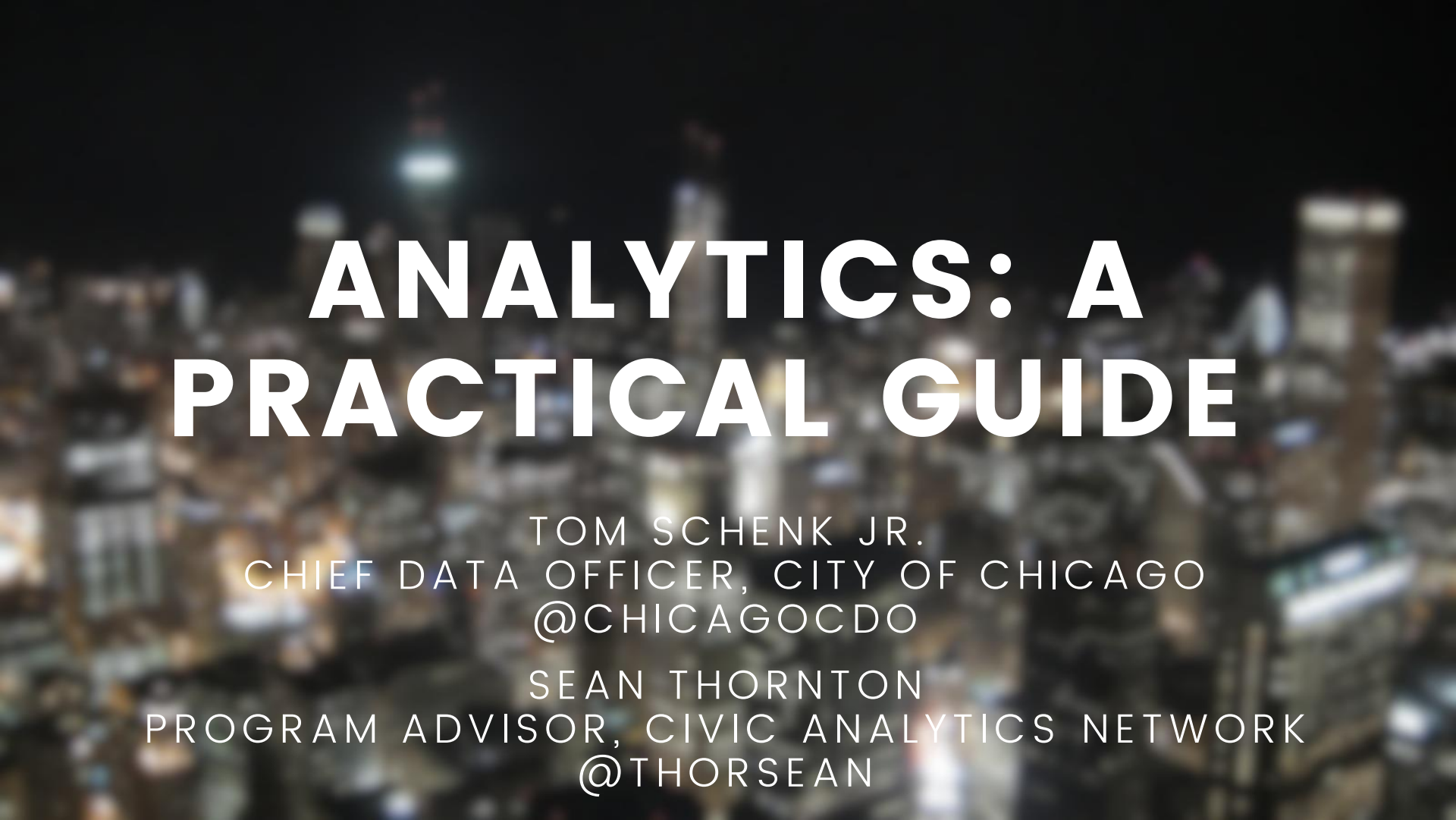


ANALYTICS: A PRACTICAL GUIDE

TOM SCHENK JR.
CHIEF DATA OFFICER, CITY OF CHICAGO
@CHICAGOCDO

SEAN THORNTON
PROGRAM ADVISOR, CIVIC ANALYTICS NETWORK
@THORSEAN

Source: techplan.cityofchicago.org

APPLIED ANALYTICS GUIDE

The "Applied Analytics Guide" is a 10-step process that the City of Chicago follows to identify, analyze, and address important problems with the advanced use of data. When implemented, the process can lead to new operational methods that enhance operational efficiency and effectiveness, as well as improve quality of life for residents.

1. IDENTIFY PROBLEM

SECTIONS SEARCH

Chicago Tribune

SUPPORT TODAY
Sponsor a Student

ESDAY NOV. 7, 2017 SPORTS BREAKING MOST POPULAR OPINION SUBURBS ENTERTAINMENT BUSINESS ADVERTISING 38

City falls far short inspecting restaurants

The Tribune visits City of Chicago sanitarian Kimberly Franklin as she does a health inspection of a Subway restaurant in the South

Optimizing Inspections

The objective for this research project is to order inspections to increase the speed of finding critical violations at retail food establishments.

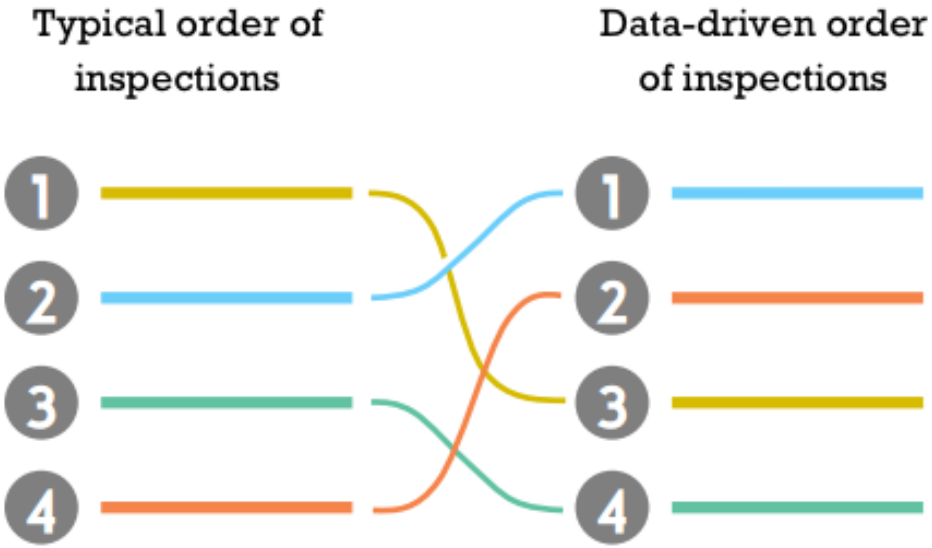


Image adapted from Michael Mooney's *Little Chicago* (CC-BY 2.0).

2. RANK PROBLEMS

Research Question Evaluation Criteria:
Predictive Analytics Projects

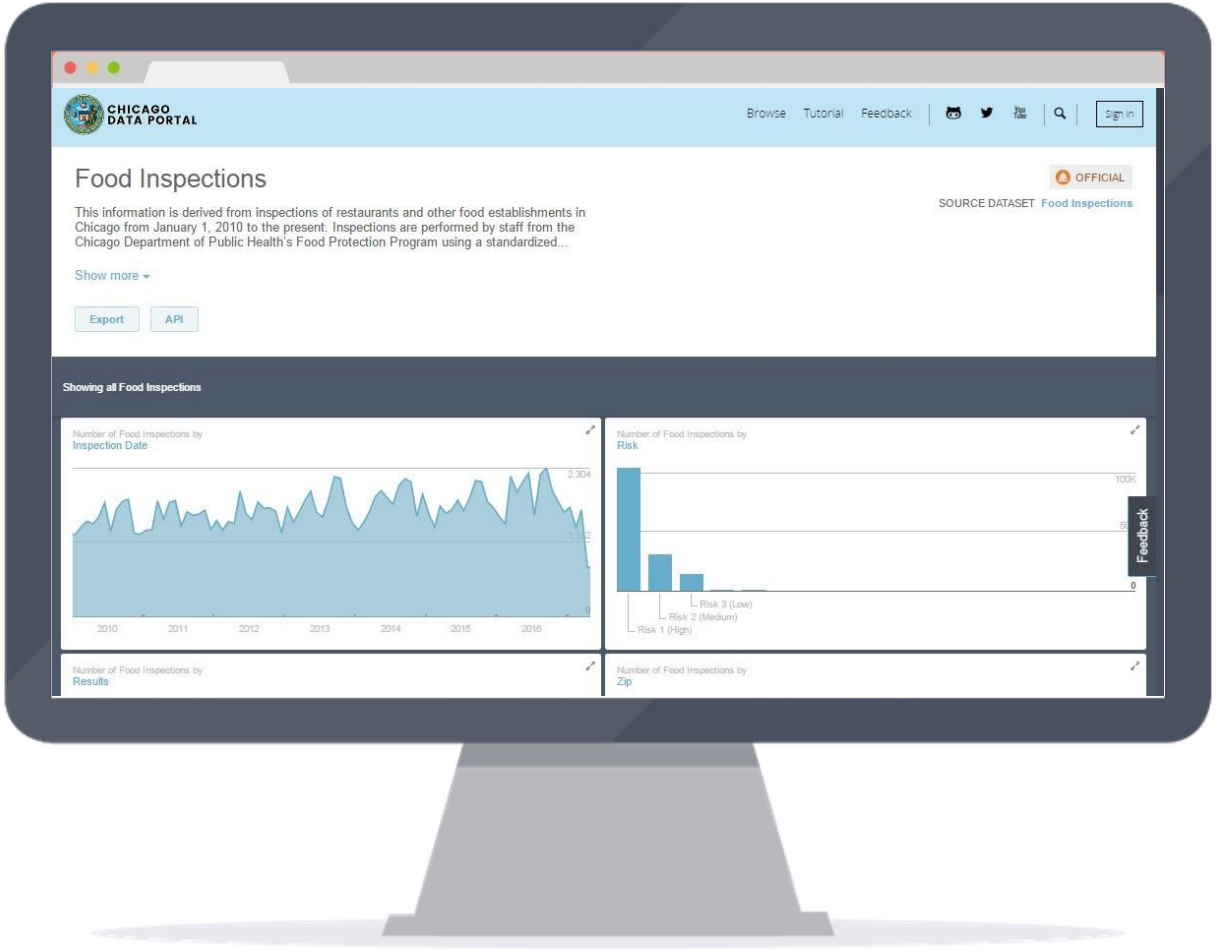
HARVARD Kennedy School
ASH CENTER
for Democratic Governance and Innovation

chicago.github.io • dev.cityofchicago.org • @ChicagoCDO • @ThorSean

Criteria	Question	Response
1. Data Readiness	Is data associated with this RQ available and/or ready to be analyzed?	• Available • Partly available • Not available
2. Policy Alignment	Does this RQ align with City policy goals—from both the Mayor's Office and applicable departments?	• Yes • No
3. Operational Impact	As a pilot, what level of opportunity would this RQ provide for positive operational outcomes—i.e., a reduction in time/cost when compared to current operations?	• High • Medium • Low
4. Resident Impact	As a pilot, what level of opportunity would this RQ directly provide benefit to the residents of Chicago?	• High • Medium • Low
5. Level of Use	If this pilot were to become a fully operational tool in your department, what level of use would it receive?	• Weekly • Monthly • Quarterly • Annually
6. Potential for Replication	Would implementing this RQ as a pilot provide a model that can be reused for other operational areas of your department, or elsewhere in the city?	• Yes • Unsure • No
7. Operational Change	If this pilot were to become a fully operational tool in your department, how drastically would it alter current operations?	• (Open-ended response)

3. DEEP DIVE

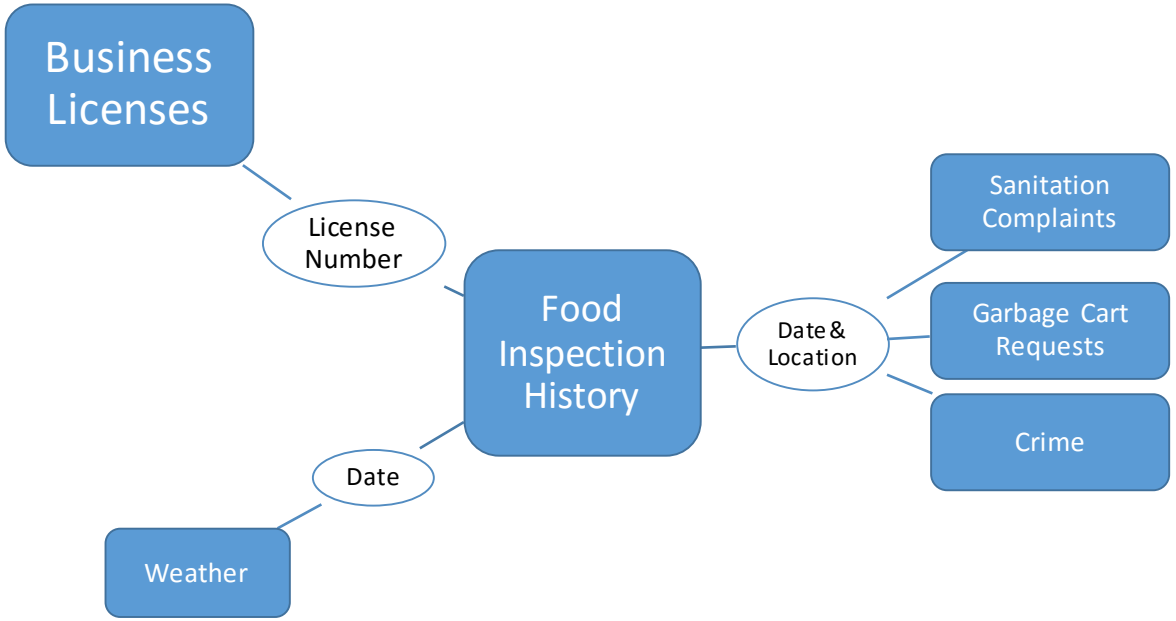
4. DATA ACQUISITION



Using #opendata

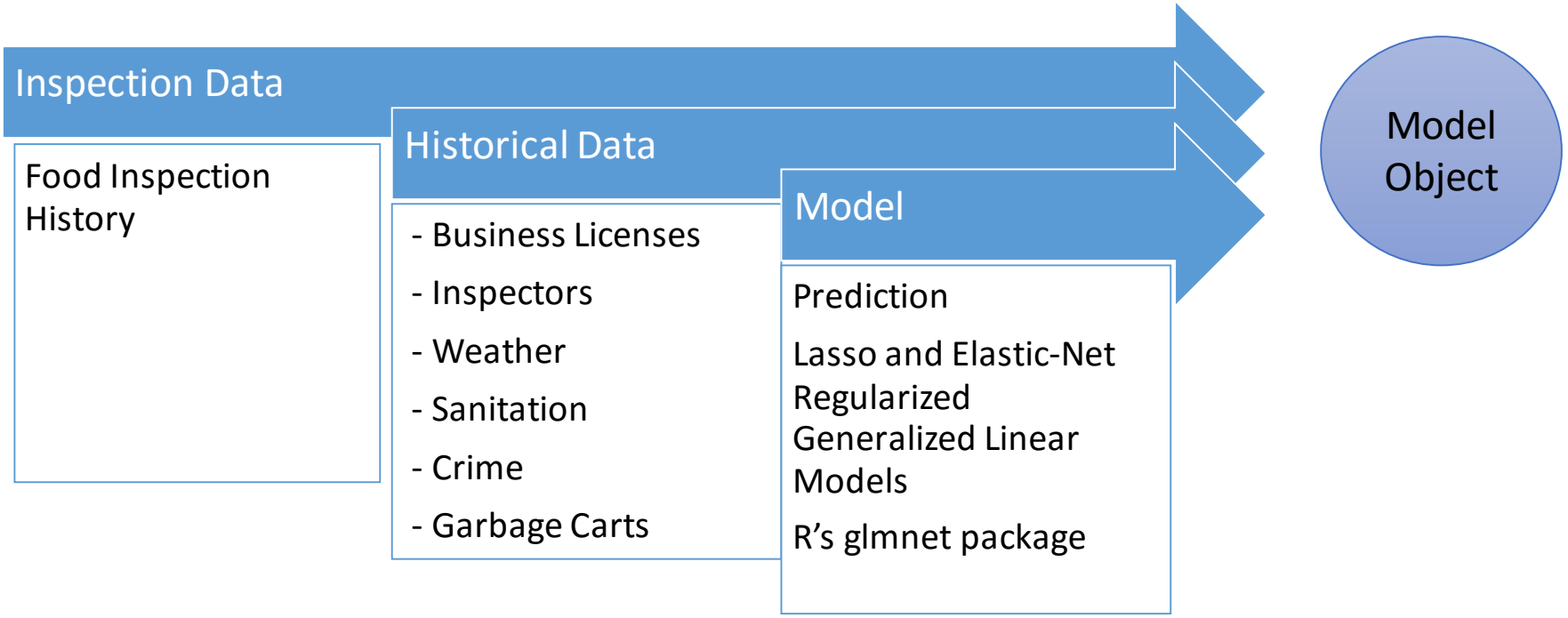
Chicago leveraged the #opendata portal to share data with external researchers, leveraging the city's premiere method of sharing data and saving time on data-sharing agreements to create #predictions.

DATA SOURCES



5. MODEL DEVELOPMENT

MODEL


Inspection Data
Food Inspection History
Historical Data
- Business Licenses
- Inspectors
- Weather
- Sanitation
- Crime
- Garbage Carts
Model
Prediction
Lasso and Elastic-Net Regularized Generalized Linear Models
R's glmnet package
Model Object

The model predicts the likelihood of a food establishment having a critical violation, a violation most likely to lead to food borne illnesses. Over a dozen #opendata sources were used to help define the model. Ultimately, ten different variables proved to create #predictions of critical violations.

Significant Predictors

Establishments with previous critical or serious violations

Three-day average high temperature

Nearby garbage and sanitation complaints

Nearby burglaries

Whether establishments has tobacco or alcohol license

Length of time since last inspection

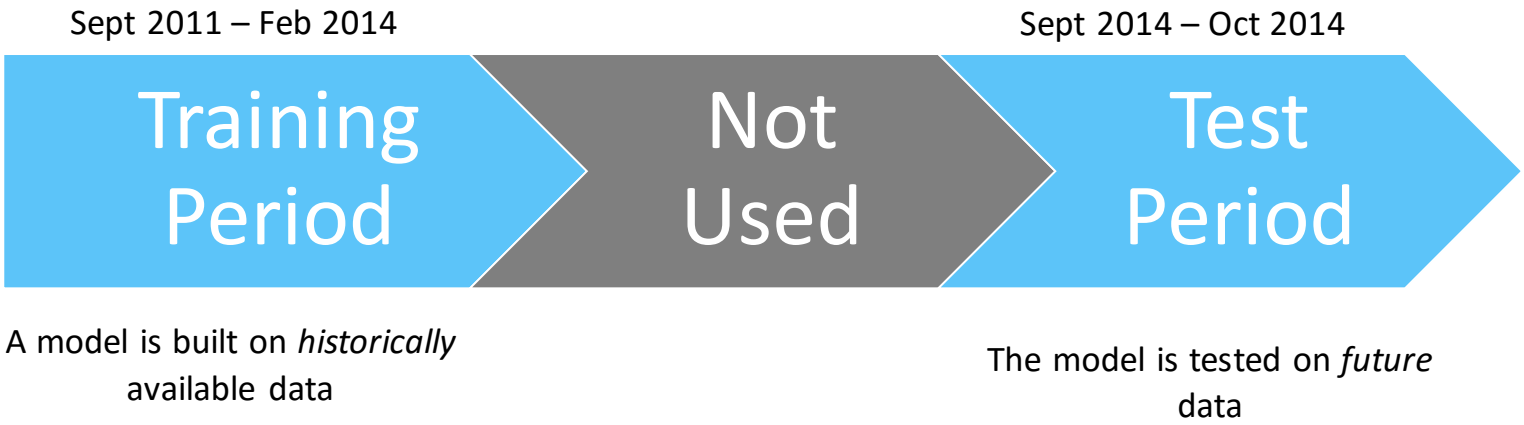
Length of time establishment has been operating

Inspector assigned

6. MODEL REVIEW

TEST / TRAIN FRAMEWORK

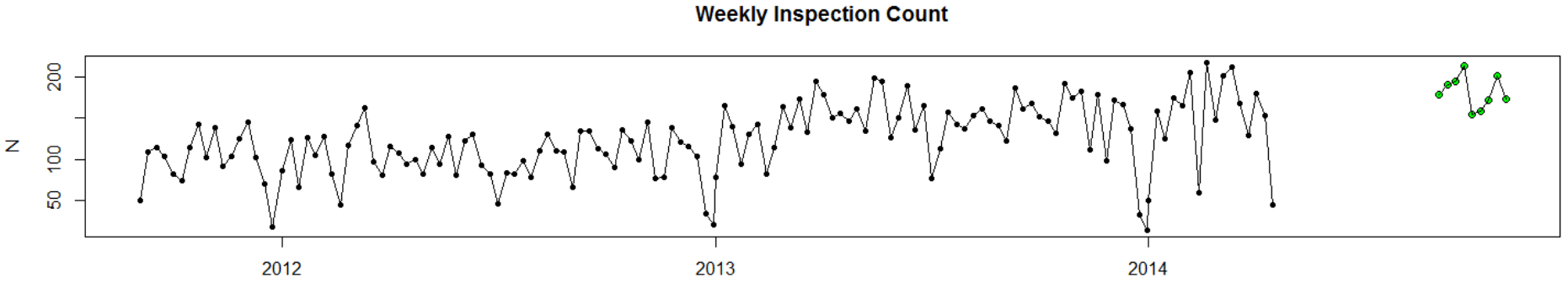
- The first model was built using data prior to 2014, and tested in early 2014
- The second model was built in response to the first
  - Completed in the summer of 2014
  - Tested in November based on actual inspection results from September and October



7. PILOT & EXPERIMENT

TEST / TRAIN FRAMEWORK

- The first model was built using data prior to 2014, and tested in early 2014
- The second model was built in response to the first
  - Completed in the summer of 2014
  - Tested in November based on actual inspection results from September and October

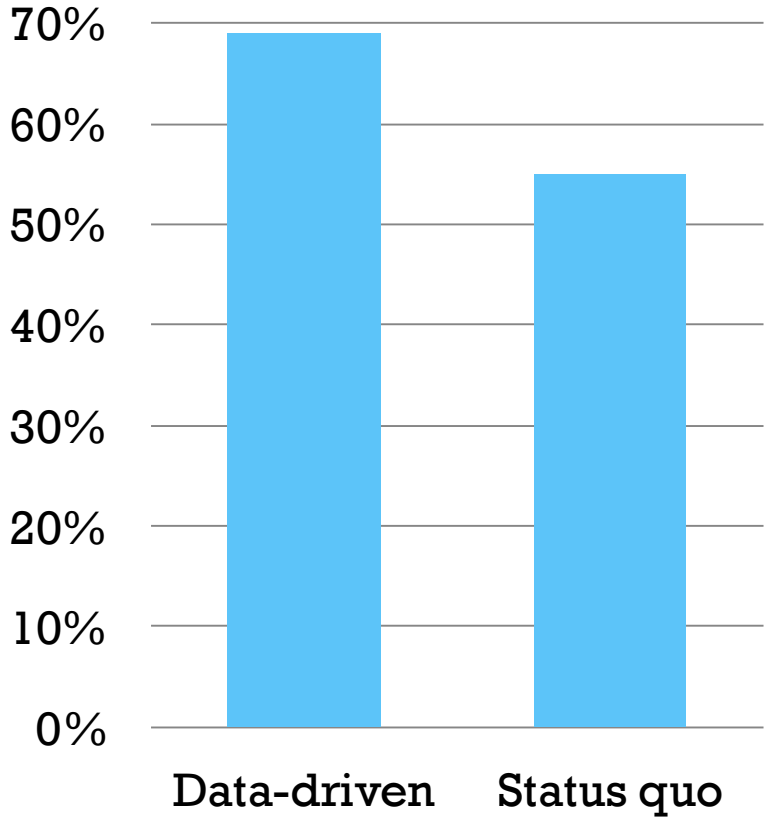


A model is built on *historically* available data

The model is tested on *future* data

8. PILOT EVALUATION

Critical violations



The #predictions revealed an opportunity to find deliver results faster. Within the first half of work, 69% of critical violations would have been found by inspectors using a data-driven approach. During the same period, only 55% of violations were found using the status quo method.

IMPROVEMENT

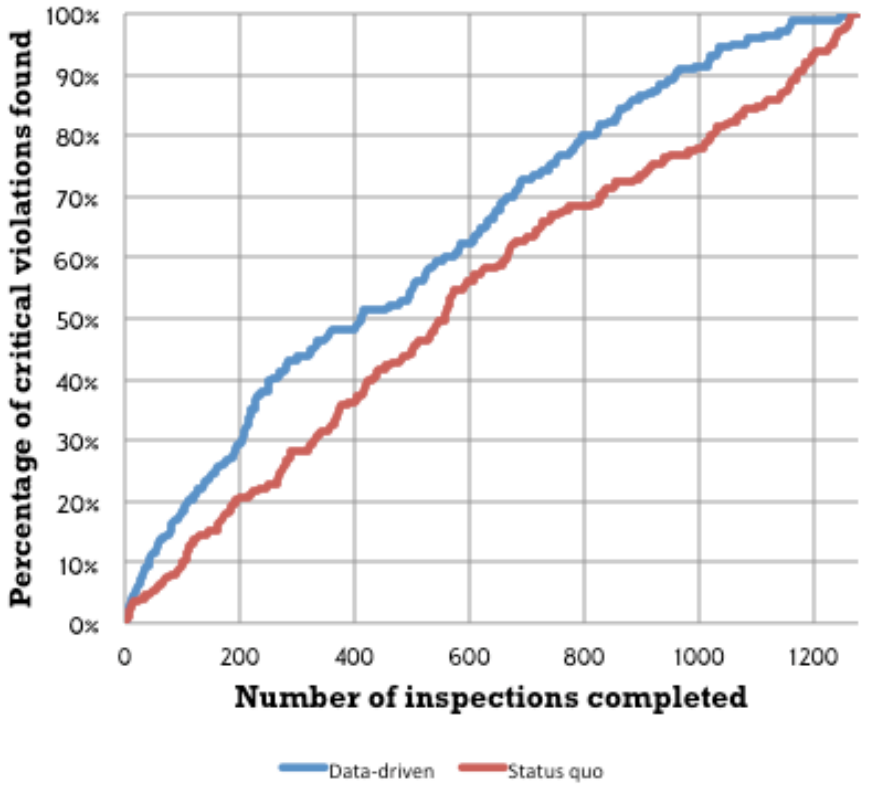
7 days

The food inspection model is able to deliver results faster.

After comparing a data-driven approach versus the status quo, the rate of finding violations was **accelerated** by an average of 7.4 days in the 60 day pilot. That means the #predictions led to more violations would be found sooner by inspectors.

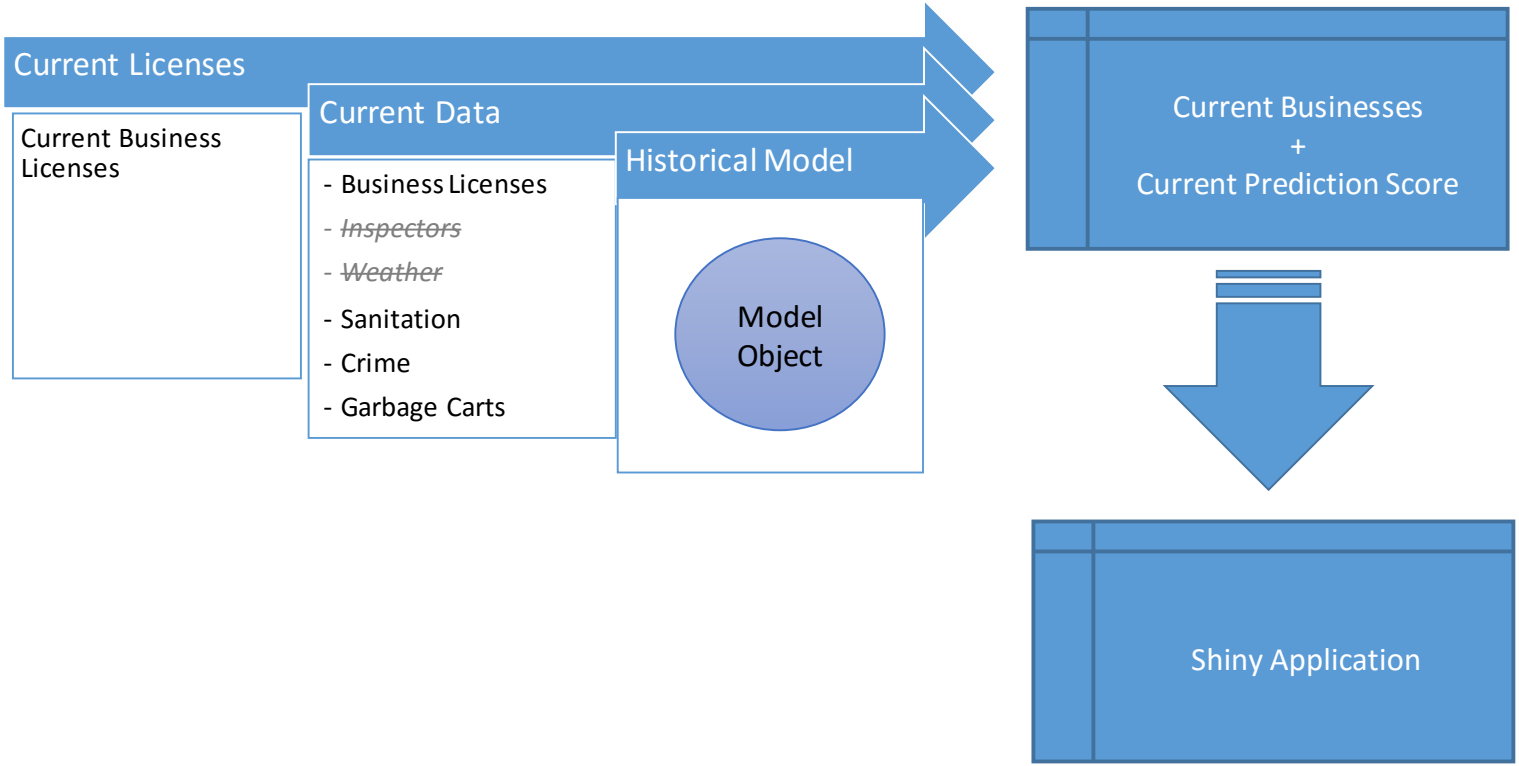
OPTIMIZING FOOD INSPECTIONS

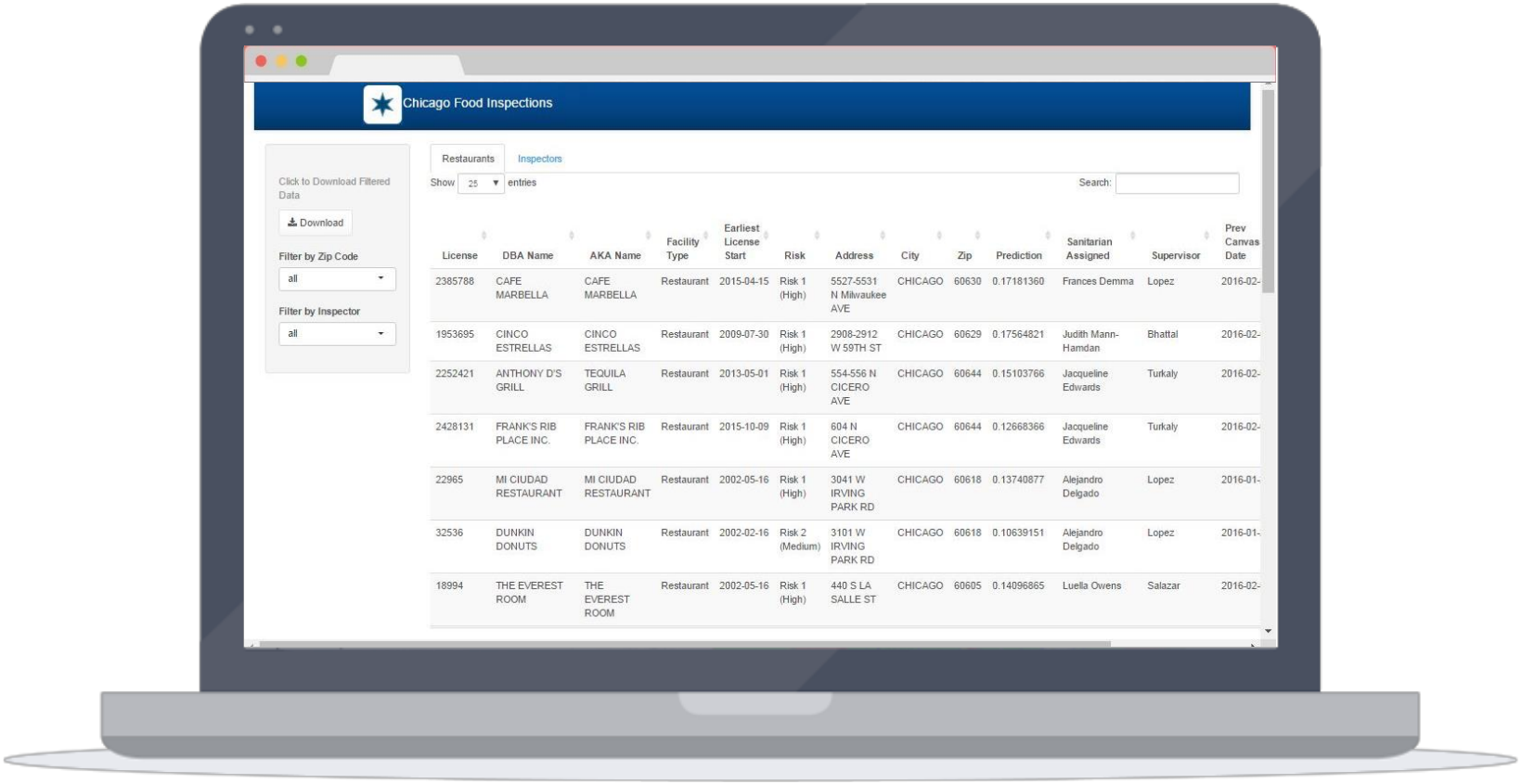
Discovering critical violations sooner rather than later reduces the risk of patrons becoming ill, which helps reduce medical expenses, lost time at work, and even a limited number of fatalities.



9. OPERATIONALIZING MODEL

PREDICTION AND APPLICATION





The data science team has built a website which lets CDPH prioritize inspections based on projected risk.

10. CHECK-IN

http://chicago.github.io/food-inspections-evaluation/

PARTNERSHIPS TO FIT YOUR NEEDS

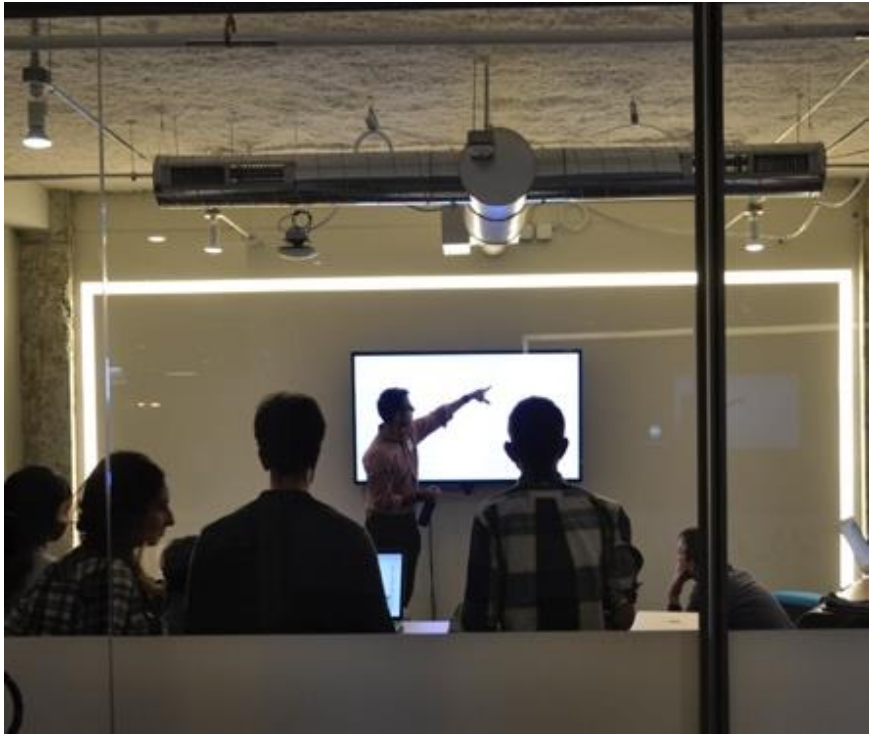
Community

Chicago has a large, vibrant, **productive**, civic community. This is led by Chicago residents interested in technology and society. Smart Chicago Collaborative and non-profits provide assistance and city officials regularly engage in meetups and other activities. This group has produced several helpful apps.


Open Gov Hack Night
happy 1 year!

CIVIC COLLABORATION

DePaul Student Interns



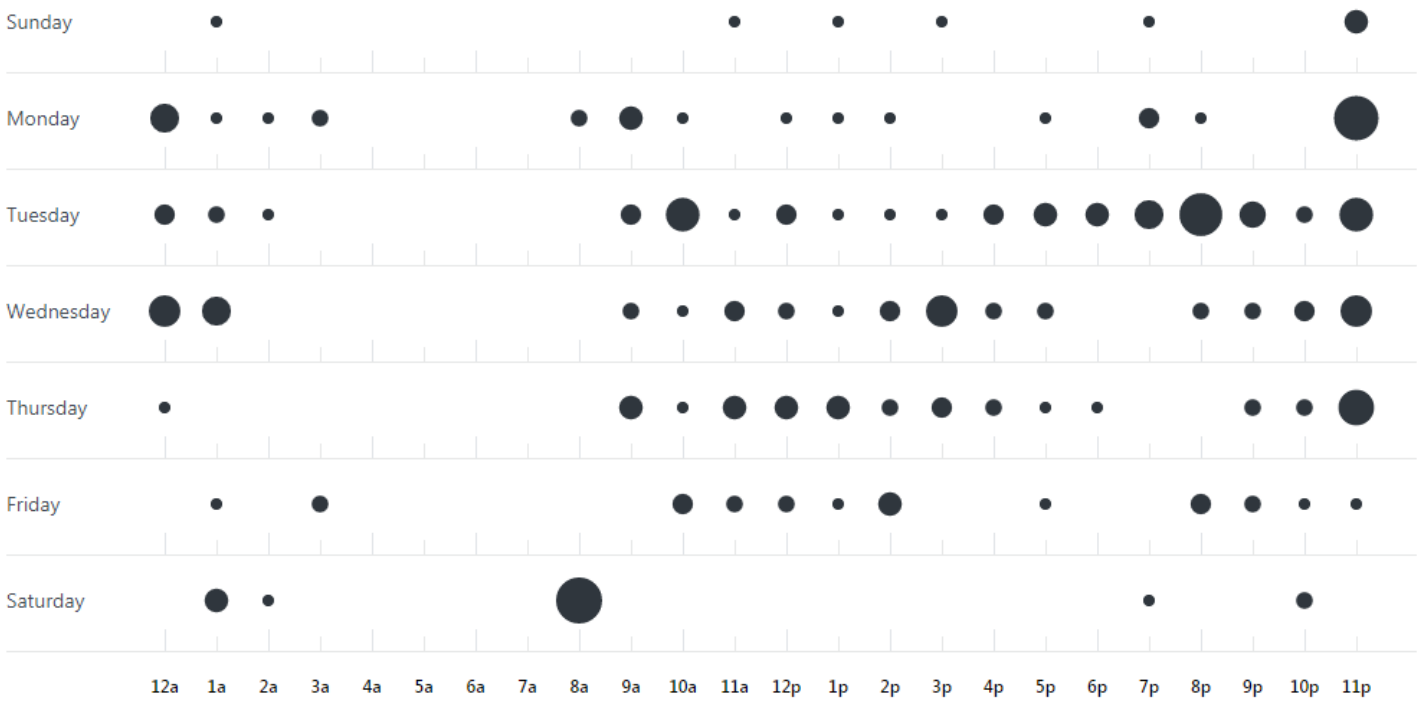
ChiHackNight


CLEAR WATER
FORECASTING CHICAGO'S WATER QUALITY

1,000 HOURS

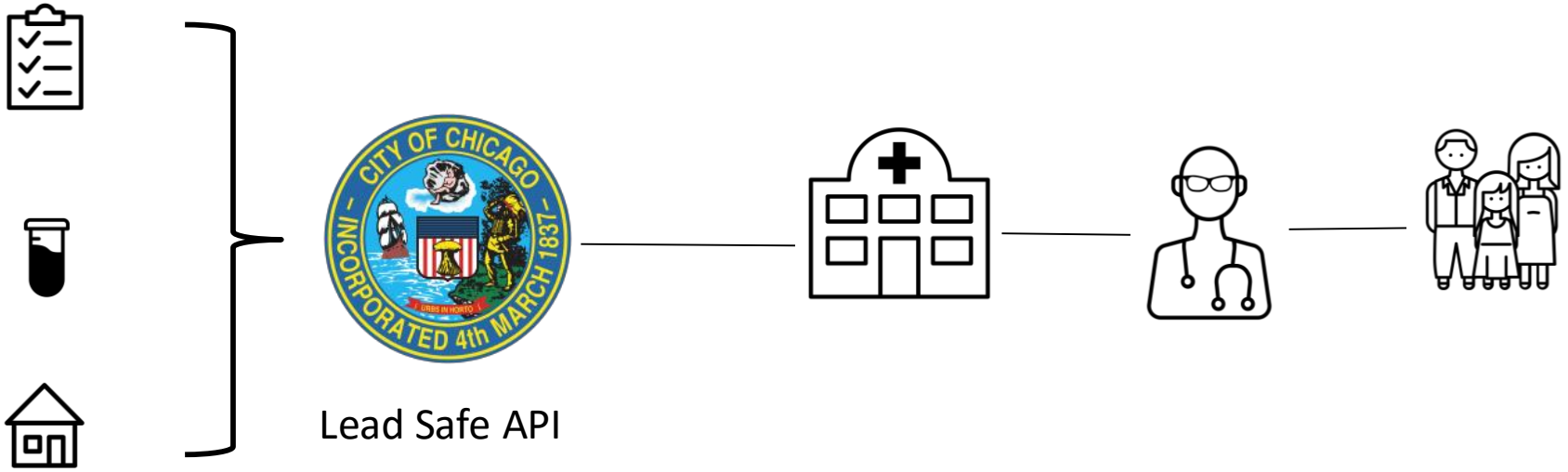
Total hours dedicated to this project through volunteers, Chi Hack Night, and students.

GITHUB PUNCH CARD



FIND OPERATIONAL PARTNERS

LEAD SAFE API



THANK YOU

Tom Schenk Jr.
Chief Data Officer, City of Chicago
Sean Thornton
Program Advisor, Harvard University

Contact Info:

@ChicagoCDO
tom.schenk@cityofchicago.org

Websites:

data.cityofchicago.org

digital.cityofchicago.org

opengrid.io

techplan.cityofchicago.org

speakerdeck.com/tomschenkjr