THE UNIVERSITY OF CHICAGO

Computation Institute

Argonne NATIONAL LABORATORY

URBAN CENTER FOR COMPUTATION AND DATA

Toward an "Urban Science"

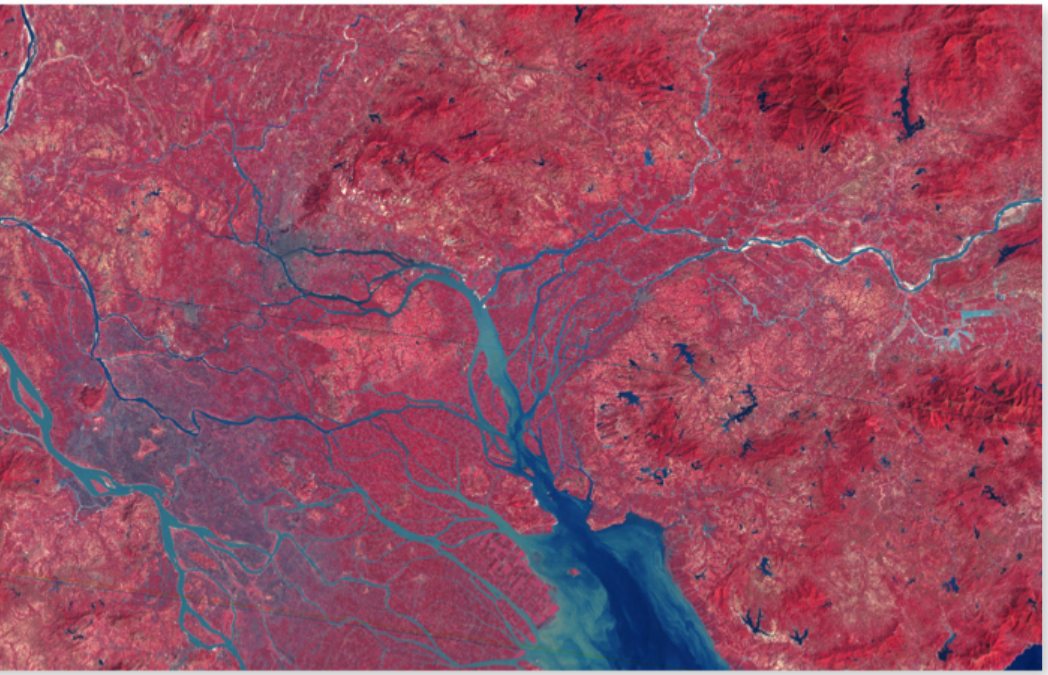
Charlie Catlett

Senior Computer Scientist, Argonne National Laboratory
Director, Urban Center for Computation and Data
Senior Fellow, Computation Institute of the University of Chicago and Argonne National Laboratory
Visiting Artist, School of the Art Institute of Chicago

August 2014
Pheasant Run

UrbanCCD is an initiative of the Computation Institute, a joint institute of the University of Chicago and Argonne National Laboratory

Rapid Urbanization



In 2025:

70%

of Chinese people will live in cities with 1M or more people.

And by 2030...

221

Chinese cities will have 1M or more people.

China will *add*

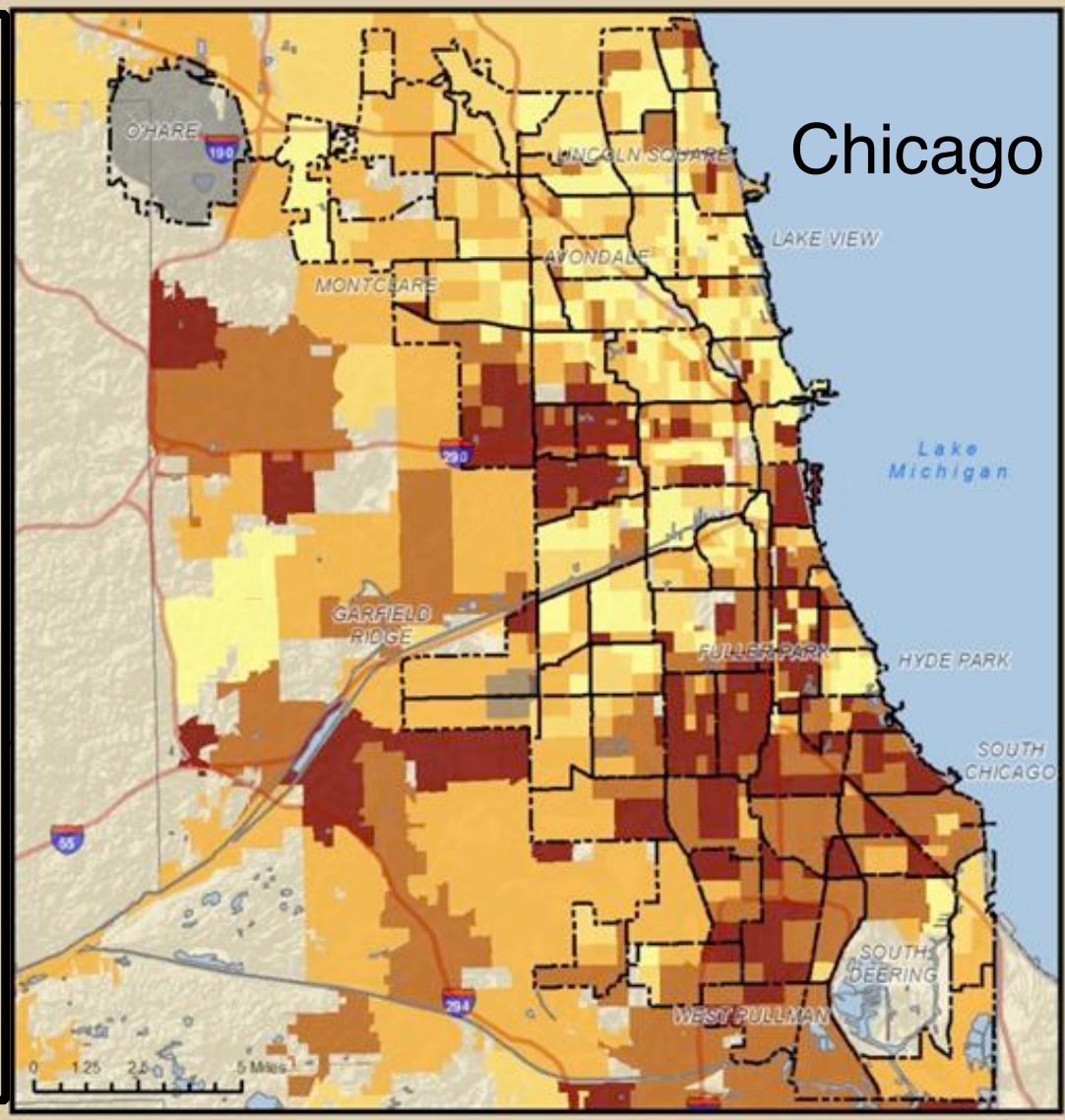
400

million city dwellers

....requiring the construction of one New York City every year for several decades

LANDSAT Images - NASA, Karen Seto (Yale)

Source: Foreign Policy Magazine, Sep/Oct 2010, "Megacities," Richard Dobbs (McKinsey Global Institute)



Source:
VCU Center on
Society and Health

Life Expectancy in Years

- Average Life Expectancy = 69.2 Years
- Average Life Expectancy = 74.5 Years
- Average Life Expectancy = 78.6 Years
- Average Life Expectancy = 81.8 Years
- Average Life Expectancy = 86.7 Years

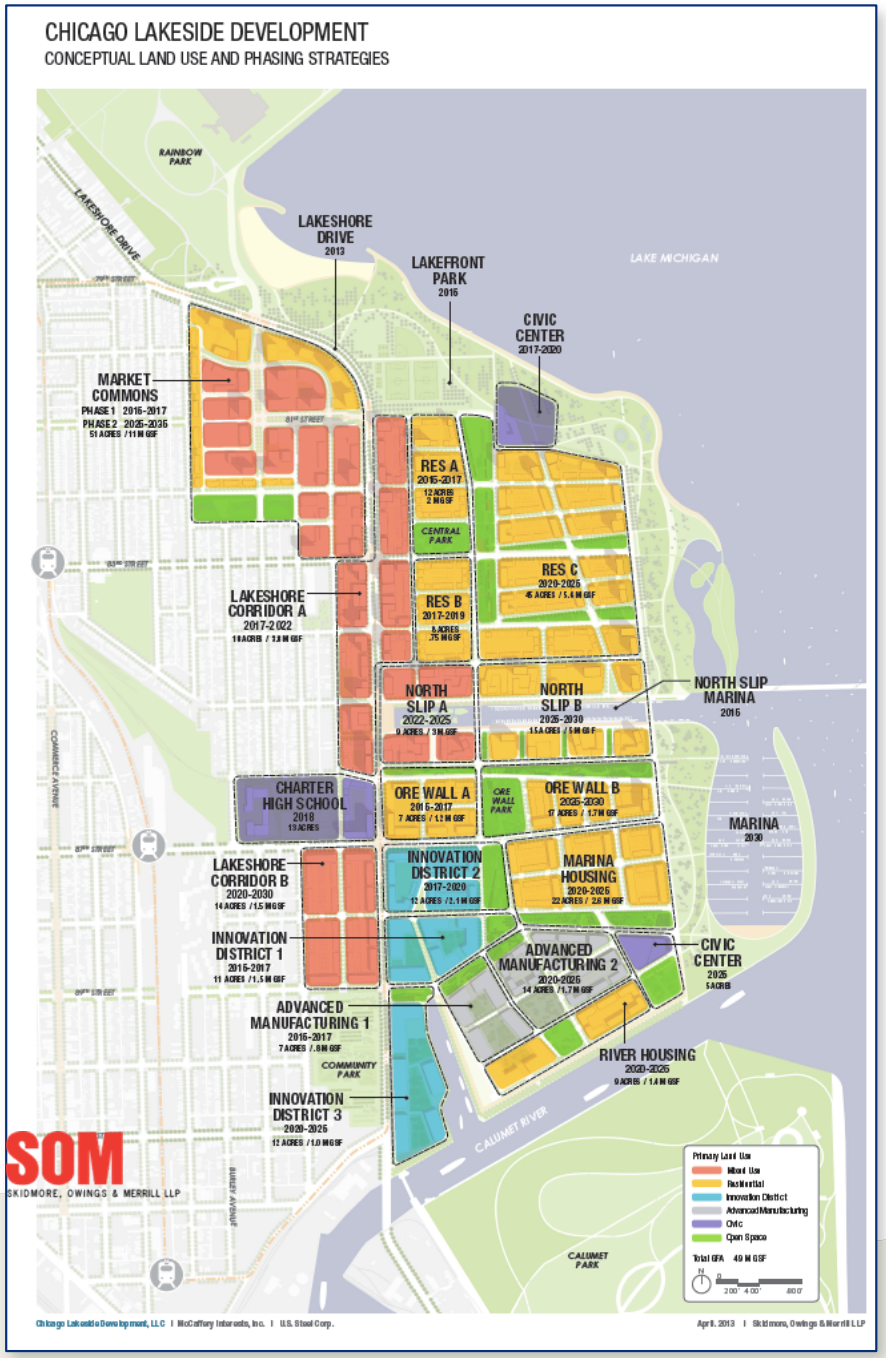
- City of Chicago
- Airports
- Interstate
- Water

Data Sources:
City of Chicago, 2010; 2009-2014 Estimates Premium Package, GeoLytics, Inc., New Brunswick, NJ, 2008; Esri, Inc., 2009; VCU Center on Human Need, 2011

	Environment	Infrastructure	People
Years	Design and Planning		
Days	Information-Driven Operations & Policy		
Minutes	New Insights; New Interactions		
Impact	Education, Training, Community Engagement		

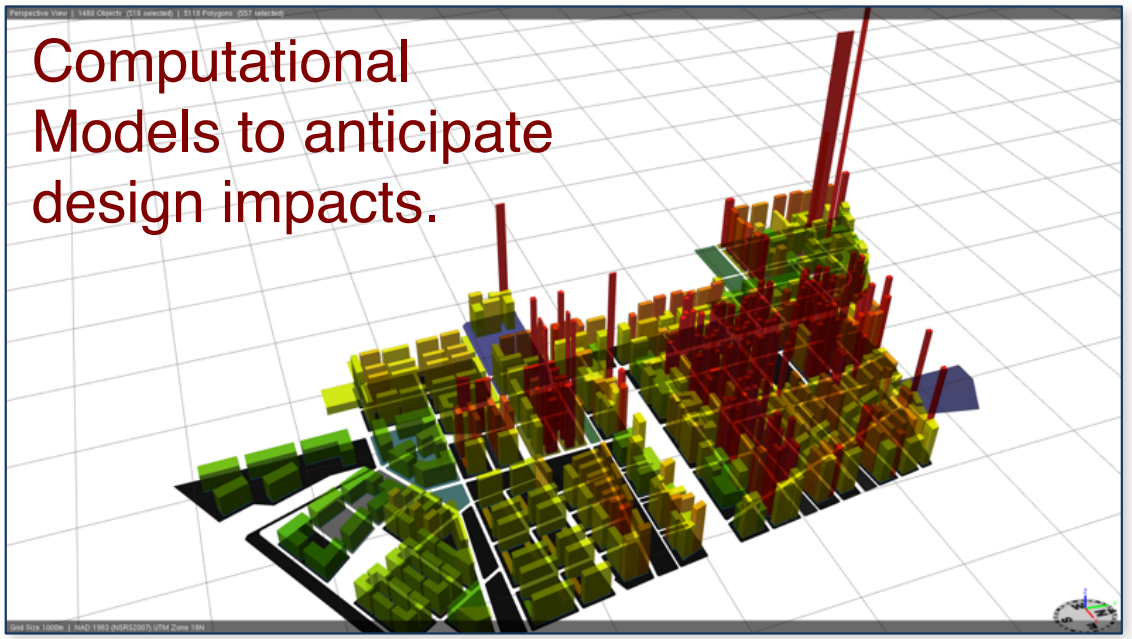
URBAN CENTER FOR COMPUTATION AND DATA

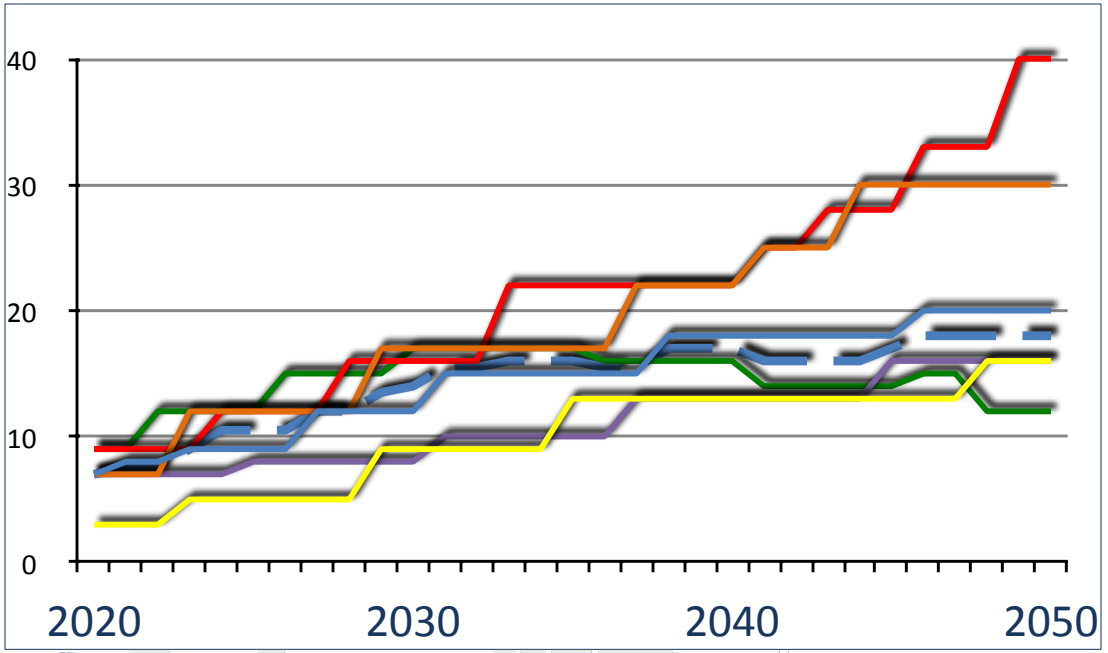
Site plan (zoning, phasing)



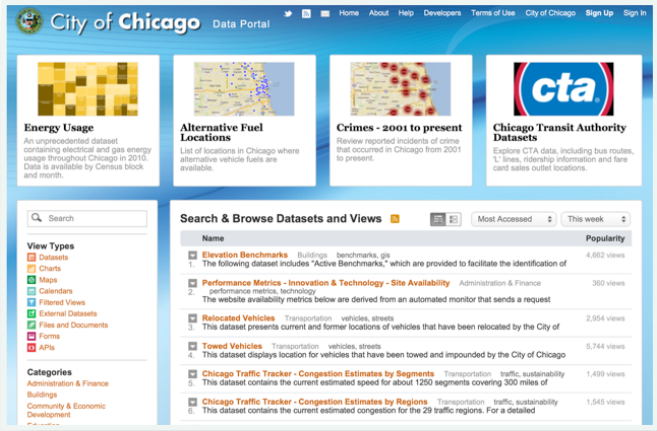
McCAFFERY Interests

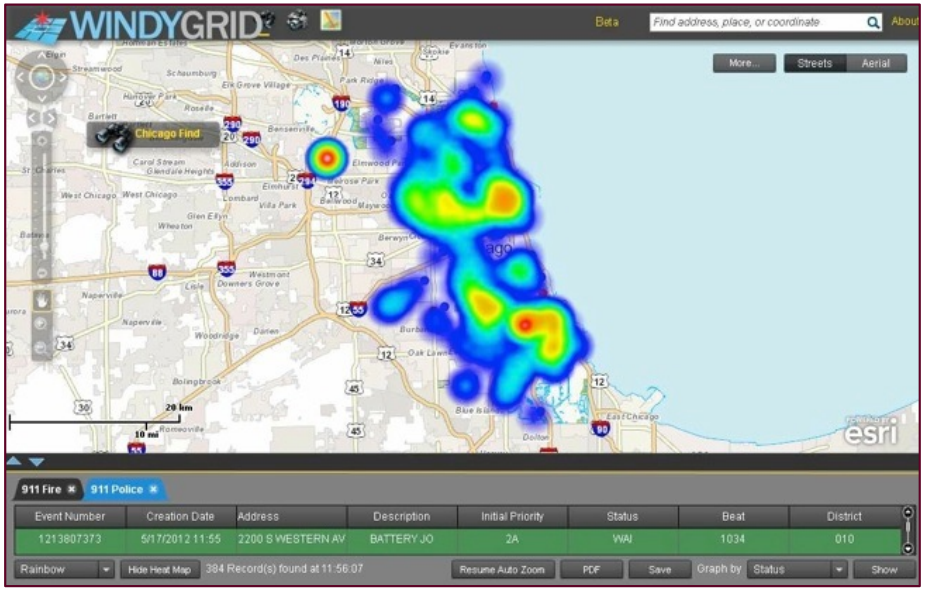
Computational Models to anticipate design impacts.

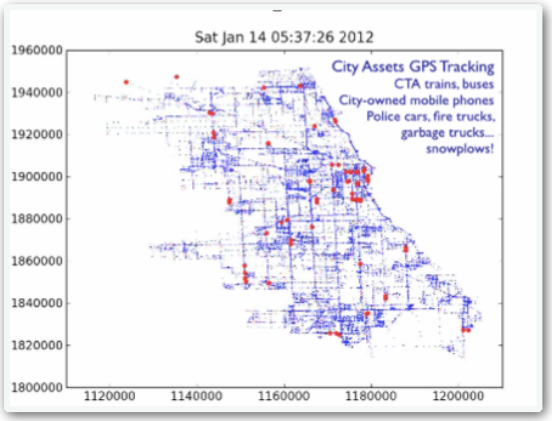


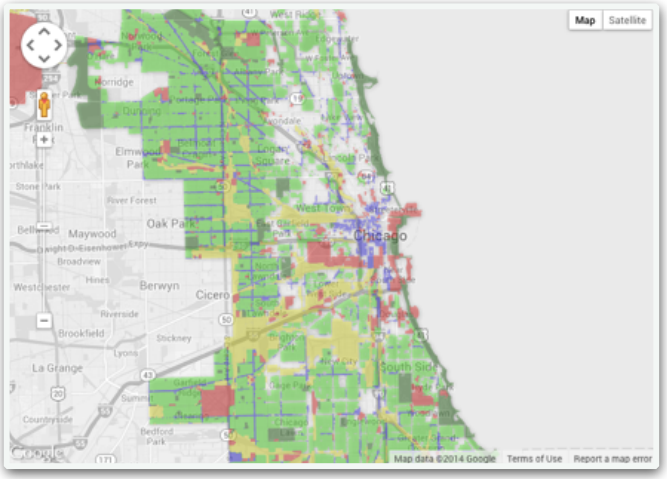


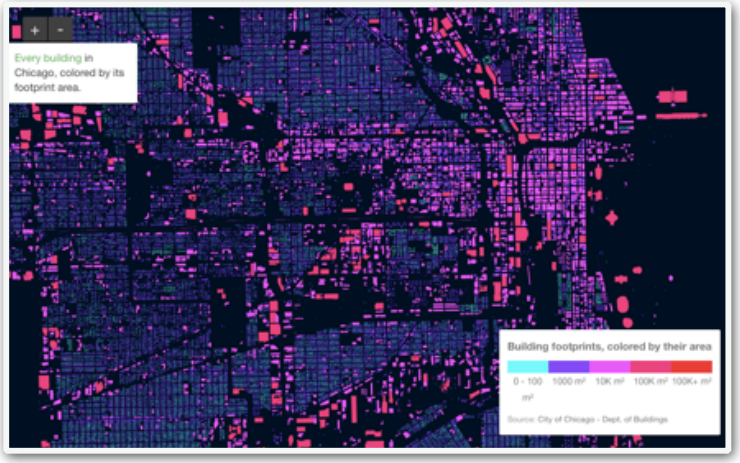
URBAN CENTER FOR COMPUTATION AND DATA

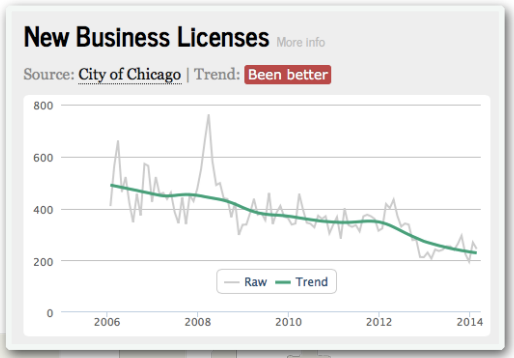












CITY OF CHICAGO – INCORPORATED 4th MARCH 1837

SEAL OF THE CITY AND COUNTY OF SAN FRANCISCO

City of Memphis

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

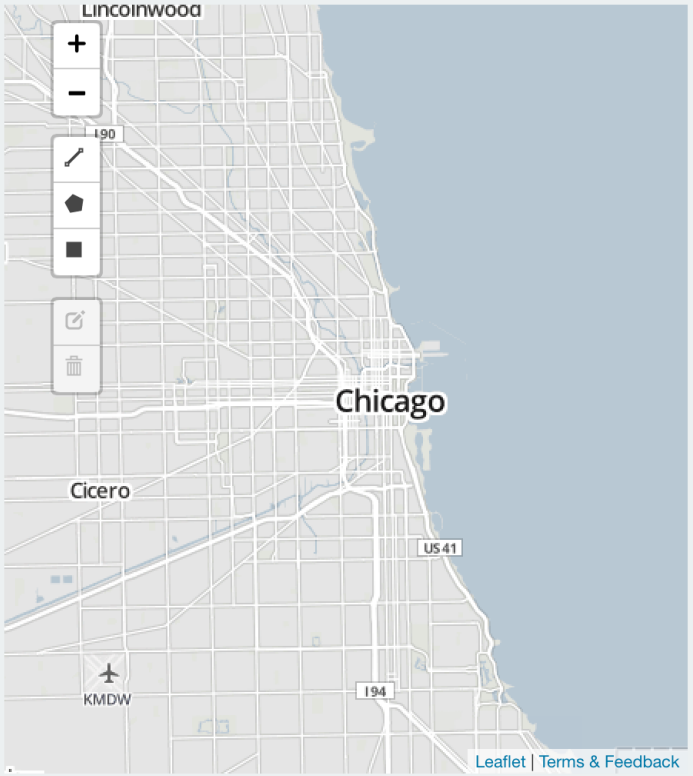
URBAN CENTER FOR COMPUTATION AND DATA

- Predicting high rates of EMS calls and contributing factors
- Populate a unique building ID across datasets
- **Creating 'neighborhood health index'**
- **Predicting movements in the neighborhood health index**
- **Predicting locations of abandoned buildings / vacant lots**
- Visual recognition of neighborhood improvement or retrograde
- Route optimization for routine city vehicle routes (snow plows, garbage collection, tree trimming) and for emergency routes as well.
- Estimate increases in crime incidents
- Prediction of restaurants which will fail food inspections
- **Estimate economic health ("micro-GNP") of neighborhoods and sub-neighborhoods**
- Generate industrial profiles for neighborhoods
- Financial fraud detection (from city transactions)
- Payment error detection (from city transactions)
- Measuring satisfaction with agencies through social media
- Identify delays in fulfilling 311 service requests
- M/WBE (minority/women owned) companies that attempt to skirt procurement regulations

**(June 2013) From Brett Goldstein, City of Chicago Chief Data Officer and Commissioner, Dept. of Innovation and technology.*

URBAN CENTER FOR COMPUTATION AND DATA

1. Draw an area



2. Select date range and aggregation

Start Date 01/22/2014

End Date 07/21/2014

Aggregate by Month

About

This is a prototype for demonstrating geospatial and time aggregation across multiple Chicago open datasets.

Datasets available

- Crimes - 2001 to present
- CDPH Environmental Complaints
- 311 Service Requests - Graffiti Removal
- 311 Service Requests - Garbage Carts
- 311 Service Requests - Rodent Baiting
- 311 Service Requests - Pot Holes Reported
- 311 Service Requests - Sanitation Code Complaints
- 311 Service Requests - Alley Lights Out
- 311 Service Requests - Street Lights - One Out
- 311 Service Requests - Street Lights - All Out
- Food Inspections
- Building Violations
- Building Permits
- 311 Service Requests - Vacant and Abandoned Buildings Reported

URBAN CENTER FOR COMPUTATION AND DATA

Dataset:

Crimes - 2001 to present

Date range:

7/14/2013 - 7/14/2014

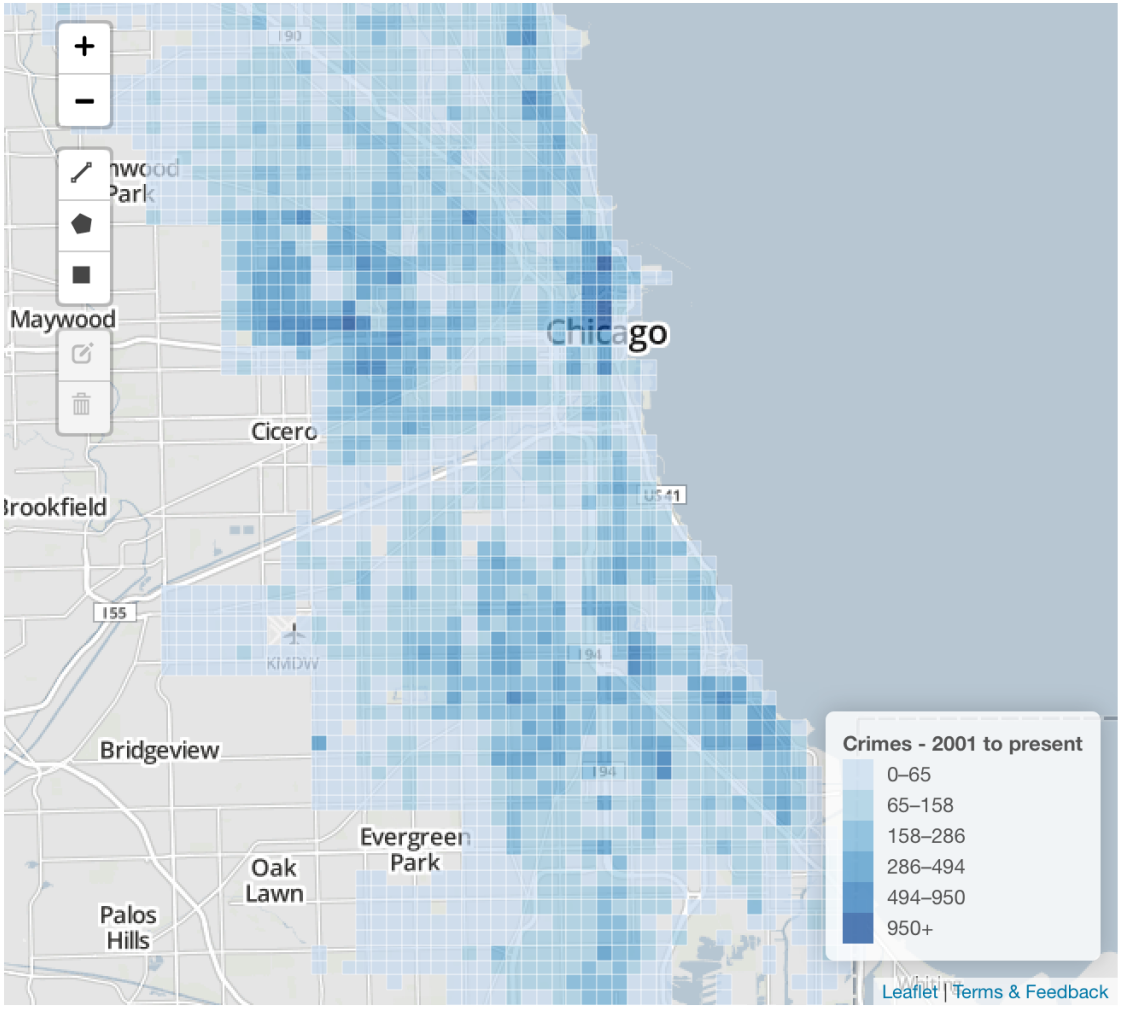
Spatial resolution:

~500m

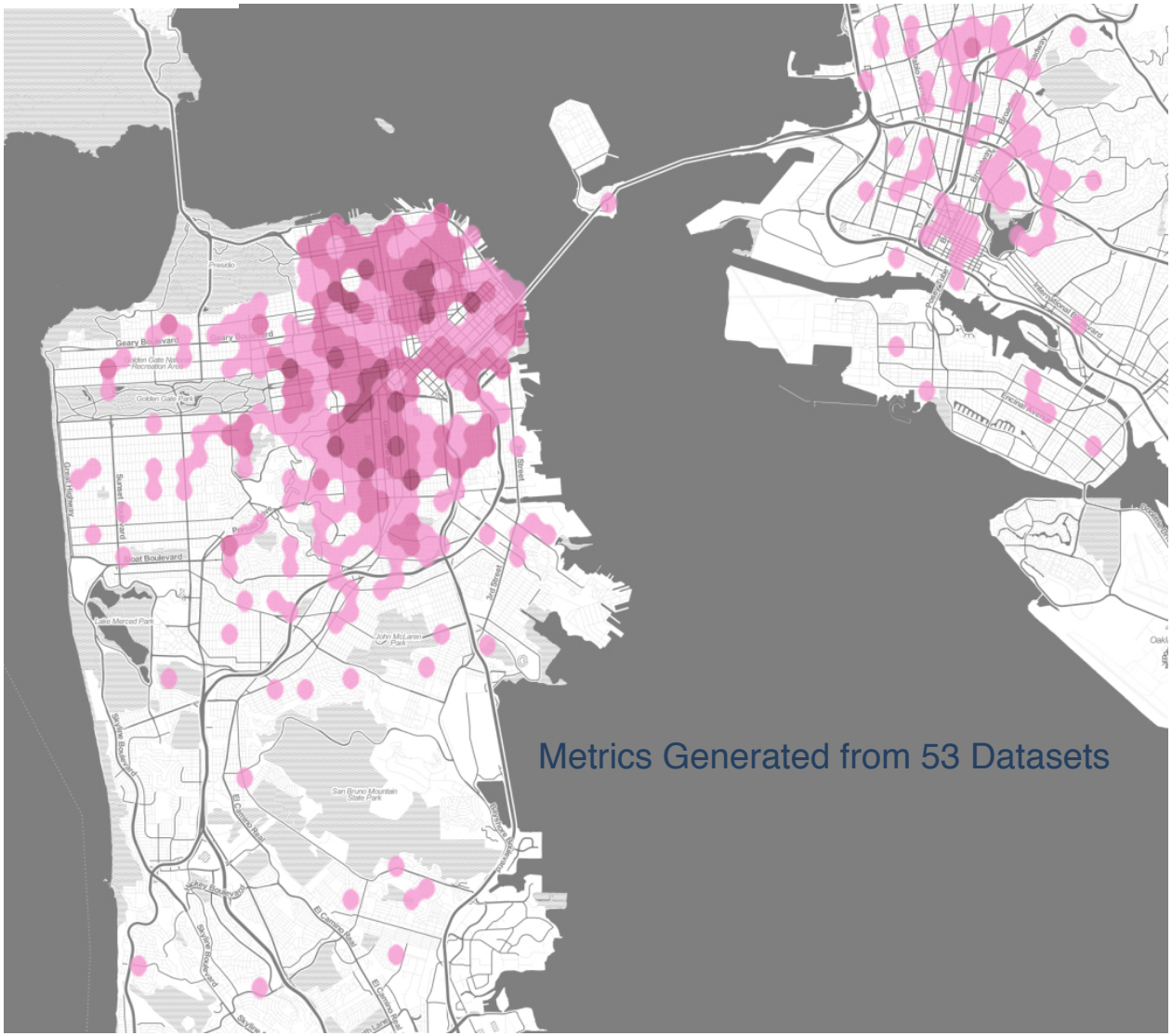
Crimes - 2001 to present

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present, minus the most recent seven days. Data is extracted from the Chicago Police

Show more



URBAN CENTER FOR COMPUTATION AND DATA



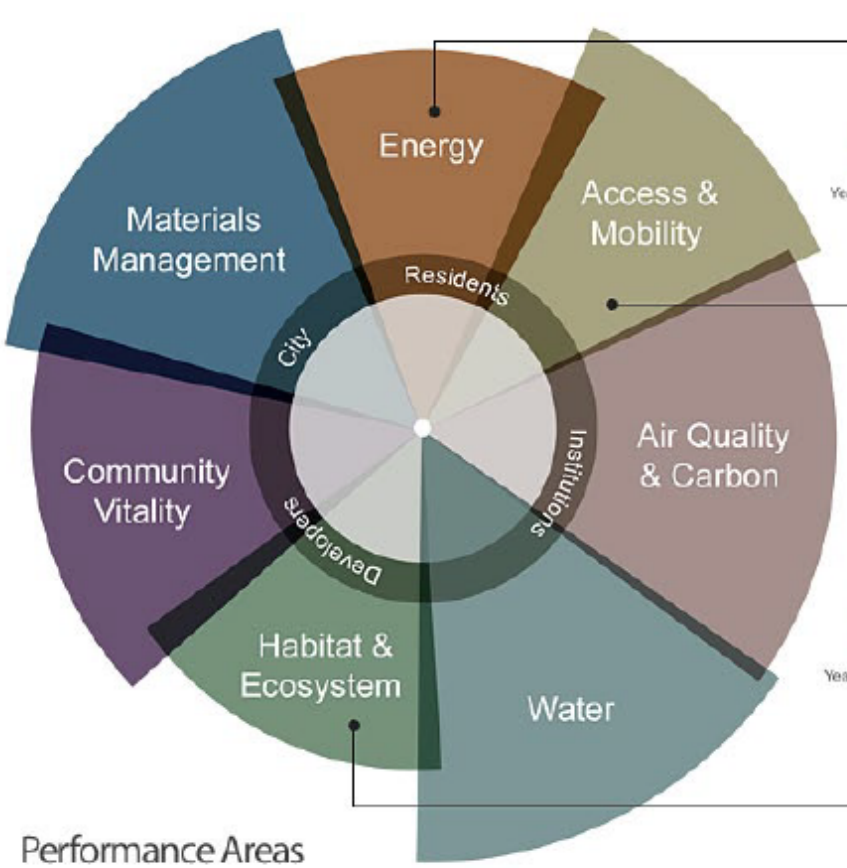
- Built Environment
- Neighborhood Assets
- Housing and Rental Prices
- Building-level Energy Use
- Solar installations
- Renewables generation
- Longitudinal Surveys
- Employment records
- Waste tonnage by block
- Transportation
- Traffic density
- Air quality
- Emissions

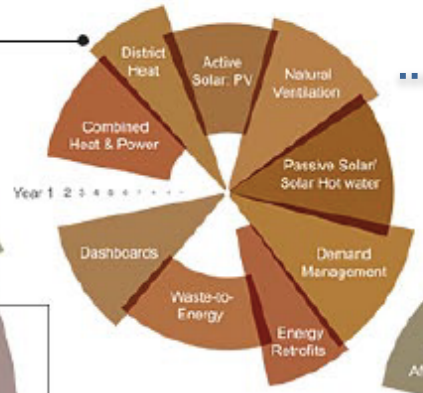
- *Energy*
- *Materials Management*
- *Access and Mobility*
- *Air Quality and Carbon*
- *Water*
- *Habitat and Ecosystem*
- *Community Vitality*

Developing methods to “measure” composite sustainability factors…

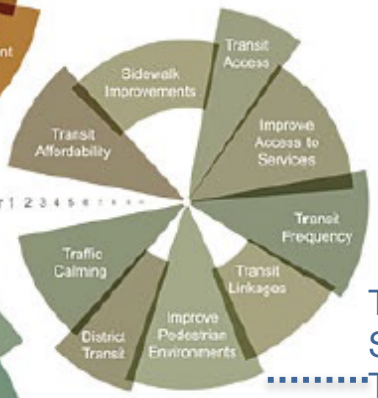
...and to identify “neighborhoods” that are “similar” (control groups for evaluation).

Matt Gee, UChicago
Kate McGee, City of San Francisco

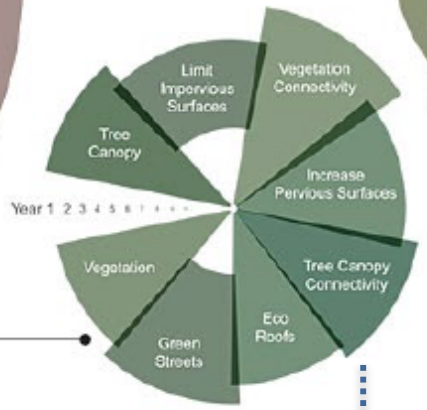




Combined Heat & Power
District Heat
Active Solar PV
Natural Ventilation
Passive Solar Hot Water
Demand Management
Energy Retrofits
Waste-to-Energy
Dashboards



Transit Affordability
Sidewalk Improvements
Transit Access
Improve Access to Services
Transit Frequency
Transit Linkages
Improve Pedestrian Environment
District Transit
Traffic Calming



Tree Canopy
Limit Impervious Surfaces
Vegetation Connectivity
Increase Pervious Surfaces
Tree Canopy Connectivity
Eco Roofs
Green Streets
Vegetation

Source: Central Corridor Eco-District: An Introduction
(March 2012, City of San Francisco)

Negative Impacts

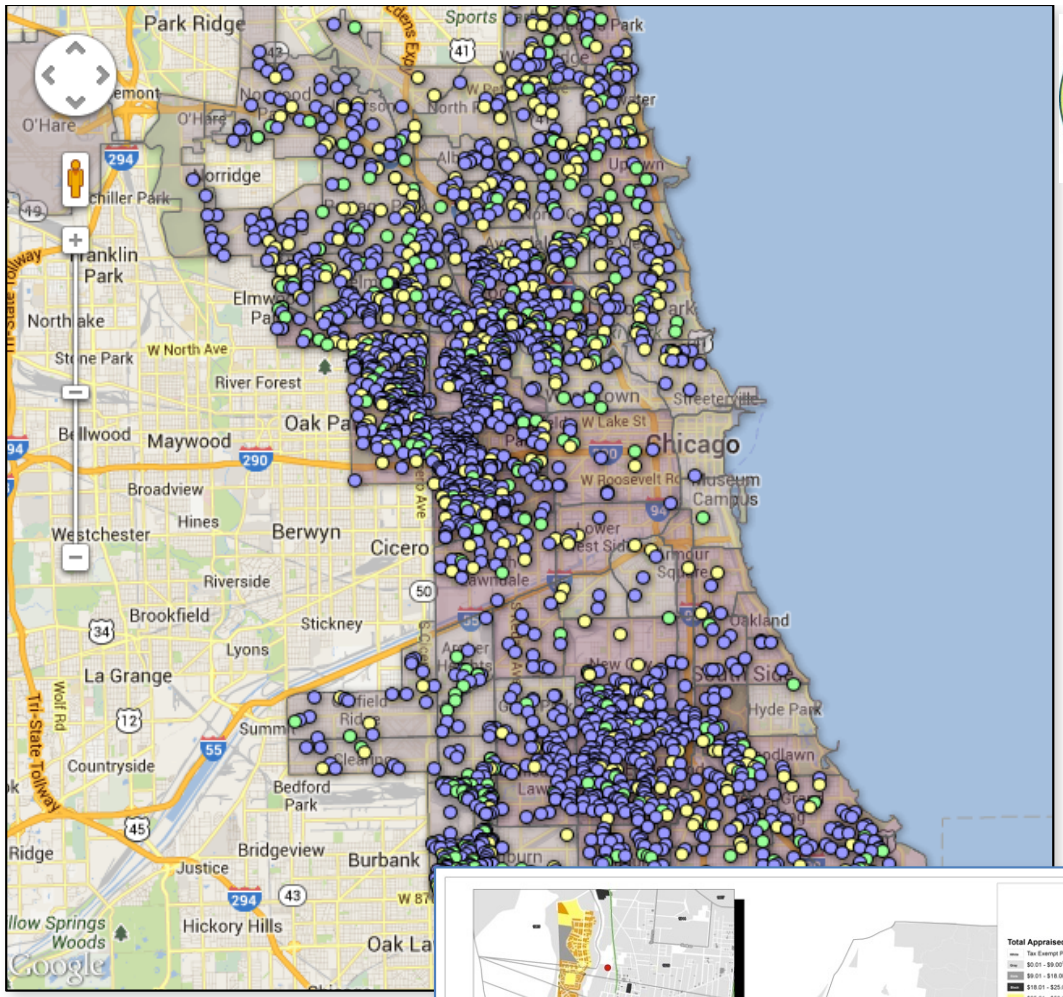
- Lowers property value
- Crime
- Lost taxes

Legal Obstacles

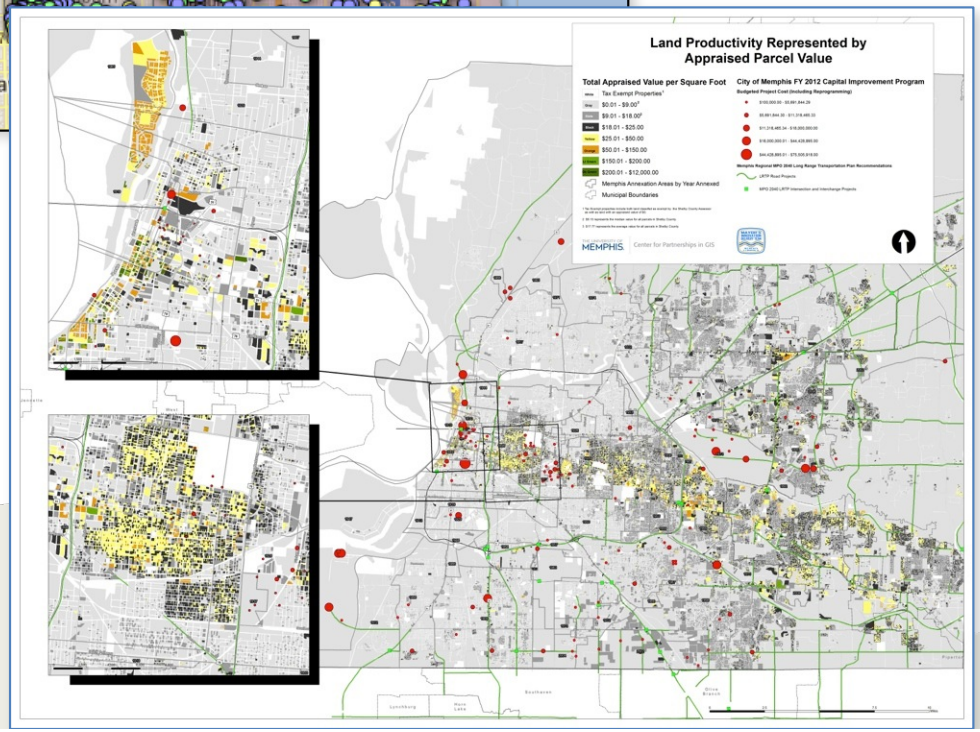
- Unknown owner
- Unpaid bills
- Contaminated land

Guiding Indicators

- Housing stability
- Affordability
- Vacancy



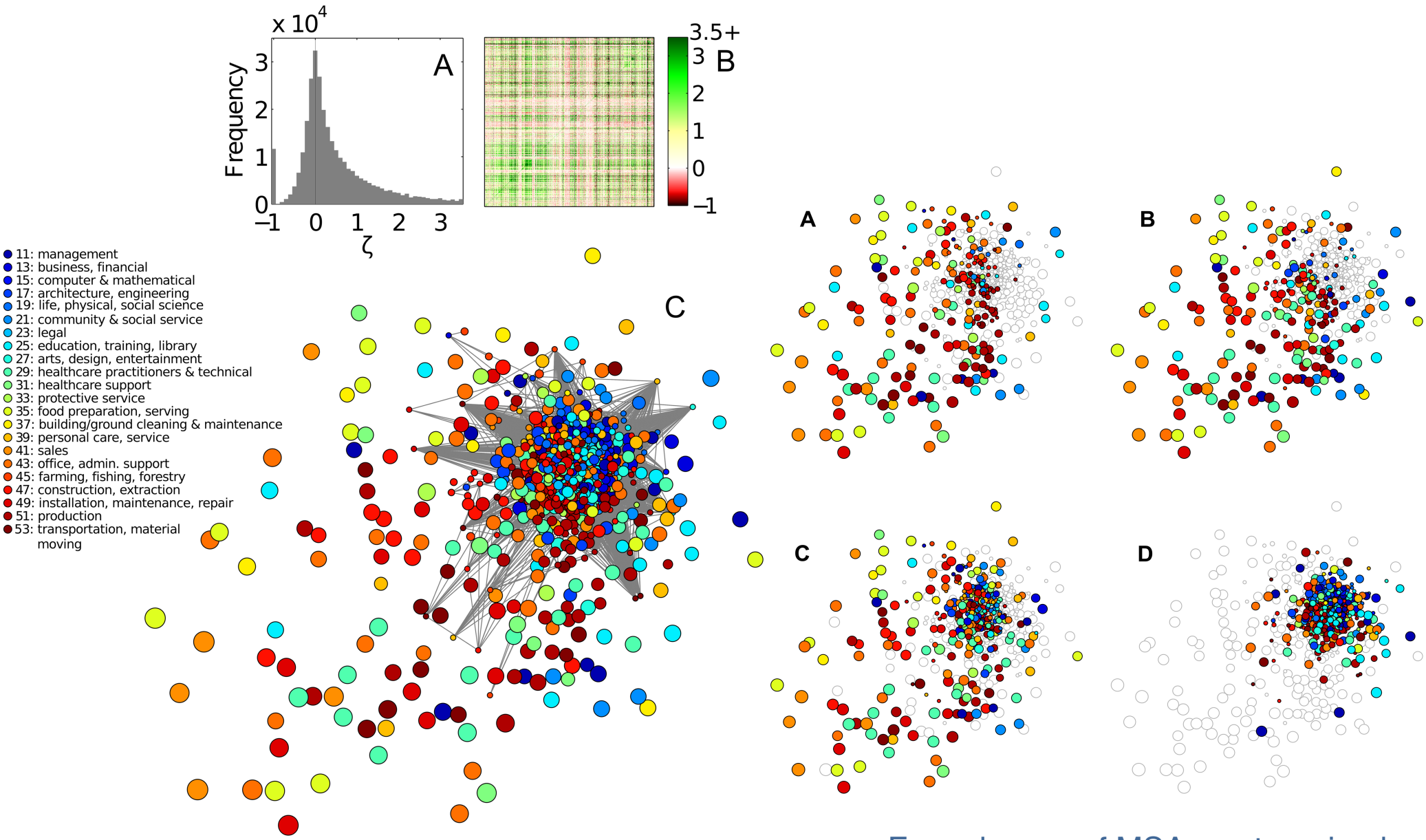
Chicago – optimize investments in vacant property.



Memphis– optimize investments in infrastructure.

URBAN CENTER FOR COMPUTATION AND DATA

Matt Gee, UChicago



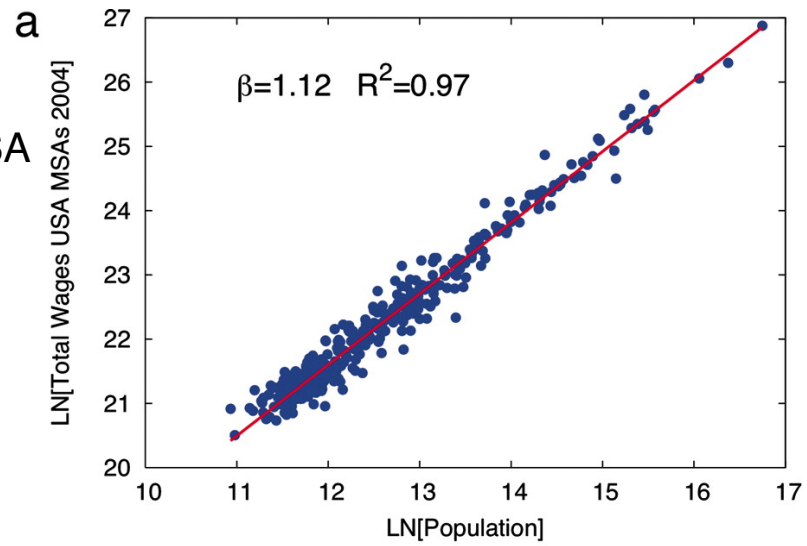
Four classes of MSAs, categorized by per capita GDP

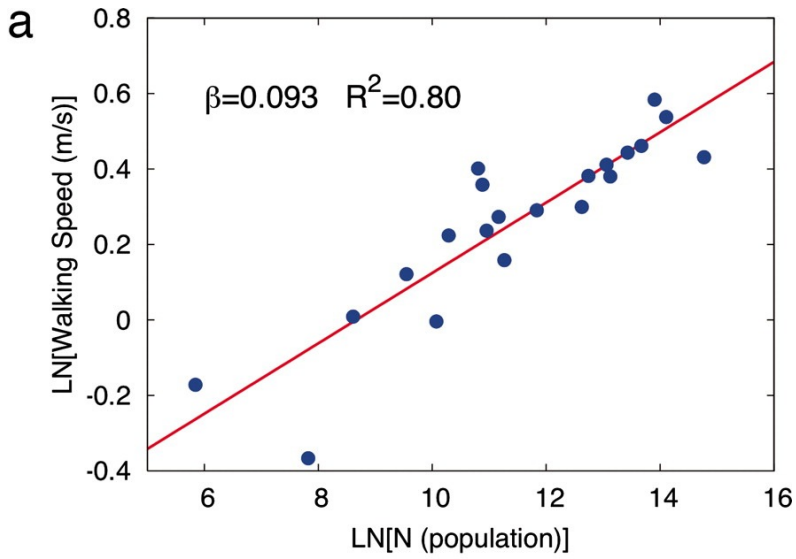
Muneepeerakul, R., J. Lobo, S. T. Shutters, A. Goméz-Liévano and M. R. Qubbaj. 2013. Urban economies and occupation space: Can they get "there" from "here"?. PLOS One 8(9):e73676. DOI: 10.1371/journal.pone.0073676.

URBAN CENTER FOR COMPUTATION AND DATA

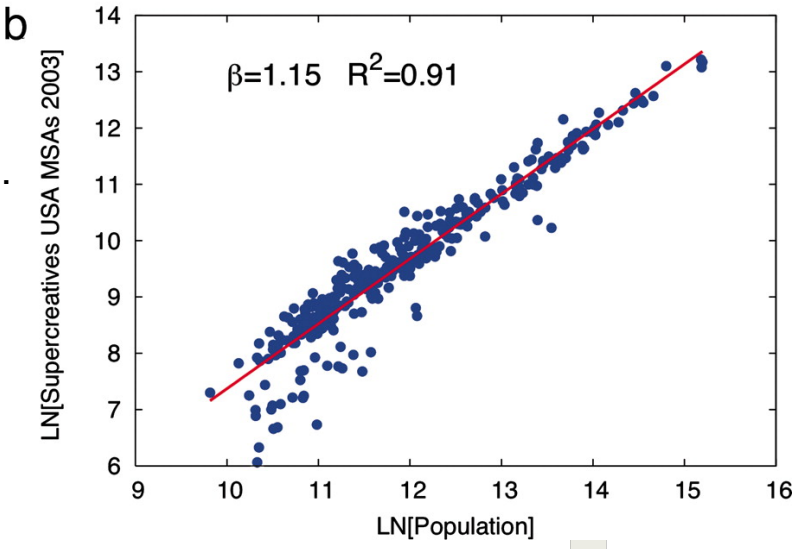
The pace of urban life increases with city size in contrast to the pace of biological life, which decreases with organism size.

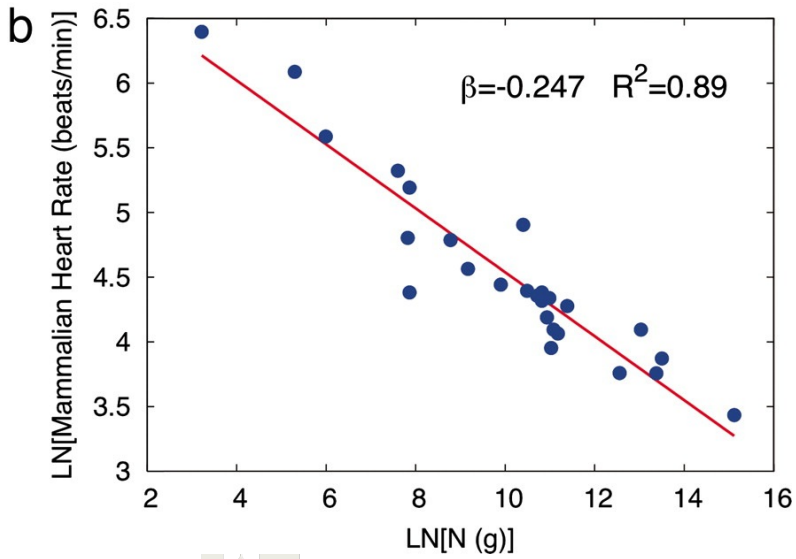
Wages vs. MSA Population (2004)





"Supercreative employment" vs. MSA Population (2003)





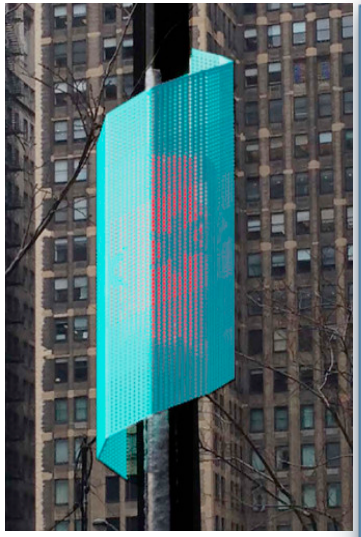
Growth, innovation, scaling, and the pace of life in cities

Bettencourt L M A et al. PNAS 2007;104:7301-7306

©2007 by National Academy of Sciences

URBAN CENTER FOR COMPUTATION AND DATA

University of Chicago,
Argonne National Laboratory,
The School of the Art Institute of Chicago,
and The City of Chicago



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.

Academic Partners:

Arizona State University
Clemson University
MIT
Illinois Institute of Technology
Northwestern University
Northern Illinois University
Purdue University
University of Illinois-Chicago
University of Illinois Urbana-Champaign
University of Notre Dame

Industry Collaborators

Cisco
Intel
Motorola Solutions
Qualcomm
Schneider Electric
Zebra Technologies

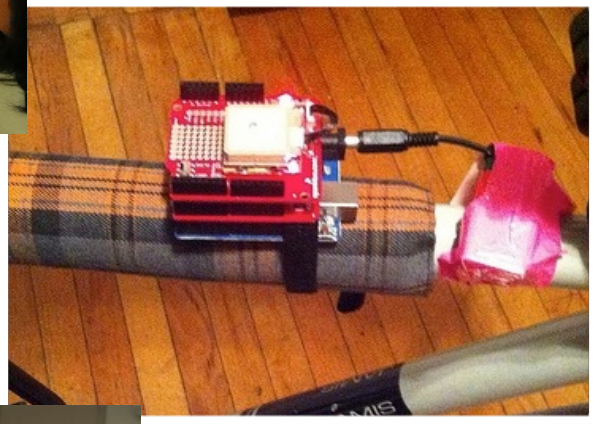
Why is asthma more prevalent in some parts of the City? Is there a link to infrastructure? Environment?

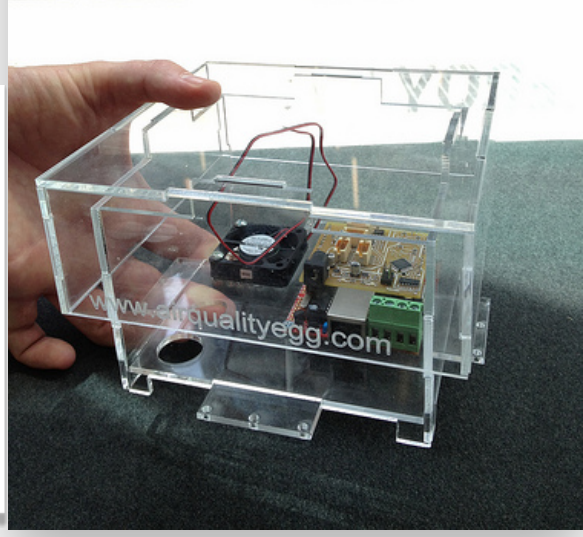
How is air quality affected by traffic flow and weather?

Would real-time pedestrian and vehicle flow data enable safer, more efficient streets?

Can place-based educational games keep teens interested in school?

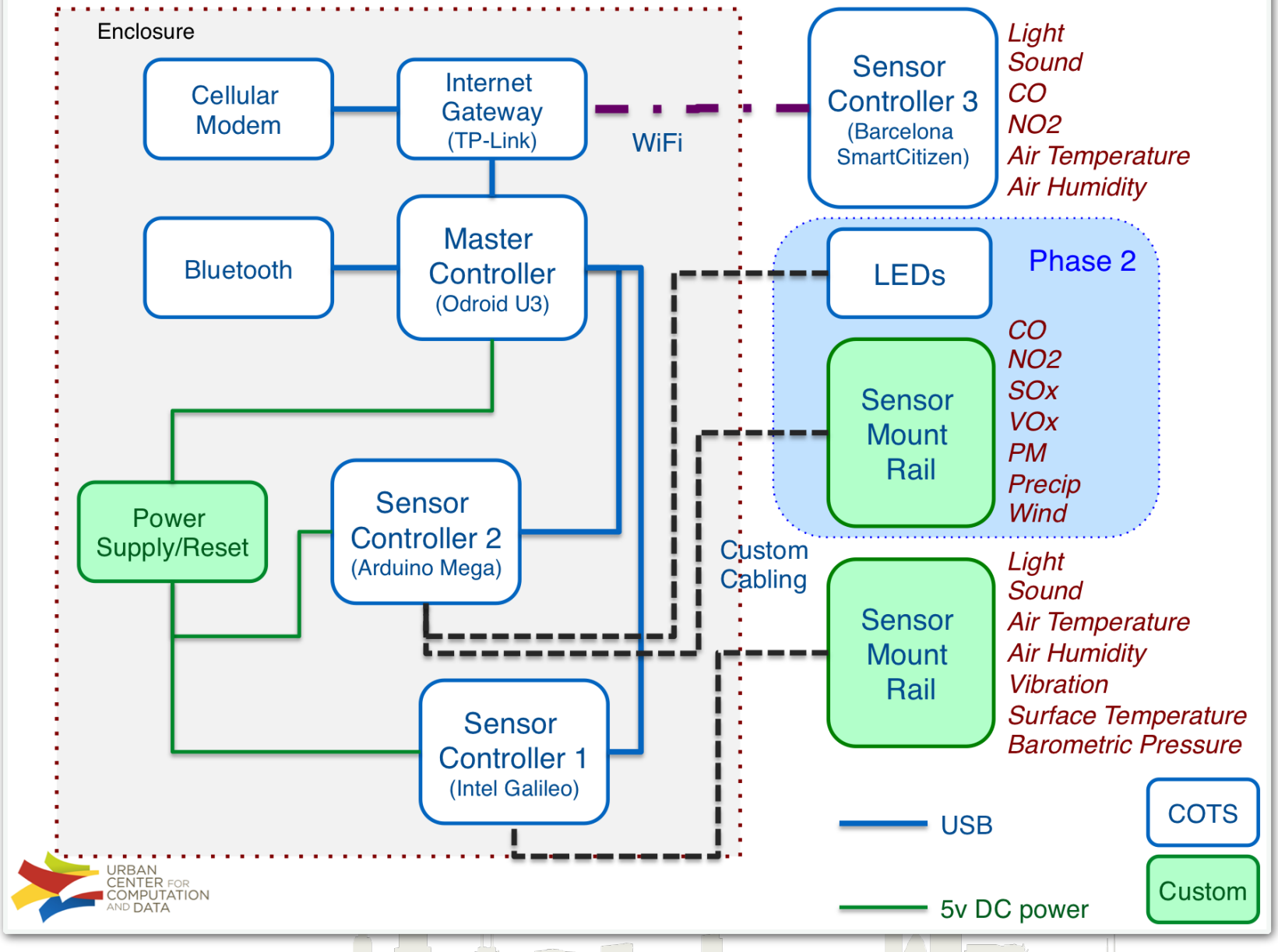
URBAN CENTER FOR COMPUTATION AND DATA





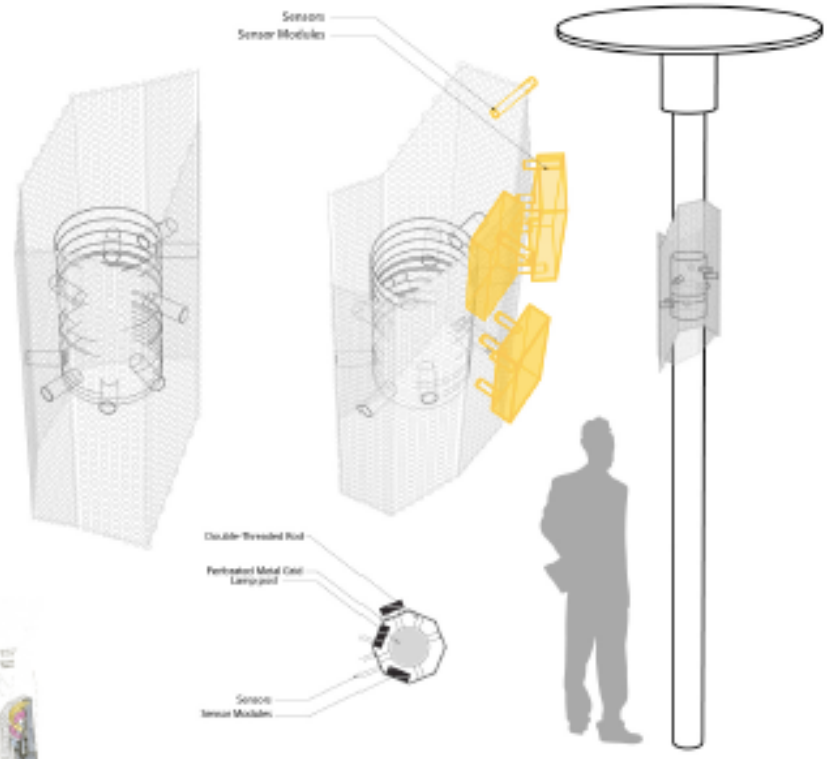
URBAN CENTER FOR COMPUTATION AND DATA

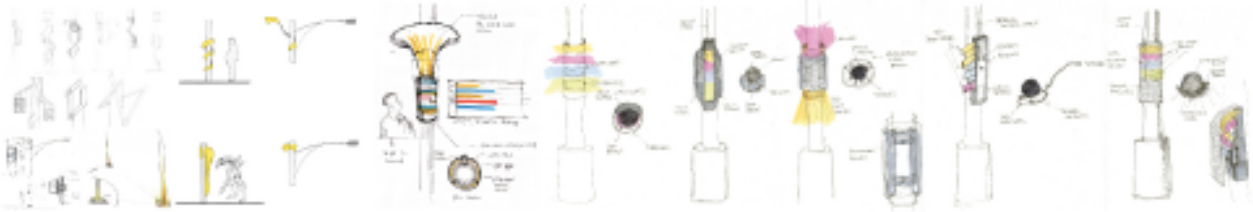
Confidential – For Partner Discussions Only



URBAN CENTER FOR COMPUTATION AND DATA

The School of the Art Institute of Chicago and the University of Chicago have developed a customizable design to *engage* and *inform* citizens. Training workshops with neighborhood youth began in July 2014.

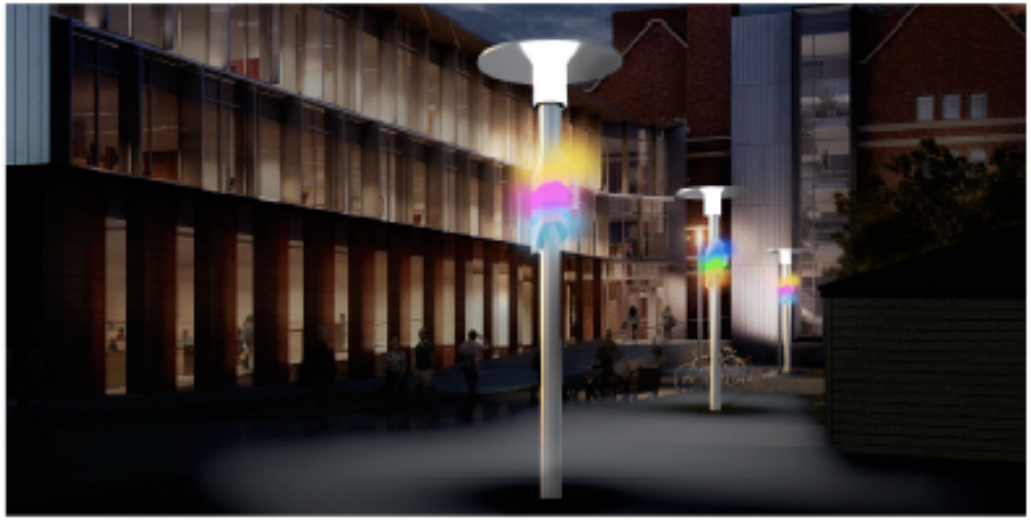




EARLY PROTOTYPE

SCREEN PATTERN PROTOTYPES

RENDERED CONCEPT



URBAN SENSING NODE
UNIVERSITY OF CHICAGO

Douglas Pancoast (Dec 2013)

ARCH/INARCH 6112
NODES/NETWORKS/INTERACTIVITY
FALL 2013

SAIC

UrbanCCD is a partner in City of Big Data a public exhibit of the Chicago Architecture Foundation (May 2014 – Sep 2015). The exhibit includes a 50% scale operational AoT node (insets).

URBAN CENTER FOR COMPUTATION AND DATA

SCHMIDT FAMILY FOUNDATION

The Eric & Wendy Schmidt Data Science for Social Good Summer Fellowship 2013

Improving Communities through Data-Driven Land Banks

THE UNIVERSITY OF CHICAGO

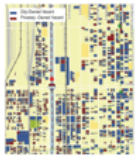
Sophia Alice, Evan Misshula, Skyler Whorton, and Tom Plagge

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 decay. The Cook County Land Bank is a new agency tasked with putting them back to use.

The Problem

The Land Bank has several tools at its disposal to redevelop abandoned buildings. They can clear title, forgive back taxes, combine parcels, and hold land tax-exempt until demand for the property recovers. However, since there are over 100,000 vacant residential addresses in Cook County (2012 Q4, HUD/USPS), the Land Bank must be selective in its acquisitions.



Vacant properties in a particularly distressed portion of Cook County (CMAP Green Healthy Neighborhoods plan).

To help the agency determine which distressed properties to acquire – and what to do with them - we are developing prototype mapping and analytical tools that:

- Estimate housing demand and affordability in the community.
- Assess neighborhood stability.
- Identify nuisance properties.
- Evaluate the economic impact of fixing up and selling the property, or demolishing it

The agency will use these community and property scores to help guide their strategy and operations as they begin acquiring properties next year.

Data

- We used housing market data provided by the Institute of Housing Studies at DePaul University and Cook County. These data include real estate transactions, mo property assessments, foreclosures and propert boundaries throughout the county, tagged by g and property type.

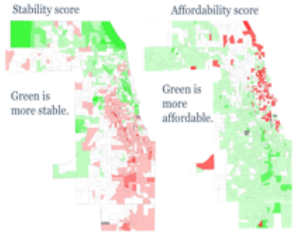
- We also incorporated 311 and crime report dat City of Chicago, vacancy data from HUD/USPS data from HMDA, and economic data from the

This work was done during

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 & Winston (2010), and depends on property values (V), transaction volume (V_1,V_2) mortgages to owner occupants (M), and the prevalence of high-cost lending (H). The affordability score (A) is based on income (I) and median property sale price (P) in each census tract.

$$S = 0.6V + 0.3V_1 + 0.1V_2 + 0.2M - 0.4H \qquad A = \int_{P/4}^{\infty} f(I)dI$$



Property Scores

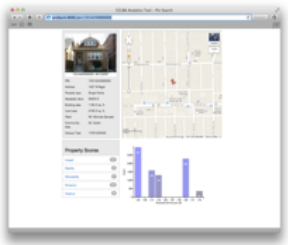
We are also developing scores for individual properties based upon their nuisance values to their neighbors and their economic impact on the community. The former is based on 311 and crime reports.

For the latter, we are using a hedonic pricing model that takes into account 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 basic demographic and economic characteristics of its community, each foreclosure that occurs within 1/8 mile of a property has approximately a -2% effect on its price.

Web Application

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



Conclusions

The Cook County Land Bank plans to attack the problem of abandoned buildings in a nimble, data-driven way. Having all of the relevant information in one place will save the staff time, and summarizing the information in a handful of meaningful, digestible scores 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 to the Land Bank will be prototypes, intended primarily to guide the board's strategic discussions. As the agency moves into operation, the indicators and scores can be calibrated against real results, and can be revised to reflect changing strategies or market conditions.

The New York Times Technology | Personal Tech | Business Day

Bits

JULY 25, 2013, 7:47 AM | Comment

A Summer of Data Hacking Social Problems

By STEVE LOHR

FACEBOOK
TWITTER
GOOGLE+
SAVE
E-MAIL
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.

Robert Kozloff

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.

2013 Fellows

WORLD BANK GROUP

Pecan Street Inc.

2014 Data Science for Social Good Partners

Sarah Abraham
Statistics
University of Michigan

Julius Adebayo
Engineering
MIT

Everaldo Aguiar
Computer Science
University of Notre Dame

Jeff Alstott
Psychiatry
University of Cambridge

Nasir Bhanpuri
Biomedical Engineering
University of Southern California

Cindy Chen
Engineering
Purdue University

Matt Conway
Geography
University of California, Santa Barbara

Nick Eng
Mathematics
University of Pennsylvania

Dylan Fitzpatrick
Computer Science
Carnegie Mellon University

Ben Green
Mathematics
Yale University

Chris Bopp
Computer Science
University of Colorado

Joe Brew
Public Health
University of Copenhagen

Nadya Calderon
Visual Analytics
Simon Fraser University

Scott Cambo
Computer Science
Cornell University

Alejandra Caro
Economics
Carnegie Mellon University

Matthew Heston
Computer Science
Illinois Institute of Technology

Madian Khabsa
Computer Science
Pennsylvania State University

Vanessa Ko
Political Science
McGill University

Himabindu Lakkaraju
Computer Science
Stanford University

Andrew Landgraf
Statistics
Ohio State University

Peter Landwehr
Computation
Carnegie Mellon University

Christopher Lazarus
Applied Mathematics
Instituto Tecnológico Autónomo de México

Jeff Lockhart
Computer Science
Fordham University

Alex Loewi
Public Policy
Carnegie Mellon University

Subhabrata Majumdar
Statistics
University of Minnesota

Miguel Perez
Architecture
School of the Art Institute of Chicago

Carlos Petricioli
Computer Science
Instituto Tecnológico Autónomo de México

Eric Potash
Mathematics
Northwestern University

Layla Pournajaf
Computer Science
Emory University

Andrew Reece
Psychology
Harvard University

Robert Manduca
Public Policy
MIT

Isaac McCreery
Mathematics
Oberlin College

David Miller
Psychology
Northwestern University

Philip Ngo
Computer Science
Harvard University

Diana Palsetia
Computer Science
Northwestern University

James Savage
Economics
University of Melbourne

Tracy Schifeling
Statistics
Duke University

Carl Shan
Statistics
University of California, Berkeley

Raphael Stern
Engineering
University of Illinois

Stephen Suffian
Computer Science
Villanova University

Sarah Tan
Statistics
Cornell University

Misha Teplitskiy
Sociology
University of Chicago

Sabina Tomkins
Computer Science
New York University

Rafael Turner
Computer Science
University of Chicago

Vrushank Vora
Mathematics
University of Chicago

Ellery Wulczyn
Computer Science
Stanford University

Zhou Ye
Computer Science
Johns Hopkins University

Sam Zhang
Computer Science
Swarthmore College

2014 Data Science for Social Good Fellows

URBAN CENTER FOR COMPUTATION AND DATA

Argonne NATIONAL LABORATORY

THE UNIVERSITY OF CHICAGO

Computation Institute

www. UrbanCCD .org

C @ ANL.GOV