



URBAN CENTER FOR COMPUTATION AND DATA

Charlie Catlett, Director

Senior Fellow, Computation Institute of the University of Chicago and Argonne National Laboratory
Senior Computer Scientist, Argonne National Laboratory



Computation Institute



THE UNIVERSITY OF CHICAGO



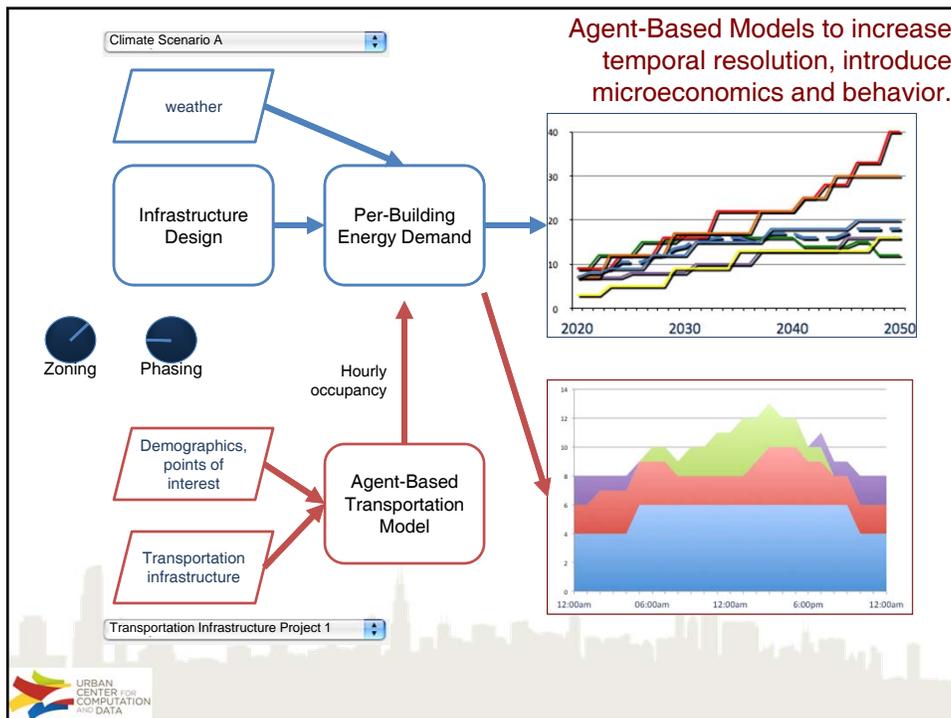
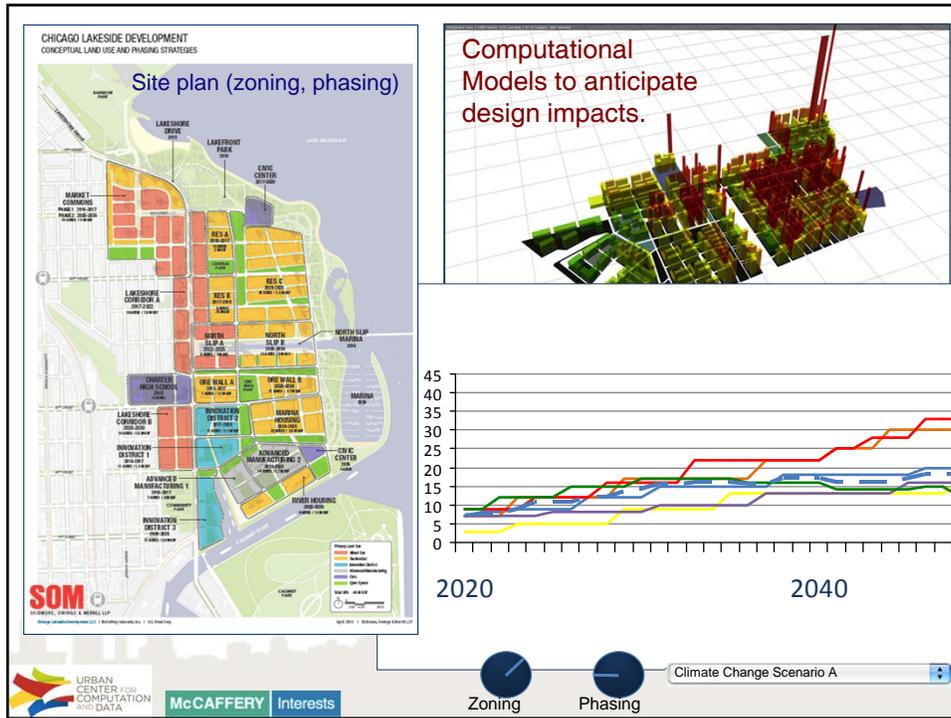
Argonne NATIONAL LABORATORY



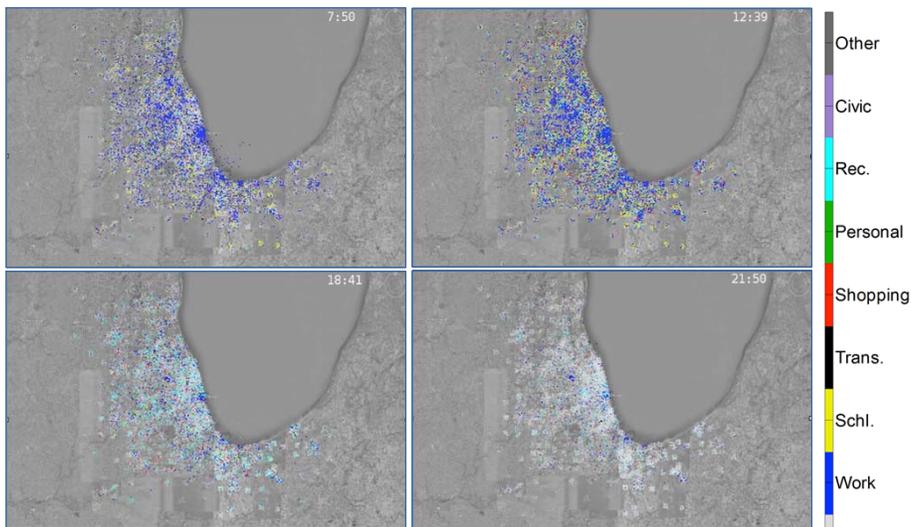
MacArthur Foundation



	Environment	Infrastructure	People
Planning and Design	Science-based design and planning. <i>Timescales: years to decades.</i>		
Proactive Policy and Intervention	Evidence-based measurement and predictive analytics. <i>Timescales: hours to months.</i>		
Optimization and Interaction	Decision-support and new ways to interact with the built environment. <i>Timescales: seconds to hours.</i>		
Application	Education, Training, Community Engagement		



Chicago Human Activity Patterns: Weekday



Produced by: S. Jiang, J. Ferreira, M. Gonzalez (2011)
 Data Source: CMAP Travel Tracker Data, 2008.
 Reference: Jiang, S., J. Ferreira, and M. González. 2012.
[Clustering Daily Patterns of Human Activities in the City.](#)
Data Mining and Knowledge Discovery. Volume 25, Number 3, Pages 478-510

Calibrating and Validating requires measured data.

13 Universities
Four science teams

Urban heat islands
Air quality
Pedestrian flow
Vehicle flow
Climate boundary layer
New interactions
...

URBAN CENTER FOR COMPUTATION AND DATA

THE CITY OF CHICAGO TECHNOLOGY PLAN

Initiative 3

Implement policies and infrastructure to allow for urban technology experimentation

The City will implement policies and basic infrastructure that make Chicago friendly to technology experimentation, allowing Chicago to become a global leader in environmental sensing, spectrum research, and wireless connectivity, while enabling researchers to develop solutions to city problems.

#Jobs #Savings #Services #STEM

Urban sensing—collecting and using data from sensors in public urban spaces—is essential to the next generation of data science. By implementing access policies that respect individual privacy and installing basic infrastructure (including platforms with power/connectivity), Chicago will become a leader in this emerging field. In addition, Chicago looks to position itself at the forefront of advanced wireless research and development.

These policies and infrastructure will enable researchers to collect data at little cost to the City, will help attract technology companies and STEM talent, and could increase R&D money spent in Chicago.

Additionally, results from this experimentation can be used to help to solve city problems. Chicago expects to have these policies in place within the next six months, and basic infrastructure will be available to approved researchers shortly thereafter.

High-frequency phenomena require new sources of measured data.

Open Data 1.0 – data portals to enable step one – visualization, mapping, correlating....

City Databases Sensors Video GPS Social Media

Carnegie Mellon University **Heinz College** Bloomberg Philanthropies NSF

Boston Research Map Powered by WorldMap Sign In | Create Map | View Map | Help

Overlays

- Place Locations
- Transportation
- Health & Human Ecology
- Society & Demographics
 - Income in Boston
 - Foreclosure Rate, January 2010
 - Percent White, 2000 (Census Tract)
 - High School Drop Out Rate (by District)
- Census 2010
- Census 2000
- American Community Survey (2005-2009)
- Historic Maps
- Local Projects
- Parcels
- Political Boundaries and Areas
- General
- Base Maps
 - OpenStreetMap
 - Google Hybrid
 - ESRI World Imagery
 - Google Terrain
 - Google Satellite

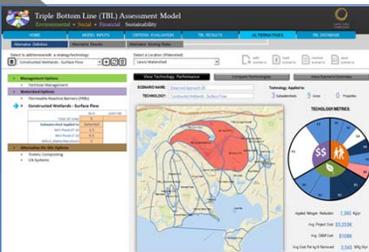
Robert J. Sampson and Daniel O'Brien, Harvard University
Boston Area Research Initiative



Metrics Generated from 53 Datasets

- Built Environment
- Neighborhood Assets
- Housing and Rental Prices
- Building-level Energy Use
- Solar installations
- Renewables generation
- Longitudinal Surveys
- Employment records
- Waste tonnage by block
- Transportation
- Traffic density
- Air quality
- Emissions

Developing methods to “measure” social/economic factors such as “neighborhood identity” or “access and mobility...”



...and to identify “neighborhoods” that are “similar” (control groups for evaluation).


 Matt Gee, UChicago
 Kate McGee, City of San Francisco

Negative Impacts

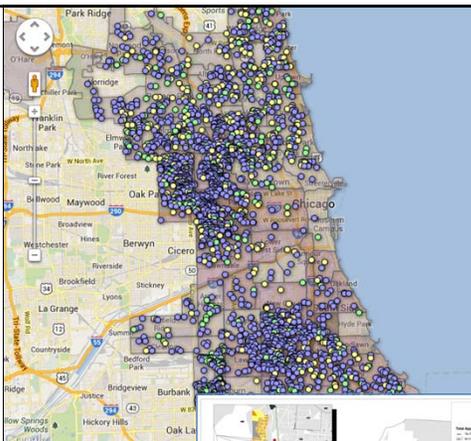
- Lowers property value
- Crime
- Lost taxes

Legal Obstacles

- Unknown owner
- Unpaid bills
- Contaminated land

Guiding Indicators

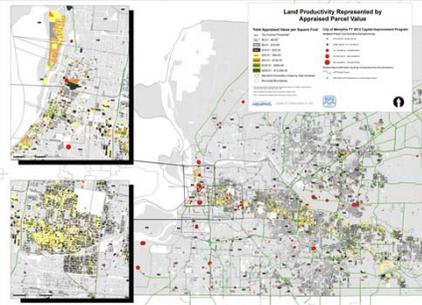
- Housing stability
- Affordability
- Vacancy



Chicago – optimize investments in vacant property.




Memphis—optimize investments in infrastructure.




 Matt Gee, UChicago

Background: Crime Prediction in Chicago



Since 2009, we have been working with the Chicago Police Department (CPD) to predict and prevent emerging clusters of violent crime.

Our new crime prediction methods have been incorporated into our **CrimeScan** software, run twice a day by CPD and used operationally for deployment of patrols.




Proactive (vs. reactive) policy – finding leading indicators.

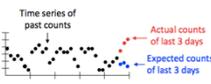
CrimeScan: Cluster Detection

We aggregate daily counts for each leading indicator at the block level, and search for clusters of nearby blocks with recent counts that are significantly higher than expected.

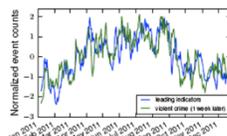
Imagine moving a circular window around the city, allowing the center, radius, and temporal duration to vary.



Is there any spatial window and duration T such that counts have been significantly higher than expected for the last T days?



Results: Exploratory Analysis

Considering all subsets of census tracts within each of the 77 neighborhoods of Chicago, 28 different potential predictors, and a 1-week lag, we found a correlation of $r = .786$ between violent crime and a subset of 12 leading indicators, for 10 census tracts in the West Englewood neighborhood.

Total run time for all 77 neighborhoods was **2.1 hours**.

Daniel B. Neill (Carnegie Mellon University - Event and Pattern Detection Laboratory),
Brett Goldstein (fmr Chicago CIO/CDO; Fellow, Harris School of Public Policy, Computation Institute, UChicago)

Garbage Cart Black	Inspection	No Building Permit &	Restaurant Complaint	Animal Abandoned
Maintenance	Vacant/Abandoned	Construction Violations	Bulk Pickup	Water Quality
Sanitation Code	Building	Missing Lid/Grate	Unwanted Animal	Sewer Odor/Bad Odor
Violation	Sewer Cave In	Block Party Request	Blue Recycling Cart	Dead Bird
Building Violation	Inspection	Recycling Pick Up	Animal Bite	Animal In Trap
Stray Animal	Garbage Pickup	Nuisance Animals	Dumpster Task Force	Animal Business
Dead Animal Pick-Up	Debris Removal	Animal - Inhumane	Inspection	Animal Fighting
Sewer Cleaning	Fly Dumping	Treatment	Demolition Inspection	

Rodent Infestation



Liquor license (decreased likelihood)
Low-risk v. high-risk gas station (both less likely than non-gas station)
Tobacco license
Neighborhood
Common names (e.g., Tom's market) in store name



Black Market Cigarettes

Food Safety Inspections





Tom Schenk, City of Chicago

Critical violation found in prior inspections
Sanitation code complaints through 311
Rodent sighting reported through 311
Request for garbage carts (per-ward)
Three-day moving avg of high temp before the inspection
Type of facility (e.g., restaurant versus grocery store)

Improving Communities through Data-Driven Land Banks

Sophia Alice, Evan Misshula, Skyler Whorton, and Tom Plagge

Summary

The foreclosure crisis led to an explosion of abandoned properties in Cook County—properties that can destabilize neighborhoods, depress tax revenues, drain government resources, and attract crime and drugs. The Cook County Land Bank is an agency tasked with putting these back to use.

The Problem

The Land Bank has several banks at its disposal to redevelop abandoned buildings. They can clear title, begin back taxes, resolve liens, and hold land until they are sold for the property owners. However, there are very few resources to assist residential addresses in Cook County (see the HUD/CESR). The Land Bank needs to be effective in its acquisition.



Identify vacant or underutilized properties.
Identify vacant or underutilized properties.

Community Scores

We are measuring the health of neighborhood real estate markets along several dimensions, including stability and affordability. The stability score (S) is based on Walker's Finance (2013), and depends on property values (V), transaction volume (T), mortgage to owner occupancy (M), and the production of high-end rentals (R). The affordability score (A) is based on income (I) and median property sale price (P) in each census tract.

$$S = 0.4V + 0.3T + 0.1M + 0.2R$$

$$A = \frac{I}{P}$$


Web Application

We are incorporating the scores and maps we developed into a Django web application with a PostgreSQL database and PostGIS extensions. The server data will be available via an API that can be kept synchronized with the Land Bank's internal and inventory systems.



Property Scores

We are also developing scores for individual properties based on their income values in their neighbors and their economic impact on the community. The latter is based on 2010 census reports.

For the latter, we are using a hedonic pricing model that takes in several property and neighborhood characteristics. Based on historical data, we can estimate the percentage by which a foreclosure, vacancy, or demolition in a given community will affect the surrounding property values.

Our preliminary model indicates that, controlling for the local demographic and economic characteristics of the community, each foreclosure that occurs within a 1/8 mile of a property has approximately a 1% effect on its price.

Conclusions

The Cook County Land Bank plans to attack the problem of abandoned buildings in a similar, data-driven way. Having all of the relevant information in one place will save the staff time, and automating the calculations in a flexible, extensible, application will help make the bank's decisions clear and transparent. A systematic approach to property acquisition will also allow the agency to evaluate the impact of the strategies it pursues.

Future Work

The application and algorithms we provide in the Land Bank will be improved, focused primarily to guide the bank's strategic decisions. And, the agency reserves the option, the indicators and scores can be adjusted against real results, and can be used to reflect changing strategies or market conditions.



SCHMIDT FAMILY FOUNDATION

The New York Times | Technology | Personal Tech | Business Day

Bits

JULY 26, 2013, 7:41 AM | Comment

A Summer of Data Hacking Social Problems

By STEVE LOHR

- FACEBOOK
- TWITTER
- GOOGLE+
- SAVE
- EMAIL
- SHARE
- PRINT

The idea, Rayid Ghani recalled, grew out of his experience speaking to computer science students at elite schools like Carnegie Mellon, Stanford and the University of Chicago. President Obama had just won his re-election bid last fall. And Mr. Ghani, chief scientist for the campaign, was on a kind of explanatory victory tour, describing how cutting-edge data analysis and computing tools gave its side an edge.



Rayid Ghani, chief scientist for President Obama's re-election campaign.

For Mr. Ghani, the Obama campaign demonstrated how those tools could be used to influence people in fields beyond the well-known commercial ones, like search, social networks and online advertising. And beyond politics, he would tell the students, were a host of social challenges in health care, education and urban development where their skills could be put to good use, working with nonprofits, civic groups and local governments.





URBAN CENTER FOR COMPUTATION AND DATA

7