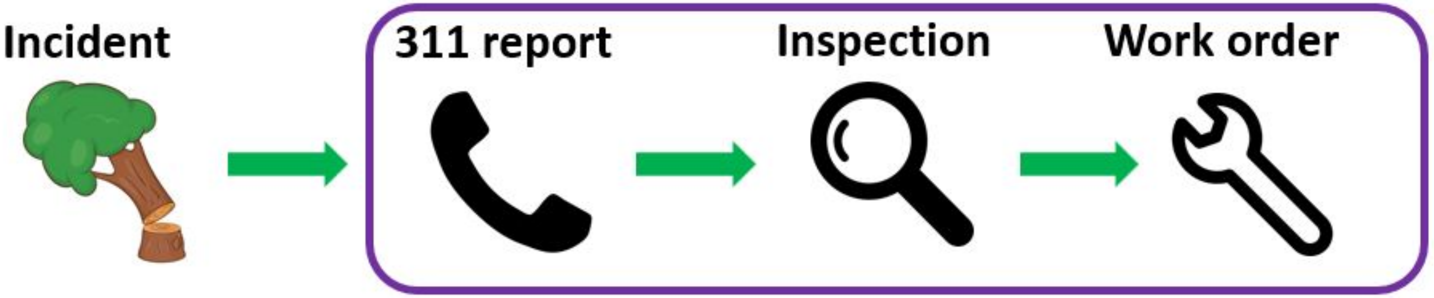
Efficiency: Reporting rates higher for more urgent incidents

Equity: Reporting rates vary substantially by neighborhood

	Manhattan	Queens
High risk hazards	2.5 days	4.7 days
Medium risk tree damage	15 days	28 days
Low risk minor issue	112 days	209 days



Auditing agency decisions in entire pipeline



Questions: Is the agency inspecting the right reports? Are decisions efficient & equitable?

Challenge: Modeling capacity-constrained decisions under uncertainty

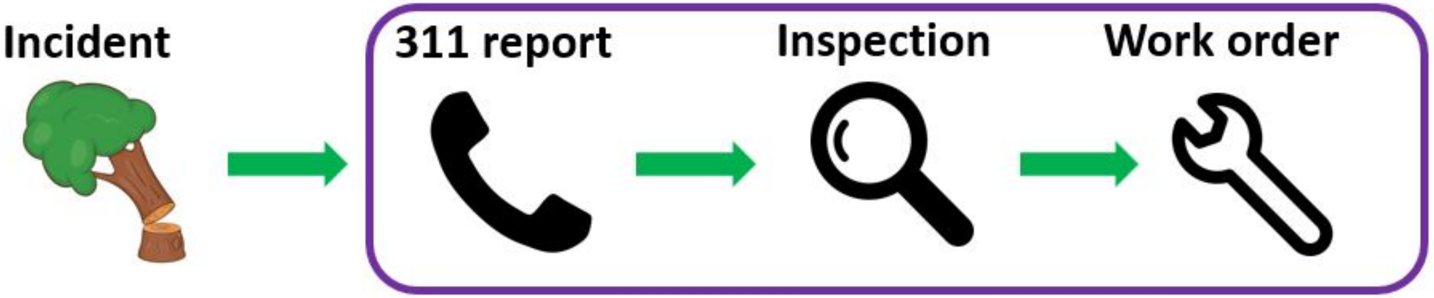
Method

(1) Use ML techniques to estimate incident risk given report characteristics

(2) Compare "optimal" set allocation decisions with empirical ones

"End-to-end Auditing of Decision Pipelines"
w/ **Benjamin Laufer** and **Emma Pierson**

Improving agency decisions



Question: Can we "optimally" re-prioritize inspections and work orders?

Challenge: Want "simple" policies that don't require maintaining an ML model

Method

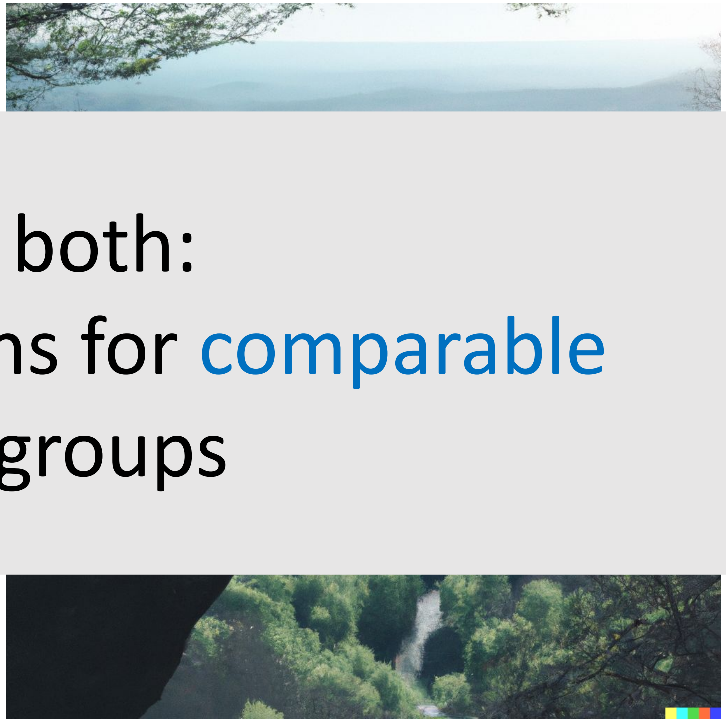
(1) Use ML techniques to [robustly] estimate incident risk given report characteristics
(2) Come up with "service level agreements" for how quickly to address reports

"Making Inspection Decisions: Designing service level agreements"
w/ **Zhi Liu**

Discussion: ML + Operations

Machine learning

Operations



Fairness requires both:
We want to compare decisions for comparable incidents/people/groups

Analyze individual incidents

(characterizing uncertainty, representing data)

Make global decisions

(comparing incidents, allocation under capacity constraints, modeling incentives)

Another example: Recommendation systems

Old school ML view: Predict match between single item and user pair

But there are many global properties of recommender systems

- How users/items affect each other [competition effects]
- How users affect what items are produced [supply-side equilibria]
- How can we recommend *sets* of items [diverse recommendations]

Joint work with: Christian Borgs, Wenshuo Guo, Meena Jagadeesan, Michael I. Jordan, Karl Krauth, Lydia Liu, Laura Mitchell, Jacob Steinhardt, Gourab K Patro, Lorenzo Porcaro, Qiuyue Zhang, Meike Zehlike

Questions?

ngarg@cornell.edu